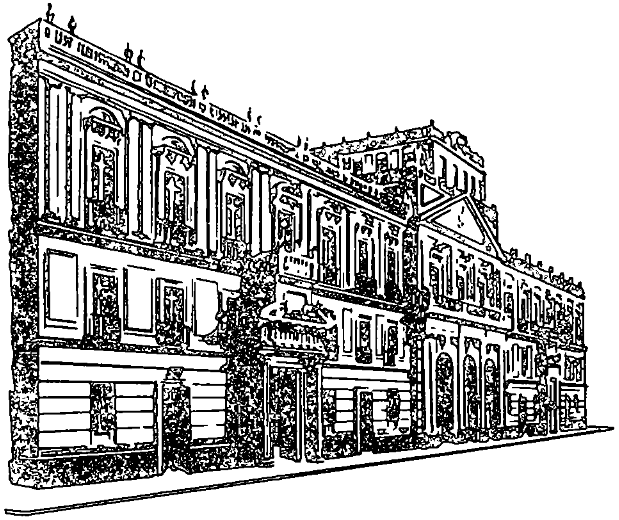


INSCRIPCIONES

CENTRO DE EDUCACION CONTINUA DE LA DIVISION DE ESTUDIOS SUPERIORES DE LA FACULTAD DE INGENIERIA, U. N. A. M.

Palacio de Minería Calle de Tacuba No. 5
México 1, D. F.

Horario de oficinas:
Lunes a viernes de 9 a 18 h.

Cuota de inscripción $ 4,500.00

La cuota de inscripción incluye:

- una carpeta con las notas de los profesores
- bibliografía sobre el tema
- tiempo de computación
- servicio de cafetería
- comidas

Requisitos

- Pagar la cuota de inscripción o traer oficio de la empresa o institución que ampare su inscripción, a más tardar una semana antes del inicio del curso
- Llenar la solicitud de inscripción

Para mayores informes hablar a los teléfonos

521-40-20 521-73-35 512-31-23

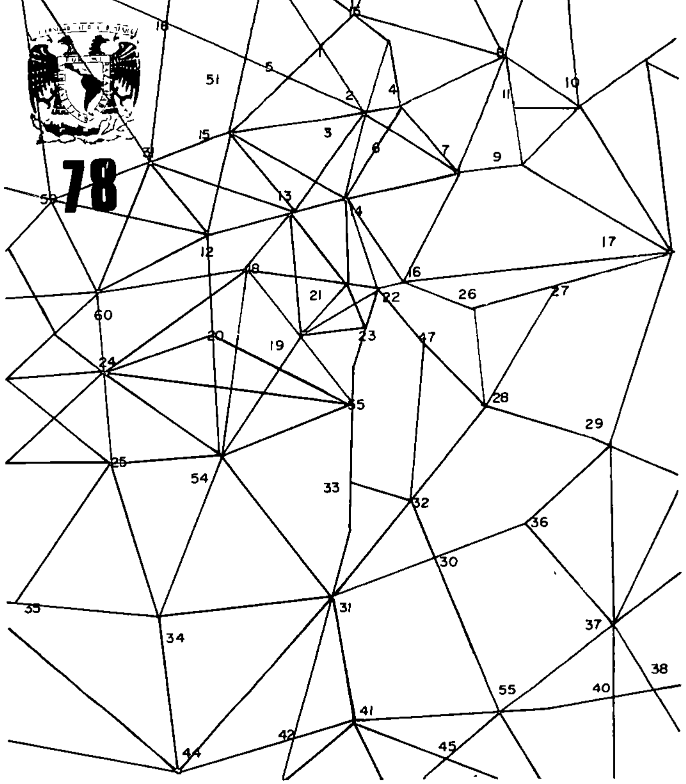
CONSTANCIA DE ASISTENCIA

Las autoridades de la Facultad de Ingeniería de la U.N.A.M., otorgarán una constancia de asistencia a los participantes que concurran regularmente y que realicen los trabajos que se les asignen durante el curso.

CIRCULA LIBRE DE PORTE
POR VIA DE SUPERFICIE
Y DENTRO DEL TERRITORIO NAL.
ART. 17 LEY ORGANICA DE LA U N A M

centro de educación continua
división de estudios superiores
facultad de ingeniería, u n a m

Palacio de Minería
Calle de Tacuba No. 5
México 1, D.F.



MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE

CURSO INTENSIVO TEORICO-PRACTICO

DURACION: 45 h
FECHAS: del 23 al 28 de enero
HORARIO: lunes a viernes de 9 a 13 y de 14 a 18 h
sábado de 10 a 14 h

Coordinadores: Ing. Agustín Paulin y
Arq. Alejandro Villanueva Egan.

En colaboración con la Sección de Planeación, DESFI, UNAM.

centro de educación continua
división de estudios superiores
facultad de ingeniería, u n a m

ANTECEDENTES

Los modelos de simulación representan actualmente una herramienta indispensable para el planificador en el análisis de los efectos de programas y proyectos sobre la estructura y dinámica de los sistemas urbanos, dándole así mayor creatividad y racionalidad a la planeación de los asentamientos humanos.

En este curso, el concepto de paquete integral de modelos será analizado y se realizarán pruebas de laboratorio para familiarizar al participante en su manejo y evaluación.

OBJETIVOS

Mostrar los fundamentos para el diseño e implementación de este tipo de modelos, así como los criterios para su utilización en el análisis de políticas, programas y proyectos.

Ofrecer una experiencia práctica en el manejo de modelos urbanos calibrados en México.

A QUIEN VA DIRIGIDO

A profesionistas de equipos interdisciplinarios encargados de la planificación de áreas metropolitanas y de proyectos específicos de transporte y uso del suelo.

TEMARIO

1. PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE
2. MODELOS DE SIMULACION DEL USO DEL SUELO
3. MODELOS DE SIMULACION DEL TRANSPORTE
4. PAQUETES INTEGRALES DE MODELOS
5. METODOS PARA LA CALIBRACION DE MODELOS
6. UTILIZACION DE MODELOS EN EL ANALISIS DE POLITICAS, PROGRAMAS Y PROYECTOS
7. APLICACIONES EN MEXICO Y LATINOAMERICA
8. PRUEBAS DE LABORATORIO CON MODELOS Y BANCOS DE DATOS DE LA ZONA METROPOLITANA DE LA CIUDAD DE MEXICO

CONFERENCISTAS INVITADOS

DR. STEPHEN H. PUTMAN
DR. FREDERICK DUCCA JR.

Planning Sciences Simulation Laboratory.
Departamento de Planeación Urbana y Regional.
Universidad de Pennsylvania.

PROFESORES

ARQ. ALEJANDRO VILLANUEVA EGAN
M. EN C. JOSE LUIS SOBERANES
ING. AGUSTIN PAULIN

LABORATORIO

MAT. MARIO RODRIGUEZ GREEN
ARQ. BRAULIO HORNEDO
ARQ. OLIVERIO GONZALEZ

MESA REDONDA

Participantes:
SUB-SECRETARIA DE ASENTAMIENTOS HUMANOS
DEPARTAMENTO DEL DISTRITO FEDERAL
INSTITUTO AURIS

NOTA:

Para las prácticas de laboratorio se contará con el banco de datos y la programoteca del Laboratorio de Planeación, DESFI, y la computadora del Centro de Cómputo, UNAM.

CUPO LIMITADO

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

A LOS ASISTENTES A LOS CURSOS DEL CENTRO DE EDUCACION CONTINUA

Las autoridades de la Facultad de Ingeniería, por conducto del Jefe del Centro de Educación Continua, otorgan una constancia de asistencia a quienes cumplan con los requisitos establecidos para cada curso. Las personas que deseen que aparezca su título profesional precediendo a su nombre en la constancia, deberán entregar copia del mismo o de su cédula a más tardar el SEGUNDO DIA de clases, en las oficinas del Centro con la señorita encargada de inscripciones.

El control de asistencia se llevará a cabo a través de la persona encargada de entregar las notas del curso. Las inasistencias serán computadas por las autoridades del Centro, con el fin de entregarle constancia solamente a los alumnos que tengan un mínimo del 80% de asistencia.

Se recomienda a los asistentes participar activamente con sus ideas y experiencias, pues los cursos que ofrece el Centro están planeados para que los profesores expongan una tesis, pero sobre todo, para que coordinen las opiniones de todos los interesados constituyendo verdaderos seminarios.

Es muy importante que todos los asistentes llenen y entregen su hoja de inscripción al inicio del curso. Las personas comisionadas por alguna institución deberán pasar a inscribirse en las oficinas del Centro en la misma forma que los demás asistentes, entregando el oficio respectivo.

Con objeto de mejorar los servicios que el Centro de Educación Continua ofrece, al final del curso se hará una evaluación a tráves de un cuestionario diseñado para emitir juicios anónimos por parte de los asistentes.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE INGENIERIA
DIVISION DE ESTUDIOS SUPERIORES
CENTRO DE EDUCACION CONTINUA
DIRECTORIO GENERAL

REGISTRO DE ASISTENTES Y PROFESORES.

OMBRE DEL CURSO: ____________________ FOLIO ☐☐☐☐☐ (1 – 5) CLAVE ASOC. ☐☐ (6 7)

☐☐☐☐☐☒ (8 – 13) ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ (14 – 41)
8 13 14 NOMBRE(S) APELLIDO PATERNO APELLIDO MATERNO 41

REG. FED. CAUS. ☐☐☐☐☐☐☐☐☐☐ (42 – 51) CED. PROF. ☐☐☐☐☐☐☐ (52 – 58)

TEL. PARTICULAR ☐☐☐☐☐☐☐ (59 – 65) TEL. OFICINA ☐☐☐☐☐☐☐ (66 – 72) EXTENSION ☐☐☐☐ (73 – 76)

MARQUE CON UNA CRUZ

ASISTENTE ☐ PROFESOR ☐ ☐ 77 [1] 80

☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐
8 DOMICILIO PARTICULAR (CALLE. NUMERO Y No. INTERIOR) 41

☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Z.P. ☐☐
42 COLONIA 71 72 73

____________________ ☐☐
ESTADO 74 75

____________________ ☐☐ ____________________ ☐☐ [2]
TITULO PROFESIONAL 76 77 ESPECIALIDAD 78 79 80

☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐
8 DOMICILIO DE OFICINA (CALLE, NUMERO Y No. INTERIOR) 41

☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Z.P. ☐☐
42 COLONIA 71 72 73

____________________ ☐☐ [3]
ESTADO. 74 75 80

ASOCIACIONES A LAS QUE PERTENECE.

____________________ ☐☐ ____________________ ☐☐

____________________ ☐☐ ____________________ ☐☐

FCEC-01

MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE

FECHA	DURACION	TEMA	PROFESOR
Enero 23	9 a 12 h	Planeación integral del uso del suelo y el transporte. Panorama conceptual. Utilización de modelos.	Dr. Stephen H. Putman
	12 a 13 h	Descripción de la Programoteca y los trabajos que se realizan en el Planning Sciences Simulation Laboratory.	Dr. Frederick Ducca
	14 a 16 h	Descripción de los modelos y las bases de datos que serán usados en el curso	Arq. Alejandro Villanueva Egan
	16 a 18 h	Mesa Redonda: Simulación Uso del Suelo	
Enero 24	9 a 11 a.m.	Modelos de Simulación del Uso del Suelo. Evolución. Modelos de Interacción Espacial.	Dr. Stephen H. Putman
	11 a 13 h	Aplicación de modelos de Usos del Suelo (comercial, industrial).	Dr. Frederick Ducca
	14 a 16 h	Fundamentos teóricos de los Modelos de Uso del Suelo. Teoría de William Alonso. Modelo de Wingo Teoría de Richard Muth Modelo de Herbert-Stevens	M. en C. José Luis Soberanes
	16 a 18 h	Laborátorio	
Enero 25	9 a 11 a.m.	DRAM-Calibraciones y Resultados	Dr. Stephen H. Putman
	11 a 13 h	Calibraciones	Dr. Frederick Ducca
	14 a 16 h	Modelos de Transporte. Modelos de generación, distribución, elección y asignación. Modelos probalísticos y determinísticos.	

Modelos de Simulación para la Planeación Integral del Uso del Suelo y el Transporte

Fecha	Duración	Tema	Profesor
Enero 25	16 a 18 h	Laboratorio	Arq. Alejandro Villanueva Ega Arq. Braulio Hornedo
Enero 26	9 a 11 h	Comparaciones DRAM-EMPIRIC Paquetes integrales de modelos del uso del suelo y el transporte	Dr. Stephen H. Putman
	11 a 13 h	Laboratorio	Dr. Frederick Ducca
	14 a 18 h	Laboratorio. DRAM	Dr. Frederick Ducca, Arq. Alejandro Villanueva Egan Arq. Braulio Hornedo Mat. Mario Rodríguez Green
Enero 27	9 a 11 a.m.	Utilización de Modelos en el Análisis de Políticas, Programas y Proyectos.	Dr. Stephen H. Putman
	11 a 13 h	Laboratorio	Dr. Frederick Ducca
	14 a 16 h	Prácticas de Laboratorio.	Dr. Frederick Ducca, Arq. Alejandro Villanueva Egan, Arq. Braulio Hornedo y Mat. Mario Rodríguez Green
	16 a 18 h	Mesa Redonda. Conclusiones Participantes, DDF, AURIS, SAHOP.	
	18 a 18:30 h	Clausura	

'edcs. 23, I, 78.

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE

JUEGO DE SIMULACION "CLUG"
USO DEL SUELO Y TRANSPORTE
INFORMACION PARA EL JUGADOR

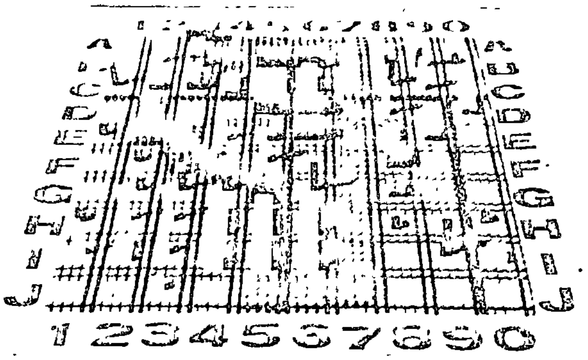
ARQ. ALEJANDRO VILLANUEVA EGAN

ENERO, 1978.

CLUG

****

EL JUEGO DEL USO DEL SUELO EN LA COMUNIDAD.



DESCRIPCION GENERAL:

CLUG ES UN MODELO QUE PRETENDE REPRESENTAR ALGUNOS DE LOS FACTORES FUNDAMENTALES QUE AFECTAN LAS DECISIONES DEL USO DEL SUELO EN UNA COMUNIDAD Y SU MEDIO-AMBIENTE EN TAL FORMA QUE SE SITUA AL JUGADOR EN UNA POSICION EN LA QUE TIENE QUE OPERAR CON LAS - RESTRICCIONES IMPUESTAS POR ESOS FACTORES. AL DESARROLLARSE EL JUEGO, LOS JUGADORES ENCONTRARAN QUE HAN SIDO OMITIDOS FACTORES IMPORTANTES, PERO LA FLEXIBILIDAD EN EL DISEÑO DEL JUEGO LE PERMITE AÑADIR ASPECTOS QUE SEAN SUGERIDOS POR EL USUARIO PARA HACER EL JUEGO MAS APROPIADO A SUS NECESIDADES.ESTA VERSION DE - CLUG PUEDE OPERARSE SOBRE UNA BASE COMPLETAMENTE MANUAL CON SOLAMENTE DOS OPERADORES Y DE 3 A 18 PERSONAS. CLUG ES LO SUFICIENTEMENTE FLEXIBLE Y COMPLEJO PARA PROVEER UN INSTRUMENTO PEDAGOGICO ADECUADO CON ALGUNAS APLICACIONES DENTRO DEL TERRENO PROFESIONAL EN LA PLANEACION. AL MISMO TIEMPO, PROVEE UNA BASE RELATIVAMENTE SIMPLE PARA EL ENTRENAMIENTO DE PROFESIONISTAS EN EL USO E INTERPRETACION DE TECNICAS HEURISTICAS DE JUEGO DE MANERA QUE PUEDAN APRENDER MAS RAPIDAMENTE A USAR OTROS MODELOS. EN RESUMEN,PUEDE DECIRSE QUE EL CLUG ES UN INSTRUMENTO QUE PERMITE AL TOMADOR DE DECISIONES O AL ESTUDIANTE DE PLANEACION URBANA, - PERCIBIR TODOS LOS ASPECTOS DEL FENOMENO URBANO A LA VEZ, EN TAL FORMA QUE PUEDA OFRECER SOLUCIONES MAS COMPREHENSIVAS.

DIVISION DE ESTUDIOS SUPERIORES
FACULTAD DE INGENIERIA, UNAM
SECCION DE PLANEACION

CLUG

JUEGO PARA LA SIMULACION DE SISTEMAS URBANOS

TIPO DE INVERSION	COSTO DE CONSTRUCCION	GANANCIA POR AÑO	NO. DE UNIDADES DE EMPLEADOS NECESARIOS	SALARIOS POR AÑO	PRECIOS TOPE		
					SUPER MERCADO	GRAN ALMACEN	OFICINAS
Gran Industria	96 000	48 000	4	24 000	—	—	4 000
Pequeña Industria	48 000	22 000	2	12 000	—	—	2 000
Supermercado	24 000	*	1	6 000	—	—	1 000
Gran Almacén	24 000	*	1	6 000	—	—	1 000
Oficina	36 000	*	1	6 000	—	—	—
Residencia Sencilla	12 000	6 000	-	—	2 000	1 000	—
Residencia Doble	30 000	12 000	-	—	4 000	2 000	—
Residencia Triple	48 000	18 000	-	—	6 000	3 000	—
Residencia Cuadruple	72 000	24 000	-	—	8 000	4 000	—

El Ingreso por TL, GA y O dependen del número de clientes obtenidos y el precio cargado. Un ingreso bruto de $10 000 $15 000 puede ser supuesto antes de que esas unidades operen.

DIVISION DE ESTUDIOS SUPERIORES
FACULTAD DE INGENIERIA, UNAM
SECCION DE PLANEACION

CLUG

JUEGO PARA LA SIMULACION DE SISTEMAS URBANOS

COSTOS DE TRANSPORTACION POR CUADRA AL AÑO

VIAJES DESDE:	VIAJES HACIA:				
	CENTRAL	OFICINA	EMPLEO	COMPRAS (TL)	COMPRAS (GA)
Industria Pesada (GI)	$4 000	$ 400	-	-	-
Pequeña Industria (PI)	$2 000	$ 200	-	-	-
Tienda Central o Supermercado	-	$ 100	-	-	-
Residencias por unidad (RI)*	-	-	$ 300	$ 200	$ 100

* Para R2, R3 o R4, multiplicar por 2, 3 o 4, respectivamente.

CLUG COMPRA DE TERRENOS

EQUIPO: ______________________________

LOCALIZACION

DEL TERRENO: __________ : __________

OFERTA $ ____________________

CLUG ACUERDOS COMERCIALES

CELEBRADO ENTRE LOS EQUIPOS: ____________________

____

POR CONCEPTO DE: ______________________________

____

____

____________________	____________________
Representante	Representante

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE

DESCRIPCIONES DE MODELOS

ARQ. ALEJANDRO VILLANUEVA EGAN

ENERO, 1978.

Palacio de Minería Calle de Tacuba 5, primer piso. México 1, D. F.

TOMM :

(TIME ORIENTED METROPOLITAN MODEL). 1.964.

I.- TOMM es un Modelo de Simulación diseñado para pronosticar el uso de la tierra, población y empleo, en los Condados de Pittsburgh y Allegheny, Pennsylvania.
Fue formulado originalmente por Lowry; Crecine desarrolló una versión revisada y posteriormente, Teplitz realizó una segunda revisión.
Debido a la estructura del Modelo, este puede adaptarse para ser aplicado a otras áreas de estudio.

II.- ESTRUCTURA CONCEPTUAL DEL MODELO :

II.1.- TEMA DEL MODELO TOMM :

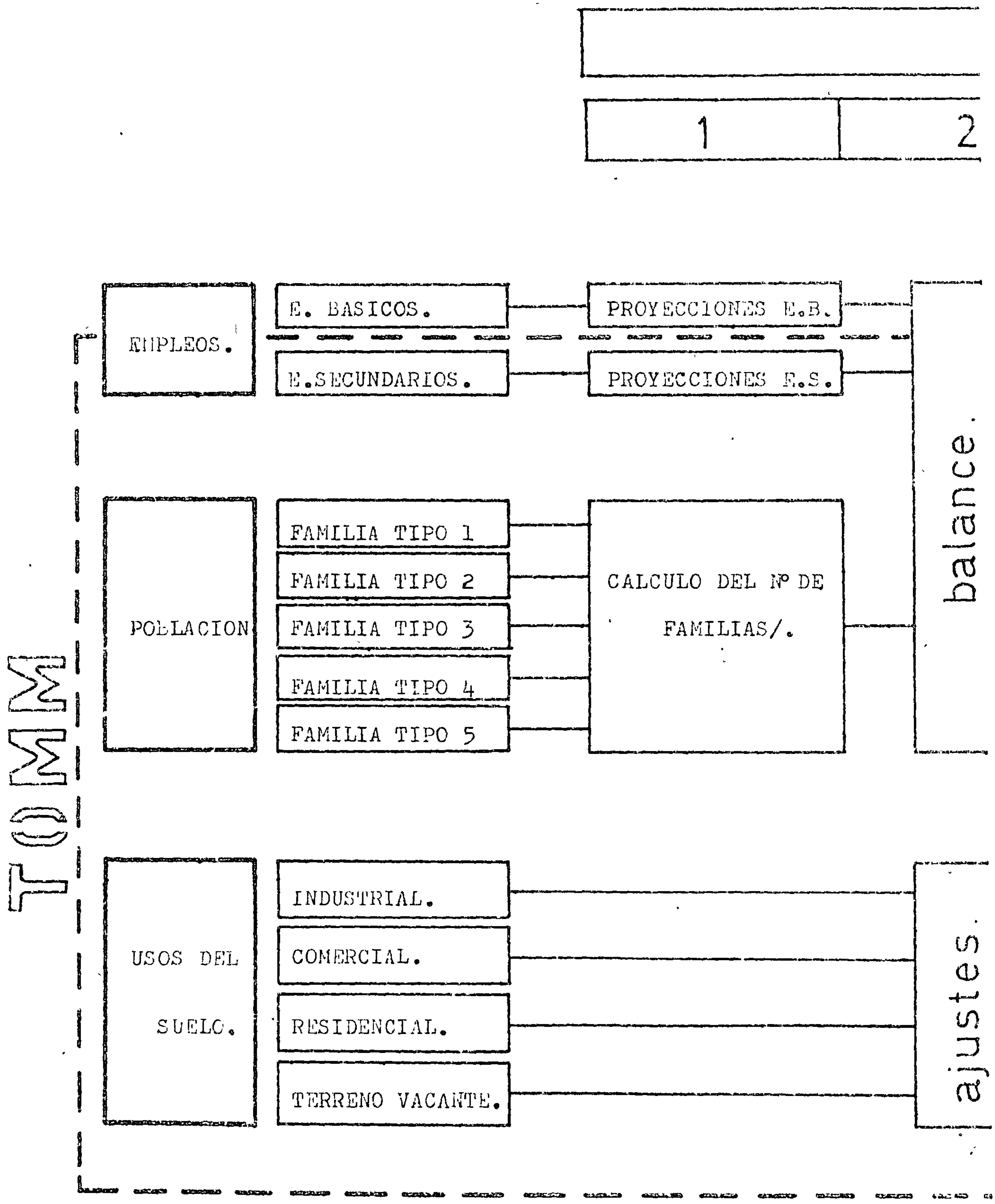
Los aspectos considerados son :

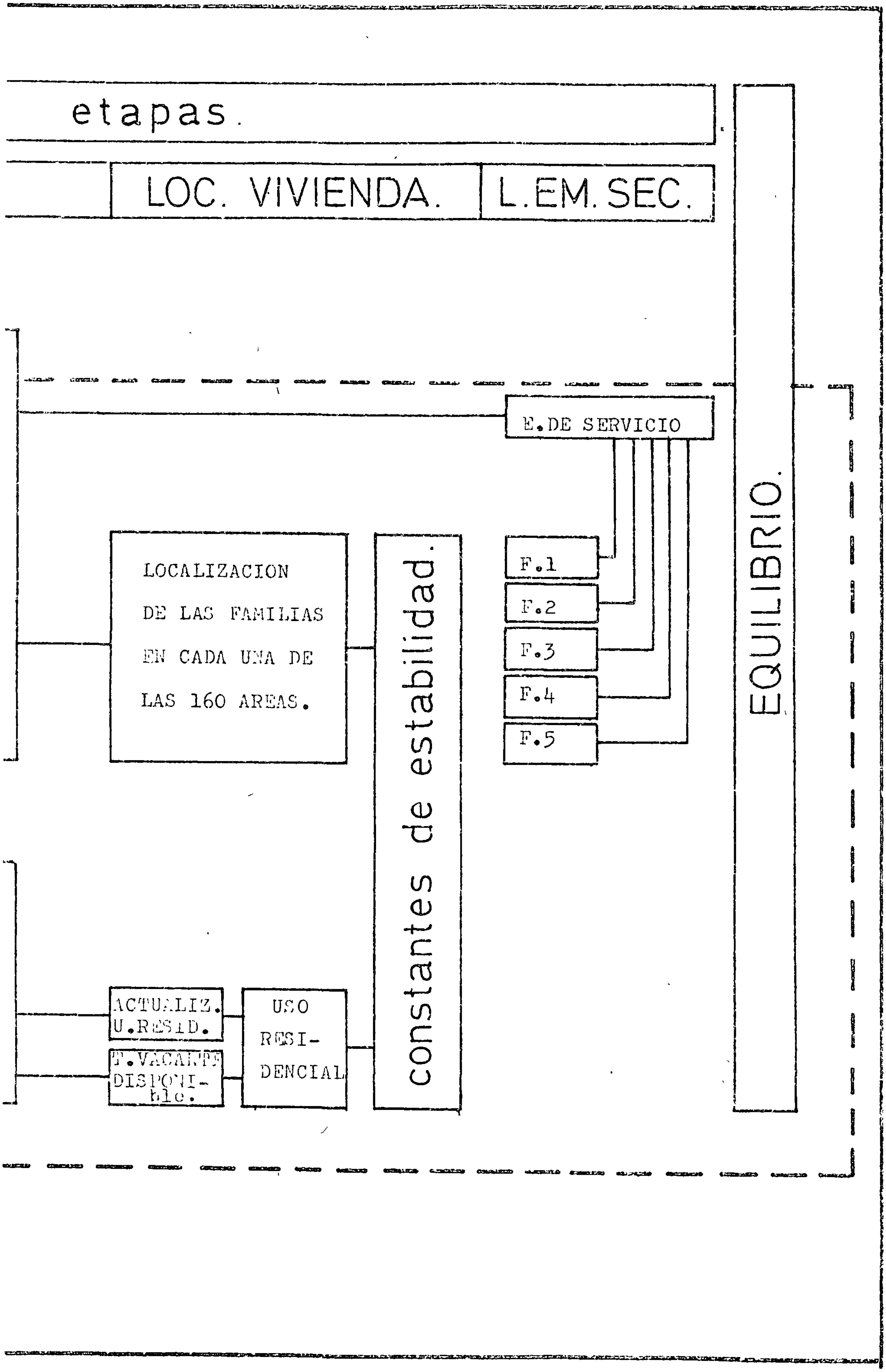
- Uso de la tierra.
- Población.
- Empleo.

El Modelo divide el área Metropolitana en 160 áreas, para cada una de las cuales considera,la población medida por el número de familias clasificadas en cinco categorías (alto nivel, parejas mayores, clase media-baja, transicional y nivel bajo). Estas categorías se derivan de un análisis de factores de variables censales.

TOMM reconoce dos tipos de empleo : básico y secundario. Todo empleo de manufactura, cuerpos corporativos, Universidades y otras Industrias es considerado como básico. Todos los otros empleos mayormente comercio y negocios, son considerados como secundarios y se consideran para servir solamente el área del modelo. El empleo base es considerado como fuerza motora del crecimiento de población.

El usuario del Modelo provee proyecciones de "empleo básico" derivadas externamente para cada área de terreno. El modelo calcula entonces el número correspondiente de viviendas para las cinco


1
2
TOMM
EMPLEOS.
E. BASICOS.
PROYECCIONES E.B.
E.SECUNDARIOS.
PROYECCIONES E.S.
POBLACION
FAMILIA TIPO 1
FAMILIA TIPO 2
FAMILIA TIPO 3
FAMILIA TIPO 4
FAMILIA TIPO 5
CALCULO DEL Nº DE FAMILIAS/.
balance.
USOS DEL SUELO.
INDUSTRIAL.
COMERCIAL.
RESIDENCIAL.
TERRENO VACANTE.
ajustes.


etapas.
LOC. VIVIENDA.
L.EM.SEC.
E.DE SERVICIO
F.1
F.2
F.3
F.4
F.5
LOCALIZACION DE LAS FAMILIAS EN CADA UNA DE LAS 160 AREAS.
constantes de estabilidad.
EQUILIBRIO.
ACTUALIZ. U.RESID.
T.VACANTE DISPONI-ble.
USO RESI-DENCIAL

categorías y proyecta los niveles secundarios de empleo para balancear la población y las dos clases de empleo.

Como el modelo añade ó remueve viviendas y empleo secundario área por área, hace contínuamente ajustes proporcionales en el "uso de la tierra", bajo las categorías de :

- Industria.
- Comercio.
- Residencia.
- Terreno vacante.

II.2.- FUNCIONES DEL MODELO TOMM :

TOMM es un modelo de equilibrio. No pronostica ó proyecta, puesto que no tiene un mecanismo para originar empleo ó población que cambie de año en año. Se basa en proyecciones de empleo básico, población y niveles de empleo secundario preperadas externamente y localiza estos en las 160 áreas.

El valor de este modelo es, su habilidad para mostrar de qué manera será distribuido entre las diferentes áreas de la región considerada, un crecimiento que ha sido asumido.

La función del modelo es determinar patrones de crecimiento y no montos de crecimiento.

II.3.- TEORIA DEL MODELO TOMM :

Para localizar viviendas en el área metropolitana, TOMM debe considerar una teoría de localización. Esta teoría es, que la población debe localizarse de manera tal, que tenga cerca un centro de empleo. Este concepto de accesibilidad es expresado por una "ley gravitacional" : el atractivo residencial de un área varía inversamente con su distancia de un centro de empleo. El atractivo total de un área es entonces, la proporción medida de sus distancias de todos los centros de empleo, lo cual en términos matemáticos se expresa de la manera sigte.:

$$Aj = \sum_{i=1}^{160} \frac{Ei}{Dij^{a}} \qquad (1)$$

Donde : Aj : es el atractivo del área j.
Ei : es empleo total en el área i.
Dij : es distancia entre el área i y la j.
a : exponente proveniente de datos empíricos.

Habiendo establecido el número de viviendas localizadas en cada zona, el modelo las divide ahora en los cinco tipos familiares. Cada tipo de familia es asumido para tener un valor preferencial para cada uno de los otros tipos, así que su preferencia total para un área es, la suma medida de sus preferencias para las viviendas ya en el sitio, tal como se expresa en la sigte. ecuación :

$$N_{1,j} = \left(\sum_{k=1}^{5} W_{1,k} \cdot N'_{1,j} \right) \cdot H_j \qquad (2)$$

Donde : $N_{1,j}$: es el Nº.de familias del tipo 1 en el área j.

$N'_{1,j}$: Nº. $N_{1,j}$ durante el período de tiempo previo.

$W_{1,k}$: preferencia del tipo de familia 1 para el tipo de familia k, obtenida de datos históricos.

H_j : Nº.total de familias localizadas en la j.

Después de añadir ó remover las familias en un área, TOMM actualiza su uso de la tierra residencial. La tierra vacante disponible es convertida en uso residencial, a la misma densidad promedio como en el terreno residencial existente. Si la tierra vacante no es disponible, la densidad residencial en esa área será incrementada, hasta el límite permitido por las ordenanzas de zonificación.

TOMM aplica algunas constantes de "estabilidad" ó limitaciones de cambio. El modelo asume que no sería razonable tener más que una cierta fracción del movimiento de población, ó cierto porcentaje de tierra convertida de un uso a otro, en un período de tiempo.

Si los cambios de población y uso de la tierra exceden estos límites de estabilidad, dichos cambios serán detenidos.

Una vez ubicadas todas las viviendas, el próximo paso de TOMM es, añadir el empleo secundario requerido para servirles.

Una regla gravitacional similar a la utilizada para localizar las familias (viviendas), localiza el "empleo secundario" a cada área, basado en su distancia de los centros de población y empleo.

Cada tipo de familia tiene diferentes necesidades de servicio, así, unos requerimientos de área variarán con su tamaño y tipo de población. Esta relación se expresa de la manera sigte. :

$$S_j = \sum_{i=1}^{160} \frac{d \cdot E_i + \sum_{l=1}^{5} C_l \cdot N_{l,i}}{P_1 + P_2 \cdot D_{ij} + P_3 \cdot D_{ij}^{2}} \qquad (3)$$

Donde : S_j : empleo secundario requerido por el área j.
E_i : empleo (secundario más básico) requerido en i.
C_1 : empleo secundario requerido por el tipo de familia 1.
$N_{1,j}$: número de familias de tipo 1 en el área i.
$D_{i,j}$: distancia entre el área i y j.
P_1, P2, P3, d : son constantes obtenidas de datos empíricos.

Sobre el número de empleados secundarios localizados en un área es impuesto un requerimiento agrupado. TOMM trata de mantener un balance entre el empleo y la población. El empleo al por menor, debe ser suficiente para servir adecuadamente a la población, y la población debe suplir el monto necesario de empleados. Para lograr este balance, deben satisfacerse dos condiciones :

$$HH = E \,.\, FC \qquad (4)$$
$$E = EB + HH \,.\, B \qquad (5)$$

La ecuación (4) es el requerimiento de que, el total de familias (HH) es igual al total de empleo (E) por el número de familias requeridas para suplir un "miembro de mano de obra" (FC).

La ecuación (5) dice simplemente que, el empleo total es igual al empleo básico (EB) más el número de viviendas (HH) multiplicadas por el número de empleados secundarios (B) requeridos para servir a una familia.

II.4.- METODO DEL MODELO TOMM :

El método del modelo es, la solución interactiva del juego de ecuaciones no lineares anteriores, con sus restricciones.

En el modelo se realiza un ciclo a través de su programa, distribuyendo y redistribuyendo el empleo secundario y las familias, hasta alcanzar un equilibrio.

NIVELES DE ABSTRACCION : ILUSTRACIONES DEL MODELO TOMM :

I.- El ESQUEMA CONCEPTUAL GENERAL :

En el fondo de TOMM se encuentra una teoría de interacción humana, un esquema conceptual general utilizado para explicar ó el tiempo ó el costo de distancia vencida en dicha interacción humana, tal como se puede observar en el gráfico anexo. (Ver pág. sigte.)

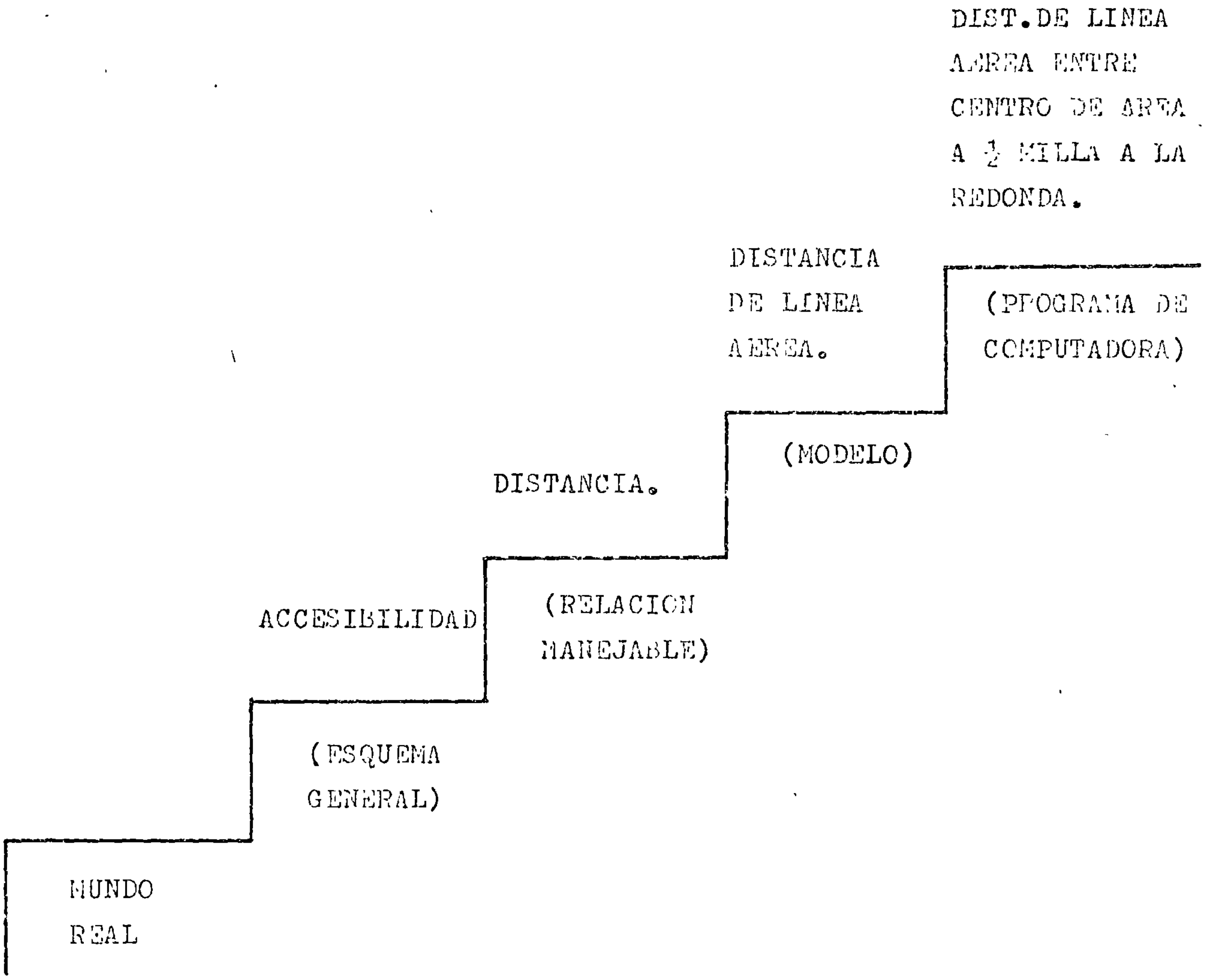
VERIFICACION Y VALIDACION DE LA SIMULACION URBANA :

La verificación consiste en chequear el sistema abstracto para consistencia interna para asegurarse de que no han sido hechos errores lógicos en su estructura,y que no han sido omitidas ó distorsionadas relaciones esenciales. Esto puede realizarse, examinando primero cada concepto ó proposición a través de todos los niveles de abstracción.

Posteriormente, pueden juntarse varias combinaciones de conceptos relacionados y repetir el proceso, a fin de observar que sus relaciones permanecen intactas. En la verificación del modelo mismo, debe hacerse un chequeo especial en sus parámetros y en el método de seleccionar las variables típicas.

La validación está examinando el sistema abstracto del modelo, para ver si produce aproximaciones razonables y satisfactorias del mundo real. Una validez del sistema abstracto es.determinada por su habilidad de predecir el comportamiento del sistema de mundo real, el cual puede juzgarse mejor mediante la comparación entre los resultados (producto) y los datos del mundo real. Puesto que ningún modelo puede ser perfectamente válido, la cuestión esencial llega a ser : qué tan de acuerdo con la realidad debe ser el producto para uso práctico en la planificación ?

Un producto del modelo de simulación urbana está usualmente en la forma de datos "de serie de tiempo" (ej : distribución de un fenómeno en un punto en el tiempo). Para validación de datos en series de tiempo, el análisis puede aplicar directamente varias pruebas estadísticas, enfocadas en el número de puntos, su tiempo, dirección y caracte-



NIVELES DE ABSTRACCION EN LA APLICACION DEL CONCEPTO DE ACCESIBILIDAD DEL MODELO TOMM.

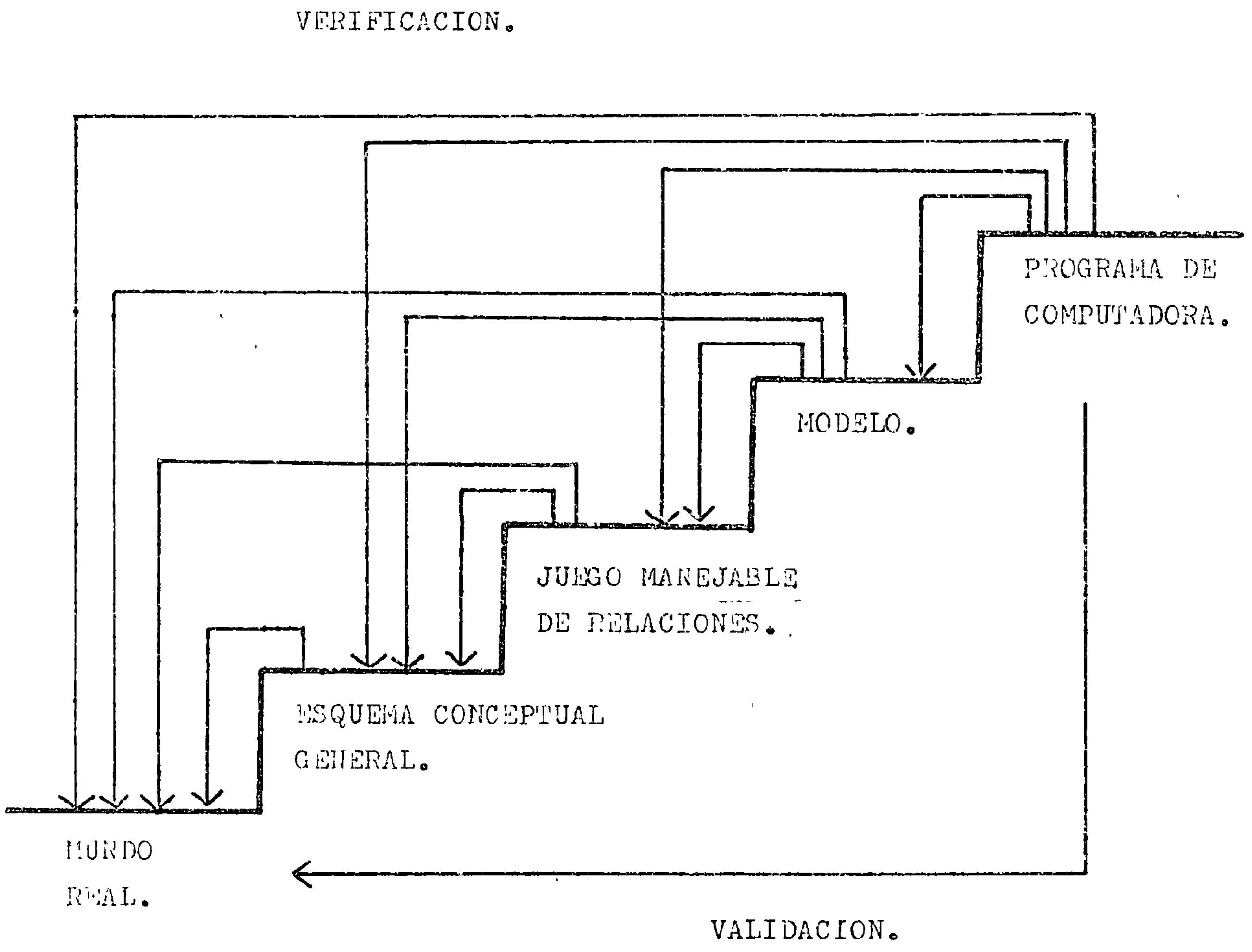
rísticas de distribución de los datos producto, su amplitud, variación acerca del significado y otros rasgos distintivos. Apropiados métodos de estadística incluyen la prueba "chi-cuadrada", análisis de variación, regresión, análisis factorial y espectral y pruebas no paramétricas.

La validación de datos modelados requiere más procedimientos elaborados. Debe trabajarse en una técnica para la descripción de fenómenos espaciales, antes de que el analista pueda aplicar técnicas estadísticas. La Ecología animal y vegetal sugiere algunos métodos para esto. Un procedimiento típico es, diagramar el fenómeno como una serie de puntos en un plano y los datos correspondientes del mundo real, superponiendo luego una malla regular sobre el plano. Entonces pueden utilizarse tres técnicas básicas para describir la distribución del fenómeno.

Contando el número de puntos de datos en un área limitada, es el procedimiento más simple. Una vez descritas las distribuciones, pueden ampliarse las pruebas estadístibas listadas anteriormente.

El gráfico anexo (ver pág.sigte.) ilustra los pasos ó puntos a los cuales la verificación y validación pueden tomar lugar teóricamente. Al menos en principio, la verificación puede ser tentativa entre cada nivel de abstracción, trabajando desde abajo, mientras que la validación es esencialmente una comparación del producto final con el mundo real.

Conjuntamente, la verificación y la validación dicen al analista qué tan bien diseñado y útil es el modelo de simulación. Solamente despúes de saber esto, puede utilizarse apropiadamente como un "pseudo-laboratorio" para experimentos urbanos.



TRABAJO DE VERIFICACION Y VALIDACION BAJO LOS NIVELES DE ABSTRACCION.

ILUSTRACIONES DE VERIFICACION Y VALIDACION DEL MODELO TOMM :

Es prácticamente imposible por supuesto, documentar una verificación y validación completa de un modelo de simulación representando un sistema tan complejo de población, empleo y uso de la tierra, como lo hace TOMM.

La verificación sigue contínuamente a través de cada paso de abstracción y a través del tiempo del uso del modelo. La validación es probada usualmente solo después de que el modelo es programado, y por su naturaleza no puede ser nunca completa.

No hay cosa semejante a un modelo perfectamente válido, y cada modelo eventualmente llega a ser menos válido como la estructura ddl mundo real y los cambios de fenómeno.

I.- VERIFICACION :

La verificación puede ser concebida como una escalera de abstracción "caminando hacia abajo", comparando cada nivel con el inferior. El primer paso es, quitar las fallas del modelo. Deben buscarse primero los errores del programa de la computadora, tales como, variables que deben ser añadidas dos veces, ó el uso de suscripciones impropias.

Muchos de estos errores se revelan a sí mismos durante la corrida del programa, cuando se trata de dividir entre "O" ó generar números no reales, ó simplemente "faltas ó paradas".

Otros errores pueden haberse realizado en la traducción del modelo a un programa. Puede haber fallas en las condiciones limitantes para evitar el número de habitantes que llegan a ser negativos, fallas en la distindión entre variables contínuas y discretas, u olvidar de establecer condiciones iniciales para una variable.

Estos son descuidos que el analista toma normalmente como datos, pero que la computadora no puede entender. No puede apreciar lo absurdo de un área de terreno poblado por un número negativo de familias.

Bajando por los niveles de abstracción, se busca como próximo paso, los errores en la lógica del proceso mismo de modelación. Cuando la teoría manejable ó el juego de relaciones era convertido en un modelo, el analista puede haber fallado en la captura de un elemento ó aspecto del sistema. Semejantes errores pueden revelarse a sí mismos solamente en casos especiales ó para algunas combinaciones de valores de las variables, y así pasar inadvertidas a través de muchas corridas de modelo que parecen razonables. Un error semejante fue encontrado en TOPM después de varias corridas. La relación de densidad de población a distancia del centrode la ciudad, no había sido incorporada apropiadamente en el modelo, y bajo ciertas circunstancias, el modelo tenuido a forzar la movilización de habitantes de los suburbios, creó una distribución de población completamente irreal.

Desafortunadamente para la verificación, los errores lógicos tienden a ser los más difíciles de encontrar y su búsqueda forza al analista a repensar las relaciones teóricas acerca de las cuales puede sentirse inseguro. Por otro lado, errores de programación menos críticos, tienden a llamar la atención inmediatamente cuando el programa está corriendo. Para vencer esto y llegar al fondo de la verificación, el único recurso es una larga serie de corridas con un modelo de trabajo, en el cual el analista experimenta, hasta que pueda anticipar confidencialmente su comportamiento. Es entonces el momento para introducir un cambio cuidadosamente seleccionado y medido, para comparar el nuevo producto con el obtenido corridas de prueba previas. Una vez que el analista esté seguro, y entiende el comportamiento del modelo bajo la primera modificación, añade una segunda y corre de nuevo, luego una tercera y así sucesivamente.

El modelo resultante de este largo proceso de verificación es muy diferente de aquel con el cual se había comenzado.

II.- VALIDACION :

La validez de un modelo de simulación es determinada por la precisión de sus predicciones. Solamente el tiempo es el factor que puede decir esto. Comparar el pronóstico con la realidad eventual, no es una prueba de validez real. Pueden surgir condiciones que no están incluidas en la teoría original del modelo para hacer su pronós tico ~~ine~~ inexacto, aunque la primera concepción del modelo era válida. Por el contrario, un modelo pobremente concebido puede dar un pronóstico exacto accidentalmente, realizando una cosa correcta para una razón incorrecta.

Tal vez el método de validación más común es la predicción retrospectiva. En simulaciones más grandes esto es posible solamente para los años 1.950 y 1.960, para lo cual hay disponibles adecuados datos censales. El método es, insertar datos históricos para algún año anterior y predecir los datos de 1.950, comparando los resultados con los del Censo de 1.950. El proceso es repetido con una predicción para 1.960. Cuando los resultados de ambas pruebas son comparados, puede ganarse alguna comprensión en la validez del modelo.

En la validación del TOMM, este proceso no puede ser utilizado, porque los datos censales de 1.950 eran inadecuados. Los datos de empleo base más recientes eran para 1.960, por lo cual fue necesaria una predicción inter-censal.

El mayor insumo del TOMM son, proyecciones de empleo básico preparadas externamente. Afortunadamente, semejantes proyecciones eran disponibles para 1.960 hasta 1.966. Con estos insumos, fue realizada una proyección del número de viviendas y localización entre las 160 áreas para 1.966. Debido a la falta de datos actuales de las familias para 1.966, fueron utilizados datos sustitutivos del año 1.967, los cuales habían sido colectados por muestreo para propósitos muy diferentes. Por esto, fue necesario basar la prueba de validación en una comparación de las proyecciones del modelo de 1.960 a 1.966, con datos basados en el muestreo de las familias para el 1.967.

Un análisis regresivo produjo :

Viviendas proyectadas por TOMM para 1.966 =
= .8235 viviendas actuales de 1.967, con un
R^2 de .7354.

Esta es una proyección de distribución razonablemente buena. Otra prueba produjo .2905, de nuevo un valor razonable, considerando la naturaleza del modelo.

La prueba de la "chi-cuadrada" indica la inapropiabilidad relativa de semejantes métodos estadísticos para probar modelos urbanos. Con 160 grados de libertad (N-1 ó N 160-1 ; hay 160 proyecciones individuales y 1 proyección externa a TOMM), el resultado es aproximadamente 21, una cifra semejantemente apropiada que se sale del rango de la mayoría de las tablas.

Los grados del concepto de libertad han sido aplicados erroneamente, por lo cual la prueba del chi-cuadrado no es realmente aplicable.

Es una lástima que los métodos estadísticos disponibles para la prueba de validez del modelo de simulación urbana, parecen ser inadecuados en algún grado. La mayoría de las relaciones urbanas tienden a ser cualitativas, interdependientes y no muy claras.

Los modelos de simulación son los únicos modelos prácticos para uso en los esfuerzos de planificación urbana de algún tamaño y complejidad. Los modelos analíticos son limitados, por falta de teoría y por dificultades de formulación y solución.

Si se van a modelar sistemas urbanos, debe continuarse todavía por un tiempo, con modelos de simulación, a pesar de sus limitaciones.

MODELOS GRAVITACIONALES.

EL PUNTO DE VISTA DE PROBABILIDAD :

Para desarrollar y comprender los modelos gravitacionales se adopta aquí un punto de vista de probabilidad bastante simple.

Se supone una región metropolitana de población P, y se divide esta región en sub-áreas. Se supone también que se conoce el número de desplazamientos interiores que realizan dentro de la región sus habitantes, este número es T= constante. Se establece que no existen diferencias de gastos, renta, distribución, estructuras ocupacionales, etc. entre las sub-áreas.

El problema es, determinar el número de desplazamientos q que se originan en la sub-área i y terminan en la j, suponiendo que un desplazamiento no implica ni tiempo, ni costo (la fricción de la distancia es cero).

En esta situación hipotética se puede establecer que, para un individuo representativo de la sub-área i, el porcentaje de sus viajes que terminan en la sub-área j es igual a P_j/P,

siendo : P_j : población de j y

P : población total.

Además, por la hipótesis de fricción cero, el número de viajes que la persona de la sub-área i realiza, es el promedio del número de desplazamientos per cápita para la región metropolitana entera, y este promedio es igual a $T/P = K$. Por esto, el número absoluto de desplazamientos que un individuo representativo de la sub-área i hace a la sub-área j es : $K(P_j/P)$.

Como existen P_i individuos en i, el número de desplazamientos que estos P_i realizarán hacia la sub-área j será :

$$T_{ij} = K \frac{P_i \, P_j}{P} \qquad (1)$$

De este modo se ha obtenido una estructura del volúmen de desplazamientos entre sub-áreas, esperados ó hipotéticos para la región metropolitana.

El siguiente paso consiste en determinar el efecto que la distancia real que separa a dos sub-áreas, pueda tener sobre el número de desplazamientos que se producen entre ellas.

Si se obtienen los datos corrientes sobre el número de desplazamientos entre cada una de las sub-áreas de una región metropolitana, se representarán estos por I_{ij}.

Se establece una relación entre el radio I_{ij}/T_{ij} y la distancia d_{ij} (ajustando una curva por mínimos cuadrados ó cualquier otro método), se supone en este caso que se encuentra la sigte. relación lineal :

$$\text{Log} \frac{I_{ij}}{T_{ij}} = a-b \log d_{ij} \qquad (2)$$

haciendo c= antilog a

$$\frac{I_{ij}}{T_{ij}} = \frac{c}{d_{ij}^{\,b}} \qquad ó \qquad I_{ij} = \frac{c\,T_{ij}}{d_{ij}^{\,b}} \qquad (3)$$

Sustituyendo T_{ij} (ecuac.1) en la ecuación (3), e introduciendo la constante $G = cK/P$, donde c,K y P son las constantes ya definidas, obtenemos :

$$I_{ij} = G \frac{P_i P_j}{d_{ij}^{\,b}} \qquad (4)$$

Esta sencilla relación puede ser aceptada para describir el volúmen de desplazamientos (actuales) en el interior de la región metropolitana (describe la interacción de los habitantes como una función de las poblaciones de las sub-áreas y la variable distancia, cuando esta interacción se refleja en desplazamientos).

Nota : esta relación deducida desde este punto de vista es esencialmente el modelo gravitacional, aunque este no ha sido desarrollado de esta forma.

Si se quiere hallar la interacción de una región con más de una de las otras regiones, se puede utilizar la sigte. fórmula :

$$I_{i1} + I_{i2} + \dots I_{in} = G \frac{P_i P_1}{d_{i1}^{b}} + G \frac{P_i P_2}{d_{i2}^{b}} + \dots + G \frac{P_i P_n}{d_{in}^{b}}$$

$$\sum_{j=1}^{n} I_{ij} = G \sum_{j=1}^{n} \frac{PiPj}{dij^{b}} \qquad (5)$$

Si se saca a P_i como factor en la parte derecha y se dividen ambos miembros :

$$\frac{\sum_{j=1}^{n} I_{ij}}{P_i} = G \sum_{j=1}^{n} . \frac{P_j}{d_{ij}^{b}} \qquad (6)$$

Esta relación muestra la interacción con todas las áreas en términos per cápita ó más estrictamente, en términos por unidad de masa. La interacción definida así se le designa potencial de i (iV).

$$iV = \frac{\sum_{i=1}^{n} I_{ij}}{P_i} = G \sum_{j=1}^{n} \frac{P_j}{d_{ij}^{b}} \qquad (7)$$

HIPOTESIS STEWART - ZIPF :

En las ciencias naturales existen leyes tales como las que gobiernan la densidad, la presión y la temperatura de los gases. Stewart razonó que, en la interacción de las unidades sociales pueden subyacer relaciones similares, las cuales es posible descubrir mediante la investigación de grandes agregados de tales unidades. Basánsose en la física de Newton, presentó tres conceptos primarios, y siguiendo la fuerza gravital, Stewart define la fuerza demográfica.

Cuando la población de las ciudades i y j, que se designa por P_i y P_j se toma como masas relevantes, la fuerza demográfica f es:

$$F = G\,\frac{P_i P_j}{d_{ij}^{\,b}} \qquad (8) \qquad b=2$$

Stewart desarrolla un segundo concepto que corresponde a la energía gravital y que se denomina "energía demográfica".

$$E = G\,\frac{P_i P_j}{d_{ij}^{\,b}} \qquad (9) \qquad b=1$$

Estas dos expresiones (8)y(9) tienen en el d_{ij} como exponentes a b=2 y 1 respectivamente.

El tercer concepto de Stewart es, el de potencial demográfico (que corresponde al de potencial gravitacional). El potencial demográfico producido en un punto i por una masa j, que puede designarse por iV_j , se define en un tiempo determinado como la masa en j, ó sea Pj dividida por la distancia existente :

$$iV_j = G\,\frac{P_j}{d_{ij}} \qquad iV = G\sum_{j=1}^{n} \frac{P_j}{d_{ij}} \qquad (10)$$

Los trabajos de Zipf se asemejan mucho a los de Stewart por lo cual no se tratarán por el momento.

CUESTIONES BASICAS DE LOS MODELOS GRAVITACIONALES :

En esta parte se va a discutir cuestiones que aparecen particularmente en la aplicación del modelo. Un aspecto se refiere a las variables masa y distancia.

En los estudios empíricos, la masa ha sido medida de diferente forma, aquí se ha utilizado la población como medida. Sin embargo, si se estudia la migración intermetropolitana, el empleo ó la renta pueden ser unos indices más significativos; la medición de la masa que se emplee depende del problema a estudiar, de los datos disponibles, etc.

La distancia se ha medido de forma similar. En este caso, se ha considerado una medida física a lo largo de una recta, no obstante, si se lleva a cabo un estudio sobre tráfico metropolitano, la distancia en tiempo puede ser igual ó más importante que la distancia en kilómetros. Si se analiza el problema de localización industrial, el costo de transporte es más significativo que la distancia física.

De la misma forma que la medición de la masa, la medición de la distancia depende del problema a abordar, de los datos disponibles y de otras consideraciones anexas.

Una cuestión básica que permite eliminar las hipótesis de homogeneidad hechas al principio, es la que se refiere a la aplicación de las "ponderaciones a las masas".

Suponiendo que se estudia el volúmen de viajes de primera clase ó de lujo. Es razonable esperar que un área con alta renta per cápita tendrá un volúmen mayor de tales viajes, que un área de menor renta. Una forma de corregir tal factor es, multiplicar la población de cada sub-área, por su renta per-cápita. De esta manera la ecuación (4) se convierte en :

$$I_{ij} = G \frac{(W_i P_i)(W_j P_j)}{d_{ij}^{\,b}} \qquad (11)$$

y la ecuación (7) será :

$$V = G \sum \frac{W_j P_j}{d_{ij}^{\,b}} \qquad (12)$$

Es posible ponderar la masa con más de un factor. En tal caso W_i y W_j representarían ponderaciones compuestas (promedios que reflejan la relativa importancia de los distintos factores ponderados).

Más difícil que la selección de las ponderaciones ó medidas de masa y distancia es, la elección de exponentes para las variables en los conceptos de energía potencial y demográfica. Desde un punto de vista teórico se puede presumir que el exponente de la variable distancia d_{ij} debería ser 1 ó 2, según sea energía ó potencial. Pero numerosos estudios empíricos realizados por diferentes investigadores no apoyan estos exponentes.

Aunque es evidente que el exponente de d_{ij} no es necesario que sea 1 ó 2, todavía no se ha hecho un estudio definitivo de la cuestión. En las ecuaciones básicas (4) y (5) estos exponentes son la unidad. Sin embargo, investigadores muy profundos han encontrado que la potencia a la cual es elevada la masa, podría ser otra que la unidad.

Ej. : Carrothers observa que factores tales como las economías de aglomeración (desaglomeración) implican que el exponente que debe aplicarse a cualquier masa, es función de esta masa. En tal situación, los exponentes de masas distintas serán diferentes; las ecuaciones (4) y (5) se convertirán en las sigs. :

$$I_{ij} = G \cdot \frac{W_i(P_i)^{\alpha} \cdot W_j(P_j)^{\beta}}{d_{ij}^{\,b}}$$

$$iV = G \sum_{j=1}^{n} \frac{W_j(P_j)^{\beta}}{d_{ij}^{\,b}}$$

Mientras no se continuen investigaciones acerca de este problema, esto queda aun sin resolver.

PROBLEMAS EN EL USO DE LOS MODELOS GRAVITACIONALES SIMPLES :

Ya se han tratado algunas cuestiones básicas en relación con las ponderaciones, exponentes y formas de las funciones. Con respecto a los problemas relativos al uso del modelo gravitacional en sus formas más simples, debe tenerse presente que estos modelos pueden utilizarse tanto con fines descriptivos, como de proyección.

Un problema común para ambos tipos de utilización es, el referente al grado en que cualquier conjunto (masa integral, población ó agregado significativo) es fraccionado, clasificado en sectores ó desagregado.

La noción de modelo gravitacional, particularmente la desarrollada por Stewart, corresponde a una masa relativamente grande, compuesta por multitud de unidades individuales. Dentro de tal masa es razonable suponer que las irregularidades, peculiaridades e idiosincrasias de cualquier unidad individual ó pequeño grupo de unidades, se anulan ó compensan. En tal situación la cláusula es válida en cierto grado; por consiguiente, se puede justificar una concentración en dos variables básicas, la distancia y la masa, y en factores que pueden sintetizarse en ponderaciones y exponentes, excluyendo las otras variables; por Ej. : para estudios de tráfico, cuando el volúmen total de tráfico se desagrega por clase de medios, ó por finalidad de desplazamientos, por tipo de ciudad u otra clasificación, las peculiaridades de cada categoría tienden a ser más manifiestas y dominantes, y la amplitud con la cual un modelo gravitacional describe ó explica cualquier efecto (disminución de volúmen,etc.) regular,tiende a reducirse. De esta forma se plantea un problema básico : por un lado, parece conveniente desagregar y estratificar con miras a distinguir entre distintos exponentes ó ponderaciones que podrían emplearse para proyectar ó describir las distintas categorías. Por otro lado, el modelo gravitacional como técnica para describir ó explicar el volúmen de desplazamientos con un propósito particular, tiende a ser menos fiable que para cualquier estrato particular de una masa, y si la desagregación es llevada demasiado lejos, los datos deducidos pueden llegar a tener una significación reducida.

Puede decirse que la desagregación se hace aconsejable cuando es posible obtener información adicional y precisión, y además, cuando tal desagregación no destruye en grado alguno la significación intrínseca y la estructura unitaria interna de la masa ó población.

Bajo estas circunstancias, será útil emplear diferentes exponentes, ponderaciones, etc.

Cuando se intenta emplear el modelo gravital como algo más que una técnica descriptiva, surgen todavía problemas de características todavía más difíciles. Un obstáculo básico para su utilización proyectiva consiste en la carencia de cualquier teoría, para explicar los valores ó funciones que se asignan a las ponderaciones y exponentes.

Comunmente, la justificación del modelo es simplemente, que la interacción entre dos poblaciones cualquiera puede suponerse que está en relación directa con su tamaño, permaneciendo todo lo demás igual; y puesto que la distancia implica fricción, inconvenientes y costo, tal interacción puede suponerse inversamente relacionada con la distancia.

MODELO DE OPORTUNIDAD ACCESIBILIDAD.

En este modelo, la distribución espacial de una actividad es vista como la evaluación sucesiva de oportunidades alternativas para los sitios, los cuales han sido ordenados jerarquicamente en función al tiempo de traslado de un centro urbano al sitio donde se quiere localizar la actividad.

Las oportunidades son definidas como : el producto de tierra disponible y la densidad de la actividad (unidades de actividad por unidad de área).

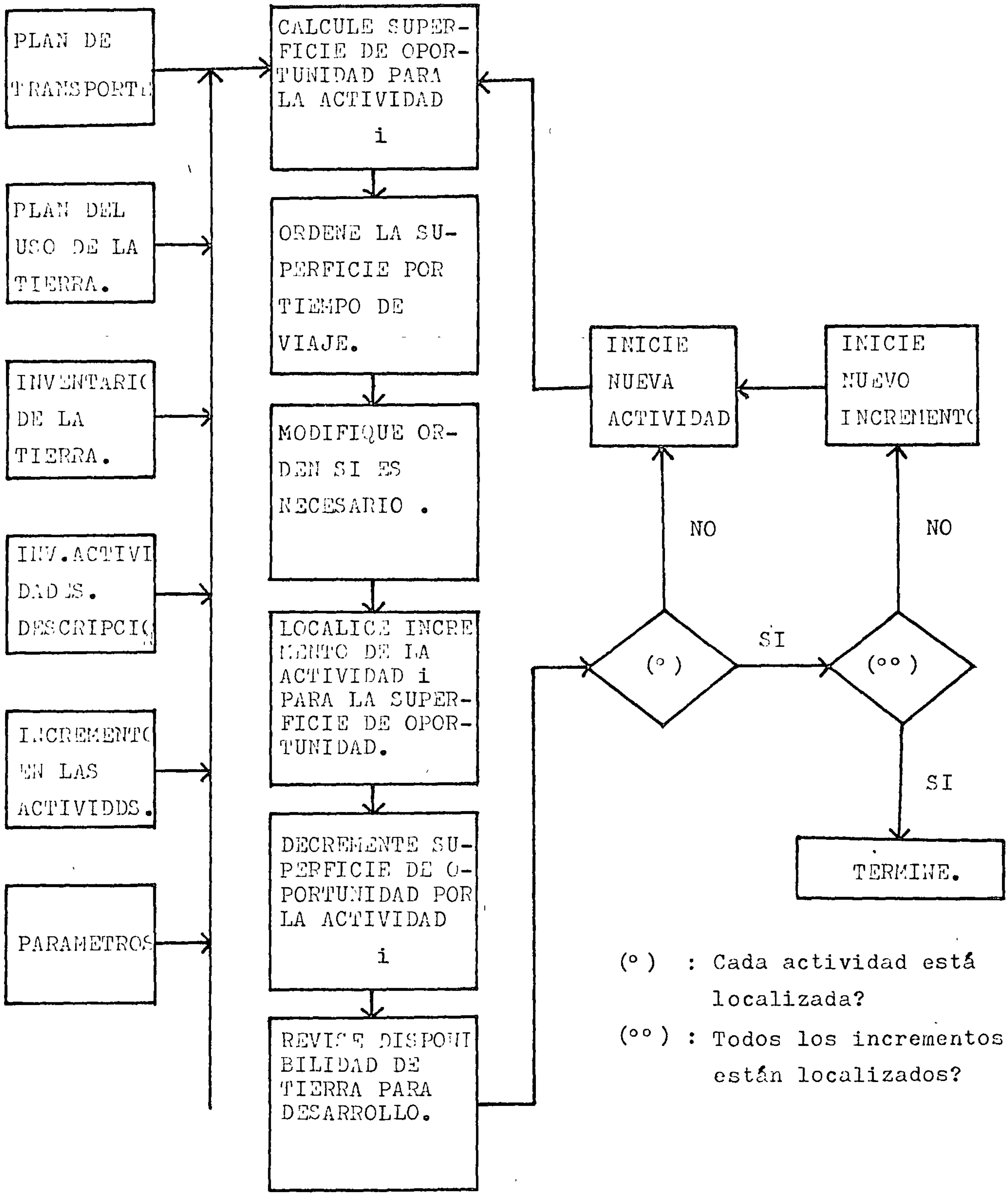
$$A_j = A \left(e^{-lO} - e^{-l(O+j)} \right)$$

donde : A_j : actividad a ser localizada en la zona j.
A : suma total de las actividades que se deben localizar.
l : probabilidad de que una actividad sea situada con un nivel de oportunidad dado.
O : oportunidades que preceden a la zona j.
O_j : oportunidad en la zona j.

El uso de una exponencial negativa, permite que sean localizados primero en la superficie de oportunidad, los puntos de máximo acceso (máxima oportunidad).

Esta presunción es consistente con las observaciones empíricas y con la mayoría de las teorías económicas que tratan del uso de la tierra.

SECUENCIA DE OPERACION DEL MODELO DE OPORTUNIDAD ACCESIBILIDAD :



(°) : Cada actividad está localizada?

(°°) : Todos los incrementos están localizados?

OPORTUNIDADES :

El concepto de una oportudidad para asentar una unidad de actividad envuelve el área de tierra y la medida de la intensidad ó densidad del uso de la tierra.' En un análisis de intensidad deberán considerarse posibles cambios de esta, debido a cambios en los costos de transporte, en los costos de construcción, cambios en los requerimientos para localizar actividades y cambios en las preferencias y gustos de los posibles usuarios.

Por estas dificultades es bastante difícil simular dentro del modelo, la intensidad del uso de la tierra. La densidad ha sido introducida exógenamente dentro del modelo; esta independencia permite tomar densidades alternativas derivadas de suposiciones ó planeadas preferiblemente.

PROBABILIDAD DE ASENTAMIENTO "1" :

Este parámetro es la probabilidad de que una actividad sea establecida ó este situada a una unidad de oportunidad. Para una superficie de oportunidad dada, valores grandes de "1" harán que el área urbana sea más concentrada; valores pequeños de "1" la harán más esparcida.

LA NOCION DE ACCESOS :

Los patrones de asentamiento son muy sensitivos a las facilidades de transporte. En este modelo se enfatiza que las zonas urbanas tenderán a desarrollarse más rápida y concentradamente, por los caminos de mínimo tiempo en las redes de transporte.

La noción de accesibilidad que se desarrolla aquí, se debe entender como facilidades para el transporte y se toma en consideración dentro de los parámetros del modelo.

Este concepto de accesibilidad no debe ser confundido con los conceptos gravitacionales de accesibilidad (atractividad).

Con todas las limitaciones de los procedimientos de los procedimientos de localización, es posible incorporar los efectos de los cambios en los medios de tranporte y en los accesos, logrando así una retroalimentación entre el uso del suelo y los sistemas de transporte.

El modelo es útil como una base para predecir la distribución futura de la ciudad y sus habitantes, así como también, la generación de viajes, igualmente, es posible evaluar las políticas de uso de la tierra, los controles de la densidad, etc.

BASS.

(BAY AREA SIMULATION STUDY) 1.962.
CENTER FOR REAL ESTATE AND URBAN ECONOMICS.
UNIVERSIDAD DE CALIFORNIA , BERKELEY.

I.- OBJETIVOS :

- Análisis del impacto probable sobre los usos del suelo futuros, en el área de la Bahía de San Francisco, debido a cambios en el nivel de empleo.
- Análisis del impacto de grandes inversiones en transporte, sobre la estructura y distribución de los usos del suelo urbano, en el área de la Bahía de San Francisco.

II.- ESTRUCTURA DEL MODELO :

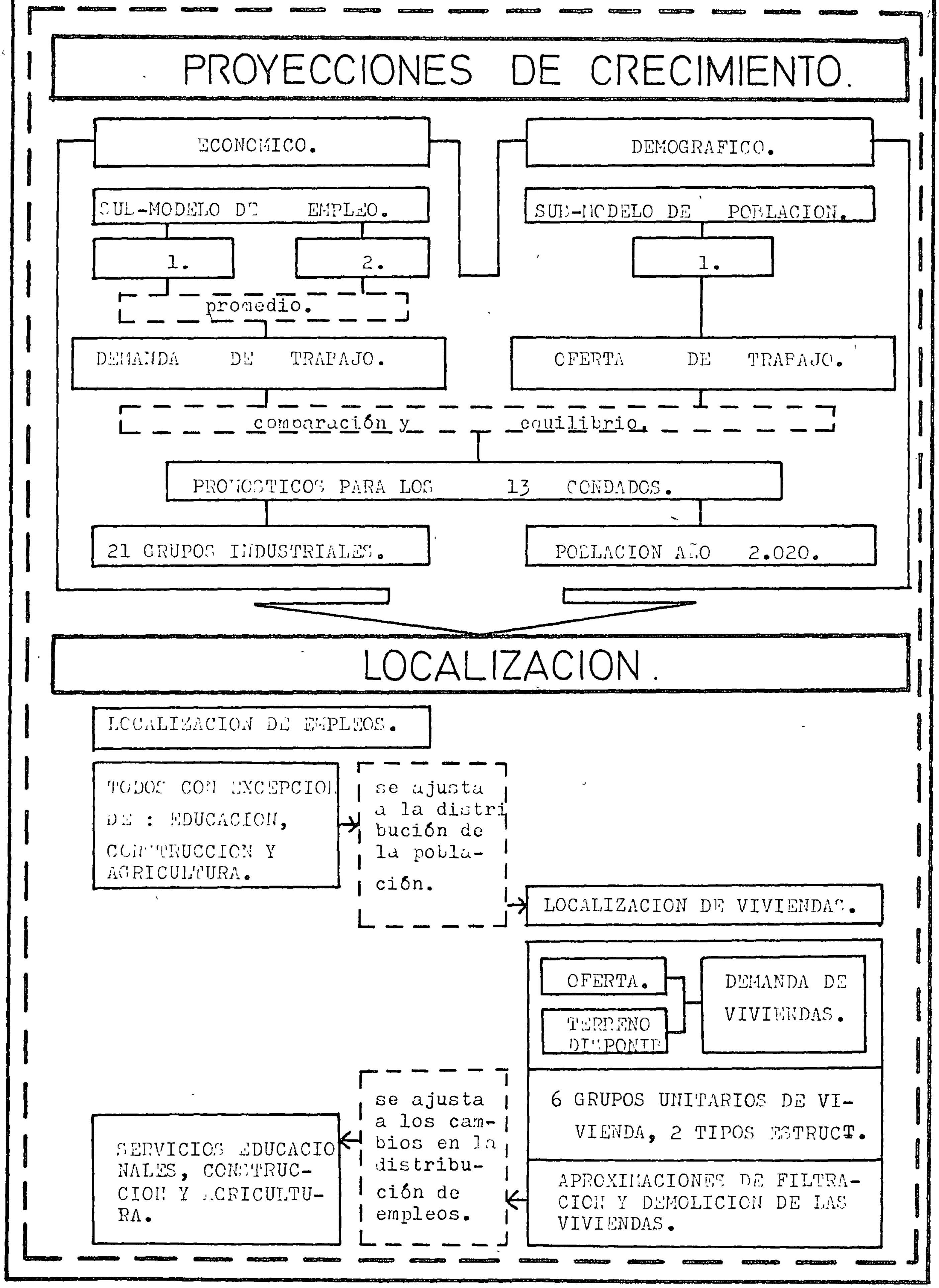
En el modelo BASS, dos sub-modelos de empleo y uno de población generan las proyecciones globales del área de empleo y población para interacciónes de 5 años, proyectando hasta el año 2.020.

Los dos pronósticos de empleo son promediados a fín de obtener la demanda de trabajo, la cual es comparada entonces, con la oferta de trabajo generada por el modelo de población. La demanda y la oferta son equilibradas y los pronósticos de empleo para los 21 grupos industriales, así como tambíén, un pronóstico de población para el año 2.020, son obtenidos para los 13 condados del área de la Bahía.

BASS diferencia entre las determinantes localizacionales de los diversos grupos industriales, haciendo uso de 6 sub-modelos de localización de empleo.

Un sub-modelo de localización residencial equilibra la oferta de viviendas y tierra utilizable con la demanda de viviendas gene-

BASS.


PROYECCIONES DE CRECIMIENTO.
ECONOMICO.
DEMOGRAFICO.
SUB-MODELO DE EMPLEO.
SUB-MODELO DE POBLACION.
1.
2.
1.
promedio.
DEMANDA DE TRABAJO.
OFERTA DE TRABAJO.
comparación y equilibrio.
PRONOSTICOS PARA LOS 13 CONDADOS.
21 GRUPOS INDUSTRIALES.
POBLACION AÑO 2.020.
LOCALIZACION.
LOCALIZACION DE EMPLEOS.
TODOS CON EXCEPCION DE : EDUCACION, CONSTRUCCION Y AGRICULTURA.
se ajusta a la distribución de la población.
LOCALIZACION DE VIVIENDAS.
OFERTA.
TERRENO DISPONIB
DEMANDA DE VIVIENDAS.
6 GRUPOS UNITARIOS DE VIVIENDA, 2 TIPOS ESTRUCT.
APROXIMACIONES DE FILTRACION Y DEMOLICION DE LAS VIVIENDAS.
se ajusta a los cambios en la distribución de empleos.
SERVICIOS EDUCACIONALES, CONSTRUCCION Y AGRICULTURA.

rada por los pronósticos de población. El sub-modelo utiliza 6 grupos unitarios de vivienda, basados en dos tipos estructurales (uni y multifamiliar) y tres valores ó niveles (alto,medio y bajo). El submodelo incorpora aproximaciones de filtración y demolición de las viviendas.

Las localizaciones de empleo, con la excepción de los servicios educacionales, construcción y agricultura, son determinadas a priori de la localización de residencias. Esta secuencia supone implícitamente que, el empleo de servicios se ajusta a la distribución de población, tanto en el sitio de trabajo, como en el de vivienda que existía al comienzo de un período de 5 años, durante el cual la localización residencial se ajusta a los cambios en la distribución de empleo durante ese período.

III.- METODOLOGIA :

Los métodos para la localización del empleo para las 6 localizaciones industriales son los sigs. :

1.- EMPLEO AL POR MENOR : distribuido mediante una combinación de un modelo de potencial gravitacional y una medida de atractividad derivada de un análisis de regresión.

2.- MANUFACTURAS, FLETES, BODEGAS, VENTAS AL MAYOR : distribuido por una medida de atractividad derivada de las sub-áreas.

3.- FINANZAS, SEGUROS, INMUEBLES, EDUCACION Y GOBIERNO : distribuidos a condados y sub-áreas importantes sobre la base de una extrapolación simple de las tendencias históricas, modificadas en base a juicios. El residuo es distribuido a otras sub-áreas en base a su población.

4.- SERVICIOS : asignados con un modelo de regresión en en cual el empleo es una función de la densidad y accesibilidad.

5.- CONSTRUCCION : distribuido a sub-áreas como una función de la cantidad de nuevo empleo y nueva vivienda.

6.- AGRICULTURA, MINERIA, TRANSPORTES, COMUNICACIONES, MILITARES : se distribuye el empleo total derivado del modelo de proyección, a sub-áreas en la misma proporción que la actual.

Los cuatro primeros de esto sub-modelos se explican posteriormente con mayor detalle. Los últimos dos se explican a sí mismos. La construcción es relacionada con nuevas viviendas y nuevos empleos y es localizada a sub-áreas sobre la base de los requerimientos de las actividades mencionadas. El nivel de actividad agrícola es un residual, y el empleo agrícola es reducido en cada sub-área, en proporción a la extracción de terreno agrícola de la sub-área.

1.- SUB-MODELO DE LOCALIZACION 1 : COMERCIO AL POR MENOR :

La proporción de nuevo empleo de comercio al POR MENOR en cualquier zona j, es una función del porcentaje del nuevo empleo regional localizado en j, el porcentaje de la población total en j y la atractividad relativa de j.

$$\mathrm{ALLOC}_j^r = \frac{((\alpha\, \mathrm{NE}_j / \sum_j \mathrm{NE}_j) + (\beta \mathrm{POP}_j / \sum_j \mathrm{POP}_j) + \mathrm{RAI}_j)}{2} \Delta \mathrm{EMP}^r$$

donde : ALLOC_j^r : nuevo empleo al menor en j.
NE_j : nuevo empleo total en j.
POP_j : población total en j.
RAI_j : índice de la atractividad relativa de j.
$\Delta \mathrm{EMP}^r$: nuevo empleo al menor en toda el área.
α, β : constantes exógenas que suman 1.

RAI es una función de la demanda potencial para comercio al por menor en la zona y de la adecuación comercial del sitio.

La demanda potencial es la diferencia entre la demanda actual en la zona (empleos) y la demanda esperada.

La demanda esperada es calculada por un modelo gravitacional. La adecuación comercial del sitio es obtenida a partir de una ecuación de regresión, relacionando la cantidad de empleo al por menor, con la accesibilidad de la población, accesibilidad industrial, densidad de la zona y las cantidades y localización de otros tipos de empleo.

$$RAI_j = \sqrt{DP_j\ (1 + 0.5\ (\ CSS_j / \sum_k CSS_k))}\ +$$

$$+ \sum_i \sqrt{DP_i\ (\ 1 + 0.5\ (\ CSS_i / \sum_k CSS_k))}\ .$$

donde : DP_j : demanda potencial.
CSS_j : adecuación comercial del sitio.

2.- SUB-MODELO DE LOCALIZACION 2 : MANUFACTURAS, FLETES, BODEGAS, VENTAS AL MAYOR :

Los incrementos en el empleo son localizados por medio de índices medidos por zona, de los factores de localización calculados separadamente para 11 grupos industriales. La zona con el marcador más alto para un grupo industrial, recibe una empresa de tamaño promedio de ese grupo industrial. Si se van a localizar más empleos, se recalculan los índices, tomando en cuenta los cambios surgidos por la primera localización. De nuevo la zona con el marcador más alto, recibe una empresa.

Este proceso continúa hasta que se ha localizado todo el empleo. Una vez distribuido, el empleo es convertido en uso del suelo, mediante coeficientes de absorción de terreno.

Los marcadores de las zonas (S_j^k) se calculan de la manera siguiente :

$$S_j^k = \sum_i W_i^k I_j^i$$

donde : K : número de grupo industrial.
j : zona.
i ; factores de localización importantes para el grupo industrial k.
S_j^k: marcador de zona para el grupo industrial k.
W_i^k: medida del factor i para el grupo k.
I_j^i: índice para el factor i en la zona j.

Para determinar las medidas del "iésimo" factor para cada uno de los 11 grupos de empleo, fue utilizado el análisis de regresión..

Entre las variables consideradas están :

- localización de otras empresas industriales en

el área.
- Acceso de ferrocarril.
- Acceso de carretera.
- Terrenos vacantes.
- Bibliotecas, Restaurantes, Densidad, etc.

Los expertos en localización industrial alterarán las medidas en base a juicios, cuando se considere necesario.

Un índice para el "iésimo factor" en la zona j se calcula de la sigte. forma :

$$I_j^i = (X_j^i - MIN^i) / (MAX^i - MIN^i)$$

donde : I_j^i : índice para el factor i en la zona j.

X_j^i : magnitud actual del factor i para la sub-área. j.

MIN^i : valor mínimo de X para todas las j.

MAX^i : valor máximo de X para todas las j.

En el modelo BASS, la migración industrial intra-área es reconocida e incorporada en el modelo. La tasa a la cual emigra el empleo de una zona es función de la densidad de la zona.

La densidad está definida como la suma de población y empleo por acre. Si la densidad de una zona es mayor que 10, entonces alguna industria emigra del área sobre la base de datos históricos.

La migración de empleos es añadida al grupo de nuevos empleos a ser distribuidos para cada industria. Los terrenos dejados vacantes por la migración, son añadidos al inventario de terrenos disponibles en la zona.

Cuando se proyecta una declinación en el empleo en algún grupo industrial, se utiliza una técnica diferente para determinar donde debe decrecer el empleo.

El porcentaje del total de decline en un grupo industrial a ser distribuido en cada zona se calcula por la sigte, ecuación :

$$PLDECL_j^k = (EMP_j^k \cdot DEN_j^{\frac{1}{2}}) / \sum_m (EMP_m^k \cdot DEN_m^{\frac{1}{2}})$$

El número de empleos perdidos del tipo k en la zona j se determina por :

$$EMPLOS_j^k = PLDECL_j^k \cdot (1.2 \Delta EMP^k)$$

donde : $EMPLOS_j^k$: número de empleos del tipo k perdidos en la zona j.

$PLDECL_j^k$: porcentaje de declinamiento en la industria k en la zona j.

ΔEMP^k : declinamiento proyectado en empleos de tipo k.

DEN : densidad.

Como $\sum_j EMPLOS_j^k$ = 120 % del total de declinamiento proyectado en la industria k, BASS distribuye un 20 % equivalente, utilizando el modelo de crecimiento industrial.

3.- SUB-MODELO DE LOCALIZACION 3 : F.I.R.E. : FINANZAS, SEGUROS, INMUEBLES, EDUCACION Y GOBIERNO :

La descomposición geográfica para este algoritmo localizacional es :

- Condados.
- Sub-áreas importantes en cada condado.
- Sub-áreas residuales.

Estos tipos de empleo son localizados en los condados de acuerdo a las proporciones actuales, con modificaciones en base a juicios.

La misma combinación de historia y sitio es utilizada para distribuir el empleo a las sub-áreas importantes dentro de los condados. El empleo restante es distribuido a las otras sub-áreas sobre la base de su población.

4.- SUB-MODELO DE LOCALIZACION 4 : SERVICIOS :

El empleo de servicio es distribuido a las zonas como una función de un índice de atractividad relativa y de la población.

$$ALLOC_j^s = \left(\alpha RAI_j + \beta \frac{POP_j}{\sum_i POP_i} \right) \Delta EMP^s$$

donde : $ALLOC_j^s$: nuevo empleo de servicio localizado en la zona j.

RAI_j : índice de atractividad relativa.
POP_j : población en la zona j.
ΔEMP^s : nuevo empleo de servicio en el área.
α, β : constantes exógenas con suma 1.

El RAI es obtenido a partir de una ecuación determinada mediante análisis de regresión.

Se estiman diferentes ecuaciones de RAI, incluyendo diferentes variables independientes para :

- Comida, bebida.
- Servicios personales.
- Servicios misceláneos.
- Servicios médicos.

SUB-MODELO DE LOCALIZACION DE VIVIENDAS :

La simulación de localización de vivienda puede verse como un paso hacia una réplica explícita del mercado de vivienda.

El modelo separa los tratamientos de oferta y demanda.

En el lado de la oferta, BASS introduce innovaciones importantes tales como la demolición y filtración de las viviendas.

Se supone una tasa de demolición de la vivienda existente determinada en forma global para toda el área (4 % de la vivienda existente por 5 años, obtenido en investigaciones sobre la tasa de demolición en San Francisco en años recientes).

La tasa de demolición no es constante para todos los tipos de casas. La distribución de las demoliciones depende del valor de las unidades de vivienda de la zona, la proporción de estas unidades que pertenecen al tipo multifamiliar y la densidad de desarrollo de la zona. Para viviendas unifamiliares, la proporción de nuevas demoliciones en la zona se determina de la manera sigte. :

$$DR_j^s = (DD_j)^{1/4} \cdot (PM_j^{\frac{1}{2}} / HV_j)$$

donde : DR_j^s : proporción de demolición de unidades unifamiliares en la zona j.
DD_j : densidad de desarrollo en la zona j.
PM_j : unidades múltiples expresadas como un porcentaje de la vivienda total en la zona j.

HV_j : valores de la vivienda en j ó (2.valor alto + valor medio)/total de unidades de vivienda.

La parte de demolición de unidades múltifamiliares depende de la proporción de viviendas solas :

$$DR_j^m = (DR_j^s)^{\frac{1}{2}}$$

La demolición afecta la oferta (cambiando el número de viviendas disponibles) y la demanda (alterando el número de habitantes en busca de viviendas.

La filtración del inventario de viviendas no altera el número de viviendas disponibles, pero cambia la distribución del valor.

El modelo supone que en cada período de tiempo, el 20 % de toda la vivienda de valor alto se convierte en vivienda de valor medio, y el 20 % de toda la de valor medio, se convierte en vivienda de valor bajo. En este caso, no se hace distinción entre unidades uni ó multifamiliares.

La tasa de filtración difiere entre zonas. La tasa de filtración de unidades unifamiliares en la zona j depende del porcentaje de unidades multifamiliares en j, y el valor de la vivienda en j :

$$FIL_j^k = S^k (PM_j^{\frac{1}{2}} \cdot HV_j)(H_j^k)$$

donde : FIL_j^k : unidades netas filtradas de la vivienda del tipo k en j.

S^k : factor escalar para la vivienda del tipo k.

PM_j : porcentaje de unidades multifamiliares en j.

HV_j : valores de la vivienda en j ó (2.Valor alto + valor medio)/total de unidades de vivienda.

H_j^k : número de unidades de vivienda del tipo k en j.

La tasa de filtración de unidades multifamiliares en una zona depende solamente de los valores de la vivienda.

$$FIL_j^k = S^k (1 / HV_j) (H_j^k)$$

La oferta potencial total de la construcción de nueva vivienda en una zona, depende de la oferta de terrenos, pendiente, atractividad, distribución del valor de la vivienda, proporción de unidades unifamiliares y densidad.

La oferta de terrenos es la suma de los acres agrícolas, terreno vacante utilizable y terrenos vacantes por migración o demolición.

La densidad de una zona está definida como el promedio de la densidad (definida externamente) de la zona y las zonas circundantes.

La construcción potencial de nueva vivienda es particionada en unidades potenciales uni y multifamiliares y luego, en unidades de valor alto, medio y bajo.

La primera partición es realizada promediando dos relaciones : el porcentaje actual de unidades unifamiliares y el porcentaje potencial. El porcentaje potencial está determinado por la densidad de desarrollo de la zona.

La segunda partición depende de tres proporciones : la partición de valor existente, la densidad de desarrollo y la pendiente del terreno.

Los coeficientes de absorción del terreno utilizados para determinar el número de unidades potenciales de nueva vivienda, varían con el tipo y valor de la vivienda, y entre zonas.

La demanda total de nueva vivienda es la suma de las familias forzadas a cambiarse por causa de la demolición, y el incremento de familias proyectado por el modelo de población.

La demanda es particionada en base a juicios, en tipos de vivienda y se supone un declino secular en la proporción de unidades unifamiliares.

El último paso en el modelo de distribución residencial es, la determinación de la localización de la construcción de nueva vivienda. La proporción de construcción potencial de nueva vivienda desarrollada en una zona, depende del acceso al empleo.

VISION PANORAMICA :

La innovación más interesante de BASS es, la introducción de la demolición y filtración en el modelo del mercado de vivienda. No hace intentos para modelar el comportamiento detrás de estos procesos, pero trata de simular los resultados.

Los cambios en el inventario de vivienda tienen grandes implicaciones en la forma urbana y estos cambios usualmente son omitidos de los modelos urbanos.

BASS localiza a las familias sobre la base de un acceso global al empleo. Un enfoque alternativo es, distribuir los empleados de un sitio de trabajo particular sobre la base de acceso a ese sitio. (VER BATS).

El rasgo más inquietante de BASS es el uso repetido de ecuaciones arbitrarias, para distribuir el empleo y las familias. Estas ecuaciones no son estimadas vía técnicas estadísticas, sino que son definidas en base a juicios.

Aunque esta metodología puede dar buenas predicciones, probablemente añade poco el entendimiento de las determinantes del comportamiento de la forma espacial de las áreas urbanas.

* * *

BATS :

(BAY AREA TRANSPORTATION STUDY).

INTRODUCCION :

LA Comisión del Estudio de Transporte del Area de la Bahía ha introducido un estudio en tres etapas, acerca de planes alternativos de uso de la tierra y transporte para el área en referencia.

- La primera etapa, es un extenso inventario del empleo, población, uso de la tierra y patrones de tráfico.
- La segunda, envuelve el desarrollo del modelo para la evaluación de una amplia gama de usos alternativos de la tierra y redes de transporte. El objetivo de esta etapa es, limitar las selecciones ó preferencias a un pequeño número de alternativas factibles.
- En la tercera etapa, estas alternativas son evaluadas de manera más detallada. Esta etapa ha de generar un juego más complejo de productos (outputs), los cuales permiten una evaluación más detallada y desagregada de las redes de transporte.

BATS reconoce tres problemas de localización :

- Localización del empleo básico,
- Localización del empleo de servicio, y
- Localización de viviendas.

La primera localización es manipulada por un modelo de localización del EMPLEO BASE - BEMOD (BASE EMPLOYMENT ALLOCATION MODEL). Los otros dos problemas son manipulados en un modelo de proyección del USO DE LA TIERRA - PLUM (PROJECTIVE LAND-USE MODEL).

PLUM requiere el "producto" del BEMOD (Ej. : la localización de todo el empleo base) para localizar el empleo de servivio y las áreas de vivienda. La estructura general del modelo es mostrada en el esquema anexo en la página sigte. En este momento, ni el BEMOD ni el PLUM han sido completados, sin embargo, ya ha sido recibido suficiente material de William Goldner , Director de Investigación de BATSC, para describir se estructura actual en detalle considerable.

METODOLOGIA :

LOCALIZACION INDUSTRIAL :

El empleo es considerado de servicio, si su localización espacial es determinada por la localización espacial de las familias, los mercados y las concentraciones diarias de población.

El empleo es básico, si su localización espacial es determinada por rutas de transporte inter-regional, recursos, vínculos inter-industriales y economías de aglomeración.

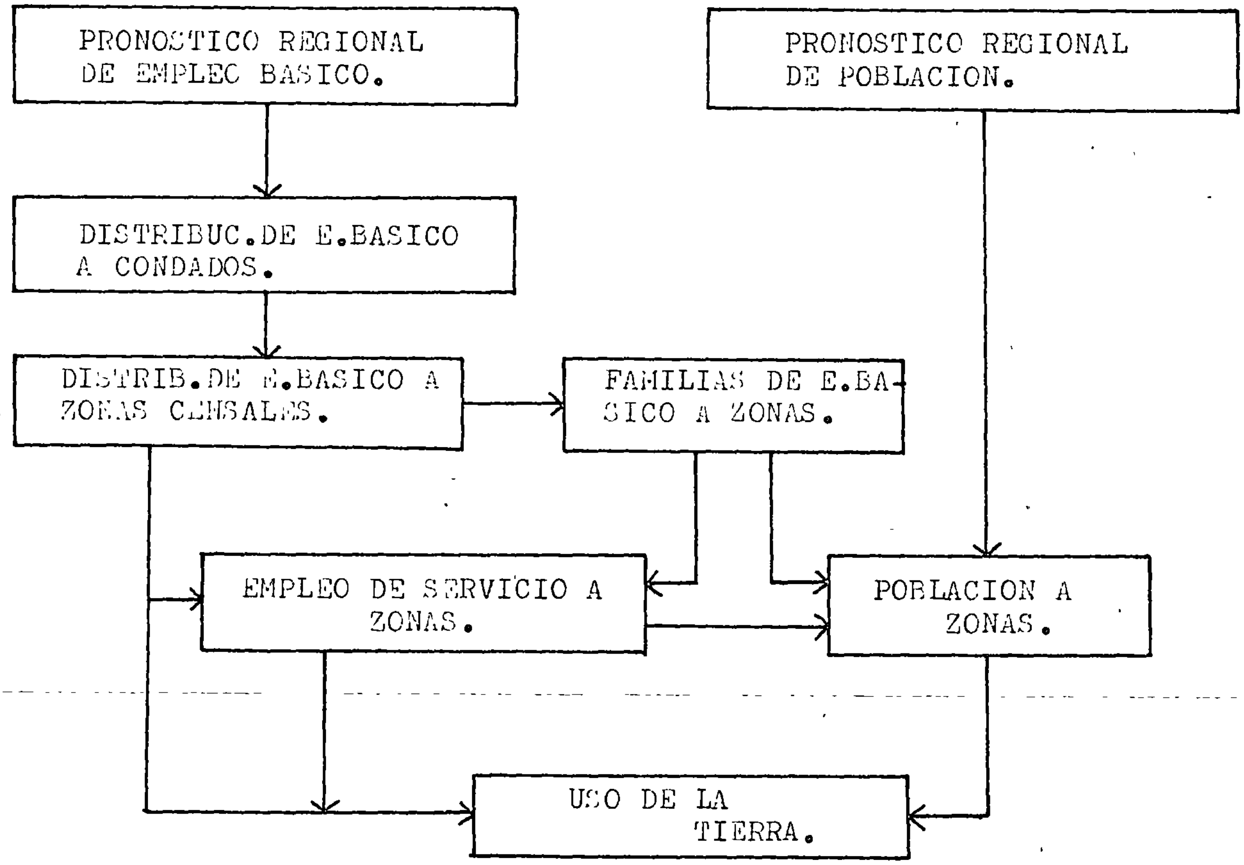
En el BEMOD, el empleo básico está dividido en 8 grupos industriales (dividiendo a su vez en sector manufacturero en 5 sub-grupos) :

- MANUFACTURA (industrias de nueva tecnología, oficina central, industrias intermedias, metales fabricados, petroquímicas).
- TRANSPORTE.
- VENTA AL POR MAYOR.
- COMUNICACIONES.
- OFICINAS.
- GOBIERNO ESTADAL Y FEDERAL.
- AGRICULTURA.
- MINERIA.

La localización espacial de cada uno de estos grupos industriales, envuelve un proceso en dos etapas :

- 1 - Los incrementos de empleo industrial son ubicados a Condados, utilizando un modelo de "repartición" y "cambio". Entonces, estos incrementos son distribuidos entre áreas censales, utilizando análisis de regresión.

DIAGRAMA SINTETIZADO PARA EL MODELO LATS.



El modelo de repartición y cambio requiere proyecciones del crecimiento del empleo de cada una de las industrias básicas para la región de la Bahía. BEMOD utiliza análisis regresivo en los datos 1.950-65, para estimar las desviaciones del crecimiento de las industrias en el Condado. Las variables independientes en este análisis son :

- Densidad,
- Rata de crecimiento,
- Acceso intra-regional, y
- en algunas industrias, el empleo.

Además de esto, es utilizada una rutina de juicio especial, a fín de ubicar la localización única del empleo. Ejemplos de estas localizaciones únicas son : colegios, universidades, campos al aire libre.

El producto de esta fase son, los empleos totales para - cada uno de los 12 grupos industriales básicos, en cada uno de los 9 Condados del área de la Bahía. (Ver esquema de la pág.sigte.).

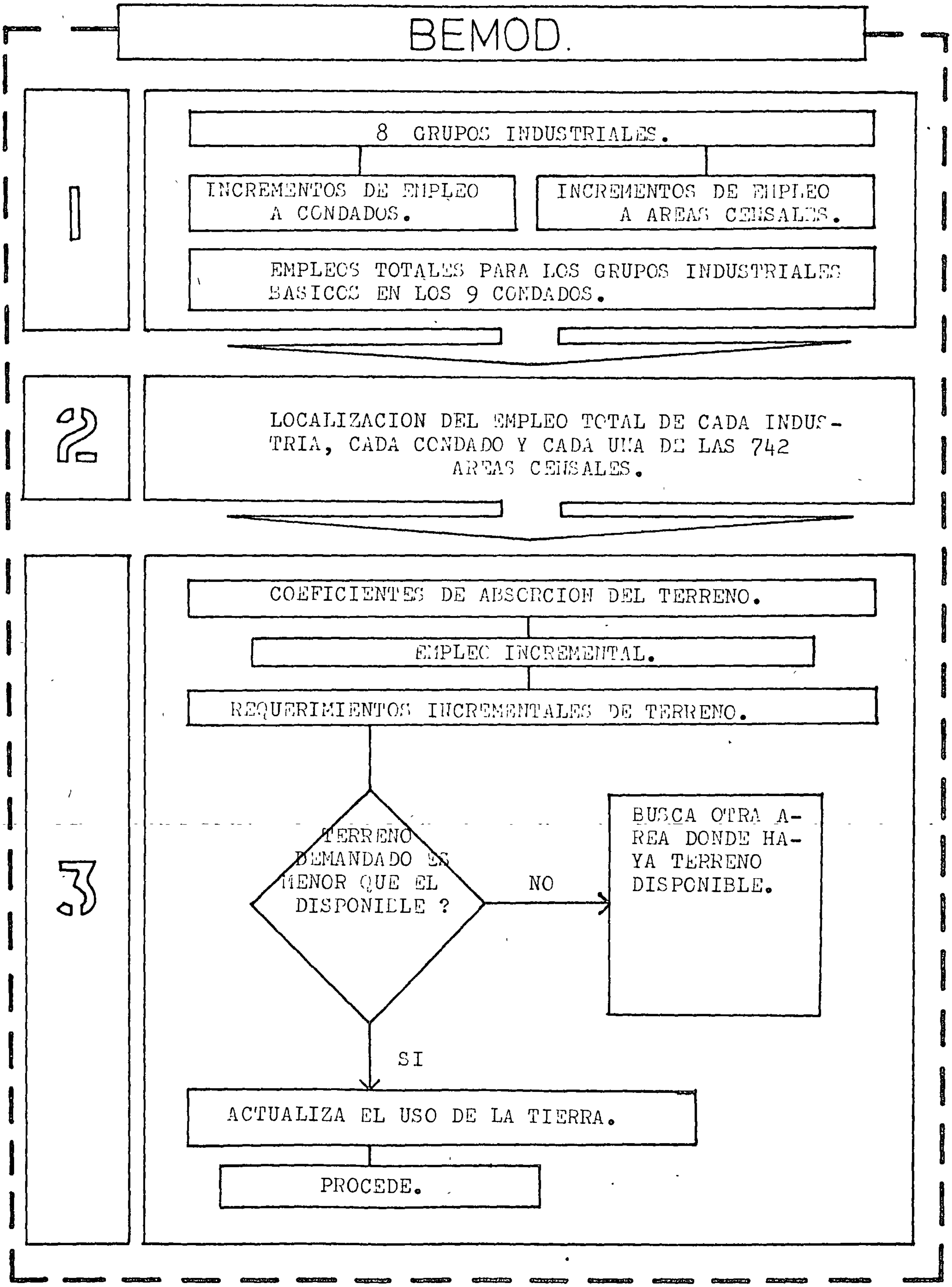
- 2 - La segunda fase de BEMOD, localiza estos totales de Condado, a cada una de las 742 áreas censales, La rutina se basa en un análisis regresivo de sección cruzada, utilizando datos de 1.964. Cada uno de los grupos de empleo básico utiliza como variables independientes, algunos sub-juegos de las sigs. 8 variables :
  - Pendiente (0-5%).
  - Elevación principal del área.
  - Presencia de agua.
  - Presencia de vías férreas.
  - Accesibilidad a la población 1.965.
  - Densidad de empleo.
  - Uso de la tierra.
  - Repartición del área y empleo del Condado.

Los valores "β" producidos por la regresión, son mantenidos constantes durante el período de proyección 1.968-90.

La variable dependiente de esta regresión es :

$$Z_{ij} = (E_{ij} / L_j) / (E_{ik} / L_k).$$

BATS A:


BEMOD.
1
8 GRUPOS INDUSTRIALES.
INCREMENTOS DE EMPLEO A CONDADOS.
INCREMENTOS DE EMPLEO A AREAS CENSALES.
EMPLEOS TOTALES PARA LOS GRUPOS INDUSTRIALES BASICOS EN LOS 9 CONDADOS.
2
LOCALIZACION DEL EMPLEO TOTAL DE CADA INDUSTRIA, CADA CONDADO Y CADA UNA DE LAS 742 AREAS CENSALES.
3
COEFICIENTES DE ABSORCION DEL TERRENO.
EMPLEO INCREMENTAL.
REQUERIMIENTOS INCREMENTALES DE TERRENO.
TERRENO DEMANDADO ES MENOR QUE EL DISPONIBLE ?
NO
BUSCA OTRA AREA DONDE HAYA TERRENO DISPONIBLE.
SI
ACTUALIZA EL USO DE LA TIERRA.
PROCEDE.

donde : i : clase de industria básica.
j : área censal.
k : condado.
E : empleo.
L : territorio total ocupado por la industria básica.

El valor proyectado de "Z" es sustituido entonces, en la ecuación sigte. :

$$\Delta E_{ij}^{t+1} = Z^{*}_{ij} \Delta E_{ik}^{t+1} (L_j^{t} / L_k^{t}).$$

donde : Z^{*}_{ij} : es el valor estimado.

El empleo es dado por :

$$E_{ij}^{t+1} = E_{ij}^{t} + \Delta E_{ij}^{t+1} .$$

donde : el dato t indica un período de tiempo.

Los coeficientes de absorción del terreno utilizados para convertir empleo incremental en requerimientos incrementales de terreno, son específicos.

Si el terreno incremental demandado, es menor que el terreno disponible en el área, EMOD simplemente actualiza el uso de la tierra de empleo y procede. Si el incremento es mayor, el cambio de empleo de la industria con el "Z_{ij}" más bajo, es removido del área y es localizado en otras, donde haya disponibilidad de terreno.

EMPLEO DE SERVICIO Y VIVIENDAS - PLUM :

Tanto la localización del empleo de servicio como la de viviendas, son determinadas por PLUM, el cüal utiliza las proyecciones de empleo y las localizaciones del empleo base.

Además, PLUM hace uso de información exógena de agencias locales de planificación.(Por Ej. : información para la determinación de los límites máximos en la cantidad de tierra disponible).

La idea básica detrás de todas las localizaciones del PLUM es, que allí está alguna función que da la probabilidad de que un individuo trabajando en "i" vivirá a "t" minutos de "i" ó comprará en un comercio a "t" minutos de "i". La función de distribución asumida para todos los casos es la sigte. :

$$P_t = e^{a-\beta/t} .$$

donde : P_t : probabilidad de un individuo viviendo a menos de "t" de su lugar de trabajo.

A fín de determinar la probabilidad para algún intervalo "t" a "(t + K)", es necesario evaluar la diferencia entre la probabilidad acumulada en (t + K) y t.

Formalmente esto es :

$$P_{(t,t+k)} = P_{t+k} - P_t = e^{a-\beta/(t+k)} - e^{a-\beta/t} .$$

donde : $P_{(t,t+k)}$: probabilidad de un individuo viviendo en el intervalo t a (t + k).

Estas funciones son ajustadas separadamente para cada uno de los 9 Condados, con datos obtenidos de entrevistas realizadas en las viviendas. Las funciones de vivienda- trabajo, vivienda- comercio y trabajo- comercio, fueron estimadas separadamente.

Las funciones estimadas para los condados, son aplicadas entonces a las zonas del condado, a fín de derivar tres matrices de probabilidades para cada uno de los viajes a cada zona.

En lugar de ubicar el empleo de servicio, PLUM hace uso de una variante en "la base de técnica múltiple". En lugar de relacionar empleo de servicio con empleo base, PLUM lo relaciona con la población base, esto es, empleo básico más las familias de los empleados básicos. PLUM obtiene lo último, distribuyendo los empleados básicos en zonas residenciales, de acuerdo a una matriz de probabilidades de distribución vivienda a trabajo "P_5", y aplicando luego, la proporción histórica de la población que no está trabajando en cada zona.

Estas dos etapas son :

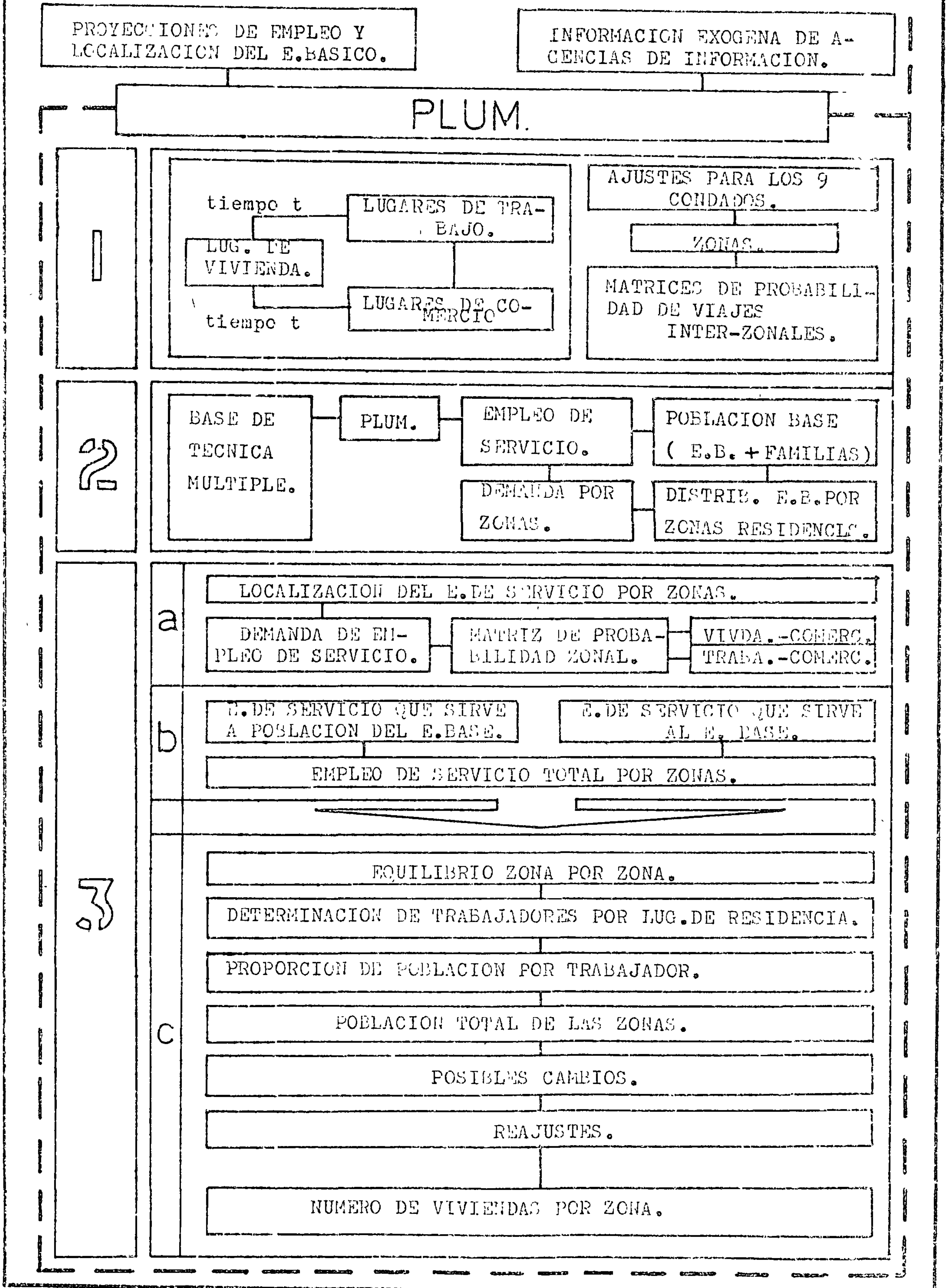
$$r_1 = P_5 \, e_1 .$$

donde : r_1 : vector de viviendas de los empleados básicos por zona.

P_5 : matriz de las probabilidades trabajo a vivienda por zonas.

e_1 : vector de empleados base por zona.

BATS B:


PROYECCIONES DE EMPLEO Y LOCALIZACION DEL E.BASICO.
INFORMACION EXOGENA DE AGENCIAS DE INFORMACION.
PLUM.
1
tiempo t
LUGARES DE TRABAJO.
LUG. DE VIVIENDA.
LUGARES DE COMERCIO
tiempo t
AJUSTES PARA LOS 9 CONDADOS.
ZONAS.
MATRICES DE PROBABILIDAD DE VIAJES INTER-ZONALES.
2
BASE DE TECNICA MULTIPLE.
PLUM.
EMPLEO DE SERVICIO.
POBLACION BASE (E.B. + FAMILIAS)
DEMANDA POR ZONAS.
DISTRIB. E.B.POR ZONAS RESIDENCI.
3
a
LOCALIZACION DEL E.DE SERVICIO POR ZONAS.
DEMANDA DE EMPLEO DE SERVICIO.
MATRIZ DE PROBABILIDAD ZONAL.
VIVDA.-COMERC.
TRABA.-COMERC.
b
E.DE SERVICIO QUE SIRVE A POBLACION DEL E.BASE.
E.DE SERVICIO QUE SIRVE AL E. BASE.
EMPLEO DE SERVICIO TOTAL POR ZONAS.
c
EQUILIBRIO ZONA POR ZONA.
DETERMINACION DE TRABAJADORES POR LUG.DE RESIDENCIA.
PROPORCION DE POBLACION POR TRABAJADOR.
POBLACION TOTAL DE LAS ZONAS.
POSIBLES CAMBIOS.
REAJUSTES.
NUMERO DE VIVIENDAS POR ZONA.

$$q_1 = (L - I) r_1 .$$

donde : q_1 : población de empleo base no trabajadora.
L : matriz diagonal de población por empleado. y por zona.
I : matriz unitaria.

El multiplicador es determinado entonces como :

$$K = E_3 / (1_{e_1} + 1_{q_1}) .$$

donde : K : multiplicador base.
E_3 : empleo no básico total.en el área (obtenido exógenamente).
1_{e_1} : empleo base total.
1_{q_1} : población total relacionada con el empleo base, que no se encuentra trabajando.
1 : vector unitario.

Aplicando el multiplicador base a la población no trabajadora relacionada con el empleo base, a zonas residenciales y empleo base a zonas de empleo, genera demandas para el empleo de servicio por zonas :

$$d_{4.1} = K_{q_1} .$$

$$d_{2.1} = K_{e_1} .$$

donde : $d_{4.1}$: denota el vector de demanda por zonas para la población que sirve a la población base no trabajadora.
$d_{2.1}$: vector de demanda por zonas para el empleo de servicio que sirve al empleo base en sus lugares de trabajo.

Se asume que el mismo multiplicador genera ambas demandas, tanto de la vivienda, como del trabajo.

El próximo paso es, localizar el empleo de servicio por zonas. Esto se realiza multiplicando los dos vectores de demanda para el empleo de servicio, por la matriz de probabilidades zonales para viviendas a comercio y trabajo a comercio. Tal como se ha descrito anteriormente, cada una de estas matrices de probabilidades es derivada de funciones separadas de localización. Los cálculos se muestran como sigue :

$$e_{4.1} = P_4 d_{4.1}$$
$$e_{2.1} = P_2 d_{2.1} ,$$

donde : $e_{4.1}$: vector del empleo de servicio que sirve a la población relacionada con los empleos base (poblac.no trabajadora).

$e_{2.1}$: vector del empleo de servicio que sirve al empleo base.

P_4 : matriz de probabilidades de vivienda a comercio.

P_2 : matriz de probabilidades de trabajo a comercio.

Finalmente, el empleo de servicio total por zonas es obtenido por sumatoria del empleo basado en el trabajo y en la vivienda.

$$e_3 = e_{4.1} + e_{2.1} .$$

donde : e_3 : empleo de servicio total.

Estos valores son equilibrados zona por zona, de acuerdo a las proyecciones del área, suplidas exógenamente a PLUM.

$$C_{(1)} = E_3 / 1_{e_3} .$$
$$e_3' = C_{(1)} e_3 .$$

donde : 1_{e_3} : suma del vector del empleo de servicio.
e_3' : vector ajustado del empleo de servicio.

El vector del empleo total en los lugares de trabajo, es obtenido por adición del empleo de servicio ajustado al empleo base determinado exógenamente.

$$e_6 = e_1 + e_3' .$$

donde : e_6 : vector del empleo total en su lugar de trabajo.

Dado el empleo total en su lugar de trabajo, es posible ahora reaplicar la matriz de probabilidades de vivienda a trabajo , P_5, y determinar los trabajadores por lugar de residencia, r_6 .

Aplicando la proporción de población por trabajador por cada zona, L, es posible determinar la población total no trabajadora, q_6, por lugares de residencia. Estas operaciones son mostradas de la manera sigte. :

$$r_6 = P_5 e_6 .$$

$$q_6 = (L - I) r_6 .$$

Es necesario entonces, ajustar esta población total no trabajadora con la dada exógenamente al modelo. El factor de corrección es aplicado a cada población no trabajadora de las zonas, y la población no trabajadora ajustada es añadida a los trabajadores por lugares de residencia, para determinar la población total de las zonas.

Estos pasos son sintetizados como :

$$C_{(2)} = Q_6 / 1_{q_6} .$$

donde : Q_6 : área total de la población no trabajadora determinada exógenamente.

1_{q_6} : suma de las zonas de población no trabajadora.

$$q_6' = C_{(2)} q_6 .$$

donde : q_6' : vector ajustado de la población no trabajadora.

$$n_6' = r_6 + q_6' .$$

donde : n_6' : vector ajustado de la población total.

Un cambio en la población en cada zona, sin cambio en el número de trabajadores, sugiere un cambio en la población por trabajador (L). Luego, asumiendo un tamaño familiar constante, los cambios sugerirían un cambio en los trabajadores por vivienda, (F). Este es el último valor ajustado que es utilizado para calcular el número de viviendas en cada zona.

$$h = F r_6 .$$

donde : h : vector de viviendas en cada zona.
r_6 : vector de empleados base por lugar de residencia.
F : matriz diagonal de viviendas por empleado base.

Con el empleo base, el de servicio y las viviendas localizadas por zona, el próximo paso es, aplicar coeficientes de absorción de terreno a cada una de las actividades y llevar los "records" del uso de la tierra. El terreno no utilizable es sustraido primero del terreno total. Incluye tierras naturalmente inadecuadas, tales como, tierras inundadas ó con demasiada pendiente ó tierra en la cual se prohibe su uso debido a políticas públicas.

El uso de la tierra del empleo base es suministrado a PLUM y aceptado sin cambio. El uso de la tierra del empleo de servicio es asumido primero que el uso residencial. El terreno residual es disponible para el uso residencial.

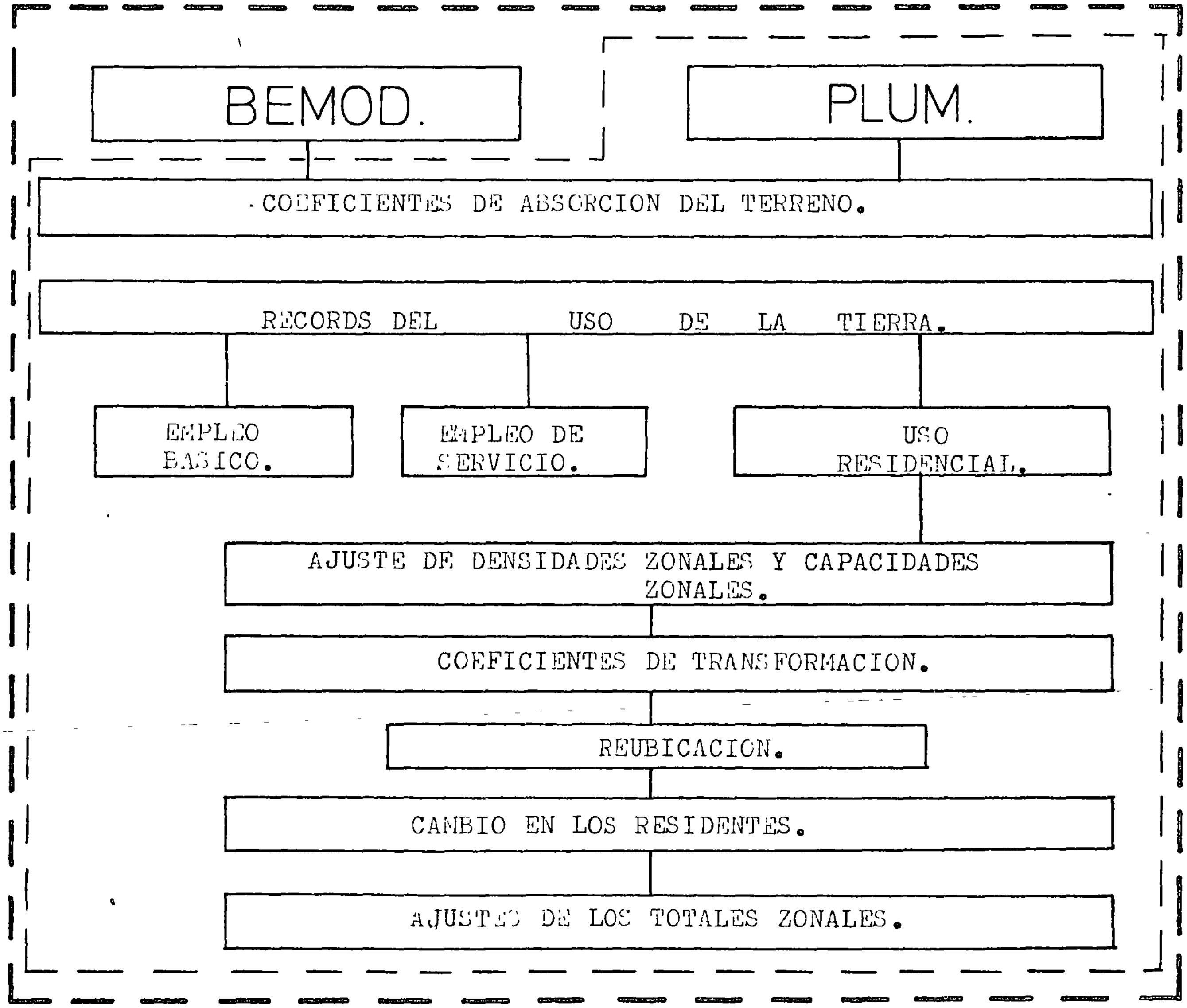
Como es aparente de la descripción de la localización de viviendas, la disponibilidad de tierra no es considerada como una obligación. Es posible que, dado el coeficiente de absorción de tierra para viviendas, puede ser distribuida en alguna zona, más tierra de la que es disponible actualmente.

PLUM tiene una rutina para ajustar la tierra localizada ó destinada a uso residencial con la tierra disponible. Primero, la capacidad presente en términos del número de viviendas es definida dividiendo el inventario presente de tierra residencial y vacante, entre la matriz de absorción de terreno residencial :

$$C^* = A_5^{-1} (a_5^* + a_8^*) .$$

donde : C^* : vector de las capacidades residenciales de las zonas.
a_5^*: vector del inventario actual de terreno residencial.
a_8^*: vector del inventario actual de tierra vacante.
A_5 : matriz diagonal de los coeficientes de absor-

BATS C:


BEMOD.
PLUM.
COEFICIENTES DE ABSORCION DEL TERRENO.
RECORDS DEL USO DE LA TIERRA.
EMPLEO BASICO.
EMPLEO DE SERVICIO.
USO RESIDENCIAL.
AJUSTE DE DENSIDADES ZONALES Y CAPACIDADES ZONALES.
COEFICIENTES DE TRANSFORMACION.
REUBICACION.
CAMBIO EN LOS RESIDENTES.
AJUSTES DE LOS TOTALES ZONALES.

ción de terreno residencial.

Luego, son definidos dos vectores de capacidad de utilización : el primero, mide la capacidad de utilización inicial y el segundo, la utilización después de que las viviendas han sido localizadas por PLUM.

$$y_i^{*} = h_i^{*} / C_i^{*} \quad .$$

$$x_i = h_i / C_i^{*} \quad .$$

donde : h_i^{*} : viviendas iniciales en la zona i.
h_i : viviendas proyectadas en la zona i.
y_i^{*} : capacidad de utilización actual en la zona i.
x_i : capacidad de utilización proyectada en i.

Mientras que los elementos del vector y^{*} deben ser, menores que ó iguales a 1, los elementos del vector x pueden ser "0" ó algún valor positivo. Cuando algún elemento de x es mayor que 1, es localizada más tierra residencial a esa zona, de la que es disponible.

En lugar de ajustar la distribucuíón espacial proyectada de las unidades de vivienda y sus requerimientos de tierra asociados con la tierra disponible en cada zona, PLUM define primero dos transformaciones de los elementos de x :

$$y^{\circ} = 1 - e^{-(e^{x_i}-1)} \quad .$$

$$y_i^{\circ\circ} = 1 - e^{-(e^{2x_i}-1)} \quad .$$

Tanto y_i° como $y_i^{\circ\circ}$, son siempre mayores que ó iguales a "0" y menores que 1. También, excepto cuando $x_i=0$, $y_i^{\circ\circ}$ es mayor que y_i°.

Las densidades residenciales zonales y las capacidades zonales, son ajustadas en el modelo para reflejar cambios en la demanda residencial. Si y_i° es mayor que y_i^{*} , la proporción de capacidad desarrollada inicialmente, la densidad residencial zonal es ajustada como sigue :

$$G_{i5}' = G_{i5} e^{m\,(y_i^{\circ} - y_i^{*})} \quad .$$

donde : G_{i5} : densidad residencial original en i.
m ! coeficiente de transformación de densidad.
G_{i5}' : densidad residencial ajustada en i.

Si y_i^o es menor que y_i^* , G_{i5} es mantenida constante.

Los coeficientes de transformación de densidades son derivados para cada uno de los 9 condados en la región, utilizando análisis regresivo de sección cruzada.

El vector G_5 es utilizado para definir una capacidad zonal ajustada, C'es $G_5'(a_5^* \ a_8^*)'$. Utilizando esta capacidad ajustada y las mediciones previamente derivadas de la proporción de la capacidad desarrollada, y' y y^{oo}, son derivados dos vectores de localización de viviendas : $h^{o'} = y^o C^*$ y $h^{oo} = y^{oo} C^*$.

El vector h^{oo} es considerado una localización "de máximo límite", y es utilizada en la sigte. forma :

$$W = 1 \frac{1h - 1h^o}{1h^{oo} - 1h^o} .$$

W, es una escala, la cual es utilizada para derivar un vector nuevo de proporciones de desarrollo zonal,

$$x' = W \frac{h^o}{C'} .$$

Los elementos de este vector son transformados entonces, para derivar las proporciones finales de desarrollo zonal :

$$y_i' = 1 - e^{-(e^{x_i' - 1})} .$$

La localización zonal de viviendas es determinada por :

$$h'' = y'C' .$$

Esta rutina de redistribución, cambia por supuesto la configuración espacial de los residentes empleados y los residentes que no están trabajando. Esto requiere la re-calculación y ajuste de estas variables, a fín de hacer consistentes los totales zonales.

VISION PANORAMICA :

Los modelos BATS han sido diseñados para ajustar dos propósitos frecuentemente conflictivos de modelación de usos de la tierra. Primero, los modelos fueron diseñados para ser un instrumento de planificación inmediatamente útil. Segundo, han sido diseñados para permitir la relativamente fácil introducción de los resultados de su programa contínuo de investigación, en el comportamiento a ser modelado.

Un área en la cual la investigación sería valiosa es, en la localización del empleo de servicio.

Parece cuestionable aplicar un multiplicador de área a todos los empleados básicos y todas sus familias. Igualmente, parece cuestionable sugerir que el mismo multiplicador se sostendría ó mantendría para trabajadores y sus sitios de trabajo y no trabajadores y sus residencias. Esta suposición implica que, los empleados hacen todos sus gastos en sus sitios de trabajo y que estos gastos generan el mismo multiplicador, así como el resto del consumo per-cápita de la familia, en los sitios de residencia. Esto es obvio que los resultados de investigaciones futuras de desagregación del multiplicador por tipo de trabajador, sitio de trabajo, ingreso familiar, tamaño familiar y localización residencial, pueden ser fácilmente adaptados por insumo dentro de PLUM.

El modelo podría incorporar fácilmente, los resultados de investigación, utilizando la misma clase de desagregación para las funciones de localización de vivienda a trabajo, trabajo a comercio y vivienda a comercio.

PLUM introduce cambios en la densidad residencial, con cambios en la extensión del desarrollo de la tierra, pero no toma en cuenta los cambios en la densidad de empleo. Esta asimetría sugiere que deben ser fructíferas extensiones mayores en esta área.

---

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE

APENDICE BIBLIOGRAFICO

DR. STEPHEN PUTMAN
DR. FREDERICK DUCCA JR.

ENERO, 1978.

Palacio de Minería Calle de Tacuba 5, primer piso. México 1, D. F.

Transpn Res Vol 9 pp 187-202 Pergamon Press 1975 Printed in Great Britain

URBAN LAND USE AND TRANSPORTATION MODELS: A STATE-OF-THE-ART SUMMARY*

STEPHEN H PUTMAN
University of Pennsylvania, Philadelphia, PA 19174 U S A

(*Received* 15 *April* 1974 *in revised form* 22 *November* 1974)

Abstract—The paper assumes that the general relationship between transportation and land-use may be defined in terms of three primary components (1) Economic activity (i e. employment), (2) Demographic activity and (3) Transportation facilities After a brief introductory section, the second section of the paper contains a review of the state of the art of models of economic activity location i e intra-urban employment forecasting The third section of the paper contains a review of the state of the art of models of demographic activity location, i e intra-urban population forecasting The final section of the paper begins with a brief description of urban transportation network models The section and paper are then concluded with a discussion of possible integrated land-use and transportation model packages

1. INTRODUCTION

The processes which determine the location of activities in metropolitan areas have been of some interest to planners and scientists of various disciplines for almost as long as there have been such areas for them to puzzle over However, it was not until the development of electronic data processing equipment and dissemination of the associated information processing capabilities that it became possible to even consider empirical testing of some of the theories which have been developed regarding the location of these activities Beginning in the 1950's with metropolitan transportation studies, there have been numerous attempts, with varying degrees of complexity, and widely varying degrees of success, to forecast these phenomena By the mid 1960's, a number of rather large efforts involving the use of computer simulation to forecast metropolitan dynamics, had been undertaken

Partly because of the over-optimistic outlook of their creators, and partly due to the unrealistic expectations of their potential users, many of these early efforts were not particularly successful The situation was further compounded by a basic antipathy towards the funding of basic research in this area, and a consequent necessity for basic research to be squeezed into purportedly on-line, applied modelling efforts One point, however, which was made by the early proponents of these techniques was that the inherent complexity of metropolitan systems and the response of those systems to exogenous influences such as the national or regional economy and/or local policy actions, was so great as to make it inevitable that these techniques would be necessary for any thorough, reasonable, analysis of the problems. That this point was true seems to be borne out by the more recent resurgence in urban or metropolitan simulation modelling endeavors

It is said that if one, or in this case many, do not learn from history then they will be condemned to repeat it This is very much the case when it comes to urban simulation models Unfortunately, many modelling efforts are very poorly documented Subsequently the details of these efforts are forgotten, and at their conclusion, with the scattering of the people who did them, a great deal of information is frequently lost It is with the hope of helping various persons and agencies to avoid repeating some of the past errors or blind alleys in urban modelling that reviews such as this are prepared Consequently, the intention here is not to provide specific point by point comparisons of various models and the precise equations by means of which their forecasts are made Rather, brief reviews of the general structures of a number of different urban models are presented Some models are presented in more detail than in others, when it seems that there are points of interest worth noting for future investigation In general, however, the intention has been to give a flavor of what has been done and some notion of the chronological order and general flow of thought in this field over the past decade

The remainder of the paper is divided into three sections The first of these, Section 2, is a description of the state-of-the-art of models of economic activity location, that is, intra-urban employment forecasting This is given first in as much as it tends to be the driving force behind household location in most urban simulation models Section 3 of the paper is a review of the state-of-the-art of models of household location The final section of the paper, Section 4, after giving a very brief description of urban transportation network models, describes the recent development of an integrated land use and transportation package

---

*Presented at the Second Intersociety Conference on Transportation Denver, Colorado, September 1973 This paper was prepared over a period of several months and was therefore partially supported by several contracts The U S Department of Transportation, Federal Highway Administration, Urban Planning Division contract No DOT-FH-11-7843 resulted in a report entitled "The Interrelationships of Transportation Development and Land Development", part of which is summarized in Section 4 of this paper Other support was under U S Environmental Protection Agency, Office of Research and Monitoring contract No R-801-373 to the Institute of Environmental Studies entitled "Review and Pilot Project in Urban Modelling", and under National Science Foundation Grant GI-38978 to the Department of City and Regional Planning, entitled "Laboratory Testing of Predictive Models An Evaluation of Cost Effectiveness"

2. INTRA-URBAN EMPLOYMENT LOCATION MODELS

Introduction

The processes which result in the spatial distribution of employment in urban areas are among the more important phenomena in the urban system. To date, with the possible exception of the retail sector, attempts to model these processes have been only partially successful. Many attempts at urban modelling sidestep the problem by assuming that large portions of, or all, urban employment activity is exogenously (to the model) determined. The basis for this assertion is frequently not well developed, often being the result of rather intuitive feelings of the model builder about basic and non-basic employment classifications.

The determination of which types of employment are to be included in the model, and how they are to be handled is a key decision, usually made early in employment modelling efforts. Consequently the question of employment types is critical, if for no other reason than that based on their definitions some sectors will be inappropriately modelled or even excluded from consideration. Many classifications stem from economic base theory and implementation of economic base studies, and as such are defined in terms of the industry's output. In this section of the paper, after a brief discussion of this classification problem, a number of the more important modelling attempts are reviewed.

If it is assumed that every employment type may be considered as performing three primary activities: (a) procuring, (b) processing, (c) distributing, then by characterizing the ways in which these functions are performed by different employment types, another basis for classification may be derived (Hoover, 1948).

The primary activities may be defined as follows:

1. The procurement aspect includes everything associated with obtaining the required material inputs and getting them to the location where they are to be used.
2. The processing aspect includes all the costs of performing whatever processing function the employment performs. For example, in a manufacturing sector it would include cost and availability of land, labor and capital for production. In a service sector, it would include the same costs but would undoubtedly be dominated by the labor cost.
3. The distribution aspect includes all the costs associated with the sale and delivery of the product or function performed.

At the same time, and for each employment type, it is necessary to consider whether any or all of these elements of activity are *intra-urban location* determinants. That is, to classify an employment type according to this scheme, it must first be determined whether the particular aspect of employment activity (i.e. procurements, processing, distribution) is of locational importance, then it must be determined if there are sub-area to sub-area differences within urban areas which will affect the performance of this activity.

Before proceeding with further discussion, a note regarding the modelling of employment location as opposed, e.g. to sales or value-added location, is in order. There are obvious disadvantages concomitant with the use of numbers of employees as a measure of activity levels, most of which stem from the nonlinearities in each employment type's employee/output relationships (i.e. inherently different production functions). Consequently, when these forecasts of economic activity, as represented by employment totals, are used in turn to generate other forecasts, certain inconsistencies are introduced into the results. Despite this employment is the measure most often used, for the single overwhelming reason that it is the only measure for which data are available. If other data were to become available, then research into their use as a more appropriate means for modelling the intra-urban location of economic activity may be justified. For the present and the immediate future, employment will have to continue to be the unit of measurement.

The following discussion assumes that (a) *the problem being addressed here is that of small-area intra-urban employment location and that* (b) *citywide totals of each employment type are exogenously determined*. These restrictions have a substantial impact on the nature of the problem. By restricting the problem to *intra*-urban location, inter-urban or inter-regional comparisons are obviated: that is, many location decisions consist first of selecting a region or a metropolitan area in which to locate, and second, of selecting a specific site within the chosen area. The discussion prosented here focuses only on the latter of the two parts of the problem. The first part of the problem is an important part of *regional* employment models and is discussed elsewhere (Putman, 1969, 1970). In addition, it is assumed that within the urban area, the size of the sub-areas being considered is relatively small, being of the order of census tracts.

The portion of the proposed classification scheme which deals with intra-urban differences with regard to distribution, refers to the location of "markets" and the importance of their location to the employment sector being considered. Existing models of urban employment activity are frequently classified by the sectors which they attempt to model, e.g. retail, manufacturing. Drawing upon the discussion above, it can be seen that an alternative classification of existing models can be made by analyzing the means by which "markets" are considered in them. Under this scheme, *there are some employment types whose location within the urban area will be market-sensitive and some whose location will not*. The models reviewed here are considered according to this classification.

Models of Non-Market-Sensitive Location

First, it must be recognized that the models discussed here are not necessarily non-market-sensitive in their constructs, but rather they are considered by their authors to be capable of producing forecasts of non-market-sensitive employment location.

Most intra-urban employment models have as their operational procedure a "continuous function allocation" process. These models, with few exceptions, have basically market-sensitive constructs, they are mentioned here because they sometimes include non-market-sensitive employment types as well as market-sensitive types. One exception is the LINTA model (Seidman, 1969, pp. 135–138), developed to forecast the location of manufacturing employment (which is predominantly non-market sensitive) by small area. This model first spreads the regional growth rate of a given activity over all sub-areas. Then, based on sub-area specific measures of "attractiveness," the model reallocates the growth to different sub-areas in proportion to their attractiveness. While tests of this model were moderately successful, its high level of sectoral aggregation and its inability to consider discrete facility location are important deficiencies. This is particularly the case where (e.g. in a wholly urban region) the economic attributes (i.e. determinants of growth) of sub-areas are sufficiently similar as to result in the model's being capable of forecasting only small regular changes in all sub-areas rather than the discrete irregular changes which so often occur.

Another continuous function allocation model designed specifically for forecasting non-market-sensitive employment is Nathanson's BEMOD (Nathanson, 1970). This model was developed at the Bay Area Transportation Study to provide basic employment inputs to their PLUM population and non-basic employment model (BATS, 1968). Based upon certain attributes of each sub-area, a sector (employment type) specific score or index is calculated for the sub-area. When all sub-area scores have been computed, the regionwide growth or decline for each employment type is allocated to the sub-areas as a function of the "score" and available land. Excess allocations are handled by iterating the procedure, with modified land availabilities (Stevens and Wheaton, 1970, p. 3). This model is likely to suffer from the same problems as does LINTA while perhaps performing somewhat better over-all.

The EMPIRIC model is another continuous function allocation model, similar in principle to the LINTA model which, in fact, was developed from the EMPIRIC and POLIMETRIC model efforts. The EMPIRIC model's construct is described by Hill in the following way (1965, p. 113):

> The change in the subregional share of a located variable in each subregion is proportional to: one, the change in the subregional share of all other located variables in the region; two, the change in the subregional share of a number of locator variables in the subregion, and three, the value of the subregional shares of other locator variables.

The comments made above regarding LINTA are equally applicable here.

The Loewenstein model (Loewenstein, 1967, pp. 3–6), in the words of the author:

> combines a trend-based analysis of industrial sector growth within districts in the Philadelphia region with an allocation technique based on the attractiveness of each district relative to all other districts in the region for each industry under consideration.
>
> the critical notion of this process is that it ascertains the importance of a number of locational determinants for a variety of industries as well as the availability of these determinants in a number of areas. The areas are ranked according to the degree to which they possess these characteristics and an apportionment is made.

In this model, which was never actually made operational, the employment types were two-digit S.I.C. classes. The question of simultaneous versus sequential location of different industry types is not discussed in the model description, the impression being given that the location of each industry will be independent of the location of each other industry, thus bypassing the problem. Each of the location factors was weighted according to its importance to the locating industry. The model deals with allocation of employment increases or decreases as a continuous function allocation, and as such there is no constraint on the allocations except that of employment density by sub-area. This does present the possibility, though operationally it may never occur, that over longer projection periods a uniform density by industry types may be the predominant characteristic of industrial location as forecast by this model. There is no consideration in the model of the possibilities of discrete facility location. While it probably would not be useful to try to implement this model, consideration of the list of factors which are postulated as influencing intra-urban industrial location for use in another model might prove rewarding.

In summary, some continuous function allocation models are capable of producing reasonably good forecasts of non-market-sensitive employment location in some circumstances. However they lack the ability to deal with discrete facility location. This is likely to be a significant problem when dealing with non-market-sensitive employment types, and on the occasion where they attempt to forecast significant changes in the location of non-market-sensitive industry types, these models are open to criticism.

The most promising direction for future development of models of non-market-sensitive location appears to be that taken by those few models which attempt to forecast discrete facility locations. Several such models are discussed in detail below.

Goldberg describes his model in the following way (1967, p. 1):

> [It] functions by identifying for each industry group a set of location factors it needs in finding a site. For each site (in the case at hand for each census tract) the factors are measured and added together to give a score for each census tract for each industry group. The tract with the highest score for a given industry group receives a unit of employment. The scores are then usually recalculated and the allocation continues as above until all projected employment is allocated.

In this model the employment classes were aggregations of two-digit S.I.C. classes into six manufacturing subgroups. One of the typical problems of this type model is the sequential rather than simultaneous treatment of industry location. In this model the problem was side-stepped by grouping the industry types such that a different set of relevant location factors (of a total set) could be defined for each. This avoided the problem of competition between sectors by precluding it, and at the same time, vitiated the need to weight the location factors since only the "relevant" factors are considered in any case. The most serious problem with the model construct is its unit of allocation. When there is a forecast region-wide increase in a given employment type the model allocates this increase, to subareas in the region in clumps called "units of employment " Each unit is equal to the average facility size for that type employment in the region This represents a relatively unsatisfactory compromise between continuous function allocation (i e any number of employees of a given type to a sub-area) and discrete facility location (i e location of discrete facilities in a sub-area according to a known size distribution) Finally this model handles employment decreases in a rather peculiar fashion The regionwide percentage decrease for a given employment type is multiplied by 1 2 (i e 120%), and a continuous function allocation of the decreases is made, where the function used is a total activity density index, weighted by the specific employment types, of the sub-area The excess decrease of 20% resulting from the above is then allocated back to sub-areas as if it were a normal 20% regionwide increase of that employment types This procedure implies that those sub-areas which, if a region's employment of a given type were growing would grow the most, are also the ones that, if the region's employment of that type were declining, would decline the least While this is not an unreasonable assumption, the documentation of the model is not clear as to why this particular procedure was adopted Despite these shortcomings, some of which stem primarily from data inadequacies, the model has been made operational and has produced reasonable forecasts

Another discrete facility location model, Putman's INIMP model, is described by its author (1967, p 203),

> uses an allocation procedure for describing a portion of projected city-wide employment changes (positive or negative) among existing facilities and upon reaching certain critical values of the magnitudes of allocated increases or decreases switches to a location technique for the addition or deletion of facilities comprising the remainders of the city-wide projections

In this model, too, the employment types were two-digit S I C classes, and the areal unit was the census tract

The allocation portion of the model is a simple continuous function allocation procedure Increases or decreases dealt with in the allocation portion of the model are simply proportional to the portion of the city's employment of each type in each area While this did not produce unreasonable results when the model was used, this is definitely a weakness of the model This portion of the model ceases to function after certain capacity constraints in terms of S I C specific levels of employment density have been met

The location portion of the model, which takes over when the capacity constraints are exceeded, locates clumps of new employment in areas according to a set of measures which indicate the likely preference of the industry for the area, based upon known past location behavior The size of these clumps is determined stochastically from the prior distribution of facility sizes for each industry Location of clumps of employment decrease occur when the city-wide percentage decline of a particular industry reaches certain magnitudes called "shutdown percentages" When this occurs, facilities are selected, stochastically depending upon the magnitude of the percentage decrease, for deletion Again in the discussion of the model no mention is made of the question of sequential versus simultaneous location and it is a fault in the model construct that the model follows a sequential (by S I C number) location procedure

The flexibility of this model construct is greater than others, but consequently, it is significantly more difficult to establish the parameter values (e g capacity levels and shutdown percentages) In addition the model uses only four measures of area unit acceptability This number should probably be increased Despite these failings the model is conceptually quite complete and offers significant promise for future development

The Rose model (1967, pp 5–6) is summarized by the statements that,

> rather than analyzing only the net change in employment in a zone we are attempting to explain that change specifically in terms of its components natural growth of existing firms "death" and "birth" of firms and movement of firms within the region
>
> only "births" and "moves" are locationally relevant decisions—size changes (including "deaths") are a function of the industrial "climate" of the region, the type of industry the size of the firm—with few exceptions (e g Urban Renewal) the characteristics of the locations have little to do with the behavior of locationally static firms

The employment sectors used in this model, a total of fifty, were approximately the same as are delineated in the national input-output tables Though the model was not described in an operational form, it appeared to attempt simultaneous location In a later document (Rose, 1968) the allocation portion of the model is described as a linear programming formulation However, the number of sectors (fifty) and the number of areal units (647) indicates that operational simultaneity in this model may be very nearly impossible

In this model all employment forecasts are treated as discrete facility locations What is of particular interest is the means by which the facilities to be located are generated First, based upon regional employment projections and the base-year firm size distributions, regional projections of the firm size distribution are made The base-year firm size distribution is then "aged" via a Markov process, yielding an updated firm size distribution The difference between this distribution and the exogenous regional projection is the estimate of firms which were "born" or which "died" during the period If the data and model construct were available, firms making intra-area movements would be estimated here also Then, via a mathematical programming algorithm, an allocation to zones is made of the new or relocating firms In conclusion, while the model was not described in an operational form, and does appear to have only discrete facility location (i e no continuous function allocation of even small employment increments), the idea of facility "birth" and "death" via a Markov process is worth considering as an area for future research

Models of Market-Sensitive Location

Virtually all existing operational or potentially operational models of market-sensitive location are continuous function allocation models. Future progress in this area is most likely to come from further development of one or more of these models. As many of these models are rather similar in their theoretical constructs, the reviews which follow attempt to cover the principal constructs rather than describe all such models.

The retail market potential model (Lakshmanan and Hansen, 1965) takes as its premise that the size and number of retail establishments in an area is a function of the number of consumers as represented by the magnitude of their retail expenditures. This model is descended from an earlier one (Lakshmanan, 1965) and has as its central theoretical construct a measure of market potential derived from the traditional gravity model construct (Huff, 1962). This construct, in the words of the authors (1965, p. 135)

> . . . states the retail center of zone *j* attracts consumer dollars (from consumers from zone *i*)
>
> (a) in direct proportion to the consumer expenditures
>
> (b) in direct proportion to its size
>
> (c) in inverse proportion to distance to the consumers
>
> (d) in inverse proportion to competition

The model construct implicitly assumes that retail activity will locate solely as described by the above market potential function, nevertheless, it does produce reasonably good forecasts. Two important considerations are, however, omitted. First, there are other influences on retail activity location which will cause location deviations from those forecasted by the above construct. These other factors should be included in a realistic model of such location. Second, and perhaps of less importance in already developed areas, is the question of minimum facility size. That is, certain types of market-sensitive activities have important scale economy considerations which will dictate minimum sizes for new facilities. Future development of models such as this should include these considerations.

The EMPIRIC model (Hill, 1965) discussed above in the section on models of non-market-sensitive location, may also be considered to be a model of market-sensitive location. The sets of variables, affecting location, tested in the model were

Set I	Intensities of land use (densities), Zoning practices, Automobile accessibilities (within the region)
Set II	Same as Set I plus Transit accessibilities
Set III	Same as Set II plus Quality of water service, Quality of sewage disposal service

The results of the calibration tests of the model, in terms of variance "explained", were acceptable, and the model may be considered as an acceptable, operational, forecasting technique, subject to the restrictions mentioned previously.

The Harris retail trade model (1964) was developed as a component of the Penn Jersey Activities Allocation Model. Derived from selected portions of central place theory, the model attempts to model the equilibrium which results from the supply of goods and services, the demand for them and the pattern(s) of interaction between them. The model, as described by the author (1964, p. 6)

> works in the following way. Given the location of demand and any set of interaction parameters, a computer program is provided with a preliminary estimate of the location of supply. The interaction formulas are then applied and result in a hypothetical set of "arrivals" at the supply points. . . . the result of this first estimate of "arrivals" is substituted for the original hypothesis as to the distribution of supply, and the process is reiterated.

While acceptable forecasting results have been obtained from this model, the estimation of its parameters (which may be highly collinear) is a more than normally difficult problem. Because of this, the model cannot be considered to be a truly practical forecasting device. Its explicit consideration of supply and demand equilibrium is, however, unique and should be considered for future development.

The Lowry model (1964) as well as its derivatives (e.g. Crecine 1964) considers the location of market-sensitive employment simultaneously with the location of population. This model is also a market potential model, but the formulation of the concept is somewhat different. Lowry describes it as follows (1964, p. 10)

> The distribution of this retail employment among the . . . tracts depends on the strength of the market at each location. Assuming that shopping trips originate either from homes or from workplaces, the market potential of any given location can be defined as a weighted index of the numbers of households in the surrounding areas, and the number of persons employed nearby.

This model, too, has been shown to be capable of producing acceptable forecasts, and its parameters have been successfully estimated for several different areas. The model should be considered to be an operational forecasting technique.

The Bay Area Simulation Study developed several continuous function allocation models to deal with market-sensitive employment of different types (CREUE, 1968 pp. 179–234). For retail trade an estimate is made of site suitability as a function of site attributes (including accessibilities). Expected retail trade demand is developed from a simple market potential expression. The expected demand is compared to the actual demand (existing employment) to estimate potential "new" demand. Finally, a relative attractiveness index based on a combination of site suitability and potential "new" demand is calculated. The allocation of retail trade employment is then made in proportion to each sub-area's relative attractiveness. Services employment is allocated proportional to a relative attractiveness index and the population percentage in each sub-area. Finance *et al.* and Government employments are allocated by a function of past employment and population. Contract construction was allocated in proportion to new houses and new employment. In summary, these BASS employment submodels represent operational but highly empirical continuous-function allocation procedures for forecasting the location of market-sensitive employment. It would be desirable to try to develop a more general, less cut-and-try, type of procedure to make these forecasts.

It can thus be seen that there are continuous function allocation models which produce quite acceptable forecasts of market-sensitive location. All but two of these, however, omit consideration of scale economies. The Lowry model considers them in a rather crude fashion and the Harris model considers them explicitly. While this is not a critically important failing, it should be considered in future efforts of this type. A more important problem is that many of these models include employment activities such as office employment as market-sensitive when in fact they may not be. This, of course, reflects back to the classification question posed earlier. For certain employment classes these models appear to be quite adequate. For other classes, other models are necessary. Thus a model developed for forecasting all employment types would have to be capable of both continuous function allocation and discrete facility location along with the ability to determine which procedure is appropriate under what circumstances. Such a model is described in a recent work by this author (Putman, 1972).

Finally, it should be noted that while some urban residential models also forecast market-sensitive "non-basic" employment, many do not. No urban residential model includes forecasts of all employment, and in this sense, all are dependent on exogenous employment forecasts, perhaps from the types of models reviewed above, as input. In the next section of this paper the important urban residential models are reviewed.

3. INTRA-URBAN RESIDENTIAL LOCATION MODELS

Introduction

The second major component of urban activity is that of population or residential location. Most urban land use simulation models have given primary emphasis to these activities, with rather less emphasis being placed on the problems of employment location. Some of these models have, however, included retail-commercial employment location procedures and, as such, were discussed in the previous section of this paper. This section of the paper will discuss only models or those components of models which have been solely devoted to forecasting population and residential location.

As was the case in the previous section, it will be helpful to begin by attempting to classify these models in a way which allows a systematic discussion and review of them. One commonly used classification, which seems to work reasonably well, is that of behavioral models vs. non-behavioral models. The essence of a behavioral model is that its allocation procedure is in the form of an equation or set of equations which purport to describe, more or less explicitly, residential location behavior in terms of the actual determinants of that behavior. In contradistinction non-behavioral models attempt this same description principally by virtue of the correlations which may exist between and among a large number of measures of attributes of the urban system.

The behavioral models may be further divided into those which are macro-descriptive and those which are micro-descriptive. The macro-descriptive models attempt to describe phenomena observed at a rather aggregate level of detail. An example of macro-description of a behavior would be the gravity model, which seems to describe a variety of social phenomena without a genuine correspondence between its formulation and the observed behavior of individuals. The micro-descriptive models attempt to describe the same general phenomena by aggregating the results of the decision making processes of many individuals. An example of this can be found in some of the work done at the University of North Carolina on residential location (Donnelly, *et al.*, 1964). All but one of the models to be reviewed in this section can be classified as behavioral and, as will be seen, tend toward the macro-descriptive rather than the micro-descriptive.

Behavioral Models

The class of behavioral models of urban residential location is clearly dominated by the Lowry model and its many derivatives. One model in this class which is not a Lowry or Lowry derivative model should be discussed first. The Lathrop and Hamburg model, which has not received the further testing, examination, and modifications as have followed upon the Lowry work, has as its central locating concept, the notion of intervening opportunities (Lathrop, *et al.*, 1965). Simply stated, the concept first asserts that a trip maker will encounter alternative opportunities for satisfying his trip purpose as he travels. Second, the concept asserts that there is a finite probability that the tripmaker will stop at any of these possible alternatives. Further, this probability increases at each successively encountered alternative, with each prior alternative not taken. The implementation of this model requires that a measure of opportunities be constructed for each zone in the region. In addition, some measure of zone-to-zone travel time must be constructed, and based on this, each zone is rank ordered (according to this measure), from each other zone. It is then a rather straightforward matter to calculate the probability of a trips terminating in any zone from each given origin.

The opportunities measure used in the earlier development of the model was simply the product of the available land times the existing density of activity, both of which were exogenous inputs. In a later form of the model, a procedure was added to calculate the density endogenously, based on distances from the center of the region. Having calculated the opportunities measure, the amount of activity allocated to each zone is then calculated as a function of the probability of stopping in that zone during a hypothetical passage through the zone.

The model may be operated in either of two modes. It is capable of allocating all the residential activity to an essentially featureless region, i.e. the development of an equilibrium forecast, or it may be used to allocate increments of new activity to an existing base population of a region. Like the Lowry model, to be described below, the model is entirely macro-descriptive and seems to be a reasonable construct for this type of forecasting. Despite the fact that it has not received the attention and additional work that the Lowry model has, it should not be disregarded as an alternative, supplementary, description of residential location

The Lowry (1964) model with its numerous derivatives and descendents is certainly the most well known model of this type. Both the model and its derivatives have been described by Goldner (1971) and Batty (1972) The essential ingredients of the Lowry construct are that the location of residences is a function of the location of employment and the trip making behavior of employees In general, this procedure involves the use of the existing locations of employment and assumptions about work trip behavior along with a measure of their spatial separation in order to generate the spatial distribution of the residences of the employees The model also contains a procedure for the location of the retail-commercial employment which has been reviewed in the previous section of this paper, on employment models

The household spatial allocation function in the Lowry model is quite simple, being solely a function of the accessibility of potential residential locations to employment The function used asserts that the number of households which will locate in a particular zone j is directly proportional to the sum of employment in all the surrounding zones weighted by a function of the distance between each of those zones and zone j This allocation is subject to two external constraints First, the sum of all the households in all the zones must equal a given regional total Second, the density, in terms of households per acre, may be constrained to be below certain predetermined levels in certain zones The distance function used in the Lowry model is simply a negative exponential function of the straight line distance between the zones

The operation of the Lowry model, in essence, generates an instant metropolis Given the exogenously provided distribution of basic employment, the model allocates the population and non-basic employment to the various zones of the region These allocations represent an equilibrium situation which (assuming the model to be correct), would eventually come to pass if all other factors had remained constant while the equilibrium was being achieved Consequently, the model is not (nor does it purport to be) an actual forecasting procedure This is because it is not possible to associate points in real time with the model's solution, since it is not clear if and when such an equilibrium solution would come to pass Actually, the solutions produced by the model bear a strong resemblance to reality or to anticipated reality, and as such, the model provides useful insights into urban spatial processes In addition, as will be seen below, it has opened the way to a great deal of further research

One of the first Lowry derivative models and one of the most substantively significant was the Time Oriented Metropolitan Model (TOMM) This model was developed as a component of a large comprehensive model system for use in Pittsburgh, Pennsylvania (Crecine, 1964) Unfortunately, the system was never completed and the model was never calibrated The prototype versions of the model were, however, tested with fabricated data A number of later efforts have attempted to calibrate various portions of this model and have obtained rather interesting results

The substantive differences between this model and the Lowry model are important to note First, and most important was the fact that since it had been designed to produce actual forecasts it was necessary for the TOMM model to have a real time orientation. Consequently, the model was developed in an incremental form as opposed to the static equilibrium form taken by Lowry In the TOMM model the base year distributions of all activities are included as determinants of the projection year's distribution of activities The essential notion here, and the difference from the Lowry model, is that neither all households, nor all non-basic employment are free to move (change location) in any given projection period.

The second important difference between this model and Lowry is that TOMM attempts to disaggregate the locating population into several types In the TOMM model the allocation of households, by types, to zones, is done in two steps The first step consists of allocating total households to zones, and the second step involves disaggregation of the total households into household types The first allocation of total households, is rather simple and, like Lowry, is strictly a function of access to employment This allocation is done in terms of residential density First a series of land use accounting equations estimate the land available in each zone for new residential location Then, the model estimates as a function of access to employment, the projected residential density in each zone Using this density and the available land, the numbers of households are calculated By subtraction from the base year's households, the new household increment is determined

The access to employment function is exactly as in Lowry The estimated residential density in zone j is directly proportional to the employment in all the surrounding zones weighted by a trip index between those zones and the locating zone j The trip index is even simpler than that used in Lowry, being simply the straight line distance between each zone and zone j There are several constraints which operate on the allocation function, the first of which is a constraint on the maximum residential density in each zone The second constraint is the requirement that a minimum number of households per zone, corresponding to the stable households, are not allowed to move away from the zone in the projection perion Finally, there is a constraint such that the sum of all households in all zones must be equal to externally provided regional totals Allocations which violate the constraints, as in allocations which violate the density constraint in the Lowry model, are reallocated to other zones

The second step in the TOMM model allocations is the disaggregation into household types of these zonal projections of total households At this point a very important substantive difference between TOMM and Lowry appears The TOMM model, in contradistinction to Lowry, uses a measure of local amenities to determine this distribution of household types The function used, though an extremely simple one, postulates that the number of households by type in each zone is proportional to an additive function of the base year distribution of household types in the zone, and the work trip propensities of each household type. Finally, the sum of all the types of households in a given zone is constrained to equal the previously estimated numbers of total households in the zone.

Several later versions of the TOMM model were constructed. One of the first of these involved simply deleting the household work trip propensities part of the zonal household type allocation (CONSAD, 1967). In this case, the numbers of households by type in each zone are simply determined as a function of the distribution of all household types which had existed in that zone in the base period. In that version, there was a matrix of coefficients of household type-to-household type attractiveness indices. In yet a later version of the model, the household type projections were postulated to be a function of the base year distribution of all household types in that zone, and a measure of population potential and employment potential which appear to be some form of accessibility measure (Lee 1968). A number of existing proprietary model packages have one or another of these earlier forms of the TOMM model as their core component. In these, as in the case of the earlier TOMM developments, while some of the model parameters have been estimated, no truly rigorous calibrations of the models have been completed.

A number of years after the publication of the first version of TOMM, a much more sophisticated version of the model, referred to as TOMM-III, was released (Crecine, 1969). The principal difference between this version and earlier versions was a substantial increase in the complexity of the household-type allocation functions. In this model, the household type allocation is a function of ten independent variables. These variables, used in an additive linear function include three accessibility measures, housing prices, lagged household type change two indices of public facilities, an index of structural deterioration, a measure of household demand potential, and information about the base year's distribution of household types in this zone. This is an enormous list of independent variables and raises strong doubts about the likelihood of ever completing a successful calibration. Though some work has been done in this direction, it is clear that a good deal more would need to be done should someone decide to try to really calibrate this model.

Another small family of Lowry derivative models are the various forms of the Projective Land Use Model (PLUM). This family of models, which evolved from the great tangle of modelling projects undertaken in the San Francisco Bay Area in the early to mid-1960's, began with a first form almost exactly like that of the Lowry model (Goldner 1968). This first form of the model, which we shall refer to as PLUM, had in its population allocation procedure only one significant substantive difference that being the use of a rather different allocation function from Lowry's. This allocation function is of some interest and it will be useful to describe it, along with some notions of the theory implicit in these access-attractive types of allocation function, as used in the various Lowry models.

First recall that these models concern themselves with a set of trip makers and their work-to-home or home-to-work trip making behavior. Taking for example a particular origin and assuming all other variables to be equal, the theory postulates that the trips to any given destination will be proportional to the difficulty of reaching the destination and the degree to which that particular destination is capable of satisfying the trip purpose. The difficulty of reaching the destination is expressed in terms of travel time or cost. The degree to which the particular destination is capable of satisfying the trip purpose is usually expressed in terms of some measure of attractiveness, opportunities, or quantities of attractors located at the destination. Two possible variables often included in the "all other variables" class are particularly important and are therefore sometimes included in these formulations. The first of these is the question of intervening opportunities as discussed above, and the second (really just a different form of the first) is the possibility of competition amongst alternative destinations. Both of these concepts have appeared in different models in various ways.

The allocation function used in these models may be thought of as having two components. The first component is the probability of making a trip, for a given trip purpose, of a particular length. The second component is the measure of attractiveness of the destination. In the PLUM model the probability of making a trip of length t is inversely proportional to the length squared and directly proportional to an exponential function of the negative reciprocal of the length. This function is applied, in the allocation of residences, to annular rings around each given origin zone. First the probability of making a trip from the given origin to a given annulus is calculated, and then this probability is divided amongst the zones in the annulus. Based on this procedure, a matrix of trip probabilities from each zone to each other zone is calculated. With the use of a scale factor, these probabilities are applied to the zonal employment to produce the distribution of residences. These allocations, as was the case in the models previously discussed, are subject to several external constraints regarding densities and land availabilities. The PLUM model used no attractiveness measure.

An interesting aspect of this model, though probably not theoretically correct, raises a general question which is worth noting. Given the situation of forecasting from some known base year to some future year, the PLUM model is first used to make its own estimate of land use allocation for the base year. These estimates are then subtracted from the known base year distribution and the residuals, which are the variances not captured by the model's equations are saved. The model using its own estimates of the base year as inputs then makes a forecast of the future year allocations. To this forecast, the previously calculated residuals are added in an attempt to preserve the unexplained variance from the base year. The interesting question raised here has to do with the fact that all of these models, if run recursively over several time periods, tend to lose or "wash out" the unexplained variations from their forecasts. Thus, one might observe more systematic variation from time period to time period as a long run, involving several recursions, of one of these models was made The question of the importance of the variance not captured in the model equations tends to be overlooked in discussions of the use of these models This is an area where further research is necessary not only (as is currently the case) towards reducing the amount of unexplained variance, but also towards coping with such unexplained variance as inevitably will be present

A more recent version of the PLUM model, which will hereafter be referred to as IPLUM, was developed as an incremental model In this form, the basic employment inputs are changes from the base year to the projection year, and all of the model's projected allocations are also in terms of changes from the base year (Goldner, 1972) These changes are added to the base year variables to produce the future year projections The allocation function in IPLUM is based on the same exponential expression as was used in PLUM, used in a somewhat different way An apparent attempt to introduce the notion of intervening opportunities is added to the IPLUM allocation algorithm. In this case the probability of traveling from a given origin zone to a given destination zone is equal to the probability of traveling to the ring (annulus) in which the destination zone is located, divided by the travel time from the origin to that ring

Additionally, in the IPLUM model a measure of attractiveness of the possible destinations is introduced This measure is basically a measure of residential holding capacity, a rather simple function similar to that used in the Lathrop–Hamburg model reviewed above The attractiveness is simply the product of vacant acreage times the residential density, both of which are base year inputs Having calculated the attractiveness, or as it is called in the model, the opportunities measure, it is used to weight the probabilities which are calculated in the trip probability function described above This modified trip probability matrix is then applied to the zone employments to generate the zonal residential allocations

A problem which occurs when using an available land type of opportunities measure results from the fact that zones at the fringe of urban areas tend both to be large and to have large quantities of vacant land The use of vacant land as a measure of attractiveness usually overstates the development potential of these zones This is due to the fact that in order for the development to become a reality, some level of what is known as urban infra-structure facilities, such as water and sewer lines, as well as electricity, must be available In many urban fringe zones these facilities tend not to be available and therefore to slow development to a degree not properly represented by a simple vacant area measure A remedy for this, for forecasting purposes, is to modify the opportunities measure by some function which in some way measures the developability of these fringe areas One such measure, used by Goldner, was described in terms of the proportion of developable land which has already been developed A further variation of this same idea seems to have produced improved forecasts in a much modified version of IPLUM used in the model system described in the last section of this paper Finally, none of the PLUM or IPLUM models has a truly endogenous procedure for disaggregating household types This disaggregation is done in some forms of the IPLUM model by an *ad hoc* procedure which operates after all of the allocations have been completed

The last model which will be discussed as one of the Lowry derivatives is the residential model from the Bay Area Simulation Study, BASS In particular, the comments which follow will refer to the BASS-IV version (CREUE, 1968) This model is of considerable importance in the evolution of urban land use models All of the models discussed previously are strictly demand oriented models Thus their forecasts of residential allocations are functions solely of household demand, and are only somewhat constrained, in a rather arbitrary manner, by the availability of residential land It is clear however, after even cursory consideration of the dynamics of the urban land market that the supply of housing and/or residential land will have a significant affect on the ultimate resolution of demand for residential location The BASS model was the first urban land use model to incorporate explicitly both the supply and the demand sides of the problem within the allocation procedure

In the BASS model the residential housing demand and residential housing supply are calculated separately and then matched, one to the other, to determine the final location The demand for housing, in essence, is the product of an accessibility measure and a measure of the potential supply of housing The accessibility is essentially that of residence to work place The problem in attempting to describe this procedure is that the model makes extensive use of constructed variables the calculations of which are embedded in a large, rather complex, model system The accessibility measure begins by calculating an employment measure by zone, a weighted function of both base year employment as well as a projected employment increment This employment measure is in turn used in conjunction with an exogenously provided measure of the potential supply of housing to produce what is called a measure of the proportion of total housing demand employed in a particular tract This measure is used, in turn, to calculate a further measure of accessibility to employment which involves a declining linear function of travel time from the origin zone and to the possible destination zones, used to weight the employment measure in the destination zones

Despite the confusion surrounding this model, due largely to rather inadequate documentation, the reasoning behind this procedure can be deduced and may be described as follows First note that this accessibility measure is ultimately to be used in a housing demand calculation Thus it is possible to consider the first employment measure calculated, as a measure of potential demanders of housing, including both existing employees and new employees who have their place of work, say in area *i* Subsequently, the next measure calculated may be considered as a measure of aggregate potential demanders of housing in area *i* This demand results from the previously calculated demanders, adjusted by the spatial distribution of the potential supply of housing in surrounding areas Finally, the accessibility measure for any area *i* is the result of the spatial distribution of the adjusted measure of demanders, again for all surrounding areas, and thus the accessibility measure is one of accessibility to potential demanders of housing

The measure of potential supply of housing in the BASS model is also constructed in a series of somewhat obscure steps The principal inputs involve the product of a measure of the relative suitability of land in the area for housing times the measure of the "holding capacity" of the developable land The suitability measure involves a normalized measure of existing nearby housing stocks further adjusted by the slope of the land in the area and the proxy measure of land value The measure of developable land includes previously existing vacant land as well as land made available by demolition of old housing This entire product is then further modified by exogenous measures of attractiveness, all of which results, at best, in a rather fuzzy notion of exactly what is being calculated in this part of the model Finally, the measure of potential supply of housing is multiplied by the previously calculated accessibility measure to obtain what is called the demand potential This demand potential is then normalized and converted to housing demand by zone The demand is then matched with the calculated housing supply by zone, with any excess housing demand being saved for subsequent iterations of the model

Again it should be noted that the BASS model occupies a pivotal position in the stream of development of urban land use models All the previously described models develop a measure or measures, of the attractiveness and accessibility of each zone and subsequently allocate households to these zones in proportion to the measure There is no consideration in these models as to the availability of housing for the locating households While a few efforts have been undertaken to model the housing supply, none of these had been integrated with models of the housing demand A notable supply side model was the Arthur D Little study of the San Francisco Housing Market a rather complex simulation of housing filtering, construction, and demolition (Wolfe, 1967) The Herbert-Stevens model, described below, includes housing supply as one of the determinants of household location, but does not attempt to generate estimates of housing supply within the model Consequently, the BASS model, despite many shortcomings, particularly that of somewhat incomplete empirical support for an extremely long chain of assertions in its theoretical development, was the first operational model to link explicitly the supply side to the demand side Even now, after a decade, very few other models have attempted to deal with this rather critical problem

A model with a very similar overall structure to the BASS model is the relatively recent NBER model (Ingram, *et al* 1972) While its discussion at this point is chronologically out of sequence, it is appropriate to discuss it here because it is a step towards a more micro-descriptive model and it is a somewhat better integration of all of the necessary components for a complete model of the urban housing and household system At the most general level, the notion of generating a spatial distribution of housing supply and a spatial distribution of housing demand and then matching them is contained in both BASS and NBER This notion is also implicit in the Herbert-Stevens model which predates both of these, but which accepts the housing stock as exogenous input along with an exogenously supplied household preference function The Herbert-Stevens Model will be discussed below

In the NBER model, the procedure for developing the spatial distribution of housing supply consists of three principal steps First, a vacancy sub-model estimates the housing vacancies by zone which develop as a result of households vacating their current housing Second, the filtering sub-model estimates changes in housing quality both negatively due to aging as well as positively due to renovations and repair The net result of this process is a distribution of expected prices for various housing types by zone Third, the supply sub-model for simulating demolitions, constructions, and structural transformations further modifies the distribution of housing stocks The net result of these three sub-models is the spatial distribution of housing by type and zone which is available for occupancy by the housing demanders It should be noted that while individual segments of these submodels are rather different from actual parts of BASS, there are many points at which the models are quite similar

The demand estimating portion of the NBER model also contains three sub-models First, the employment location sub-model translates exogenously developed estimates of changes in employment location into numbers of employees by place of work, income class, and education class Second, the mover's sub-model generates estimates of the numbers of housing demanders These are the sum of inter-metropolitan movers as well as new household formations Finally, the demand allocation sub-model allocates the housing demanders to the various housing types That is it determines the distribution of kinds of housing which will be demanded by the various types of households on a regionwide basis

The final sub-model in the overall system is the market clearing sub-model This sub-model is analogous to the supply and demand matching sub-model in BASS and is rather similar to the essential components of the Herbert-Stevens Model This sub-model utilizes a linear programming algorithm to find a least transport cost allocation of housing demanders working in zone *j* to housing supplies in zone *i* This allocation is done for one housing type at a time with the minimization being a function of the transport costs Transport cost is computed by mode for each income class (household type), by each residential zone, and each work-place zone The cost is the weighted sum of out-of-pocket travel cost plus travel time, which is converted to cost by weighting with the implicit wage rate of the income class for which it is being computed Having found each of these least cost distributions, the model is concluded Housing prices are up-dated and other inputs prepared for successive regressions

The proposed Herbert–Stevens model mentioned above (Herbert, J D, Stevens, B, 1960) was the forerunner of more than a decade of continuing research on developing an operational version of this concept of the dynamics of urban spatial location. The model begins by assuming an existing distribution of housing stocks as well as knowledge about personal utility or preference functions of households. The overall construct of the model is incremental in nature, the process being that of adding an increment of population to each zone. The model presumes knowledge of household budget making activity in that it makes use of a preference function to determine household location. It is assumed that each household type has a fixed budget of which a certain part must cover costs of location which include amenity cost, housing cost and transportation cost. Further, it is assumed that the household is willing to make certain tradeoffs between these. The remaining amount of money is what Harris (1963, p. 9) terms "bid-rents". This amount is what is available for land-rent. The model then uses a linear programming algorithm to locate households so as to maximize their bid-rent paying ability (not the rents actually paid however). The linear program allocation of households to the housing stocks is such as to result in a pareto-optimal distribution. That is, it would not be possible to change that distribution in order to make any given household better off without making some other household worse off with respect to the utility level received for the given bid rent.

Despite the rather brief description given above, the concepts embedded in this model are of particular importance. A considerable quantity of important research has been done on various aspects of the Herbert-Stevens model, to which the reader is referred (Harris, *et al.* 1966. Wheaton, 1973). This work includes extensions to and revisions of the theory as well as empirical estimation of the preference functions and testing of model prototypes.

The Urban Performance Model (Brown and Kirby 1971, 1971a) is an alternative macro-descriptive construct which falls somewhere between the Lowry and the Herbert-Stevens constructs. The essence of the model is that each locator makes tradeoffs between "opportunities" and "quality" and attempts to locate at a point where the highest "relative utility" is attained. This is done by first calculating measures of "opportunities" which, operationally are rather simple accessibility calculations. The opportunities are a function of access to jobs and "non-job attractiveness", where the latter turns out to be a heavily massaged surrogate measure based on non-work trip travel times. The "quality" is intended to be a measure of "the satisfactions of urban life in that location and its immediate surroundings." Operationally it appears to be a normalized measure of the proportion of different resident types in all zones, weighted by the distance from the zone of interest, *i*, to each other zone *j* and summed over all zones.

Having calculated these measures, the relative utility of location in a particular zone is calculated as the product of the opportunities and the quality, each of which is raised to an empirically derived power, called an elasticity exponent. Since the values of opportunity and quality change if a unit of activity is located in a zone, it is possible to calculate a marginal utility of location which is the difference in relative utility due to the location of a small quantity of a particular activity.

Location of activity is accomplished by calculating, for each locator, a relative preference for each zone. This relative preference is the product of the above described marginal utility and a holding capacity which is an exponential function of the product of the prior average density of development of the locating activity and an exogenously supplied quantity of land. The relative preferences are then used to rank each zone. An increment of activity is located in the zone of greatest relative preference and the entire process is iterated until all locators are located.

The documentation of this model is not sufficiently detailed to further evaluate the model, and no discussion of any empirical testing is provided. At this point it can only be said that the general construct of the UPM is somewhat interesting, but there is no indication as to its operational usefulness.

Non-Behavioral Models

While there are a number of these models which have been developed, the only one which has had sufficient use and/or testing to be worth reviewing here, is the EMPIRIC model. This model, which was discussed under both market sensitive and non-market sensitive employment models above, is also regularly used for residential forecasting. As mentioned above, the basis for such models is that there is sufficient correlation between many measures of urban attributes, and that these relationships are sufficiently stable, to allow extrapolation via this model to be reasonably accurate in most cases.

As stated above, the concept of the model

> "may be stated as follows: the change in the subregional share of a located variable in each subarea is proportional to the change in the subregional share of all other located variables in the subregion, the change in the subregional share of a number of locator variables in the subregion, and the value of the subregional shares of other locator variables." (Hill *et al.* 1966)

The model assumes an additive linear construct. Since many of the important measures of urban attributes are highly correlated and rather stable, it is usually possible to obtain reasonably good fits of this model to the data. The selection of which variables to test in fitting the model involves educated guesses as to what will work, based on the considerable prior experience which has been had in the use of this model. Criticism of the model stems principally from its lack of any genuine theoretical construct. A survey of reports based on the model's use in several cities indicates that whatever variables produce the best statistical fit, those are the ones selected for use. Often the signs of the coefficients of these variables, not to mention their definitions, do not have any reasonable theoretical relationship to what is currently known about the determinants of urban phenomena. Despite this, it must be admitted that, given the good statistical fits which are obtained, it is to be expected that this model under stable conditions, will produce reasonable forecasts for many situations. The degree to which this model's performance is either better or worse than some of the behavioral models has not yet been determined.

There are more models which could have been reviewed here. One in particular to which the reader's attention is called, has recently been completed and represents an important further step in model development (Birch, *et al.* 1973). Wilson's entropy-maximizing approach, which is more a means of analysis than an actual model, should be reviewed by prospective model users or builders (Wilson, 1970). Many other models are omitted due to lack of genuine differences from those already included, lack of documentation, lack of space in an already overly long paper, and limits on this author's time. The main stream of modelling in this area has been covered, however, along with a few peripheral excursions.

In conclusion it can be seen that there has been rather a lot of work done in this area which does indicate slow progress in unravelling the Gordian knot of urban phenomena. To the extent that prognostication is in order the Herbert–Stevens, with modifications by Harris, and the BASS-NBER streams of thought, perhaps aided by notions from Wilson's entropy-maximizing approach, are the likely source of major future developments.

In the next, and last, section of this paper a project involving the integration of a much modified version of an existing land use model with a transportation network model is described as an example of an attempt to determine some of the urban phenomena which are not captured by any of these types of model operated in isolation from the others.

4. URBAN TRANSPORTATION NETWORK MODELS AND THEIR INTEGRATION WITH LAND USE MODELS

Transportation Network Models

There are a number of computer programs or sets of programs generally known as "network packages". These generally contain procedures for accomplishing the following tasks:

1. Trip Generation: a determination of the number of trips originating from each zone of the area under investigation, often by means of multiple regression analysis based on household survey data.
2. Trip Distribution: a determination of the origin-destination pattern of the trips estimated in 1. Produces a matrix of trips from each origin to each destination, often by calibrating a gravity model to fit trip-interview data.
3. Modal Split: a determination of the number of trips from 2, which will travel on each available mode. Done in a variety of ways, sometimes independent of and sometimes jointly with 4.
4. Trip Assignment: the determination of the paths through the network(s) which will be taken by the various trips, and the subsequent assignment of the trips to those paths.

The trip-assignment component of this process is sometimes augmented by the addition of a capacity-restraint or volume-capacity function. This is an expression which describes the time to traverse a network link as a function of the design capacity of the link, its hypothetical free-flow volume, and its actual volume. The basic notion here is that as the volume on a link increases the link becomes more congested and each additional "vehicle" encounters a somewhat longer travel time for that link. When a capacity-restraint function is included in the trip-assignment procedure, it is clear that some form of iterative solution must be adopted so that the problem of some trips being assigned to an "empty" network and some to a "full" or congested network can be properly resolved. There are many ways of doing this, all of them being more-or-less arbitrary. One procedure which seems a little less arbitrary than many others involves assigning a portion of each origin's trips to all destinations and then repeating this process several times, say, loading 40%, 40%, and 20% of the trips in three passes through the assignment routine (Kuner 1973a).

All this is by way of an introduction since, in fact, very few urban land use models ever have any direct connection to the network package. The standard procedure is to utilize the "link-table" from the final network loading as the basis for finding the minimum path from each zone to each other zone. These paths are used to prepare a matrix of zone-to-zone travel times. This matrix of travel times or impedances is the usual form of transportation input to land use models.

In this final section of the paper an example of the use of an integrated land use model and transportation network package is described. The package was developed to analyze the problem of determining the feasibility of balanced development of land use and transportation facilities, and to determine whether or not land development and travel could be brought into a satisfactory balance, such that transportation improvements made could provide a continuing high level of service over time (Putman, *et al.* 1973). The essence, then, of the problem was that of controlling (i.e. altering, directing, and modifying) the spatial organization of the metropolis. In particular the concern was with spatial expansion and with the feasibility of balanced development of land use and transportation facilities. Further, if it appeared that balanced development was feasible, there was the problem of testing what means (policies) would be relevant to achieving that balance.

An Integrated Model Package

Two aspects of the problem were the principal determinants of the method of approach. The first was that of the inherent complexity of a comprehensive analysis of both transportation and land use. The second was the requirement for a self-consistent procedure for testing the sensitivity and response of transportation and land use to policies for their manipulation. Unfortunately, the state-of-the-art in simulation modelling of urban and regional phenomena continues to pose a problem for projects of this sort. In general, those models which have had operational success are, at present, rather too highly aggregated in both spatial and sectoral detail as well as in concept. Two such models, as reviewed above, are EMPIRIC and PLUM. On the other hand, those models which are more conceptually complete and which could provide highly detailed forecasts, are not yet operational

An example of one such model, also reviewed above, is the Herbert–Stevens model as subsequently modified at the University of Pennsylvania Nonetheless, it was clear from the outset, that an appropriate model package would need to be even more complete (and therefore complex) than any previous effort It was required that a complete system, i.e population and employment sectors, land use model and a complete i e. public and private, transportation network model, be assembled The overall system of models had to be highly integrated with respect to model inputs, outputs, and feedback loops. By this integration the model package would attempt to eliminate the principal failing of contemporary land-use or transportation studies, by explicitly including these feedback loops.

Typically a transportation study assumes a future land use pattern as given, and designs a transportation system to cope with it This procedure ignores the redistributive effects which are produced by the construction of the system Transportation systems obviously do not just suddenly appear, but are constructed in stages with consequent redistribution of activities all during the period of construction. The typical land use study accepts a transportation system as given, and then estimates the consequent distribution of activities while ignoring the congestive effects of that distribution on the network The model package as described above, attempts to capture the interrelatedness of the transportation system and the distribution of activities This comprehensiveness was clearly necessary to meet the research objectives of the project regarding the feasibility of balance between transportation facilities and patterns of land use, as well as the testing of policies to achieve such a balance.

The general structure of the model package was to work as follows First, the various base-year data as to the obtaining spatial distributions of activity, along with data on the characteristics of the unloaded base year transportation network were to be input These data would then be used to generate a preliminary, and probably inflated, estimate of base year trips taking place in the metropolitan area Given this preliminary estimate of trips, it becomes possible to load the future (projection year) network so that its travel characteristics i e time and cost, reflect the traffic volumes which would be on the network if there were no change in the spatial distribution of activities from the base year. These network characteristics, along with the base year data and the projection year control totals would than be used to generate a spatial distribution of activities for the projection year From this spatial distribution a new estimate of metropolitan trips would be produced These would, in turn, be loaded on the projection year transportation network The modified characteristics of the transportation networks would then be used to reallocate the projection year spatial distribution of activities. This distribution of activities would then be compared to the first estimate thereof. If there were no significant differences an equilibrium would have been reached and the model run ended. If there were significant differences, new trips would be generated and loaded on the networks and further iterations run

It is very often the case in complex simulation studies that the operational version of the model system is somewhat less comprehensive than the system originally proposed While some simplifications were necessary in the study, the principal goal of an integrated transportation model and land use model was attained The land use model used in the operational model package was the IPLUM model, which was reviewed above This model accepts the spatial distribution of "basic" employment as input and then solves for the interrelated spatial distributions of population and "non-basic" or "local-serving" employment. Regional levels of activity are exogenously determined, though this is not a particularly unusual situation, as virtually all urban land use models have this property Further, the "basic" employment location must be forecast exogenously Actually these forecasts come from a separate model, BEMOD, which is reviewed above, but BEMOD was not included in the model package. Consequently, the spatial distributions of these activities are independent of the remainder of the activities in the model package i.e. population, "non-basic" employment, and transportation networks This assumption is clearly not valid over any substantial projection period and is an aspect of the model package which should be improved. One possibility would be to integrate the modelling of all employment locations, link this to the residential model and link this overall model to the network model, but this would be very ambitious undertaking and remains several years in the future. The level of sectoral aggregation in the PLUM model is rather more gross than is desirable, but was tolerable in the preliminary development of the package

As described above, the principle determinants of spatial distributions of activities in IPLUM are access to workplaces for population, and access to customers for non-basic employment The access measures are non-linear functions of travel-times weighted by attractiveness measures of the "trip destinations" The actual allocation functions used in the IPLUM model have never been subjected to thorough statistical examination, and our early experience with model runs indicated that there was too heavy a reliance on somewhat arbitrary "correction-factors". Consequently, a thorough investigation of the model algorithms was made, and they have since been modified These modifications appear to have substantially improved the model's allocations, though attempts to evaluate these improvements have not yet been completed Of course, as is the case with all other urban land use models, IPLUM's reliability for specific small area forecasts is considerably less than that for general regional patterns

The transportation model package selected for use in the package was developed by the Planning Sciences Group at the University of Pennsylvania (Kuner, 1973) It contains programs for network coding, tree tracing, and several alternative capacitated traffic assignment procedures The principal connection going from IPLUM to the Network model is via trip generation That is, the main phenomenon which relates the spatial distribution of activities back to the transportation facilities is specified in terms of the trips between activities which make use of the transportation system Two types of trips are considered, work-trips and non-work-trips As mentioned above, in many transportation studies where land use projection models are being utilized, the work-trips are generated by separate procedures These trip patterns, being independently calculated, may not be consistent with the work-trips implicit in the land use model's patterns of residence and workplace In IPLUM the location of residences is generated by work-trip allocation functions based on transportation impedances and the locations of workplaces In so doing, IPLUM implicitly generates a matrix of work-to-home trips which, in this model package, are extracted and used as the package's work-trips Similarly, the location of "non-basic" or "local-serving" employment in the IPLUM model is generated by home-to-shop and work-to-shop allocation functions based on transportation impedances and the locations of the activities This procedure generates an implicit matrix of shopping trips which are extracted, and by a somewhat different procedure involving the use of a gravity model type algorithm for balancing the origins and destinations of the trips, are used as the package's non-work trips In the present version of the model package, peak-hour work trips and shopping trips are loaded on the networks A problem to be resolved for later versions of the model package is that of other non-work-trips besides the shopping trips Currently shopping trips are "inflated" to equal all non-work trips Additionally, the problem of mass transportation facilities was ignored There is, however, no reason to preclude its inclusion in the model package in future efforts This would involve construction of a modal-split model and a procedure for dealing with multi-modal networks in the IPLUM allocation functions Neither of these appear to present insurmountable problems for future efforts

Several alternative procedures for loading trips on the network were tested and are described elsewhere (Kuner, 1973a) The procedure used, called "Partial incremental accumulative trip method," is a modification of one developed at the Chicago Area Transportation Study. Beginning with a free-flow network, a single tree is traced (i e all the minimum paths from that origin to all destinations) Onto these paths a portion of the trips from that origin to the various destinations is loaded By use of a volume-capacity relationship, the new volumes on the links of the tree are used to calculate new travel times Then a new origin is selected, a new tree is traced, and the process is repeated When a tree has been traced for each origin, the network will have been loaded with a portion of all the trips This over-all procedure is then repeated. Currently the package loads the networks in three passes of 40%, 40% and 20% of the trips respectively. After the network loading procedures have been completed a matrix of travel times from each zone to each other zone, via the minimum path on the loaded network, is calculated for input to the land use model

The first tests of the model package were done for a small prototype urban region of eight areas, while later, full-scale runs, used San Francisco Bay area data The current form of the model system utilizes two forms of the PLUM model The first of these is essentially the PLUM "state"-variables model where the forecasts are made in terms of the values of variables at time t The second form of the model is the incremental variable version, IPLUM, where the forecasts are made in terms of changes in values of variables from time $t-1$ to time t. These changes are then added to the base values (i e those at time $t-1$) to yield the forecasts of time t variables The assemblage of models used for these runs

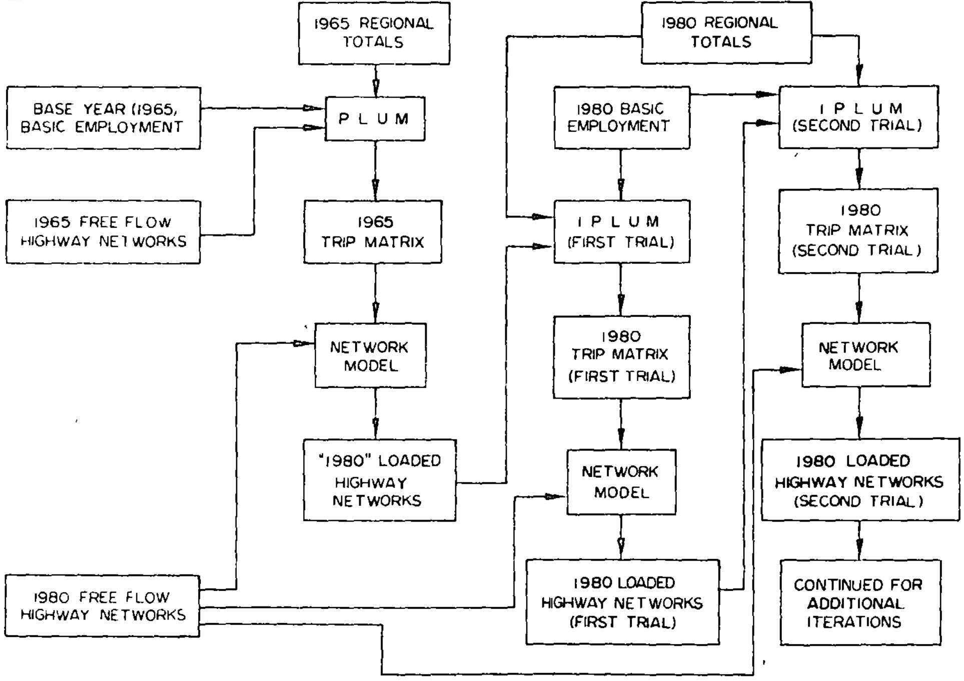


Fig 1 Schematic of experimental model runs

is shown in Fig 1 Basically the procedure involves a run of PLUM to generate base-year trip matrices for input, after loading onto a network, to the IPLUM model After this, successive iterations of IPLUM and the network model are run

If, as suggested above, important interrelationships are, indeed, customarily ignored by the typical transportation study or typical land-use study, either of which assumes fixed values for the other, then these model runs should have indicated significant and systematic differences in output from such "unintegrated" runs In particular, in terms of the degree to which accessibility shapes metropolitan areas one would have expected

(a) If land use forecasts are made based on a free-flow network or a network where flows, and therefore congestion, are significantly less than would realistically be present, the over-all spatial distribution of activities will be more dispersed than is realistic

(b) An excessively dispersed pattern of activities will produce overestimates of tripmaking and network congestion

(c) If land use forecasts are made based on overcongested networks, the over-all spatial distribution of activities will be excessively concentrated

(d) An excessively concentrated pattern of activities will produce underestimates of tripmaking and network congestion

Given these expectations, it was to be anticipated that a linked land-use and transportation model would oscillate between decentralization and centralization and consequent overcongestion and "undercongestion"

The actual results obtained from the model runs were generally as expected and produced better land use forecasts than those obtainable from a land-use model not linked with feedbacks to a transportation model This seemed to be most evident in subareas which are, in reality, experiencing rapid development Many of the individual results observed would not have been predicted by a separate land-use or separate transportation model Consequently, while there remain many areas for improvement of the model package, these results clearly demonstrated the need to approach the use of these models in a systematic way and that even though the individual models may have their deficiencies there is much to be learned from integrating them and studying the properties of the whole system

Conclusion

It is difficult, having come this far, to think of how to conclude this review In some cases a great deal of effort on a particular modelling project has been discussed here in just a few lines of text This author is at least as aware as anyone else of the advantages of hindsight in reviewing such efforts Each model builder has surely tried to do the best possible job, given the time, resources, and then current state-of-the art Perhaps in reminding model builders, and in presenting to prospective model users, of what has gone before, this work will help to improve what is yet to come

BIBLIOGRAPHY

Batty M (1972) Recent development in land use modelling A review of British research *Urban Studies*, 9, No 2, 151–177

Bay Area Transportation Study Commission (1968) *BATSC Locational Model System* BTM 217/00/00, Berkeley, California

Birch D Atkinson R, Sandstrom S and Stock L, (1973) The New Haven Laboratory A Testbed for Planning Report to the U S Dept of Housing and Urban Development, Contract H-1492

Brown A and Kirby R F (1971) *General Formulation of the Urban Performance Model* prepared by PRC Systems Science Company for the Metropolitan Council of the Twin Cities Area

Brown A and Kirby R F (1971a) Measuring urban performance *Journal of Cybernetics* 1, No 4, 32–54

Center for Real Estate and Urban Economics (1968) *Jobs, People and Land Bay Area Simulation Study (BASS)* Special Report No 6, CREUE, Institute of Urban and Regional Development, University of California, Berkeley

CONSAD Research Corporation, (1967) *Impact on Allegheny County Due to the Relocation of Residential and Commercial Activity in the East Street Valley*, prepared for the Department of City Planning, Pittsburgh, Pennsylvania

Crecine J P (1964) *A Time-Oriented Metropolitan Model for Spatial Location*, mimeo, Department of City Planning, Pittsburgh, Pennsylvania

Crecine J P (1969) *Spatial Location Decisions and Urban Structure A Time-Oriented Model* Discussion Paper No 4, Institute of Public Policy Studies, University of Michigan, Ann Arbor, Michigan

Donnelly T G Chapin F S Jr and Weiss, S F (1964) *A Probabilistic Model for Residential Growth* Institute for Research in Social Science, University of North Carolina, Chapel Hill

Goldberg M (1967) *Industrial Location Model for the San Francisco Bay Area*, paper presented at the Western Regional Science Association Las Vegas Nevada

Goldner W (1968) *Projective Land Use Model (PLUM)* BATSC Technical Report 219, Bay Area Transportation Study Commission Berkeley California.

Goldner W (1971) The Lowry model heritage *Journal of the Americal Institute of Planners* XXXVII, No 2 pp 100–110

Goldner W (1972) *Projective Land Use Model* (in three volumes Institute of Transportation and Traffic Engineering University of California, Berkeley, California

Harris B H (1963) *Linear Programming and the Projection of Land Uses* Penn Jersey Paper No 20, Penn Jersey Transportation Study, Philadelphia, Pennsylvania

Harris B (1964) *A Model of Locational Equilibrium for Retail Trade* mimeo, Institute for Environmental Studies, University of Pennsylvania, Philadelphia, Pennsylvania

Harris B Nathanson J and Rosenburg L (1966) *Research on an Equilibrium Model of Metropolitan Housing and Locational Choice Interim Report* Planning Sciences Group, Department of City and Regional Planning, University of Pennsylvania, Philadelphia Pennsylvania

Herbert J D and Stevens B H (1960) A model for the distribution of residential activity in urban areas *Journal of Regional Science* 2 No 2, 21–36

Hill D M (1965) A growth allocation model for the Boston region *Journal of the American Institute of Planners*, XXXI, No 2, 111–120

Hill D M Brand D and Hansen W B (1966) Prototype development of statistical land use prediction model for greater Boston region *Highway Research Record No* 114 51–70

Hoover E M (1948) *The Location of Economic Activity* McGraw-Hill Book Co, New York

Huff D L (1962) *Determination of Intraurban Retail Trade Areas*, mimeo Real Estate Research Program University of California Los Angeles, California

Ingram G *et al* (1972) *The NBER Urban Simulation Model* Columbia University Press, New York

Kuner P (1973) ITLUP Integrated Transportation and Land Use Package Version Computer Documentation Appendix B 1 in Putman S H, *et al*, *The Interrelationships of Transportation Development and Land Development*

Kuner P (1973a) Investigation of Traffic Assignment Algorithms Appendix C-1 in Putman S H, *et al The Interrelationships of Transportation Development and Land Development*

Lakshmanan T R and Hansen W G (1965) A retail market potential model *Journal of the American Institute of Planners*, XXXI, No 2 134 143

Lakshmanan T R (1965) A Theoretical and Empirical Analysis of Intraurban Retail Location Ph D Dissertation, Ohio State University

Lathrop G T Hamburg J R and Young G F (1965) An opportunity-accessibility model for allocating regional growth *Highway Research Record No* 102 54–66

Lee D B Jr (1968) *Household Disaggregation in Urban Models* presented at the Regional Science Association Meetings Cambridge, Massachusetts

Loewenstein L K (1967) *A Proposed Manufacturing Location Model* paper presented to the Western Regional Science Association Las Vegas, Nevada

Lowry I S (1964) *A Model of Metropolis* RM-4035RC, RAND Corporation, Santa Monica, California

Nathanson J (1970) *Basic Employment A Model for Intra-County Location of Basic Employment and Land* BATSC Technical Report 222 (Preliminary) Bay Area Transportation Study Commission, Berkeley, California

Putman S H (1967) Intra urban industrial location model design and implementation. *Papers and Proceedings of the Regional Science Association*, 19, 199–214

Putman S H (1969) An Economic Model for Forecasting Employment by Industrial Sector in Small Sub-Areas of a Large Region Ph D Dissertation, University of Pittsburgh, University Microfilms Ann Arbor, Michigan

Putman S H (1970) Developing and testing an intra-regional model *Regional Studies*, **4**, No 4, 473–490

Putman S H (1972) Intraurban employment forecasting models A review and a suggested new model construct *Journal of the American Institute of Planners* XXXVIII, No. 4, 216–230

Putman S H *et al* (1973) *The Interrelationships of Transportation Development and Land Development* (in two volumes) Institute for Environmental Studies, Department of City and Regional Planning, University of Pennsylvania, Philadelphia, Pennsylvania

Rose J (1967) *Intraregional Industrial Location Theory and Methodology* paper drafted for the Regional Planning Council, Baltimore, Maryland

Rose J (1968) *Interim Industrial Allocation Model Structure and Analysis* mimeo, Regional Planning Council, Baltimore Maryland

Seidman D R (1969) *The Construction of an Urban Growth Model* Report No 1, Vol 4, Delaware Valley Regional Planning Commission, Philadelphia, Pennsylvania

Stevens B H and Wheaton W Jr (1970) *A Review of Available Land-Use Activity Distribution Models* TM-4524 System Development Corporation, Santa Monica California

Wheaton W C Jr (1973) Ph D Thesis, Department of City and Regional Planning, University of Pennsylvania Philadelphia, Pennsylvania

Wilson A (1970) *Entropy in Urban and Regional Modelling* Pion Ltd , London

Wolfe H B (1967) Model of San Francisco housing market *Socio-Economic Planning Sciences* 1, No 1, 71–95

PUTMAN LP7

PROOFS

Calibrating a Disaggregated Residential Allocation Model—DRAM

S.H.PUTMAN
University of Pennsylvania

Introduction

As part of a research effort, sponsored by the National Science Foundation, to compare the performances of two different land-use models calibrated on the same data base, several fundamental problems in urban land-use modelling have been encountered and partially resolved. In particular, the fact that no Lowry-derivative, land-use model had ever been properly calibrated in work done in the US became abundantly clear. In order, then, to accomplish the desired comparison of different models on a common data base, it became necessary to develop a calibration procedure for these models. The development of this calibration procedure suggested, in turn, a reformulation of the model into one which appears to be much superior to the original and which is sufficiently different to justify a new name, Disaggregated Residential Allocation Model—DRAM, to differentiate it from its predecessor. A rather unique characteristic of this model, cast in entropy-maximizing form, is its multivariate attractiveness measure.

Background

The development of the Lowry model of land-use distribution (Lowry, 1964), along with that of numerous derivatives of its basic structure, has been described elsewhere (Goldner, 1971, Putman, 1975). Some years after development of these models had begun in the US, further substantial development of them was undertaken in Great Britain (Batty, 1972). Interestingly, despite the fact that the model originated in the US, some of the most fruitful work in extending the concept has been done in recent years in Great Britain. Further, and of critical importance to applications of the model, the question of estimation of the model's parameters has (as far as I know) never been propertly settled in any US work, with perhaps one exception, that of Voorhees and associates (1972). In contrast, it appears that the British work has produced rather conclusive evidence about the means by which these models may be calibrated (Batty, 1970; Batty and Mackie, 1972).

A modified version of the Incremental Projective Land Use Model (IPLUM) was used in the Integrated Transportation and Land Use Package (ITLUP). This ITLUP version of IPLUM is fully described elsewhere (Putman, 1973). In brief, the residential portion of this model allocates increments of residential locators to their places of residence in response to increments in basic employment and changes in the transportation

Author et al?

PUTMAN LP7

facilities. This response is determined by a probability function that describes the distribution of work trips, and by a measure of residential attractiveness for each potential location zone. The purpose of this paper is to outline the steps thought to be necessary for a proper calibration of the ITLUP–IPLUM model, and which ultimately led to the development of DRAM.

Virtually all derivates of the Lowry model used in the US have, as their residential-allocation function, some form of the following expression:

$$N_i = g\sum_j p_{ij}E_j \,, \tag{1}$$

where
N_i is the number of residential locators locating in area i,
p_{ij} is the probability of living in area i and working in area j,
E_j is the number of employees in area j, and
g is a scaling factor such that the sum of the N_i over all i is equal to an exogenous control total

There are often other scaling or multiplier factors to convert from employees to households and to ensure various types of internal consistency.

The probability p_{ij} is the most important component of equation (1). In the original Lowry model, the function used was

$$p_{ij} = (d_{ij})^{-1.33R} \,, \tag{2}$$

where
d_{ij} is the airline distance between the centroids of area i and area j, and
R is the number of zones in an annulus d_{ij} miles from the origin.

In various derivatives of the Lowry model, p_{ij} is modified to include measures of the attractiveness of area i. In particular, in the ITLUP form of IPLUM,

$$p_{ij} = \mathrm{f}(d_{ij})O_i \,, \tag{3}$$

where

$$\mathrm{f}(d_{ij}) = \frac{\beta}{d_{ij}^2}\exp\left(\alpha - \frac{\beta}{d_{ij}}\right), \tag{4}$$

O_i is a measure of residential 'opportunities' in i,
d_{ij} is proportional to the travel time between centroids of zones i and j, and
α, β are empirically derived parameters.

Author confirm

The measure of opportunities is basically an adjusted measure of residential holding capacity (previous level of residential density times amount of available land). The adjustment, Q_i, is a logistic-curve function

PUTMAN LP7

of the proportion of the developable land in zone i that has been developed by the end of the base time period. Thus

$$O_i = a_i^v \left(\frac{h_i}{a_i^r}\right) Q_i \,, \qquad (5)$$

where
a_i^v is the vacant acreage in zone i,
h_i is the number of housing units in zone i,
a_i^r is the residential acreage in zone i, and
Q_i is a factor that denotes the level of development.
This factor is defined as

$$Q_i = 1 - \frac{\gamma}{(1-\gamma)\exp(\delta x_i^2)} \,, \qquad (6)$$

where
γ, δ are parameters, and
x_i is the percentage of developable land area in zone i that has been developed.

The parameters of the trip function were estimated by fitting the equation to observed work-trip distributions from the San Francisco area. The parameters of the development-level function, Q_i, have not been statistically estimated nor has the complete probability function, p_{ij}, been fitted to any actual data. It was precisely this fitting which was necessary, but which had never been done (with the exception of Voorhees' attempt) in US work with derivatives of the Lowry model.

Reformulation of the model

In all of these models the essence of the residential allocations is either the work-trip (home-to-work or work-to-home) or a combination of the work-trip with measures of attractiveness of the potential residential locations. Hence a set of work trips is implied when any of these models is used to estimate residential locations, but very little use has been made of this fact in US work. Yet, it is precisely the existence of these implicit trip matrices that leads to a more satisfactory method of estimating the parameters of these models. The use of IPLUM in the ITLUP package is a particular exception to the usual practice of ignoring these implicit trips. In this case these implicit work trips are made explicit by extracting the trips from the model directly, and these trips are later used to load the transport network (Putman, 1973).

Author, et al?

It is a virtue (and perhaps, in the first instance, was the source) of the Wilson entropy-maximizing approach to the analysis of these models that the question of these trips is made explicit (Wilson, 1967). For example, based on this approach (Wilson, 1970), the Lowry model may be rewritten as

$$T_{ij} = E_j f(c_{ij}) \,, \qquad (7)$$

where

T_{ij} is the number of persons working in zone j and residing in zone i,

E_j is the number of persons working in zone j, and

c_{ij} is the impedence (usually travel time or travel cost) between centroids of zone i and zone j.

An important problem of this formulation is that there is no constraint on the sums of trips, in which case there is no reason to expect that

$$\sum_i T_{ij} = E_j \, . \qquad (8)$$

This implies that the number of employees in zone j will not necessarily be equal to the sum of the employees residing in all zones i who claim to work in zone j.

A simple residential-location model may be derived from entropy-maximizing concepts as follows

$$T_{ij} = A_i B_j O_i E_j \, \mathrm{f}(c_{ij}) \, , \qquad (9)$$

where now

T_{ij} is the number of trips between zones i and j, or the number of persons living in zone i and working in zone j,

O_i is the number of trip origins, or the number of employed persons living in zone i,

E_j is the number of trip destinations or the number of employees employed in zone j,

A_i is a balancing factor for trip origins,

B_j is a balancing factor for trip destinations, and

c_{ij} is an impedance function.

It is possible to replace the trip origins O_i by a measure of attractiveness of the origin zone, W_i. This eliminates the need for the origins balancing factor A_i, thus giving

$$T_{ij} = B_j W_i E_j \, \mathrm{f}(c_{ij}) \, . \qquad (10)$$

In order for the constraint on the sums of trip destinations, given by equation (8), to be met, we have

$$B_j = \frac{1}{\sum_i W_i \, \mathrm{f}(c_{ij})} \, . \qquad (11)$$

It is informative to substitute this expression back into the original equation (Senior, 1973), this yields

Author not in refs?

$$T_{ij} = E_j \left[\frac{W_i \, \mathrm{f}(c_{ij})}{\sum_i W_i \, \mathrm{f}(c_{ij})} \right] . \qquad (12)$$

If the term $W_i \, \mathrm{f}(c_{ij})$ is called an 'accessibility attractiveness' measure, then the fraction in equation (12) is a relative measure of the accessibility-attractiveness of zone i to zone j compared to all other zones i.

PUTMAN LP7

Further, it is clear that the total number of employed residents residing in zone i is

$$N_i = \sum_j T_{ij} . \tag{13}$$

where

$$N_i = \sum_j \left\{ E_j \left[\frac{W_i f(c_{ij})}{\sum_i W_i f(c_{ij})} \right] \right\} . \tag{14}$$

If one is willing to assert that

$$p_{ij} = \frac{W_i f(c_{ij})}{\sum_i W_i f(c_{ij})} ,$$

then equation (14) is equivalent to saying that

$$N_i = \sum_j E_j p_{ij} , \tag{15}$$

which is the same function as that of the Lowry model, and described in equation (1).

Thus it can be seen that the IPLUM allocation procedure may be considered, in the context of the entropy-maximizing formulation, as a simple residential-location model. However, IPLUM is a dynamic model in that it estimates changes in the number of residential locators as follows:

$$\Delta N_i = \sum_j (\Delta E_j) p_{ij} , \tag{16}$$

where

ΔN_i is the change in the number of employed residents of zone i from time t to time $t+1$,

ΔE_j is the change in the number of employees in zone j from time t to time $t+1$, and

p_{ij} is the probability that a person will live in zone i and work in zone j, at time $t+1$.

A question arises here as to whether Δp_{ij} might be more appropriate in the new formulation than p_{ij}. Unfortunately, the resolution of this question leads to the further question, among others, of location of in-migrants versus the location of intrametropolitan movers. In-migrants probably make their location decisions somewhat differently than the intrametropolitan movers. None of the Lowry class of models deals properly with this question. The TOMM models (Crecine, 1964; 1969) do so in a very superficial way by means of the 'stable-household' functions. It was not possible to resolve this problem in the current work, so the existing practice of using p_{ij} has been maintained for the present. Furthermore, as will be discussed below, it was the static form of the model that was finally estimated.

PUTMAN LP7

Calibration: initial discussion

To date, virtually all attempts in the US to calibrate these models have involved assorted procedures, no one of which achieved any more than a partial calibration of the allocation function. Some procedures have fitted $f(d_{ij})$, as in equation (2) or equation (4), to observed trip data without taking into account the effects of the characteristics of the origin zone or the destination zone. Other calibration attempts have fitted a function with N_i as the dependent variable and various characteristics of zone i as independent variables, thus ignoring any explicit consideration of the trip distribution. Neither of these two procedures nor any of their many variations is capable of estimating properly the parameters of such a model.

For a model expressed in the form of equation (9), the only parameter(s) to be estimated is that/are those which may be included in $f(c_{ij})$. It has been shown that, in fitting the parameters for such a model, statistics that summarized the goodness-of-fit of the work-trip distributions were much more sensitive to changes in model parameters than those that summarized the goodness-of-fit of the activity distributions (Batty, 1970). This result argues for the use, when possible, of work-trip statistics as criteria for model calibration. Other work has derived several types of statistics that have summarized the distributions of the work trips, each of which is appropriate for particular functional forms of $f(c_{ij})$ (Hyman, 1969).

Author Confirm

A problem posed by the form of the model shown in equation (10) is that W_i, the attractiveness measure, is not a directly observable or measurable variable. In one attempt to model this, the number of dwellings in zone i or the population in zone i were proposed as proxy measures of W_i (Cripps and Foot, 1969). Population was finally selected and produced quite acceptable calibration results In another attempt, usable land area in zone i was suggested as a proxy measure of W_i (Barras et al, 1971). In both of these cases, by using a single proxy variable for W_i, calibration of the model remains a matter of estimating the parameter(s) of $f(c_{ij})$.

In these cases, as well as those using the original form of the model in equation (9), the calibration process involves, (1) selecting starting values of the parameters, (2) estimating the trip distribution, (3) comparing the estimated trip distribution to the actual trip distribution, (4) revising the parameter values, and (5) iterating to find the best-fit parameter values. Work has been done on efficient means of doing this (Hyman, 1969, Batty and Mackie, 1972).

At this point, it becomes necessary to introduce a further consideration, the need to disaggregate the residential locators into types. First, we acknowledge that this disaggregation may easily be described in terms of the entropy-maximizing approach by considering T_{ij}^{kw} to be the number of employees of type w who work in zone j and live in type k housing in zone i. An appropriate set of equations and constraints can be developed

PUTMAN LP7

to cover this situation as well as several others (Wilson, 1970). Solving such a model involves an endogenous procedure for estimating the housing stock by zone. This is not a welcome prospect for our current research efforts, though clearly it is a consideration for the future. What is necessary, then, is a model in the form of equation (10), but disaggregated only by type of locator. This may be written as

$$T_{ij}^k = B_j^k E_j W_i^k f_k(c_{ij}) , \tag{17}$$

where

$$B_j^k = \left[\sum_i W_i^k f_k(c_{ij}) \right]^{-1} . \tag{18}$$

Second, it seemed desirable to investigate the use of a multivariate attractiveness measure. There is empirical evidence that the attractiveness of zone i is a function of, among other variables, the distribution of household types living in zone i (Putman, 1973). This evidence suggests that the attractiveness of a zone to a particular household type is a function of the percentage composition of household types in that zone. Further, the amount of developable land in a zone seems to be a determining factor in residential location, as does a developability factor which appears to act as a proxy variable for the extent of the available urban infrastructure. Thus W_i^k may be defined as follows:

Author et al ?

$$W_i^k = \left[\sum_g a_g^k \left(\frac{N_{ig}}{\sum_g N_{ig}} \right) \right] r_i L_i^v Q_i \tag{19}$$

where

a_g^k are parameters to be estimated,
N_{ig} is the number of households of type g in zone i (note the g household types correspond directly to the k household types),
r_i is residential density, that is households acre^{-1} in zone i,
L_i^v is the available, developable, vacant land in zone i, and
Q_i has the same definition as in equation (6).

The parameters in the expression for Q_i may be estimated independently of the rest of the model. The parameters a_g^k need to be estimated within the structure of the model, in addition, the parameter(s) of $f_k(c_{ij})$ must also be estimated in the same way.

The precise form of the model desired would be, as in the previous discussions, dynamic rather than static, that is

$$\Delta T_{ij}^k = B_j^k W_i^k (\Delta E_j^k) f_k(c_{ij}) . \tag{20}$$

To do this it would be necessary to have data for ΔT_{ij}^k and ΔE_j^k. While this work was in progress these data were not available, so it was impossible to estimate any but the static model.

In order to specify data requirements it will be helpful to write out the model in full. Thus

$$T_{ijt}^{k} = B_{j}^{k} W_{it}^{k} E_{jt}^{k} f_{k}(c_{ijt}) . \tag{21}$$

Substituting for B_j^k and W_i^k we obtain

$$T_{ijt}^{k} = E_{jt}^{k} \left\{ \frac{\left(\sum_{i} a_{g}^{k} \ N_{igt} \Big/ \sum_{g} N_{igt} \right) r_{it} L_{it}^{v} Q_{it} f_{k}(c_{ijt})}{\sum_{i} \left[\sum_{g} a_{g}^{k} \left(N_{igt} \Big/ \sum_{g} N_{igt} \right) \right] r_{it} L_{it}^{v} Q_{it} f_{k}(c_{ijt})} \right\} . \tag{22}$$

Thus the required data are

T_{ijt}^{k} the number of persons of type k employed in area j and living in area i at time t,

E_{jt}^{k} the number of persons of type k employed in zone j at time t,

N_{igt} the number of households of type g living in zone i at time t,

r_{it} residential density (households acre^{-1}) in zone i at time t,

L_{it}^{v} vacant, developable land in zone i at time t,

Q_{it} a development index, as described above, for zone i, and

c_{ijt} travel cost (impedance) between the centroids of zones i and j at time t.

A further problem occurs with the definitions of T_{ijt}^{k}, E_{jt}^{k}, and N_{igt}. E_{jt}^{k} is defined as number of persons of type k working in zone j at time t, and T_{ijt}^{k} is the number of persons of type k employed in area j and living in area i at time t. However, N_{igt} is the number of households of type g living in zone i at time t. Clearly a conversion from employees to households is necessary at some point in the process. In order to simplify the conversion of T_{ijt}^{k} to vehicle trips for use in the network model, it will be most convenient to make the conversion at the residence end of the trip. Thus a matrix for converting households of type g to employees of type k must be developed from regional data for the regions to which the model is being fitted. This was done when the model was applied to San Francisco and Minneapolis–St Paul, but the use of these regional conversion rates across the board makes it necessary to keep careful track of this conversion throughout the calibration process.

Calibration results: partial estimates

Initially, it was intended that before the complete model equation was fitted, preliminary estimates of its parameters would be developed by partial estimation of them by least-squares regression. This was later found to be unnecessary, but some of the results related to the independent fitting of the trip distribution are of some interest.

It will be recalled that in equations (2) and (4) above, the distance functions used in the Lowry model and in PLUM were given. These are but two of a vast number of functions that could be fitted to trip-making data. To test several of these, a tabulation of the first work trips from the San Francisco Home Interview data file was prepared. These trips

PUTMAN LP7

9

were tabulated according to the household and employment classes, enumerated earlier, for the 291 zone areal system. The distributions were then normalized and the resulting distributions were fitted, using a nonlinear least-squares procedure, to several different functions. The work-trip distributions took the familiar form shown in figure 1.

Of the various functions investigated, several types of the gamma distribution seemed to produce the best fits. The general form of this distribution is

$$\nu = c^{\alpha} \exp[f(c)] , \qquad (23)$$

where

ν is the number of trips, or trip frequency,

c is the trip time or cost.

The specific functions that best fitted the data were sometimes best in one household income class and sometimes best in another; no one function was best for all four income classes. The function selected for further work on this prototype was

$$\nu = c^{\alpha} \exp(-\beta c) . \qquad (24)$$

This function, known as Tanner's function, had been used in this type of model elsewhere (Cripps and Foot, 1969). The best-fit parameters for the 291 zone system in San Francisco are shown in table 1. These parameters in Tanner's function do yield the skewed, peaked, curve shown above. T1

In the calibrations described below, the San Francisco data were aggregated to a thirty-zone system, thus increasing to greater than eight minutes the three-minute, average travel time between adjacent zones of

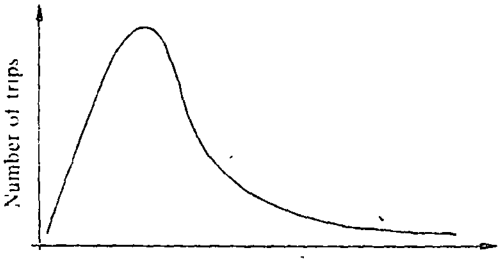


Figure 1. Typical intraurban work-trip distribution.

Table 1. Parameters of the work-trip function for San Francisco.

Income class ($)	α	β
0-4999	0·383	0·900
5000-9999	0·750	0·963
10000-14999	0·849	0·992
>15000	0·784	0·990

PUTMAN LP7 /0

the 291 zone system. At that scale all the values of α become negative and Tanner's function takes on the appearance of a simple, declining exponential function. For the Minneapolis–St Paul calibration (approximately a 100 zone system) the level of disaggregation is sufficient for α to be positive again and for the skewed peaked curve to reappear. All of this reinforces the proposition that the level of spatial aggregation, or disaggregation, has noticeable effects on the apparent functional forms of these models.

Calibration results: complete estimates for San Francisco

The preliminary estimates of parts of the model were of very little use, except that they indicated that a product formulation for W would probably yield better fits than the sum form initially proposed, consequently equation (22) was rewritten. First, let

$$W_{it}^{k} = \left[\prod_{g}\left(\frac{N_{it}^{g}}{\sum_{g} N_{it}^{g}}\right)^{a_{g}^{k}}\right]\left[\prod_{m}(X_{it}^{m})^{a_{m}^{k}}\right], \tag{25}$$

where

N_i^g is the number of type g households in area i, and

X_i^m is a measure of attribute m of zone i.

It was hoped that the attractiveness measure would continue to include intrinsic neighborhood attractiveness, as indicated by the household types located there, a measure of 'capacity' for development, and a measure of developability in terms of infrastructure. Various attributes were tested, including residential density, vacant developable land, percentage of developable land developed, and percentage of industrial (basic) land.

The variables that were finally selected are

n_i^k the percentage of the total households in zone i which are of type k,

L_i^{v} the available, developable land in zone i,

L_i^{d} the percentage of developable land in zone i which has been developed, and

r_i the residential density (households acre^{-1}) in zone i.

Thus the form of W used in the final calibrations was (using four household types)

$$W_{it}^{k} = \left[\prod_{g=1}^{4} n_i^{g}(\exp a_g^k)\right] L_i^{\mathrm{v}\,a_5} L_i^{\mathrm{d}\,a_6} r_i^{a_7} . \tag{26}$$

Note that, based on the preliminary estimates, it was decided to replace the development-level factor, Q_i, by a simple measure of existing level of development, L_i^{d}.

Then, rewriting equation (23) we get

$$T_{ijt}^{k} = E_{jt}^{k}\left[\frac{W_{it}^{k} f(c_{ijt})}{\sum_{i} W_{it}^{k} f(c_{ijt})}\right]. \tag{27}$$

PUTMAN LP7

11

Now there are two ways in which the parameters may be estimated. First, the simplest case, is by looking at the distribution(s) of the activities. In this case, $\sum_j T^k_{ijt}$ is the number of households of class k living in i, by definition. Thus

$$N^k_{it} = \sum_j E^k_{jt} \left[\frac{W^k_{it} f(c_{ijt})}{\sum_i W^k_{it} f(c_{ijt})} \right]. \tag{28}$$

Consequently it is possible to estimate the parameters in the functions W and $f(c)$, and this may be called calibration of the aggregated form of the model.

However, various authors have asserted that there are disadvantages in calibrating the aggregated form of the model. Their remedy for these problems involves calibration of the disaggregated form of the model given in equation (27). It is an unfortunate fact that in order to calibrate the disaggregated form of the model it is necessary to have a good source of trip data. In the work described here there were questions about the quality of these data. If, at some later date, these questions can be satisfactorily resolved, along with the development of an acceptable expansion of the San Francisco 'sample' to an estimate of the 'population', then a calibration of the disaggregated form may be undertaken. In the meantime, calibration has been undertaken for the aggregated form of the model only.

It is immediately obvious that equation (28) cannot be fitted to a data set by using the traditional procedures of linear or even nonlinear multiple regression. In fact the only procedures available are those which, by some efficient procedure, search for the parameters that produce the best fit of the model to the data. One such procedure is that of gradient search, which involves the following steps:

(1) definition of a criterion function to be maximized or minimized;
(2) definition of the partial derivatives of the criterion function with respect to each of the parameters;
(3) selection of a starting point (parameters) and calculation of the criterion and the derivatives, hence the gradient, at that point;
(4) alteration of the parameters as a function of the calculated derivatives and gradient, and iteration through steps (3) and (4) until a minimum or maximum has been reached.

Author ?

While this may sound like a rather lengthy and difficult undertaking, this is not actually the case. The computer software is somewhat difficult, but is available from a variety of sources, including the University of Pennsylvania. At this stage in its development it requires experienced staff for its proper use; nevertheless, once set up, the procedure is rather straightforward and results may be quickly obtained.

The San Francisco data were aggregated into a thirty-zone areal system primarily for economy of operation in the face of no previous experience

PUTMAN LP7

7 12

of the costs and difficulties of performing such calibrations. It was felt that the thirty-zone system would take less computer time to calibrate although still providing useful information about both the model and the calibration process in general.

The model to be fitted is given in equation (28) and the distance function is that of equation (24). The variables in the attractiveness measure are those that were used in equation (25). The calibration was achieved with surprisingly little difficulty, and once the programs were operating correctly, there were no significant problems encountered. An interesting point is that a broad, flat ridge in n-space was found where the search program's criterion value, R^2, was somewhat insensitive to parameter variations. This was an expected occurrence, as suggested earlier (Batty, 1970), nonetheless, with patience, a maximum was reached. The parameters found are shown in table 2.

There are a number of observations to be made about these parameter values, the most important being that, before leaping to unwarranted conclusions, it must be remembered that the household data used in these runs are from the 1960 census, whereas the land-use and employment data are from surveys conducted in San Francisco in 1965. Thus the time subscripts for these variables are not correct for the formulation of the model. The purpose of this particular effort was to explore the problems of calibration of the derivatives of the Lowry model via the Wilson entropy approach, that this is a practical procedure has been amply demonstrated.

Table 2. Best-fit parameters (exponents)—DRAM -for San Francisco (thirty zones).

Household type ($)	Household composition				Land development			Distance		r^2
	a_1^k	a_2^k	a_3^k	a_4^k	a_5	a_6	a_7	α	β	
<5000	1·90	0·40	−0·50	0·33	0·18	−0·73	−0·26	−2·06	0·57	0·91
5000–10000	0·06	1·65	−1·22	0·48	0·27	−1·50	−0·07	−1·75	0·72	0·87
10000–15000	0·14	1·09	−0·26	0·76	0·24	−1·34	−0·14	−1·76	0·76	0·90
>15000	0·72	1·00	−0·34	1·50	0·23	−1·48	−0·04	−1·64	0·48	0·93

Calibration results: complete estimates for Minneapolis–St Paul

The DRAM model was also calibrated for an available data base for the Minneapolis–St Paul metropolitan area which was divided into one hundred and eight zones. The form of the equation used was also that of equation (27), with the distance function given by equation (23). The household income classes differed from those of San Francisco in that they were income quartiles. One of the attractiveness measures, r' (residential density), was replaced by L_i^r (residential land), which produced better fits; the results of these estimates are shown in table 3.

PUTMAN LP7

The data used in this case are all from about 1970, thus resulting in parameter estimates for a static form of the model. It is interesting to note that the scaling or control-total procedures, typically used in these models after the allocations are completed, have moved, with the DRAM reformulation, deeper into the workings of the model. Referring to equation (27) it may be seen that the term in brackets on the right hand side is a proportion. Consequently each E_{jt}^k is simply allocated over all i zones, and, thus, the sum of N_i^k will be equal to the sum of E_j^k. It was mentioned above that it was necessary to convert E_j^k from employees of type k to heads of households of type k. If it is assumed that E_j^k sum to a prespecified regional employment total (or are forced to do so) then N_i^k can be forced to sum to a regional population total as part of the conversion from employees to heads of households. This, while still arbitrary, is not as arbitrary as the various forms of scaling procedure typically used in these models, which often involve altering sophisticated model estimates with rather crude pro rata procedures that vitiate the results of the model.

Table 3. Best-fit parameters (exponents)—DRAM—for Minneapolis-St Paul (one hundred and eight zones).

Household type	Household composition				Land development			Distance		r^2
	a_1^k	a_2^k	a_3^k	a_4^k	a_5	a_6	a_7	α	β	
First quartile	0·77	0 14	−0·56	−0·34	−0·03	0 15	0·89	1·04	2·18	0 93
Second quartile	0 24	0 84	−0·37	−0·15	−0·04	0·18	0·90	2·11	1·46	0·90
Third quartile	0·09	0·16	0·50	−0·08	−0·08	0·25	0·80	2·81	1·31	0 90
Fourth quartile	0·13	0·10	−0·19	0·78	−0·03	0·29	0·75	2·10	1·44	0·91

Discussion—problems of calibrations

The work described here was originally undertaken simply to explore the possibility of calibrating a Lowry-derivative model with a multivariate attractiveness measure via the Wilson entropy formulation. That this is possible has been amply demonstrated, but problems with the available data, particularly with respect to their time indices, make interpretation of the results somewhat chancy.

The general question of parameter interpretation in models of this form is worth discussing. First note that the scale of any of the variables is immaterial since the effect of the balancing factor, defined by equation (11), will be to normalize each variable in all cases. Thus parameters may be interpreted in terms of a variable which ranges from zero to one. Care must be taken to avoid having a variable reach zero if its exponent is negative, and checks should be incorporated in both parameter estimation and forecasting programs to alert the user, at the least, to this situation if it should arise.

PUTMAN LP7

7/4

In figure 2 several members of the family of curves of the form $\nu = c^{\alpha}$ are plotted for different values of α. The range in which we are particularly interested is from $c = 0$ to $c = 1$. Taking first the case of $\alpha \geqslant 0$, we see that for any α the value of ν is less than unity. Thus any variable c_i, for which the estimated α is less than unity, will have an attenuating effect on the region's share of households in area i. This attenuation gradually diminishes as the value of c_i increases from zero to one. It is important to note that the intuitive expectation of a variable that has a positive exponent as an amplifying variable is not quite correct here. In the case of a variable whose range is zero to one, a positive exponent implies decreasing attenuation, with increases in the magnitude of the variable to its limit of unity.

Author: line missing?

values of c, with the amount of amplification decreasing as c increases to its limit of one. Again, the intuitive notion of a variable with a negative

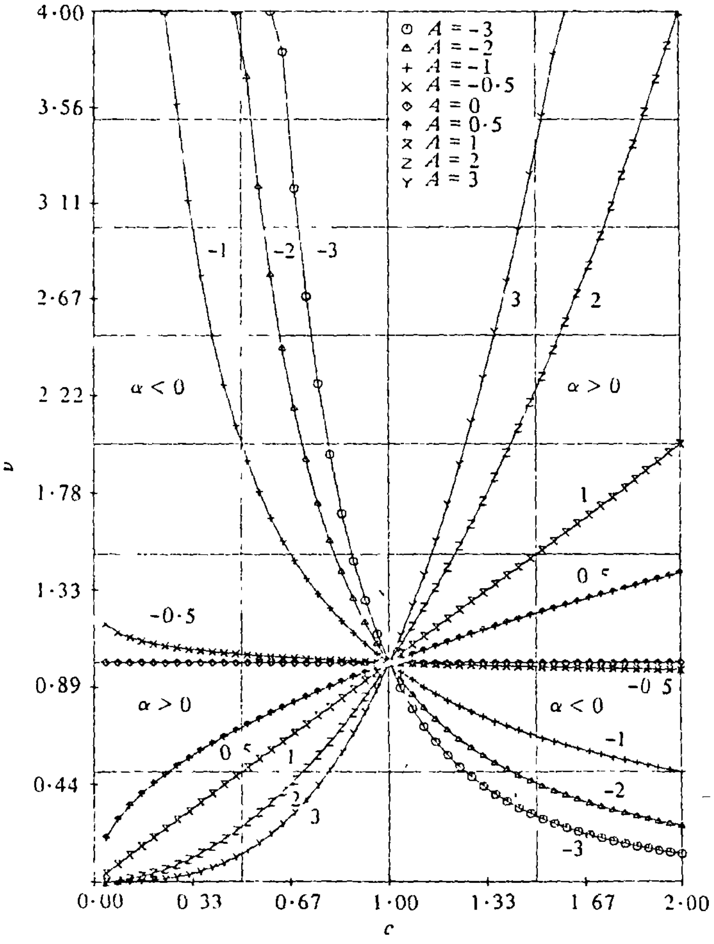


Figure 2. Plots of $\nu = c^{\alpha}$ for $0 < c < 2$ and $\alpha = -3, -2, -1, -0{\cdot}5, 0, 0{\cdot}5, 1, 2, 3$.

PUTMAN LP7

15

exponent being an attenuating variable is not quite correct. For the case of variables whose range is zero to one, a negative exponent implies decreasing amplification as the variable goes from zero to its limit of unity.

In the static situation, it makes more sense to consider each zone vis-à-vis all other zones. For a variable with a positive value of α, all other variables being equal, one would expect greater values of the dependent variable to be associated with greater values of the independent variable that has a positive α Similarly, smaller values of the dependent variable would be expected to be associated with greater values of the independent variable that has a negative α. This reasoning also holds for the situation in which the particular independent variable increases or decreases. Nonetheless, it must be remembered that interpretation of the model's parameters does involve the notions that a decrease in attenuation produces increases and a decrease in amplification produces decreases, and that these are, to a certain degree, counter-intuitive.

In this same connection the use of the exponential-product form of the model caused some operating difficulties. These arise when one or another of the independent variables approaches zero. It may easily be seen in figure 1 that near zero the function $\nu = c \exp \alpha$ becomes rather volatile for all nonzero values of α. Consequently the Minneapolis-St Paul data were rerun with all the independent variables, with ranges from 0·0 to 1·0 shifted to the range 1·0 to 2·0 by simply replacing L_i^q, say, by $(1 \cdot 0 + L_i^q)$; these results are shown in table 4. Although there are some noticeable changes in the coefficients, as compared to the results in table 3, the overall patterns of coefficients are virtually identical. In this form both the problem of instability as the variables approach zero and the problem of the counter-intuitive operation of the exponents are remedied.

It is very difficult to refrain from speculating about the substantive implications of the parameters obtained in these estimations; nonetheless, this would be the wisest policy at this time. One cannot, however, resist the temptation to call attention to the household-composition variables and the interesting speculations that the reader may wish to make about them. Two questions are posed here which should be explored during

Table 4. Revised best-fit parameters (exponents)—DRAM—for Minneapolis-St Paul (one hundred and eight zones).

Household type	Household composition				Land development			Distance		r^2
	a_1^k	a_2^k	a_3^k	a_4^k	a_5	a_6	a_7	α	β	
First quartile	2·92	0·62	−1·71	−1·82	−0·10	0·55	0·83	0·92	2·14	0·89
Second quartile	1·51	2·04	−1·36	−1·57	−0·06	0·65	0·85	2·24	1·36	0·88
Third quartile	0·03	0·45	1·06	−0·64	−0·09	0·60	0·87	2·84	1·32	0·89
Fourth quartile	−0·54	−0·55	−0·06	1·33	−0·07	0·63	0·88	2·48	1·52	0·86

PUTMAN LP7

7 16

further work with the model. First, with regard to these parameters of the household composition in each zone, is there an apparent preference amongst household types for 'equals' or 'betters', that is, higher income classes? Further, if this preference appears, is it a preference for the amenities with which they are associated? Second, having seen how a change in the size of the areal unit changes the shape of the travel function, how does such a change affect that part of the model that deals with attractiveness? To the extent that the household compositions are representative measures of a complex of variables, their meaning may be lost on large areas. For example, the representation of neighborhood, which may show up at a small area level, may disappear when the areas are aggregated to larger zones.

Another set of questions that must be resolved during further work with this model is concerned with the interaction between the 'travel' parameters and the 'attractiveness' parameters. In these experiments one might first constrain the attractiveness parameters to zero and observe the fit of data to the travel function only, within the construct of the model. Then the reverse could be explored by constraining the travel parameters to zero and observing the fit of data to the attractiveness function only. This information might have been obtained from the independent fitting of the two parts of the model formulation described earlier. However, the functions used were not quite correct, nor were the data.

In retrospect it seems that the earlier, independent estimation of portions of the model that was performed for the San Francisco data was unnecessary in terms of estimating starting values of parameters for the complete model. The knowledge obtained about the appropriate functional forms to be used in the complete model was, however, a worthwhile result. In future calibration work with this model it will probably be more efficient to begin with the complete form of the model, perhaps omitting some of the attractiveness variables, or at least constraining their parameter values for the first few runs, while initial values of the other parameters are determined. This procedure seemed to work reasonably well for the Minneapolis–St Paul data. Finally, it should be noted that the use of r^2 as the criterion for parameter fitting is not clearly the best criterion for functional forms like DRAM. The use of maximum-likelihood criteria is being investigated for future work.

In conclusion, the initial tests of this model formulation are quite promising. The model appears to be capable of providing direct spatial allocations of households, by several types, without the need for complex input variables or involved sets of constraints and adjustments that are usually found at the tail-end of land-use models. Work is currently underway to reevaluate much of the work described here and to produce a more final and definitive form and calibration of DRAM.

PUTMAN LP7

Acknowledgements. This work was supported by the National Science Foundation, and, in part, by the US Department of Transportation, via Grant APR 73-07840-A02, "Development of an improved transportation and land use model package".

A large portion of the computer work that was necessary to produce these results was done by F Ducca, a graduate student in the Department. Earlier work on fitting portions of the models was done by C Sawyer and R Mathie, both graduate students in the Department.

References

Barras R, Broadbent T A, Cordey-Hayes M, Massey Doreen B, Robinson Krystyna, Willis J, 1971 "An operational urban development model of Cheshire" *Environment and Planning* 3 115-234

Batty M, 1970 "Some problems of calibrating the Lowry model" *Environment and Planning* 2 95-114

Batty M, 1972 "Recent developments in land-use modelling: a review of British research" *Urban Studies* 9 (2) 151-177

Batty M, Mackie S, 1972 "The calibration of gravity, entropy, and related models of spatial interaction" *Environment and Planning* 4 205-233

Crecine J, 1964 "A time-oriented metropolitan model for spatial location", Department of City Planning, Pittsburgh, Pennsylvania

Crecine J, 1969 "Spatial location decisions and urban structure: a time-oriented model", DP-4, Institute of Public Policy Studies, University of Michigan, Ann Arbor, Michigan

Cripps E L, Foot D H S, 1969 "The empirical development of an elementary residential location model for use in sub regional planning" *Environment and Planning* 1 81-90

Goldner W, 1971 "The Lowry model heritage" *Journal of the American Institute of Planners* 37 100-110

Hyman G M, 1969 "The calibration of trip distribution models" *Environment and Planning* 1 105-112

Lowry I, 1964 "A model of metropolis" RM-4035-RC, Rand Corporation, Santa Monica, California

March L, 1971 "Urban systems. a generalized distribution function" in *London Papers in Regional Science 2 Urban and Regional Planning*, Ed. A G Wilson (Pion, London) pp 157-170

Putman S H, 1975 "Urban land use and transportation models a state-of-the-art summary" *Transportation Research* 9 187-202

Putman S H, 1973 "The interrelationships of transportation development and land development", 2 volumes, Department of City and Regional Planning, University of Pennsylvania, Philadelphia, Pa

Voorhees A M. and Associates, 1972 "Application of the urban systems model (USM) to a region—North Central Texas", prepared for North Central Texas Council of Government

Wilson A G, 1967 "A statistical theory of spatial distribution models" *Transportation Research* 1 253-269

Wilson A G, 1970 "Disaggregating elementary residential location models" *Papers of the Regional Science Association* 14 103-125

Author not cited in text Author names of other authors!

Author [illegible]?

Environment and Planning A, 1976, volume 8, pages 613-623

A model of shopping center location

F W Ducca, R H Wilson
Department of City Planning, University of Pennsylvania, Philadelphia, Pennsylvania, USA
Received 1 September 1975, in revised form 21 April 1976

Abstract The purpose of this paper is to present a model of retail employment which includes planned shopping centers in the employment-allocation function The differences between planned shopping centers and unplanned areas are explored The model to project retail employment is presented Multiple sources of demand are allowed and an equilibrium between employment and demand is assumed A procedure to locate additional planned shopping-center areas is derived The calibration procedure involves a hill-climbing algorithm, a criterion function, and the partial derivatives of the allocation function Calibration results indicate a good parameter fit and support the hypothesis that planned shopping centers attract retail employment

Introduction

The most frequently used type of model for predicting retail employment is the gravity model, for example, Huff (1962), Lakshmanan and Hansen (1965), Fidler (1967), BALFLO (Harris, 1964), the retail component of the Lowry model (Lowry, 1964), and the Lowry-model derivatives (Goldner, 1971) During the last fifteen to twenty years planned shopping centers have become a very significant part of the technology of retail trade (Sternleib, 1962, Simmons, 1964, Cohen, 1972) It is contended that, with the movement of population to the suburbs and the rapid growth of planned shopping centers, the effect of such centers should be incorporated into location models of retail employment It is further contended that these models have not adequately allowed for the effects of planned shopping centers.

Planned shopping centers—differences

Briefly the differences between planned shopping centers and unplanned shopping areas stem from the control of the former by a single management. A planned shopping center is defined as being "planned, developed and designed as a unit, at a single time, to meet the trading requirements of a specific area, and that it is developed under a single management" (Davidson, 1960) The differences between planned shopping centers and unplanned agglomerations of retail employment tend to be reflected in rental patterns, the type of stores in the centers, the location of stores within the center, and the site selection and amenities associated with the center. Each of these characteristics will be discussed

Rental patterns in a shopping center reflect the value to the renter of the center location, and may also reflect the value to the center of having a particular line of retail trade located within it Rents may be lowered to attract the type of store which will in turn attract shoppers, and may be raised for the type of store for which a shopping center is an especially desirable location

The type and number of stores allowed to locate in the center is controlled by the management. For example, pornographic book stores or fresh-fish markets are excluded, and normally no more than one jewelry store will be found there

---

Management must limit the number of shoe stores located in a center, since a planned center is a very attractive site for their location and the center can be overwhelmed by them. At least one, and usually several, large department stores normally locate in the center.

Within the center the location of stores is carefully controlled, with large department stores, which are the major attractors of trade, placed at the ends of the center, and the smaller outlets located between them so as to benefit from the trade drawn by the larger stores. It was mentioned earlier that rental patterns in shopping centers are controlled by management. This is done in such a way that large department stores usually pay a much lower rent than the smaller stores; effectively they are being paid for attracting customers, whereas the small stores, through high rents, are paying for the benefits of a site to which customers are being attracted (Simmons, 1964).

Finally, planned shopping centers are oriented toward an automotive society. Sites are carefully selected with respect to major highways, and large parking areas are always included in the site for a shopping center. These factors combine to make shopping centers highly accessible to automobile traffic.

The exact relationship of these factors in attracting retail trade is not certain. The contention of this paper is that they combine to make planned shopping centers more attractive to consumers than unplanned agglomerations of retail trade. Further, if planned shopping centers are significant, the significance will become apparent in the calibration of a model in which their size and location are included.

The APORT model

The APORT (A Projection Of Retail Trade) model was developed to allow planned shopping centers to be included as a factor in the projection of retail employment. In this model, demand is assumed to be attracted to an area by the existing employment in the area, the employment density, and the distance which must be traveled. The area of a planned shopping center is used in the calculation of the employment density, and an equilibrium distribution of retail employment is projected by using an iterative technique derived from BALFLO (Harris, 1964). A submodel was also developed to allocate the area of a planned shopping center, a procedure that is accomplished by analyzing the marginal revenue of each zone with respect to the area of the shopping center and then allocating new area for a shopping center to the zone with the greatest marginal revenue. This procedure was derived from an algorithm developed by Fidler (1967)

Location of retail employment

The following are the variables of the model:

Subscripts

i denotes the zone of origin of demand, worksite, or residence,

j denotes the zone of retail trade (zone of destination of demand).

Variables

e_j^r is the retail employment in zone j,

D_i^{er} is the demand for retail employment originating in zone i,

D_{ij}^{er} is the demand for retail employment originating in zone i arriving in zone j.

d_{ij} is the distance from zone i to zone j,

q_j is the density of retail trade in zone j,

A_j is the total planned floor space of a shopping center in zone j,

D_{*j}^{er} is the demand for retail employment arriving in zone j (where an asterisk that replaces an index denotes summation)

The variable q_j is defined as the total retail employment in the zone divided by the land area of the zone. The reason for choosing this as an attractor was that the denser the retail trade, the more likely a consumer's perception that he could satisfy the purpose of a shopping trip in the zone. A_j, the total planned area, comprises only retail floor space, it does not include external malls or parking lots.

As stated earlier, there are three factors which determine the attractiveness of the zone, retail employment, density, and distance to the zone. Of these, the internal characteristics of the zone which determine its attractiveness are retail employment and density In calibrating BALFLO, Harris (1964) found that density was effective as a measure of attraction for the Central Business District (CBD) zones but not for suburban zones with large volumes of retail trade but low densities. To compensate for this, the area of the shopping center is used as a substitute for density in areas where planned shopping centers occur. The area of the shopping center is added to density so that the combined effect will allow it to be the dominant factor in zones where planned shopping centers are located The density factor then becomes $(q_j + \delta A_j)$.

In the case of demand originating at the place of residence, D_i^{er} is measured in terms of the population of zone i or the population multiplied by the average income of zone i For demand originating at place of work, D_i^{er} is measured in terms of employment in i

The parameters of the model are as follows·

α is an exponent applied to total attractiveness,
δ is a coefficient of shopping center size, and
b is an exponent applied to distance.

The constraint to be applied to the model is that total demand is equal to total retail employment, that is, $\sum_i D_i^{er} = \sum_j e_j^r$

In addition, an equilibrium condition is required The amount of demand attracted to zone j is equal to the retail employment in zone j This is expressed by $D_{*j}^{er} = e_{*j}^{r}$, for all j.

The basic structure of the model is

$$D_{ij}^{er} = D_i^{er} B_i e_j^r (q_j + \delta A_j)^{\alpha} d_{ij}^{b} ,$$

$$D_{*j}^{er} = \sum_i D_{ij}^{er} = \sum_i D_i^{er} B_i e_j^r (q_j + \delta A_j)^{\alpha} d_{ij}^{-b} ,$$

where

$$B_i^{-1} = \sum_k e_k^r (q_k + \delta A_k)^{\alpha} d_{ik}^{b}$$

The problem is to solve the equation for a set of e_j^r such that $D_{*j}^{er} = e_j^r$, for all j, the method of solution is an iterative technique illustrated below

Define the following variables in the iterative procedure.

e_j^{rm} is the value of e_j^r on iteration m, and
D_{*j}^{erm} is the value of D_{*j}^{er} on iteration m

To solve for e_j^r, let e_j^{r1} be any distribution of the supply of retail trade subject only to the constraints

$$\sum_j e_j^{r1} = \sum_i D_i^{er} , \qquad \text{and} \qquad e_j^{r1} > 0 .$$

Then

$$D_{*j}^{er1} = \sum_i D_{ij}^{er} = \sum_i D_i^{er} B_i^1 e_j^{r1} (q_j + \delta A_j)^{\alpha} d_{ij}^{b} ,$$

and

$$(B_i^1)^{-1} = \sum_k e_k^{r1} (q_k + \delta A_k)^{\alpha} d_{ik}^{b} .$$

At this point D_{*j}^{er1} is substituted into e_j^r for the second iteration. Thus D_{*j}^{er1} replaces e_j^{r2}; this process is repeated with D_{*j}^{erm} as for e_j^{rm+1} on each iteration. The process stops where $e_j^{rm} \approx e_j^{rm+1}$, or $D_{*j}^{erm} \approx e_j^{rm}$. This distribution of e_j^{rm} is referred to as e_j^r (iteration superscripts are omitted)

Multiple strata

APORT may be used with multiple sources (strata) of demand. However, in moving from one demand source to multiple demand sources a weight parameter must be added. When multiple sources are allowed, each source has its own behavioral characteristics as reflected in δ, α, and b, with the weight parameter indicating the relative influence of each source

The formulation is illustrated below. Let n denote the stratum of demand, where n can take on the values 1 and 2; subscript 1 indicates place of residence and subscript 2 indicates place of work

The formulation then becomes

$D_i^{er(1)}$ is the demand of type 1 arising in zone i,
$D_i^{er(2)}$ is the demand of type 2 arising in zone i,
α_1, α_2 are the density parameters for types 1 and 2,
δ_1, δ_2 are the shopping-center multipliers for types 1 and 2,
b_1, b_2 are the distance parameters for types 1 and 2,
W_1, W_2 are the relative weighting of demand types 1 and 2, where

$$W_1 + W_2 = 1 \qquad \text{and} \qquad W_1 \geqslant 0\,, \quad W_2 \geqslant 0\,,$$

and

$$\sum e_j^r = \sum D_i^{er(1)} = \sum D_i^{er(2)}\,,$$

this implies that

$$\sum e_j^r = W_1 \sum_i D_i^{er(1)} + W_2 \sum_i D_i^{er(2)}$$

We then redefine e_j^r as the sum of demand attracted to zone j of type 1 and demand attracted to zone j of type 2

$$e_j^r = D_{*j}^{er} = \underset{\text{type 1}}{W_1 \sum_i D_i^{er(1)} B_i e_j^r (q_j + \delta_1 A_j)^{\alpha_1} d_{ij}^{b_1}} + \underset{\text{type 2}}{W_2 \sum_i D_i^{er(2)} B_i e_j^r (q_j + \delta_2 A_j)^{\alpha_2} d_{ij}^{b_2}}\,,$$

or, more generally,

$$e_j^r = D_{*j}^{er} = \sum_m \sum_i W_m D_{im}^{er} B_i e_j^r (q_j + \delta_m A_j)^{\alpha_m} d_{ij}^{b_m}\,.$$

This is also solved iteratively for the equilibrium distribution of e_j^r. It has been determined experimentally that with any initial estimate of e_j^r, the model quickly iterates to a solution. It has also been determined experimentally that the equilibrium solution is unique and independent of the initial estimate. The uniqueness of the solution is also supported by the work of Eilon et al. (1969) who demonstrated mathematically that in a model with a structure very similar to APORT the iterative procedure produced a unique solution

Shopping center location

In developing the model for shopping center location, two basic problems seemed immediately apparent: identification of the zone in which to locate space for a new shopping center, and how much space to locate there. On further analysis this became a problem of defining the purposes for which the model would be used

It was assumed that the model would be used in a predictive mode and that the location of new shopping centers would reflect choices made by developers This is implemented by selecting zones which maximize the amount of trade that a new shopping center will attract. The underlying assumption is that the size and location of planned shopping centers result from individual developers selecting locations so as to maximize profit. This is similar to the approach taken by Lakshmanan and Hansen (1965) and Fidler (1967).

Alternative criteria were examined, in particular that of maximizing benefit to the consumer or maximizing consumer surplus (Wilson, 1974; Coelho and Wilson, 1976). It seemed that this approach was well suited as a decision tool for controlling the location and growth of shopping centers but not appropriate as a predictive tool when developers are free to choose location and size

The original formulation can be restated as

$$D_{*j}^{er} = \sum_i D_i^{er} B_i e_j^r (q_j + \delta A_j)^{\alpha} d_{ij}^b \ .$$

Differentiating this with respect to A_j yields

$$\frac{\partial D_{*j}^{er}}{\partial A_j} = \sum_i \frac{\alpha\delta}{(q_j + \delta A_j)} \left(D_{ij}^{er} - \frac{D_{ij}^{er2}}{D_i^{er}} \right) ,$$

and

$$\frac{\partial D_{*j}^{er}}{\partial A_j} \geqslant 0 \ , \qquad \text{for} \quad A_j \geqslant 0 \ .$$

This quantity represents the marginal increase in retail trade in zone j generated by an increase in shopping center area in j

$\partial D_{*j}^{er}/\partial A_j$ is calculated for all zones in which activity in the shopping center is permitted. Zones may be bypassed because of constraints, such as zoning or lack of available land. After the zone with the greatest marginal increase has been identified then the problem arises of the amount by which the space of the shopping center in that zone should be increased.

Two alternatives were examined. allocating additional floor space until the additional profit generated was less than the cost of the space, and allocating fixed increments of a total volume of floor space of a new shopping center The first alternative is theoretically preferable but cannot be implemented in practice. The APORT model measures total demand in a zone, which cannot be disaggregated into that for a shopping center and that for a nonshopping center. Conceptually this means trips may be made to a zone because a shopping center is there but that other shopping may take place in the zone but outside the shopping center.

The second alternative was the one chosen. This procedure involves deciding first, the total additional floor space for a shopping center which the area will support, and second, the size of the increment to be added. Before describing this procedure, some discussion of the distribution of shopping centers is necessary.

The existing distribution of shopping centers and the projected distribution do not represent equilibrium distributions since the location of the shopping center is fixed. Once a center has been built, it cannot 'move' to another site. Further, although employment in the center may grow or decline, the size of the center is relatively fixed.

The question arises concerning the uniqueness of the projected distribution of shopping centers. It is quite likely that the distribution is not unique. For example, if several adjacent zones have a high value for $\partial D_{*j}^{er}/\partial A_j$, locating a center in any one of them would significantly lower $\partial D_{*j}^{er}/\partial A_j$ for each of the others. If any of these zones were used as the starting point for the algorithm, the location of succeeding allocations would not change. Further, if in the real world a 'bad' location decision is made, that site will remain a shopping center owing to the large capital investment, even though it is not as profitable as an alternative. One other aspect of the location procedure needs to be examined. If two adjacent zones each have a large $\partial D^{cr}_{*j}/\partial A_j$, should two small shopping centers be located one in each zone, or one large center in a single zone? The assumptions used in the model are based on the knowledge of phenomena in the real world The construction of shopping centers is a major investment, and developers are normally well aware of the market activity of other developers Cut-throat competition does not appear to occur once a center has been built, therefore, the zone of greatest potential will receive floor space and this may result in its potential falling below other zones. However, after successive allocations are made it is possible that the original leader may again become the zone of greatest feasible potential.

Allocation procedure

The procedure begins with an examination of the potential of each zone ($\partial D^{cr}_{*j}/\partial A_j = C_j$) for trade from a new shopping center. All zones are then ranked according to their potential, and the zone with the greatest potential is the first to receive an allocation of shopping-center floor space. The zone next in rank is then examined to determine whether it is too close to the zone that has received the initial allocation (the rationale for this is that two adjacent zones may each have potential for a new shopping center, but when a shopping center is located in either zone the other will no longer be a viable location) If they are not too close, the second zone receives an allocation of floor space.

This sequential allocation is repeated until a prespecified number of zones, m, have been examined When this has been completed, the system is allowed to iterate to equilibrium with the new shopping centers The entire procedure is then repeated, beginning with reestimating the potential for shopping centers in each zone. The algorithm stops when the additional volume of floor space to be allocated has been totally distributed

A more detailed description of the algorithm is illustrated below:

A^0 is the vector of original shopping-center space,
A^1 is the vector of augmented shopping-center space,
N is the set of all zones in which shopping center location is feasible,
n are the zones which are elements of N
K is the set of zones in which new space is added,
k are the zones which are elements of set K,
m is the number of prespecified zones to be placed in set K,
S is the total amount of space to be allocated,
ΔS is an increment of space added to selected zones,
$d_{\min}$ is the minimum interzonal distance for selected zones,
d_{ij} is the distance from zone i to zone j

The procedure for locating new shopping-center space is as follows:

Step 1. Calculate C_j for all $n \in N$

Step 2. Sort the set by arranging C_j in descending order, giving C_{ni}, where i is the index after sorting

Step 3 For $N = 1$, i is assigned to K, and ΔS is added to zone i.

Step 4 For $n > 1$, (a) if $d_{kn} > d_{\min}$ for all $k \in K$, add ΔS to the area of the shopping center in n and continue, (b) if $d_{kn} < d_{\min}$ for any $k \in K$, go to next n and examine the corresponding zone Continue to repeat this step until m zones are assigned to K

Step 5 The remaining space to be allocated is S_1, $S_1 \leftarrow S - \Delta S$

Step 6 If $S_1 \geqslant m\Delta S$, go to *Step 1*, otherwise stop.

The ideal procedure for locating new shopping centers is to set $m = 1$, thus bypassing *Step 4* entirely. This is expensive in terms of computer time, since the system must be iterated to equilibrium each time *Step 1* is encountered. *Step 4* was added primarily to avoid this cost. The minimum distance test in *Step 4* ensures that two adjacent zones, each with a large C_j, are not each allocated space for a new shopping center when an allocation to either zone would cause a drop in C_j in the other zone

If S is large, m and ΔS should be adjusted so that the algorithm is applied several times For a problem consisting of thirty zones, for example, it is recommended that

$$2 \leqslant \frac{S}{m\Delta S} \leqslant 10$$

Calibration

Because the allocation function of APORT is nonlinear, a gradient-search technique was chosen for calibrating the parameters. Briefly summarized, a gradient search is a hill-climbing algorithm which will maximize a criterion function with respect to a set of parameters. It does this by first calculating the criterion and derivatives for a set of starting parameters, then altering all parameters simultaneously in the direction which will generate the greatest improvement, that is the steepest ascent, in the criterion function. With each new set of parameters the criterion is recalculated and tested against the previous value If an improvement has been made in the criterion, the parameters are again incremented and the process repeated. The process continues until there is no further improvement At that point the derivatives are recalculated, a new direction determined, and the process repeated. This continues until the criterion function is maximized.

The criterion chosen for calibration was the minimization of the sum of the squares of the deviations between the observed employment values and the projected employment values. After examining this, it was decided to transform it into a form similar to the R^2 coefficient of determination. This is done as follows

Let

e_j^r be the observed retail employment in zone j, and

$\hat{e}_j^r$ be the estimated retail employment in j

The sum of the squares of the deviations is then $\sum_j (e_j^r - \hat{e}_j^r)^2$.

This can be transformed into R^2 by dividing by the total squared deviation about the mean and then subtracting the result from unity. Thus

$$R^2 = 1 - \frac{\sum_j (e_j^r - \hat{e}_j^r)^2}{\sum_j (e_j^r - \bar{e}_j^r)^2} ,$$

where

$$\bar{e}^r = \sum_j \frac{e_j^r}{N} .$$

The gradient search requires the first partial derivation of each parameter P In this case

$$\frac{\partial R^2}{\partial P} = -2 \frac{\sum_j (e_j^r - \hat{e}_j^r)}{\sum_j (e_j^r - \bar{e}_j^r)^2} \frac{\partial D_{*j}^{er}}{\partial P} . \qquad (1)$$

The calculation of $-2\sum_j (e_j^r - \hat{e}_j^r) \big/ \sum_j (e_j^r - \bar{e}_j^r)^2$ is straightforward The problem is to calculate the partial derivatives $\partial D_{*j}^{er} / \partial P$.

As defined earlier, let B_i be the balancing factor, and let

$$F_{ij} = e_j^r(q_j + \delta A_j)^\alpha d_{ij}^b \ .$$

Then

$$D_{ij}^{er} = D_i^{er} F_{ij} B_i \qquad \text{and} \qquad D_j^{er} = \sum_i D_i^{er} F_{ij} B_i$$

Differentiating D_{*j}^{er} with respect to P, we obtain

$$\frac{\partial D_{*j}^{er}}{\partial P} = \sum_i D_i^{er}\left(\frac{\partial F_{ij}}{\partial P} B_i + \frac{\partial B_i}{\partial P} F_{ij}\right)$$

Substituting the partial derivative of each parameter into the equation yields the following results

$$\frac{\partial D_{*j}^{er}}{\partial \delta} = \sum_i D_i^{er}(F_{ij} B_i^2)\left[B_i^{-1}\frac{\alpha A_j}{(q_j + \delta A_j)} - \sum_k \frac{\alpha A_k}{(q_k + \delta A_k)} F_{ik}\right] ,$$

$$\frac{\partial D_{*j}^{er}}{\partial \alpha} = \sum_i D_i^{er}(F_{ij} B_i^2)\left[B_i^{-1}\ln(q_j + \delta A_j) - \sum_k \ln(q_k + \delta A_k) F_{ik}\right] ,$$

and

$$\frac{\partial D_{*j}^{er}}{\partial b} = \sum_i D_i^{er}(F_{ij} B_i^2)\left[B_i^{-1}\ln d_{ij} - \sum_k \ln d_{ik}\right] .$$

Combining these results with equation (1) yields

$$\frac{\partial R^2}{\partial \delta} = -2\,\frac{\sum_j (e_j^r - \hat{e}_j^r)}{\sum_j (e_j^r - \bar{e}_j^r)^2} \sum_i D_i^{er}(F_{ij} B_i^2)\left[B_i^{-1}\frac{\alpha A_j}{(q_j + \delta A_j)} - \sum_k \frac{\alpha A_k}{(q_k + \delta A_k)} F_{ik}\right] ,$$

$$\frac{\partial R^2}{\partial \alpha} = -2\,\frac{\sum_j (e_j^r - \hat{e}_j^r)}{\sum_j (e_j^r - \bar{e}_j^r)^2} \sum_i D_i^{er}(F_{ij} B_i^2)\left[B_i^{-1}\ln(q_j + \delta A_j) - \sum_k \ln(q_k + \delta A_k) F_{ik}\right] ,$$

and

$$\frac{\partial R^2}{\partial b} = -2\,\frac{\sum_j (e_j^r - \hat{e}_j^r)}{\sum_j (e_j^r - \bar{e}_j^r)^2} \sum_i D_i^{er}(F_{ij} B_i^2)\left[B_i^{-1}\ln d_{ij} - \sum_k \ln d_{ik} F_{ik}\right] .$$

These are the derivatives which are used in the gradient-search procedure.

Results and conclusions

APORT was calibrated on a data base of thirty zones for the San Francisco area using 1965 data. The data included a highway network, household income, employment, and land area. Three types of retail employment were calibrated general merchandise, food stores, and apparel The results are shown in table 1

The b parameters for food stores (SIC54) are very high At first sight this suggested a high degree of impedance for this line of trade. However, on closer examination it appeared that retail employment in SIC54 was highly correlated with the location of residence at the thirty-zone level. A more disaggregated version of the data would yield better results. This result is similar to that found by Putman (1975) in calibrating a residential-location model for the same thirty zones.

As mentioned earlier, the purpose of the shopping-center area was to substitute for density in those zones where shopping centers occur. Table 2 illustrates this for the residential parameters of SIC53 As shown, the amplified attractors vary from 0·66 to 1·60, zones with shopping centers vary from 1·08 to 1·60. Density has relatively little effect on the total attractor in the zones containing shopping centers. The largest influence of density in the shopping-center zones occurs in zone 24, where density makes a difference of 0·03 in total amplification.

The location model of the shopping center was run with the use of data collected on the San Francisco area in 1965. At that time there were 19·1 Mft^2 of planned shopping-center space in the area. It was assumed that no shopping centers would be located in the center of San Francisco, but they could be sited in the zone containing Oakland. For this projection the amount of space allocated was 3·6 Mft^2. The increment to shopping-center space, ΔS, was set at 200 kft^2.

The greatest single allocation was to a zone just south of the San Francisco center-city area. This allocation was 1·6 Mft^2, and there is some question about whether institutional constraints would allow for such a large increase, but this is a question that should be determined exogenously. There was an allocation of 800 kft^2 to both the zone containing Oakland and that containing Berkeley. Two suburban zones were allocated 200 kft^2 each.

These results are promising but several problems remain to be resolved. With respect to the location of a planned shopping center, there are two problems. First, the constraints on the location of a shopping center must be identified. Zones which are highly developed, such as downtown San Francisco, are not likely to have a large planned shopping center no matter how attractive they are. Other areas have zoning ordinances that affect the types of employment which may be located there. Zones of this type which restrict shopping centers must be identified and excluded from eligible zones. The second problem is more theoretical. One study has been done indicating that shopping centers are a change in retail technology and that this change was penetrating retail trade during the period 1960–1970 (Cohen, 1972). A model which accurately projects the location of a shopping center as a change in technology may not be the same model as should be used to project the location of a new shopping center after the change has taken place.

The employment location section of APORT needs further analyses. Based on the calibration results of SIC54, calibrations should be done at a finer level of detail to determine whether the parameters change significantly, and the criterion function used in calibration should be further evaluated. It is possible that χ^2 or maximum likelihood will provide more appropriate parameter values than R^2 (cf Batty and Mackie, 1972).

Table 1 Calibration results

Demand source	Parameters				R^2
	W	δ	α	b	
General merchandise (SIC53)					
Residential	0·61	0·82	0·15	−3·71	0·89
Worksite	0·39	0·01	0·53	−9·69	
Food stores (SIC54)					
Residential	0·80	15·0	0·37	−9·54	0·95
Worksite	0·20	14·9	0·45	−20·0	
Apparel (SIC56)					
Residential	0·58	0·0	0·51	−3·93	0·96
Worksite	0·42	0·14	0·48	−4·10	

Table 2. Effect of shopping center area on density.

Zone	Density	Area of shopping center (kft^2)	$q_j + \delta A_j$	$(q_j + \delta A_j)^\alpha$	$(\delta A_j)^\alpha$
1	1·57	0	1·57	1·07	–
2	0·178	0	0·178	0·77	–
3	0·118	0	0·118	0·72	–
4	0·11	10·38	8·90	1·40	1·39
5	0·171	8·3	6·90	1·35	1·34
6	0·268	14·7	12·2	1·47	1·47
7	0·163	7·0	5·85	1·31	1·31
8	0·299	10·4	8·24	1·40	1·39
9	0·203	15·3	12·6	1·48	1·47
10	0·177	2·64	2·32	1·14	1·12
11	0·267	25·29	20·8	1·60	1·59
12	0·171	0	0·171	0·76	–
13	0·183	0	0·183	0·77	–
14	0·170	1·85	1·67	1·08	1·06
15	0·185	26·35	21·58	1·60	1·60
16	0·252	8·4	7·07	1·35	1·34
17	0·19	0	0·19	0·77	–
18	0·167	7·3	6·09	1·32	1·32
19	0·08	0	0·08	0·68	–
20	0·211	20·63	16·96	1·54	1·53
21	0·176	4·0	3·64	1·21	1·20
22	0·166	0	0·166	0·76	–
23	0·212	0	0·212	0·79	–
24	0·296	1·8	1·75	1·09	1·06
25	0·068	0	0·068	0·66	–
26	0·163	0	0·163	0·76	–
27	0·243	19·86	16·0	1·54	1·53
28	0·11	0	0·11	0·71	–
29	0·224	0	0·224	0·79	–
30	0·177	7·01	5·87	1·31	1·31

$\delta = 0{\cdot}82$, $\alpha = 0{\cdot}15$.

Acknowledgements This paper was partially supported under US Environmental Protection Agency Office of Research and Monitoring contract number R-801-373 to the Institute of Environmental Studies, University of Pennsylvania, entitled "Review and pilot project in urban modelling". The work described was also supported by National Science Foundation Grant APR73-07840-A02, "Development of an improved transportation and land use model package" (Principal Investigator, Stephen H Putman), and in part by the US Department of Transportation Professor Britton Harris provided valuable advice in formulating the concepts embodied in this model

References

Batty M. Mackie S, 1972 "The calibration of gravity, entropy and related models of spatial interaction" *Environment and Planning* 4 205-233

Coelho J D, Wilson A G, 1976 "The optimum location and size of shopping centers" WP-136 Department of Geography, University of Leeds, Leeds, England

Cohen Y S, 1972 "Diffusion of an innovation in an urban system" RP-140, Department of Geography, University of Chicago, Chicago, Ill

Davidson F S, 1960 "Some effects of the growth of planned and controlled centers on small retailers" University of Connecticut, Storrs Conn

Eilon S, Tilley R P R, Foukes T R, 1969 "Analysis of a gravity demand model" *Regional Studies* 3(2) 115-122

Fidler, J, 1967 "Commercial activity location model" Highway Research Board, Publication TPO80808, Washington, DC

Goldner W, 1971 "The Lowry model heritage" *Journal of the American Institute of Planners* 37 100-110

Harris B, 1964 "A model of locational equilibrium" Institute for Environmental Sciences, University of Pennsylvania, Philadelphia

Huff D, 1962 "Determination of intra-urban retail trade areas" Graduate School of Business Administration, University of California at Los Angeles, Los Angeles

Kaylin S O, 1966 "Profile of the shopping center industry" *Chain Store Age* 42 E1-E62

Lakshmanan T R, Hansen W G, 1965 "A retail market potential model" *Journal of the American Institute of Planners* 34 134-143

Lowry I S, 1964 "A model of metropolis" RM-4125-RC, Rand Corporation, Santa Monica, Calif

Putman S H, 1975 "Calibrating a disaggregated residential allocation model -DRAM" paper presented at the Eighth Annual Conference of the Regional Science Association, British Section, in *London Papers in Regional Science 7* Eds P W J Batey, Doreen Massey (Pion, London) forthcoming

Simmons J, 1964 "Changing patterns of retail location" RP 92, Department of Geography, University of Chicago, Chicago (published by University of Chicago Press)

Stedman G H, 1955 "Shopping centers and local government collision or co-operation" *Journal of Retailing* 31 (2) 77-87

Sternlieb G, 1962 "The future of the downtown department store" Cambridge Joint Center for Urban Studies of the Massachusetts Institute of Technology and Harvard University, Cambridge, Mass

Wilson A G, 1974 "Retailers' profits and consumers' welfare in a spatial interaction shopping model" WP-78, Department of Geography, University of Leeds, Leeds, England

Final Report

THE INTERRELATIONSHIPS OF TRANSPORTATION DEVELOPMENT AND LAND DEVELOPMENT

Volume 1

Stephen H. Putman, Principal Investigator
Department of City and Regional Planning
University of Pennsylvania
Philadelphia, Pennsylvania 19174

June 22, 1973

(Revised September 1976)

Prepared for the U.S. Department of Transportation, Federal Highway Administration, Urban Planning Division, under contract No. DOT-FH-11-7843

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No

2. Government Accession No.

3. Recipient's Catalog No

4. Title and Subtitle

The Interrelationships of Transportation Development and Land Development

5. Report Date

June 1973 (Rev. 9/76)

6. Performing Organization Code

7. Author(s)

Stephen H. Putman, et. al.

8. Performing Organization Report No

9. Performing Organization Name and Address

Dept. of City and Regional Planning
Univ. of Pennsylvania
Philadelphia, Pennsylvania 19174

10. Work Unit No.

11. Contract or Grant No.

DOT-FH-11-7843

12. Sponsoring Agency Name and Address

Urban Planning Div.
Federal Hwy. Adm/U.S. Dept. of Trans.
Washington, D.C. 20590

13. Type of Report and Period Covered

Final Report

14. Sponsoring Agency Code

15. Supplementary Notes

Two volume report:	Vol. I	Main Report
	Vol. II	Program Documentation

16. Abstract

The purpose of the research described in this two volume report was to attempt a better understanding of the phenomena which operate to produce the heavy usage and subsequent congestion which often follows close on the heels of the construction of urban and suburban transportation facilities. It was recognized that these phenomena were the same ones which generally govern the shape and growth of the metropolis.

The method selected for analysis of these problems was that of computer simulation. Consequently, in addition to substantive conclusions about the growth of urban areas and potential policies for shaping and/or directing this growth, a package of computer models was prepared for dissemination to other workers in this field. This package of models represents the integration of land use models with transportation network models. Use of this package allows analysis not only of land use patterns and transportation flows, but of the important interrelationships between them as well.

Testing of this package on a full-scale dataset for the San Francisco metropolitan area indicates that by means of region-wide, integrated transportation planning and land use control policies, the above mentioned problems can at least be reduced and, in some cases, eliminated.

This two-volume report includes descriptions of the development and testing of the models, and substantive conclusions which resulted from this work. Copies of the computer programs for the model package are available from the sponsoring agency

17. Key Words

Land Use Models
Urban Transportation Planning
Urban Development Controls

18. Distribution Statement

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	Vol.1- Vol.2-	

Form DOT F 1700.7 (8-69)

Acknowledgements

Our Contract Manager, W. Terry Moore, of the Urban Planning Division, provided invaluable guidance and assistance.

The project staff included, at various times:

Stephen H. Putnam, Principal Investigator
Peter Kuner
Graham R. Gleave
Harvey A. Goldstein
Richard B. Rubin
Jan Williams
Charles Sawyer

Members of the Faculty of the University who contributed various amounts of time, ideas, and constructive criticism:

Department of City and Regional Planning

Prof. Britton Harris
Prof. Ann Strong
Prof. John Keene

Department of Regional Science

Prof. David Boyce

Most of the secretarial work was done by:

Pamela Riley

Prefatory Note

This report consists of two volumes.

Volume 1 consists of descriptions, at several different levels of detail, of the work done and conclusions to be derived therefrom and Appendices containing supplementary material.

Volume 2 contains documentation for the Integrated Transportation and Land Use Package of computer programs and a worked through sample problem.

As of the time of writing of this report it is planned that these computer programs will be available through the sponsoring agency.

The tables of contents for both volumes are given in each volume for the reader's convenience.

Very few changes have been made in the tect of the report for this revision. The program documentation has, however, been substantially improved.

Several papers have been published which contained updated information about the development of this model package. These are:

Putnam, S.H. "Preliminary Results From an Integrated Transportation and Land Use Models Package" Transportation Vol. 3 (1974) pp. 193-224.

Putnam, S.H. "Further Results From the Integrated Transportation and Land Use Models Package" Transportation Planning and Technology (1976) forthcoming

TABLE OF CONTENTS: VOLUME 1

I. INTRODUCTION AND SUMMARY 1

II. LITERATURE REVIEWS 5

Introduction 5

Impact of Transportation Development on Land Development 5

Land Use Control and Transportation Facilities 9

The Various Control Approaches Available 13

The Major Techniques Utilized and Proposed 15

References 19

III. METHOD OF ANALYSIS: OVERVIEW 21

Introduction and Restatement of the Problem 21

Method of Approach 25

The Operational Model Package 28

Preliminary Tests of the Model Package: Prototypes, Rounds I and II 33

Further Tests of the Model Package: Round IV, Network Measures . . 45

References 61

IV. METHOD OF ANALYSIS: FURTHER ELUCIDATION 62

Modification to IPLUM Allocation Functions 62

Possible Extensions to IPLUM 85

Transportation Networks and Trip Generation 93

Development of the "New Process" Procedure 104

References 108

V. POLICY TESTS WITH THE MODEL PACKAGE 109

Introduction 109

Strategy for Policy Tests 113

TABLE OF CONTENTS: VOLUME 1 (Continued)

Policy Test 1: Round III Runs 118

Policy Test 2: Round V Runs 126

Policy Test 3: Round VII Run. 136

Summary 140

APPENDIX A-1: by Richard Rubin
Survey of Land-Use and Development Guidance Techniques

APPENDIX A-2: by Peter Kuver
Investigation of Traffic Assignment Algorithms

TABLE OF CONTENTS: VOLUME 2

Appendix B-1

ITLUP: Integrated Transportation and Land Use Package
Version 2, Computer Documentation

Appendix B-2

ITLUP: Integrated Transportation Land Use Package, Version 3
Computer Program Documentation

Appendix B-3

ITLUP: Integrated Transportation Land Use Package
Version 3, Computer Program Test Problem

I. INTRODUCTION AND SUMMARY

The problem towards which this research has been addressed could be considered, in its narrowest sense, to be the protection of highway utility. This, however, is but a sub-section of the general problem of attempting to achieve a balance between land development and transportation facilities development. Both of these are, in turn, subsumed within the yet more general statement of the problems of the manipulation of the spatial organization of the metropolis.

A particular manifestation of this general problem which we have investigated, is the heavy usage and subsequent congestion which often follows close on the heels of the construction of new urban or suburban highways. In our view the process involved here is one of location or relocation of demand (i.e. traffic generating activities) within an urban region in anticipation of the completion of new transportation facilities. Consequently, effective means of coping with this problem clearly needs to interfere with this demand location or relocation process. Determination as to how this might best be accomplished is an extremely complex problem involving integrated considerations of transportation and land-use interrelationships.

Due to the inherent complexity of the problem it was decided to structure its analysis around a set of integrated computer models of transportation and land use. Important implications of this decision were that it would be necessary to:

1) Identify evidence as to the relationships between transportation facility development and land use,
2) Identify evidence as to the effects of land use controls on land use.
3) Translate the results of 1) and 2) into inputs to an integrated transportation and land use model package.
4) Develop the model package and test the various inputs.

It was hoped that, having performed these tasks, it would be possible to make some tentative conclusions regarding policy alternatives for dealing with the problem as discussed above.

In the course of this research effort work was done on all of the above tasks as well as several others. The results of these efforts are briefly summarized below and described in more detail in subsequent chapters of this report.

The following general conclusions have emerged from this research

1) There is abundant evidence reported in the literature of rather strong relationships between transportation facility development and land development. There is, however, no case where clear quantitative evidences of causal relationships are available. Further, there does not seem to be any prospect of obtaining such evidence since, in the strictest sense, this would require the performance of controlled experments in the real world. From a somewhat more realistic point of view, before-and-after data for situations of larger scope than the traditional highway-interchange study could be collected, and would be of considerable use in furthering our knowledge of these processes.

2) The situation with regard to evidence of relationships between land use controls and land development, i.e. the effects of land use controls on land development and vice versa, is even less satisfactory. While there is some evidence to the effect that such controls can serve to constrain or alter the land development process, it is somewhat scant. This combined with the rather ad hoc nature of most land use control policies, and their succeptibility to unplanned, irregular, change under political-economic pressure, make the assessment of their effects extremely difficult. There does seem, however, to be a small but steadily increasing social pressure on planners and legislators to improve upon

this situation. Hopefully as this situation changes over time, some useful data will be collected thus allowing a better determination of the effects of these policies.

3) It is a fact that, given a set of computer models based on a particular theory or set of theories about urban form or structure, a change in the model inputs or a constraint on the model's functions will produce a change in the model's outputs. To the extent that the model's theories properly reflect reality, and to which policy alternatives can be expressed in terms of input changes or functional constraints, the model will be capable of estimating the effects of these policies on urban form. At the present time there are no operational land use of transportation models whose internal structure is properly reflective or micro-behavioral theories of urban form and/or human decision making. All existing operational models are macro-descriptive, with varying degrees of detail. Many existing or potential transportation and/or land use policy alternatives can be expressed, rather broadly, in terms of variables to which these macro-descriptive models are sensitive. Consequently, it is possible to use such models for general analyses of such policy alternatives. The accuracy and detail of such analyses are, however, limited by the capabilities of existing models to a somewhat more general level than is in our opinion, ultimately desirable.

4) Prior to the performance of this research effort there has never been a successful integration of a sectorally and spatially disaggregated urban land use model with an appropriately detailed transportation network model. The model package produced as a substantial portion of this research effort accomplishes this integration and includes, endogenously, the necessary feedbacks from each model to the other. Based on extensive

testing of this model system it appears that the current prevalent practice of using transportation and land use models without these feedbacks, is likely to introduce significant errors into their results. On the occasion that it is necessary to perform such analyses with independent (unintegrated) models very careful attention must be paid to these errors, which will most certainly be present.

5) Finally, and perhaps most importantly, based on somewhat more limited testing of policies with this model system, it appears that significant opportunities for policy manipulation of the spatial organization of urban areas do exist. In particular, it appears that policies of <u>integrated</u> transportation planning (not limited only to new construction) and land use planning (especially in the form of land use controls) could result in a "balance" between transportation development and land development.

II. LITERATURE REVIEWS

Introduction

As a part of this project, two separate literature reviews were undertaken in an attempt to find reported evidence of the relationships between: a) transportation facilities development and land development, and b) land use control policies and subsequent land use. While much material was gathered and reviewed, the general conclusion for both of these efforts is that there is very little quantitatively described evidence of these relationships to be found. There are, however, numerous qualitatively descriptive articles which leave little question as to the existence of these relationships, but considerable uncertainty as to their precise dimensions.

It was anticipated that the readers of this report would be relatively familiar with the concepts underlying the transportation facilities - land development literature. Consequently this literature is described rather briefly, with a selected bibliography being given in Appendix A-1. In contradistinction, it was anticipated that the readers of this report would be less familiar with the general literature on land use controls. Consequently a much more comprehensive review and discussion of the general literature on land-guidance techniques was prepared and is presented in Appendix A-1. Overviews of both these topics, and a discussion of the particular relevance of land use controls to transportation facilities, are given in the remainder of this chapter.

Impact of Transportation Development on Land Development

There is a very substantial literature concerning the impacts of transportation development on land development.

As an introduction to this literature, if necessary, the reader is referred to references (25) and (26), at the end of this chapter, for annotated lists of studies.

It should, however, be noted that many of these studies are repetitious and uninformative, having been done, in many cases, to meet the funding requirements rather than as a scientific investigation. These studies, in general fall into five classes.

The first type of study is the traditional analysis of the effects or the anticipated effects of the construction of a highway bypass. The situation is usually one in which an older road ran through the center of a town or its business district. As the town grew, this older road became more congested and eventually demands for the construction of a bypass appeared. The bypass, quite naturally would have been routed around the town or business district, and the consequent question would be as to its impacts. Most of these studies focus on the effects of the bypass on local retail trade, and produce rather uniform conclusions. In general there was little or no immediate change in retail trade. Sometimes there was a slight decrease in the short run followed by a gradual recovery, and sometimes there was a slight increase. These studies are almost ritualistic in nature and are not of any particular interest to this project.[1]

The second type of study involves evaluations of impacts in a narrow corridor bordering new roadway construction, usually within urban areas.[2] These studies generally involve the comparison of land use and land value before and after the construction of the raodway. In general, land values become greater and land use tended to become more commercial along the roadway. These results are not surprising as the ugly evidence of them can easily be seen in the interminable miles of sprawling strip commercial development which surround virtually every village, town, and city in the country. This is an obvious example of where

---

[1] See (12) for more critique of these studies.

[2] Many of these are referenced in (25) and (26).

operative land use controls should be applied in conjunction with transportation facility development.

The third class of study involves the problems of development around new highway interchanges.[3] Many of these studies are like the local corridor studies described above. They involve the comparison of land use and land value before and after the construction of the interchange. The studies note the increase in the intensity of land development attendent upon the development of the interchange. This land development can be of almost any type, residential, commercial, or industrial depending on a variety of factors usually not analyzed in the studies. The studies often mention the subsequent congestion of feeder roads giving access to the interchange. Again, the evidence of these effects is all too obvious to anyone who has done even a limited amount of highway travel. Here too, coordination of land use controls with transportation facility development might have been able to reduce these blots on the landscape.

The fourth type of study encountered is of substantially more interest to this project, and represents a much expanded scope of analysis compared to the above. These studies attempt to analyze the relationship of the development of areas or regions to the construction of major highway facilities. These studies show a correlation between highway construction and regional development. Very few such studies attempt to analyze the causal links which may exist between these, though the literature on urban and regional simulation modelling (not specifically reviewed here) presupposes the existence of such links. One study in this group points out that often highway construction at the regional scale is undertaken to meet projections of current travel demand trends but do not consider the impact of such construction on that demand.[4] This is the problem to which this

---

[3] Many of these are referenced in (25) and (26)

[4] See (22).

project is particularly addressed. Another study showed that on a nationwide basis highway construction did not appear to effect manufacturing growth, but that on specific areas, especially urban areas or formerly inaccessible areas, the effect was often pronounced.[5] This finding is in agreement with our own hypothesis that over the long run any particular improvement in transportation technology becomes ubiqutous, but during the spread of this technology advantages may accrue to those areas that get it first. Other studies of this type simply describe the regional impacts of highway construction.[6] It should be noted that it is at this level of analysis that the methods to be described later in this report are most relevant.

Finally, a fifth and smaller group of studies considers the relationships between land use development and highway development, and the appropriateness of coordinated planning of both. One such article is a rather general discussion of the matter.[7] Two other articles discuss more specific matters, but nevertheless imply the necessity of this planning integration.[8] All in all this notion of comprehensiveness is conspicuous by its absence from the literature and thus clearly points up the need for the type of research done on this project.

This has been, as previously mentioned, an admittedly brief review of a large body of literature. Unfortunately there is very little in that literature which is of direct relevance to the work undertaken on this project. Further, it has been assumed that the readers of this report will be generally familiar with this literature and that an extensive review would therefore not have been a productive

---

5 See (27).

6 See (3) and (18).

7 See (21).

8 See (2) and (8).

use of the project resources. The next portions of this chapter are rather more detailed since, as mentioned above, it is assumed that the readers of this report will be somewhat less familiar with the literature reviewed therein.

Land Use Control and Transportation Facilities

Primarily since the 1950's, a rather extensive literature has accumulated on the relationships between land-use controls and transportation facilities. This discussion will briefly introduce the major ideas presented in this literature, refer to other sources for more detailed coverage, and provide a short bibliography that may be used to initiate further research. For more extensive bibliographic listings, the following sources from this chapter's references can be consulted.

a. (17) A useful and recent general bibliographic source on this subject.

b. (15) Appendix A: pp. 35-41 and (4), pp. 108-115 provide quite extensive listings. The latter is essentially a more recent and complete listing although the former does contain some sources absent in the other.

c. (28) Appendix A: p. 155-172. A very recent and fairly complete listing.

d. (24) pp. 250, 264. A listing of works dealing with the development impact of highway interchanges. Prominent studies in the area of development impact are included.

Even a cursory examination of the literature on transportation planning reveals the need for a better understanding of the interrelationship between transportation facilities and land-use controls. Transportation facilities are very expensive and hopefully long-term public investments, yet all too often they have become prematurely inadequate because of a failure to appreciate the relationship which exists between land use and transportation facilities and to plan accordingly and institute the necessary design and land use controls. This problem involves not only the safety of public transportation investments, but the welfare of land owners in areas abutting transportation facilities and of the travelling public in general.[9]

These interrelationships are mutual, land-use and thus land-use controls affect transportation service and transportation service influences land use. Various land use characteristics have very marked effects on transportation facilities in the area. These include: type of development, intensity, location, design and location of access to the use, and site design. These and other factors help to determine the nature of the traffic generated in the area, which is a principal determinant of the adequacy of the surrounding transportation facilities.[10] Transportation facilities, especially highways, in turn have a substantial impact on surrounding development and land use. Many studies have dealt with impact of freeway interchanges on surrounding land uses.[11]

This mutual interdependence has often resulted in a transportation-land use cycle, forming the basis of the problem under investigation here.[12] Overtaxed facilities (or non-existant facilities) prompt the construction of new and improved transportation facilities. This leads to better access, which prompts more intensive use of surrounding land. This more intensive use, which has all too frequently been unexpected or inadequately controlled in the past, generates more traffic. Often, this added traffic causes the premature obsolescense of the new facility. Consequently, the success of the facility in creating new access has often prompted its own obsolescense.

---

[9] (14) and (16) are two works dealing with this problem as related to highways.

[10] See (16), pp. 7-10 for a more thorough discussion of this point.

[11] See (24), for example, for a discussion of this impact, and for summaries of some recent study results.

[12] See (16), pp. 1, 8 and (6), p. 541 for further discussion of this vicious cycle.

In the literature dealing with this problem, automobile facilities are almost exclusively the object of concern. There is little written on this problem as relating to mass transportation or other, non-automobile types of transportation facilities.[13] Within the area of automotive facilities, concern is often focused on the freeway interchange area. This concentration is warranted in that interchanges are often the key to highway performance and large traffic generators tend to establish themselves near interchanges.[14]

In attempting to solve the problem and provide for an improved system of transportation facilities and land use, three goals are often forwarded:

1. To achieve a "balance" between the level of transportation service provided and the land-uses it serves (This goal is explicitly stated in the project proposal.)
2. To secure an "optimum" level of access control.
3. To reserve a sufficient amount of right-of-way for future transportation needs.[15]

The problem of balanced transportation facility protection could be approached by attempting to control land use at a detailed, very localized, level. In the immediate area surrounding the facility in question, an attempt could be made to specify the type, intensity, and location of land use, details on access to the facility from the land use, parking requirements, as well as specifications of highway design. Such an approach may be termed the "micro" approach.

It is also possible to attempt to control major traffic generators and overall traffic generation from a larger area by controlling the type, intensity, and location

---

[13] See (11) for a bibliography dealing with Mass Transit Planning.

[14] See (6), p. 542.

[15] See (23), p. 49.

of land uses. The prime motive is the control of traffic levels and traffic load characteristics for such areas, so as to be compatible with the characteristics of the transportation system in the area. This may be termed the "macro" approach.

In this study, the concern has focussed primarily on the "macro" approach. The use of the IPLUM model limited our experimentation and analysis to the effects of varying policies with respect to macro land-use controls. Detailed controls dealing with individual properties, access requirements, and highway design features cannot be easily incorporated in the current model system. However, some of the literature in the field asserts that the micro approach can have, and has had substantial impact in dealing with the problem of facility obsolescence, at least in the case of highways. "The preservation of transport utility seems to lie only partially in the control of development and more in physical design controls."[16] "Frequently, the problem is not so much the number of vehicles entering as the number of points of entry and their design."[17] It is further suggested that while overall traffic generation is important, there are several indications that macro level controls, alone, may not be sufficiently effective. Some studies have found that the impact of major traffic generators drops off precipitously past the first access point, thus suggesting the relative importance of micro as opposed to macro controls.[18] Furthermore, it is pointed out that within "macro" areas, the type and intensity of land use has little effect on overall traffic generation since the variations which do occur tend to average out over areas above a "micro" scale.[19]

---

[16] (15), p. 2.

[17] Ibid., p. 3.

[18] Ibid., p. 4.

[19] See (16), pp. 4, 39-40, 42 and (15), p. 4.

It should be noted that these comments were made with respect to highways and highway interchanges in particular. When the entire transportation system is considered, they are not likely to hold true. Furthermore, in many cases a network will be found to be entirely overloaded by intense development and micro level control will no longer suffice. A macro approach to development levels is, in these cases, much more likely to be useful.

The Various Control Approaches Available

In general, the micro approach consists of control over geometric design of the facility, access to the facility, and detailed or localized land-use controls such as those regulating parking and setback. The macro approach generally consists of less detailed and less access-point oriented land-use and development controls.

Among the non-engineering techniques used or proposed for the establishment of a desirable balance between land use and transportation service, the following grouping (mostly on a legal basis) can be discerned:

1) Eminent domain techniques
2) Police power techniques
3) Use of the law of nuisance
4) Contractual agreements. These can take place between the developer and subsequent owners or between a public agency and a land owner.
5) Licensing procedures
6) Taxation policy
7) The planning process

Both between and within these groupings, there is considerable variation in the effectiveness of the techniques in achieving the desired transportation-land use goals and in the actual use of these techniques. At present, it appears that these techniques are not often used effectively, nor for that matter, is there any clear quantitative evidence as to what their ultimate effectiveness might be. In evaluating various techniques such information as to the consistency and strength of impact is of course a crucial factor. However, also important are the various feasibility considerations:

- a) complexity and difficulty of administration
- b) relative cost to government
- c) traditional cultural acceptance
- d) legal acceptance[20]

Finally, care must be taken to be economically realistic in the exercise of controls. The feasibility and ultimate result of alternative types of control depend heavily upon existing market conditions.[21]

The influence of land use on transportation facilities is largely determined by the traffic generating characteristics associated with various land uses. As a first step in establishing this association, it has been suggested that the traditional classification of land uses into residential, commercial, industrial, etc. be replaced by a more detailed classification system in which the most detailed categories are constructed to reflect uniform traffic generating characteristics.[22] Among the important traffic-generating characteristics of land uses are:

- a) traffic generated per unit of land area or floor area
- b) the intensity and temporal distribution of traffic peaks
- c) directional orientation of traffic movement
- d) composition of traffic.[23]

---

[20] (10), p. 70.

[21] See (15), pp. 14-15 and (16), p. 100, and see (14), p. 18 for a table which lists various interchange problem solutions and evaluations of their cost, effectiveness, etc.

[22] See (16), pp. 30-37 for details.

[23] See (16), p. 9.

Studies have been conducted to determine the relationships of the values of these characteristics to various types and intensities of land use.[24] Despite growing sophistication in these studies, an acceptable level of uniformity and accuracy in determining the traffic generating characteristics of various activities has not yet been achieved.[25] Such information is clearly desirable as it would permit more effective planning of transportation and land-use systems in general, and would permit, in particular, the exploitation of variations of traffic generating characteristics among different land uses so as to achieve, through the proper mixing of land uses, more efficient combinations of transportation facilities and productive activities.[26]

The other half of the transportation-land use relationship is the impact of transportation facilities on surrounding land use. On this relationship, as mentioned above, almost all of the literature concerns itself with the impact of highway interchanges on surrounding development.[27] Naturally, since the interrelationship between transportation and land use is mutual and cyclical and not simply two independent relations, improved knowledge of both transportation impact on land use and land use control impact on transportation is necessary for successful solution of the problem and attainment of a desirable balance between land use and transportation service.

The Major Techniques Utilized and Proposed

The technique of geometric design control of highways is perhaps the most highly perfected, reliable, and widely discussed technique for dealing with the problem

---

[24] For example, see (16), pp. 32-35 and (14), pp. 5-6.

[25] See (16), p. 31.

[26] See (16), p. 44 for a discussion of the potential of use mixing to achieve "balanced" uses with respect to surrounding transportation capacity.

[27] See (24) and (16), pp. 8-9 for discussion of the future of this impact and the factors involved.

of highway dysfunction. Although this study does not directly deal with such techniques, they are important for any successful, comprehensive, solution.28 Although the techniques employed in access control and land-use control do overlap, it is helpful to conceptually distinguish the techniques used to regulate the means of access to a transportation facility (used primarily for roads and highways) from those techniques used to regulate various aspects of land use. The reasons for regulating access are well known and not too difficult to understand.[29] The purpose of regulating land use, in this context, is of course based on the traffic generating potential and characteristics of land uses.[30]

Despite its various deficiencies, zoning can be used very effectively in this regard.[31] A well established control, zoning can be used with relatively little cost to wisely locate various traffic generators and control development according to combined transportation and land use plans. Zoning as it is now practiced must be improved, and even after improvement, zoning alone will not be sufficient, but it can nevertheless be an effective component of a comprehensive control system. Subdivision control is another widely used technique for controlling land use and development vis-a'-vis transportation facilities. It is most often used in this capacity to regulate new residential development so as to provide access to adjacent roads with as little adverse affect upon the function of the road as possible. Other police power techniques used with varying degrees

---

28 For detailed discussion of highway design standards and techniques, see (8) pp. 6-12 and (16), chapters 8, 9 and 83.

29 The definitive work on the control of access to highways is (19). Also see (15), pp. 22-24 and (16), pp. 45-50.

30 Among the many works discussing the use of land-use controls in regard to highway protection are: (10), pp. 70-81; (23), pp. 39-49; and (14), pp. 12-15.

31 See (5) and (28), p. 153, and Appendix A-2 for further discussion of these and the following points.

of success to protect nearby transportation facilities are setback regulations, the official map, off-street parking regulations, and P.U.D. regulation, all of which are discussed in greater detail in Appendix A-1.

A concept that has been discussed extensively in the literature is that of the interchange zone or interchange area zoning or other controls specifically adapted for and imposed in such areas.[32] Proposals vary both in details and in the level of government in charge, but all involve the designation of a special zone or area surrounding a freeway interchange and the imposition of special controls in a comprehensive fashion on land and access in the area to protect the interchange. It should be noted that the needs of the community for access must be balanced against the need of through traffic for a free-flowing interchange area.

Another major set of techniques fall under the heading of the power of eminent domain. Except for acquisition of access rights and the actual right-of-way, little use has been made of this power to protect transportation facilities. Techniques which have been proposed, or for which more extensive use in this connection has been recommended, include excess condemnation, acquisition of development rights, and temporary acquisition followed by resale or lease-back under conditions compatible with a comprehensive transportation and land use plan.[33] The use of the last technique listed above shows great potential for dealing with the problem of balancing land use and transportation service, and is discussed further in Appendix A-1.

In dealing with the transportation problem, present land-use controls have several drawbacks. The most widely used techniques, such as zoning and subdivision

---

[32] See (19), Chapter 11; (15) pp. 32-33; (28), pp. 96-98; and (16), p. 92.

[33] See (10), pp. 73-75, 81-82 for a discussion of this technique, its advantages and questions of feasibility.

control, have difficulty in controlling new development around interchanges just as they do for development on the suburban fringe. Many land use controls, zoning the prime example, are unstable in their protective power when used to safeguard a long term investment.[34] Because of previous development and construction, vested interests, and legal complications, the success of land-use controls in achieving a desirable land use-transportation balance in developed areas is far less likely than when we have the opportunity of applying our present knowledge, planning, and control techniques to as yet undeveloped areas. However, there is great promise for upgrading present techniques such as zoning and utilizing more fully and effectively some of the eminent domain techniques discussed above. The key to success in this area is improved coordination and planning. Coordination of available techniques into a single comprehensive control proceedure is essential if the weaknesses of individual techniques are to be eliminated and controls prevented from working at cross purposes.[35]

On another level, there is need to coordinate, to an extent not previously practiced, transportation and land-use planning, balancing desired activity against available transportation service and the needs of the community against those of through traffic, in both economic and social terms. Hopefully, the results of this project will be useful in indicating the need for this delicate balancing process and the coordinated mix of techniques that may prove effective.

---

[34] See (15), Chapter 4 and (16), p. 3 for further discussion of this point.

[35] See (23), p. 49 ff.

References

1. Adkins, W.G. "Land Value Impacts of Expressways in Dallas, Houston and San Antonio, Texas." Highway Research Bulletin 227, (1959).

2. Black, Alan. "Optimizing Density of Development With Respect to Transportation Cost." Highway Research Record 207, (1967).

3. Boone, A.J. and Wohl, M. "Massachusetts Route 128 Impact Study." Highway Research Bulletin 227, (1959).

4. Boyce, Byrl N. "Excess Acquisition Revisited: Control of Land Use at the Interstate Interchange," Land Economics, Vol. 45 (August, 1969), pp. 293-303.

5. Fonoroff, Allen. "Control of Traffic Through Control of Land Use," Highway Research Record Number 137. Washington, D.C.: Highway Research Board, 1966, pp. 18-21.

6. Frankland, Bamford. "Land-Use Control at Freeway Interchanges in California," Traffic Quarterly, Vol. XIX, No. 4 (October, 1965), pp. 541-555.

7. Garrison, William L. "Supply and Demand for Land at Highway Interchanges," in (9), infra, pp. 61-66.

8. Goldenberg, M. and Keith, R. "Effect of Land Use Planning and Transport Pricing Policies in Express Transit Planning." Highway Research Record 305, (1970).

9. Highway Research Board. "Land Use and Development at Highway Interchanges," Highway Research Board Bulletin 288, Washington, D.C., 1961.

10. Horwood, Edgar M., Charles H. Graves, and Clark P. Rogers. "An Evaluation of Land-Use Control Procedures at Freeway Approaches," in (9), supra, pp. 67-82.

11. Kleiber, Michael C. "Mass Transit Planning: A Bibliography," in: Wolfgang S. Homburger (ed.). Urban Mass Transit Planning. Berkeley, California: The Institute of Transportation and Traffic Engineering, University of California, 1967, pp. 204-212.

12. Lang, A.S. and Wohl, M. "Evaluation of Highway Impact." Highway Research Bulletin 268, (1960).

13. Lemly, J.H. "Changes in Land Use and Value Along Atlanta's Expressways." Highway Research Bulletin 230, (1959).

14. Levin, David R. "The Highway Interchange Land-Use Problem," in (9), supra, pp. 1-19.

15. Marks, Harold and Salem Spitz. "A Review of Transportation Aspects of Land-Use Control," National Cooperative Highway Research Program Report 31. Washington, D.C.: Highway Research Board, 1966.

16. Marks, Harold. "Protection of Highway Utility," *National Cooperative Highway Research Program Report 121*. Washington, D.C.: Highway Research Board, 1971.

17. Mason, Joseph Barry. "A Selected Bibliography on Interchange Development and Land Use Controls, *Council of Planning Librarians Exchange Bibliography 212*. Monticello, Ill., August, 1971.

18. McKain, W.C. "Economic and Social Impact of the Connecticut Turnpike." *Highway Research Bulletin 189*, (1958).

19. Netherton, Ross D. *Control of Highway Access*. Madison, Wisconsin: The University of Wisconsin Press, 1963.

20. Neuzil, Dennis R. *Characteristics of Suburban Residential Trip Generation* (Doctoral Dissertation) Berkeley, California: The Institute of Transportation and Traffic Engineering, University of California, March, 1965.

21. Roberts, Paul O. "Model Systems for Urban Transportation Planning: Where do we go from Here?" Highway Research Record 309, (1970).

22. Shafran, I. and Wegman, F. "Influence of the Highway Network Structure on the Economic Development of W. Virginia." *Highway Research Record 285*, (1969).

23. Stanhagen, William H. "Highway Interchanges and Land-Use Controls," in (9), supra, pp. 32-60

24. Stein, Martin M. "Highway Interchange Area Development - Some Recent Findings," *Public Roads*, Vol. 35, No. 11 (December, 1969) pp. 241-250, 264.

25. U.S. Department of Transportation, Federal Highway Administration, "Economic and Social Effects of Highways", Washington, D.C., (1971).

26. U.S. Department of Transportation, Federal Highway Administration "Guide for Highway Impact Studies", Washington, D.C., (1972).

27. Wheat, L.F. "Effect of Modern Highways on Urban Manufacturing Growth." *Highway Research Record 277*, (1969).

28. Witheford, David K. and George E. Kanaan. Zoning, Parking, and Traffic. Saugatuck, Connecticut: Eno Foundation for Transportation, 1972.

29. Wohl, Martin, Bone, A.J. and Rose, Billy. "Traffic Characteristics of Massachusetts Route 128." *Highway Research Bulletin 230*, (1959).

III. METHOD OF ANALYSIS: OVERVIEW

Introduction and Restatement of the Problem

It is common knowledge that in recent decades there has been enormous public investment in highway construction. This same period has also witnessed large scale shifts in population from rural areas to urban areas and, within the urban areas, from center-city to suburb. While it has not been possible to develop complete explicit quantitative definitions of the causal relationships between these developments, it is clear that these relationships exist.

A related phenomenon is the focus of the research effort described herein. This phenomenon is observed, not infrequently, when the construction of a new section of roadway is followed, all too soon, by very heavy usage and subsequent congestion. Specifically, the nature of this process seems to be that; a) due to the perceived inadequacy of existing facilities a decision is taken in a metropolitan area to improve transport facilities (usually by road building) for a particular part of the area. Then; b) assuming that the decision is approved, in anticipation of the new roadway, land developers and/or speculators become involved with properties near the proposed route. As construction of the facility begins, some homeowners and businesses may consider and even act upon location or relocation decisions. (We do not refer here to involuntary relocation forced by the facility construction). Upon completion of the facility, additional such location decisions are made. Finally; c) in a relatively short time after completion of the facility, the demand for its use greatly exceeds the demand or the forecasts of demand which existed or were made prior to the decision to construct it. Consequently the design capacities are soon reached and often exceeded, resulting in congestion and premature obsolescence of the facility. While this is a rather general description of a very complex process, it is reasonably accurate. The principal question addressed by this research is whether it is feasible, via integrated transportation planning and land-use planning, to avoid or ameliorate the occurance of this particular phenomenon in the future. Further, it is the intention of the research described herein to analyze this process and to determine whether a) balanced development of transportation facilities and land use is feasible and b) if it is feasible what means are available to accomplish it or to at least make it more likely to take place.

There is considerable evidence indicating that the demand for highway travel is rather sensitive to changes in highway capacity. This sensitivity, as described above, frequently results in heavy utilization and congestion of new facilities soon after their construction. It might be argued that the solution to this problem in the future, as in the past, is to construct more facilities. It is possible that if this policy were followed, at some point, an equilibrium situation would indeed be reached. This conclusion can be supported by the assertion that the elasticity of demand for highway travel is finite. If, however, the "population" generating that demand is continually increasing at the same time, it is not clear that the total demand can continue to be satisfied in this manner. The limit of this strategy, at the extreme, is when so much land is converted to use for transportation facilities that land development for other purposes is restricted, with a consequent limit on trip generation and road use. It is obvious that this "equilibrium" is not the desired solution and is hopefully not feasible in any case. It is therefore clearly necessary to analyze possible "intermediate" solutions.

Should these analyses, based on the assumed interrelationships between transportation and land use, yield potentially feasible solutions, it will be necessary to subject them to further evaluation. This evaluation, leading to a definition of feasibility, would undoubtedly involve measures of private and social costs, and measures of the disparities between them. Such costs would include user's costs (e.g. taxes or tolls), pollution (of all types), costs of relocation of activity, societal costs of activity dispersion, and costs of disruptions caused by the construction process; all of which are associated with the various levels of transportation facility development. Further costs, associated with land development include environmental pollution, loss of open space, and congestion of various public facilities (which results either in deterioration of service levels or costs of expansion), in addition to the congestion of the transport facilities, as described above. Also necessary would be consideration of the political feasibility of such suggested policies as might emerge from the analyses. None of this evaluation was undertaken in this project, its principal purpose being to develop a method of analysis and to use this method for some preliminary tests of possible policies.

A balance between transportation facility development and land development implies a market equilibrium of the demand for use of the transportation facility and its supply, i.e. its cost, speed, and capacity characteristics. There are two basic alternatives abailable to the planner who wishes to modify an existing transportation and land use situation sufficiently to achieve such a balance; though it is likely that the best strategies will be mixtures of the two alternatives. The first alternative is to allow demand for transportation to fluctuate freely, with no interference, and attempt to cope with it. This would be accomplished, as in the past, by new highway construction or (hopefully with more success than has been achieved previously) by implementation of mass transportation systems. The second alternative is to attempt to restrict the demand for transportation so that facilities do not become overloaded. This could be accomplished in three ways: 1) existing transportation facilities could be made more costly to use e.g. by the imposition of tolls or by allowing congestion to develop thus imposing a time penalty on users, 2) land development controls could be imposed, thus reducing (or slowing the growth of) trip generating activities, and 3) a mixture of these two actions could be implemented. Finally, a mixture of the two basic alternatives could be attempted, where an attempt would be made to cope with a certain amount of transportation demand (by improving transportation facilities) at the same time that an attempt was being made to control its increase.

The RFP which ultimately resulted in the research described here suggested that: "the following issues have been of concern to transportation planning agencies:

1. In urban areas, it is necessary to achieve a balance between the need for mobility and the preservation of environmental quality and social stability. How can this process be incorporated into the transportation/land use planning program beginning with plan development?
2. To what extent can transportation system planning and facility capacity planning be used to influence land development? What are the available controls for accomplishing this, and how can they be applied effectively?
3. Providing too high a level of service of the auto-highway system in urban areas might result in travel demand exceeding the supply which can feasibly be provided. What should determine the minimum or maximum level of service (speed) which should be used as a criterion for planning transportation facilities in specific parts of the urban area?
4. What would be the consequence of controlling transportation facility capacities as a means of directing land use development even though such controls could mean congestion with more pollution and higher travel costs?

This study should determine the feasibility of the balanced development of land use and transportation facilities. It should also determine whether or not land development and travel can be brought into a satisfactory balance such that any transportation improvements made can provide a continuing high level of service over time."[1]

To summarize, the essence of the problem being addressed by this research is that of controlling (i.e. altering, directing, and modifying) the spatial organization of the metropolis. In particular the concern is with its spatial expansion and with the feasibility of balanced development of its land use and its transportation

---

[1] See (9).

facilities. Finally, should it appear that such balanced development is feasible, what methods are available for testing the means (policies) which may be relevant to achieving that balance.

Method of Approach

Two aspects of the problem were the principal determinants of the method of approach. The first is the inherent complexity of comprehensive analyses of both transportation and land use. The second is the requirement for a self-consistent procedure for testing the sensitivity and response of transportation and land use to policies for their manipulation. The only analysis method which stands any chance of meeting these requirements is that of computer simulation modelling. The remainder of this chapter is devoted to describing the package of models which has been developed and tested as the central focus of this research.

The state-of-the-art in simulation modelling of urban and regional phenomena poses a problem for projects of this sort. In general, those models which have had operational success are, at present, rather too highly aggregated both in areal and sectoral detail as well as in concept.[2] On the other hand, those models which are more conceptually complete and which could provide highly detailed forecasts, are not yet operational.[3] After some consideration of the project goals, it was decided to begin with an existing operational land use model and attempt to modify and improve it concurrently with efforts to integrate it with a suitable transportation network model.

The project requirements suggested, form the outset, that an appropriate model package would need to be far more complete (and therefore complex) than any now extant. What is ultimately required is a complete i.e. population and employment

---

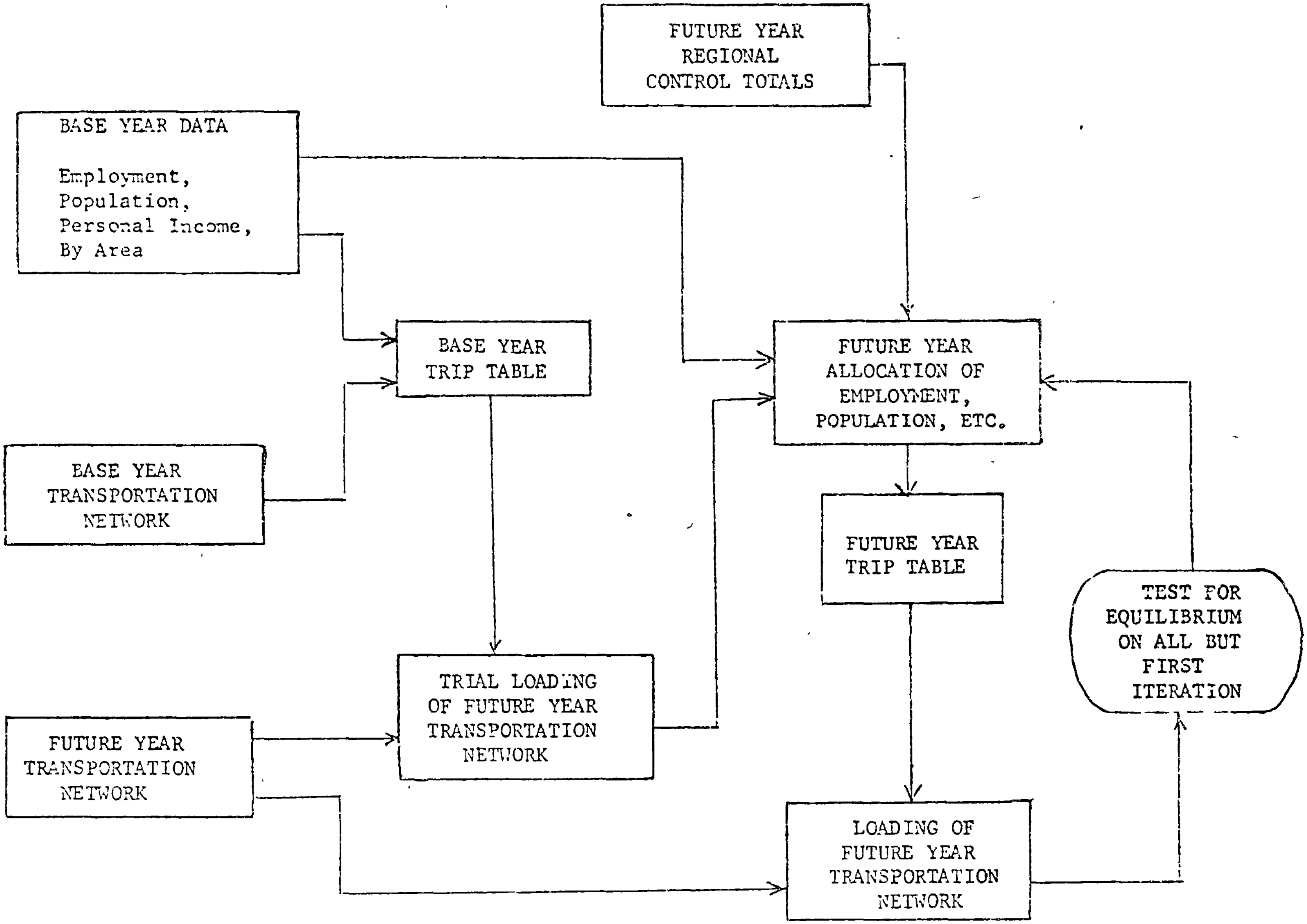
[2] See (2) and (5) for references to the PLUM and EMPIRIC models as examples.

[3] See (4).

sectors, land use model and a complete i.e. public and private, transportation network model. Further, the overall system of models would need to be highly integrated with respect to model inputs, outputs, and feedback loops. In Figure 1 a block diagram of the desired system is presented.

Following through Figure 1, first come the various base-year inputs as to the existing spatial distributions of activity, along with data on the characteristics of the unloaded base year transportation network. These data would then be used to generate a preliminary, and probably inflated, estimate of trips taking place in the metropolitan area. Given this preliminary estimate of metropolitan trips, it becomes possible to load the future (projection year) network so that its travel characteristics (i.e. time and cost) reflect the traffic volumes which would be on the network if there were no change in the distribution of activities from the base year. These network characteristics, along with the base year data and the projection year control totals would then be used to generate a spatial distribution of activities for the projection year. From this spatial distribution a new estimate of metropolitan trips would be produced. These would, in turn, be loaded on the projection year transportation network. The modified characteristics of the transportation networks would then be used to reallocate the projection year spatial distribution of activities. This distribution of activities is then compared to the first estimate thereof. If there are no significant differences an equilibrium has been reached and the model run is ended. If there are significant differences, new trips are generated and loaded on the networks and further iterations are made.

This proposed package of models attempts to eliminate the principal failing of contemporary land-use or transportation studies, by explicitly including feedback loops. Typically, a transportation study assumes a future land use pattern as given, and designs a transportation system to cope with it. This procedure ignores



GENERAL SCHEME OF INTEGRATED MODEL PACKAGE

Figure 1

the redistributive effects which are produced by the construction of the system. Transportation systems obviously do not just suddenly appear, but are constructed in stages with consequent redistribution of activities all during the period of construction. The typical land use study accepts a transportation system as given, and then estimates the consequent distribution of activities while ignoring the congestive effects of that distribution on the network. The proposed model package, as described above, attempts to capture the interrelatedness of the transportation system and the distribution of activities. This comprehensiveness is clearly necessary to meet the research objectives of the project regarding the feasibility of obtaining a balance between transportation facilities and patterns of land use, as well as the testing of policies to achieve such a balance.

The Operational Model Package

It is very often the case in complex simulation studies that the operational version of the model system is somewhat less comprehensive than the system originally proposed. While some simplifications have had to be accepted in this study, the principal goal of an integrated transportation model and land use model has been attained. The remainder of this chapter discusses the general development and testing of the model package, while several detailed descriptions of particular methodological problems are discussed in the next chapter.

The land use model selected, despite its several shortcomings, for the operational model package was the Projective Land Use Model (PLUM). The PLUM model (there are several versions of this model, two of which are used, with modifications, in this model package) is a derivative of the LOWRY model. This means that the model accepts as given the spatial distribution of "basic" employment and then solves for the interrelated spatial distributions of population and "non-basic" or "local-serving" employment. The principle determinants of these spatial distributions are access to workplaces for population and access to customers for non-basic employment. The access measures are non-linear functions of travel-times weighted by attractiveness measures of the "trip destinations".

There are several important implications of the decision to use a form of this model. First, regional levels of activity are exogenously determined. This is not a particularly unusual situation, as virtually all urban land use models have this property. Second, the selection of this model implies that "basic" employment location will be forecast exogenously. In fact, these forecasts do come from a separate model.[4] An important implicit assumption of this procedure is that the spatial distributions of these exogenously forecasted activities are independent of the remainder of the activities in the model package, i.e. population, "non-basic" employment, and transportation networks. This assumption is clearly not valid over any substantial projection period and is an aspect of the model package which should be improved as possible. One way of doing this would be to integrate the modelling of all employmebt location, link this to the residential model and link this overall model to the network model. A further implication of this choice of model is that the level of sectoral aggregation in the PLUM model is rather more gross than is desirable, though it is tolerable in this preliminary development of the package. Finally, the actual allocation functions used in the various PLUM models have never been subjected to thorough statistical examination, and our early experiences with model runs indicated that there was too heavy a reliance on somewhat arbitrary "correction-factors". Consequently, a thorough investigation of the model algorithms was made. Certain anomalies were found in the procedures used for activity allocation, which have since been modified.[5] These modifications appear to have substantially improved

---

4 These projections were from the BEMOD model, see (7), and are the same as those used by the Bay Area Transportation Study Commission (BATSC).

5 This is discussed in detail in the next chapter of this report.

the model's allocations, though attempts to evaluate these improvements have not yet been completed.

On the positive side, reports on the PLUM model did indicate that it seemed to produce reasonable estimates of the general distribution of activities in a large urban region. Of course, as is the case with all other urban land use models, its reliability for specific small area forecasts is considerably less than that for general regional patterns. The model is currently being worked on and refined by its originators and can be expected to show continued improvement over time due to both their efforts and those of this project. Finally, and not an unimportant part of the decision, was the fact that the PLUM model was available and operational at Planning Sciences Simulation Laboratory of the University of Pennsylvania's Department of City and Regional Planning prior to the initiation of the research.

The transportation model package selected for use in this project was developed by the Planning Sciences Group at the University of Pennsylvania. This package contains programs for network coding, tree tracing, and several alternative capacitated traffic assignment procedures.[6] The only deficiency of this network package as compared to others which were available is a lack of ability to handle highly detailed network specifications such as turn penalities, etc.[7] This deficiency was not relevant to the current work, in that there was no intention to conduct any investigations at that level of detail. On the other hand, the computer operation of the Planning Sciences network package is significantly more efficient than that of the FHWA package (formerly known as the BPR package).

---

[6] See Volume 2 of this report.

[7] See (8) for a description of the "FHWA Package".

The principal connection going from the PLUM and IPLUM models to the Network model is via trip generation. The main phenomenon which relates the spatial distribution of activities to the transportation facilities is specified in terms of the trips between activities, which make use of the transportation system. It should be noted that in many transportation studies where land use projection models are being utilized, the work-trips are generated by separate procedures. These trip matrices, because they are calculated independently of the land use model, may not be consistent with the work-trips implicit in the land-use model's patterns of residence and workplace. This problem is eliminated by the procedure described below.

In the PLUM and IPLUM models the location of residences is generated by work-trip allocation functions based on transportation impedances (derived from the transportation network characteristics) and the locations of workplaces. In so doing, the models implicitly generate a matrix of work-to-home trips. The work-trips used in this project's model package are extracted directly from the PLUM and IPLUM implicit work-trip matrices. It has, of course, been necessary to make certain adjustments to this set of work-trips in order to assure that their total equals the totals indicated in the San Francisco data files.[8] Currently these adjustments include auto-to-transit person trip ratios, auto occupancy rates, percentage of all work trips which occur during peak hours, etc.

The location of "non-basic" or "local-serving" employment in the PLUM and IPLUM models is generated by home-to-shop and work-to-shop allocation functions based on transportation impedances and the locations of the activities. As was the case for work-trips, this procedure generates an implicit matrix of shopping trips These trips too, are extracted from the models, though by a somewhat different

---

[8] See (1) for a description of these data.

procedure involving the use of a gravity model type algorithm for balancing the origins and destinations of the trips. Again, adjustments are necessary to insure that the projected trip totals equal the totals indicated in the San Francisco data files.

In the present version of the model package, peak-hour work trips and shopping trips are loaded on the networks. Peak hour trips were used in order to most nearly approximate the work trip travel situation because work trips are a key influence on residential location in these models. Accordingly, the shopping trips used are only those which take place during the peak hour. A problem to be resolved for later versions of the model package is that of other non-work-trips besides the shopping trips. Additionally, the problem of mass transportation facilities has, for the present, been ignored. There is, however, no reason to preclude its inclusion in the model package in future efforts. This would involve construction of a modal-split procedure and a procedure for dealing with multi-modal networks in the PLUM allocation functions. Neither of these appear to present an insurmountable problem for future efforts.

Several alternative procedures for loading the trips on the network have been tested and are described later in this report. The procedure used, called "incremental tree-by-tree loading method," is a modification of one developed at the Chicago Area Transportation Study. Beginning with a free-flow network, a single tree is traced (i.e. all the minimum paths from that origin to all destinations). Onto these paths a portion of the trips from that origin to all of the various destinations is loaded. By use of a volume-capacity relationship, the new volumes on the links of the tree are used to calculate new travel times. Then a new origin is selected, a new tree is traced, and the process is repeated. When a tree has been traced for each origin, the network is loaded with a portion of all the trips. This over-all procedure is then repeated. Currently the networks are loaded in three passes of 40%, 40%, and 20% of the trips respectively.

After the network loading procedures have been completed, a matrix of travel times from each zone to each other zone, via the minimum path on the loaded network, is calculated for input to the land use models.

Preliminary Tests of the Model Package: Prototypes, Rounds I and II

The first tests of the model package were done for a small prototype urban region of eight areas. After ironing out several minor problems, acceptable results were obtained and preparation began for full-scale runs using San Francisco Bay area data. The precise form of the model system used in these runs is described below along with a general overview of their results. More detailed descriptions of components of the model systems are to be found in the following chapter.

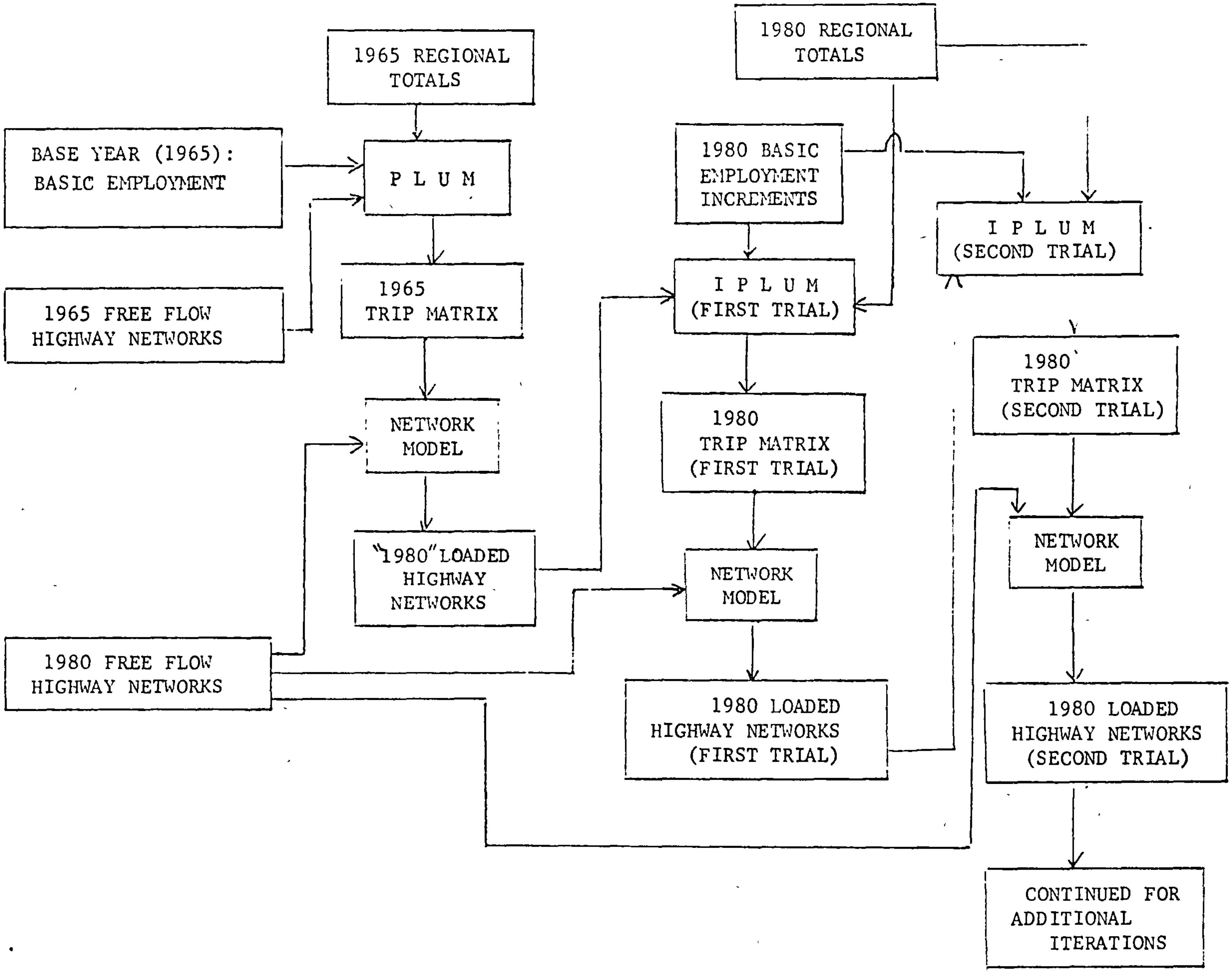
It must again be noted that there are two forms of the PLUM model. The first of these is essentially a "state"-variables model where the forecasts are made in terms of the values of variables at time t. The second form of the model (referred to herein as IPLUM) is an incremental variable version where the forecasts are made in terms of changes in values of variables from time t-1 to time t. These changes are then added to the base values (i.e. those at t-1) to yield the forecasts of time t variables.[9]

The assemblage of models used for these runs is shown in Figure 2. Basically the procedure involves a run of PLUM to generate base-year trip matrices for input, after loading onto a network, to the IPLUM model. After this, successive iterations of IPLUM and the network model are run.

Before presenting the results of these runs, it is appropriate to consider what expectations one would have as to the results of such a test run. If, as suggested earlier in this report, important interrelationships are ignored by the

---

[9] See (3).



SCHEMATIC OF EXPERIMENTAL MODEL RUNS

Figu

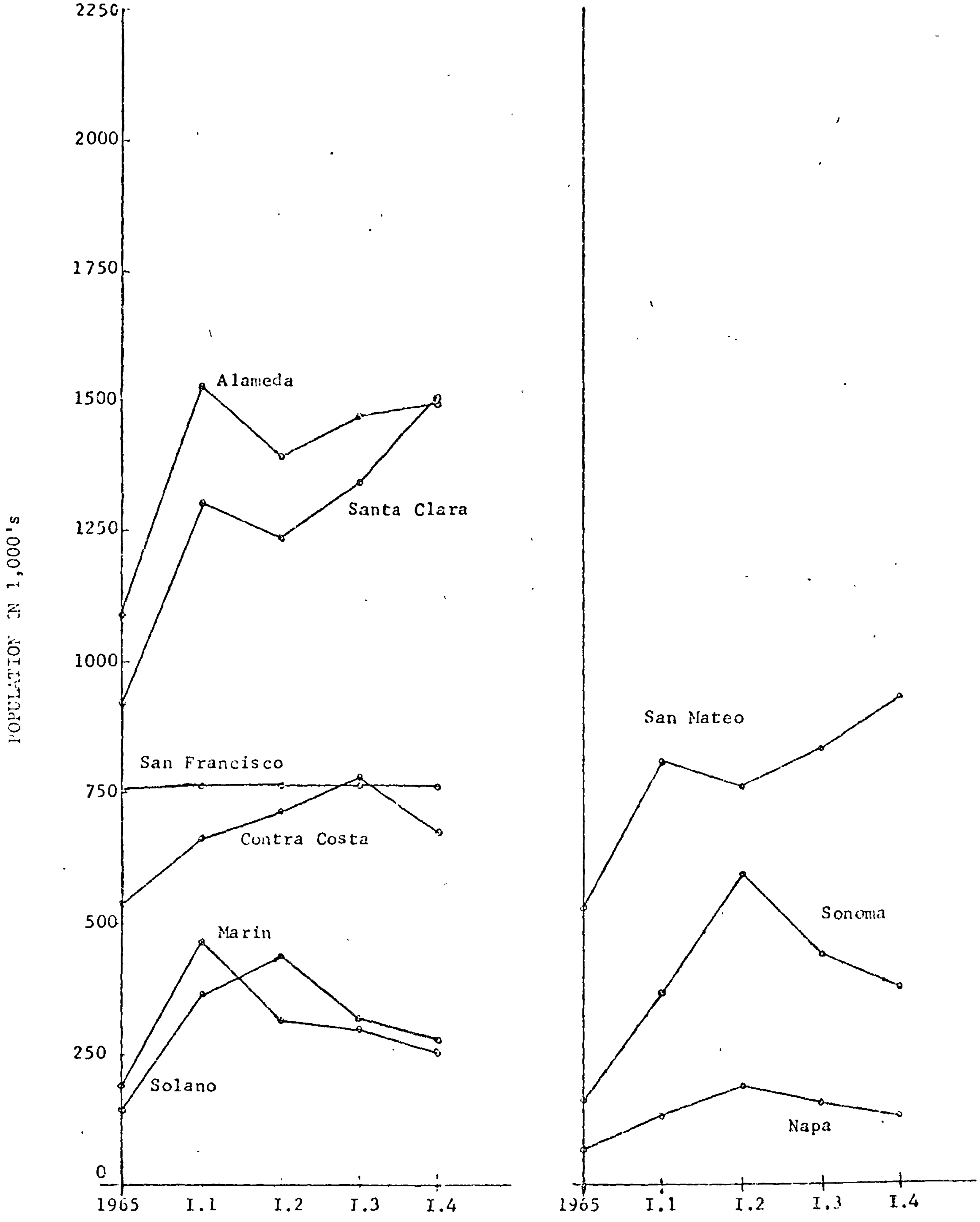
typical transportation study or typical land use study, either of which assumes fixed values for the other, then these model runs should indicate significant and systematic differences in output from a simple run of PLUM or IPLUM not integrated with a network model. These differences should be rather systematic and in accord with the following hypotheses. First, if land use forecasts are made based on a free-flow network or a network where flows, and therefore congestion, are significantly less than would realistically be present, the over-all spatial distribution of activities will be excessively dispersed. An excessively dispersed pattern of activities will, in turn, produce over-estimates of tripmaking and network congestion. Secondly, if land use forecasts are made based on over-congested networks, the over-all spatial distribution of activities will be excessively concentrated. And, in turn an excessively concentrated pattern of activities will produce under-estimates of tripmaking and network congestion. Given these expectations, it is to be anticipated that a linked land-use and transportation model would show convergent oscillation between decentralization and centralization and consequent overcongestion and "undercongestion."

The first "Round" of full model runs was done with a version of IPLUM not incroporating modifications which were subsequently made to the allocation function, and using only "work-trips" for network loading. The population estimates produced in this round are shown in Table 1 and Figure 3. An immediate conclusion which can be drawn from these results is that there is a substantial difference in the spatial distribution of population as a function of whether or not the impedance inputs are derived from "loaded" or "empty" networks. Further, it is clear that there are oscillations in the population allocations as a function of levels of congestion on the loaded networks. These conclusions reenforced our concern as to the inadequacy of either a transportation or a land use forecasting method which ignores the relationships between these two components of the problem.

		Run I.1		Run I.2		Run I.3		Run I.4	
	1965	1980 - (1)		1980 - (2)		1980 - (3)		1980 - (4)	
San Francisco	754,754	764,445	1.28%	763,414	1.14%	765,068	1.36%	765,067	1.36%
Marin	192,523	469,669	243.9	319,634	66.0	301,379	56.5	259,124	36.6
San Mateo	523,271	801,723	53.2	756,041	44.5	835,885	59.7	938,460	79.3
Sonoma	167,497	368,057	119.7	591,781	253.3	444,212	165.2	377,406	125.3
Napa	69,378	135,910	95.9	183,943	165.1	156,245	125.2	130,538	88.2
Solano	149,386	363,338	143.2	437,717	193.0	316,304	111.7	276,112	84.8
Contra Costa	534,056	666,716	24.8	720,394	34.9	778,010	45.7	672,925	26.0
Santa Clara	922,357	1,308,470	41.9	1,238,821	34.3	1,341,138	45.4	1,499,215	62.5
Alameda	1,095,236	1,528,722	39.6	1,395,805	27.4	1,469,301	34.2	1,488,690	35.9
TOTAL	4,438,458	6,407,473	44.4	6,407,471		6,407,470		6,407,474	

Round I: Base Population, Projected Total Population for 1980, and Percent Change 1965-1980.

Table 1



ROUND I - POPULATION PROJECTIONS BY COUNTY

FIGURE 3

At the zonal level of detail the general, expected, pattern of: uncongested (underutilized?) networks producing spread-out population allocation, which in turn produced congested networks, which then produced more centralized population allocation, which in turn produced less congested networks, etc., was found. At the county level of detail much of this phenomenon was obscured (See Map 1 for reference to the location of the various counties). The general performance of the model package was considered to have met our expectations. In view, however, of the obvious "holes" in the package, further, more detailed analysis of these first outputs was curtailed in favor of work on the needed improvements in the package.

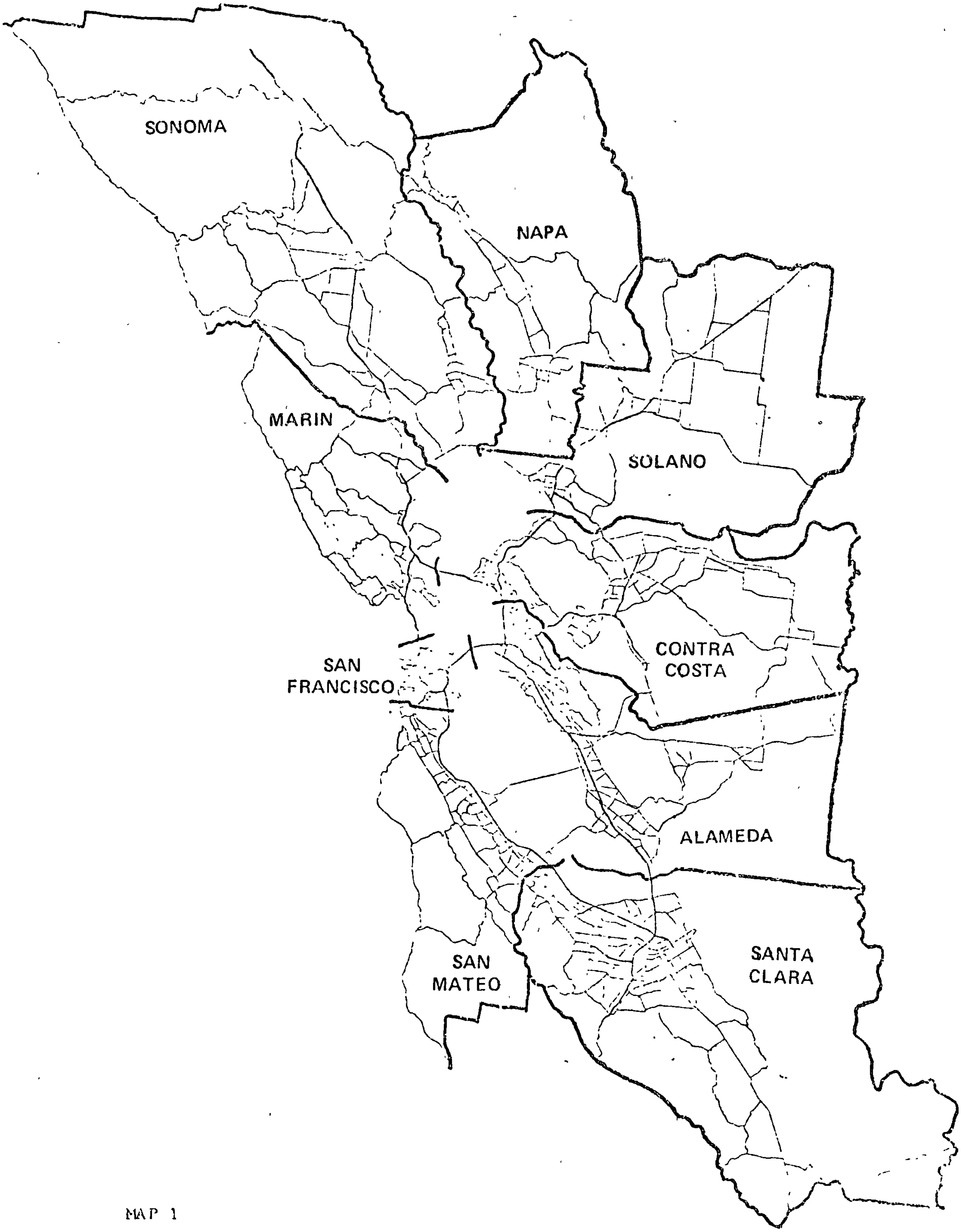
There were three principal areas of further investigation which ultimately led to the second round of model runs. First, preliminary investigations of the operation of IPLUM gave evidence of serious inconsistencies in the allocation functions. The conclusion of these investigations was that the IPLUM allocation procedure required important modifications.[10] These were implemented and resulted in an apparent improvement in the model's performance. Second, it was clear that an important proportion of total trips was being ignored by dealing only with work trips. A procedure for making explicit the home-to-shop and work-to-shop trips implicit in IPLUM was developed, along with a gravity model for distributing these trips.[11] The number of shopping trips was then scaled up to equal the total non-work trips in the region. These, plus the work trips, it was felt, would be a reasonably good estimate of the total trips on the network. Third, and finally, a procedure was developed to effect a substantial savings in computer time for runs of network package. In general this procedure involved the assumption that, after the first two iterations (runs) of IPLUM in a given round, a large proportion of

---

[10] Discussed in greater detail in the next chapter of this report.

[11] Described in more detail in the following chapter.

THE NINE COUNTY REGION AND HIGHWAY NETWORK



MAP 1

the consequent trips on the network would be relatively stable. Consequently a marginal loading of trips would be a possible substitute for the full network loading runs used in Round I.

These investigations and modifications being completed, it was possible to proceed with Round II of the model runs. The results of four iterations of the modified IPLUM, trip generation, and the modified network package incorporating the marginal loading procedure are shown in Table 2 and Figure 4. These runs were later given the suffix B and are referred to in that way. It is immediately obvious from Figure 4 that there still remained problems of oscillation from one iteration of IPLUM to the next. While this was not a particularly unlikely outcome, it was definitely not a welcome result.

A first attempt to isolate the source of these oscillations involved reverting back to the full network loading procedure. Since the marginal procedure had been instituted only after run II.2, (to prepare input for run II.3B), runs II.1 and II.2 are used for both Round II.B and II.A. Three further iterations of the network package, IPLUM, and trip generation were run, yielding runs II.3A, II.4A, and II.5A. These results are shown in Table 3 and Figure 5. It is clear from Figure 5, that the oscillations were still present in the system and the overall results were nearly identical to those obtained from Round II.B. Consequently, it was assumed that the marginal network loading procedure was not at fault.

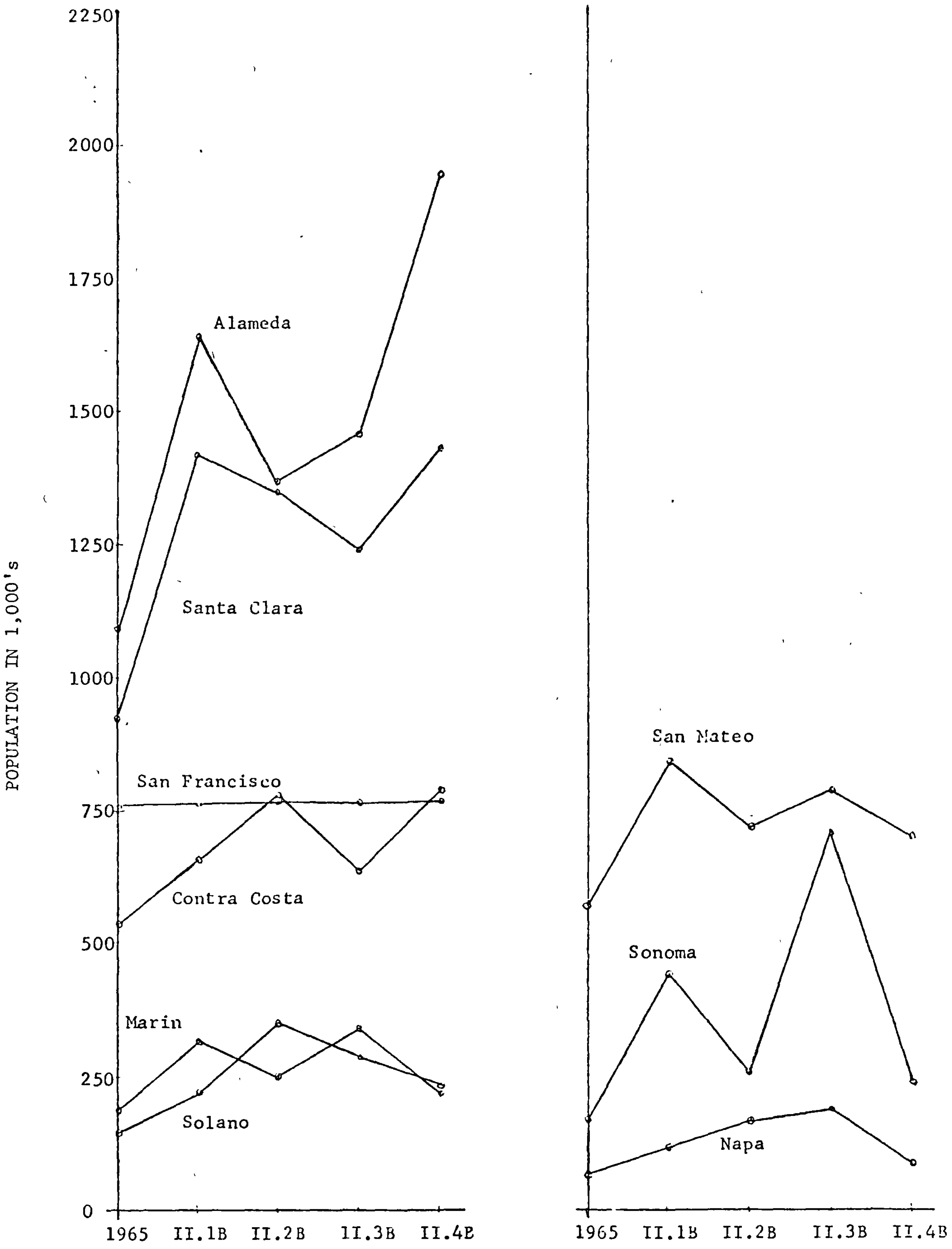
A more detailed investigation of these model outputs revealed that, though obscured at the county level of detail, the oscillations were not really county level oscillations, but were zonal oscillations. In each county the bulk of the zones remained rather stable while a few of the zones underwent rather large oscillations. For example, in Alameda county, of 69 zones, only 10-15 show significant oscillations and of these, only 2 or 3 zones show really large oscillations. In addition, it

		Run II.1		Run II.2		Run II.3B		Run II.4B	
	1965	1980 - (1)		1980 - (2)		1980 - (3)		1980 - (4)	
San Francisco	754,754	759,821	0.67%	762,455	1.0%	758,553	0.50%	760,438	0.75%
Marin	192,523	313,982	63.08	251,919	30.85	341,856	77.57	223,377	16.03
San Mateo	553,271	840,265	51.9	719,364	30.0	787,382	42.3	700,500	26.6
Sonoma	167,497	440,337	162.9	255,273	52.4	708,799	323.2	244,780	46.1
Napa	69,378	118.582	70.9	162,268	133.9	190,562	174.7	93,982	35.5
Solano	149,386	223,074	49.3	351,312	135.2	288,630	93.2	229,551	53.7
Contra Costa	534,056	655,638	22.8	771,037	44.4	635,465	19.0	783,855	46.8
Santa Clara	922,357	1,412,839	53.2	1,348,370	46.2	1,239,020	34.3	1,429,129	54.9
Alameda	1,095,236	1,643,003	50.0	1,364,748	24.6	1,457,279	33.1	1,941,722	77.3
TOTAL	4,438,458	6,407,472	44.4	6,407,476	44.4	6,407,468	44.4	6,407,474	44.4

Round IIB*: Base Population, Projected Total Population for 1980, and Percent Change 1965-1980

(*IIB Series, Partial Loading Between IPLUM Runs)

Table 2



ROUND II.B - POPULATION BY COUNTY, 1980

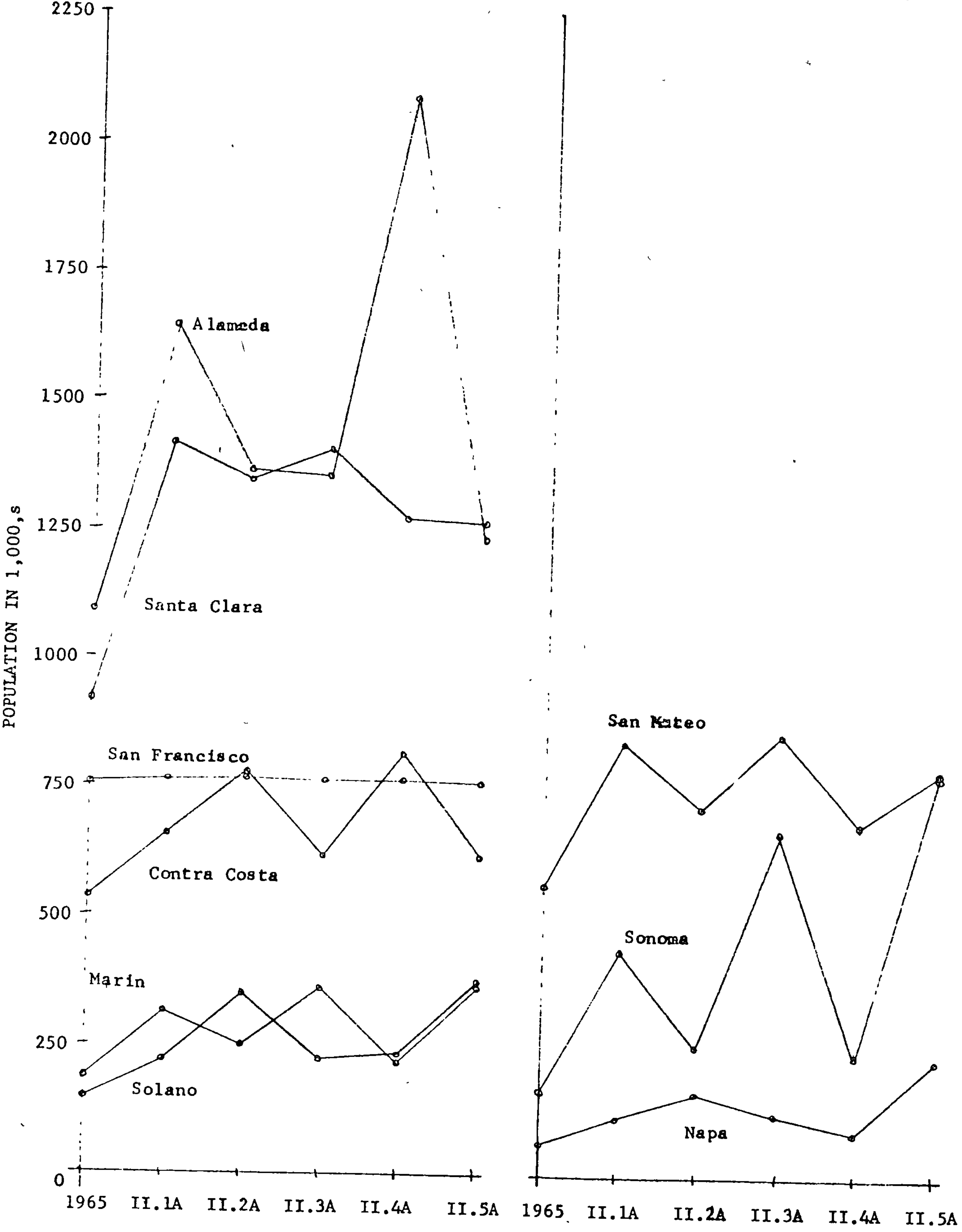
FIGURE 4

		Run II.1		Run II.2		Run II.3A		Run II.4A		Run II.5A	
	1965	1980 - (1)		1980 - (2)		1980 - (3)		1980 - (4)		1980 - (5)	
San Francisco	754,754	759,821	0.67%	762,455	1.02%	759,217	0.59%	760,607	0.78%	758,369	0.48%
Marin	192,523	313,982	63.1	251,919	30.9	360,995	87.5	220,628	14.6	362,523	88.3
San Mateo	553,271	840,265	51.9	719,364	30.0	858,236	55.1	678.678	22.7	781,160	41.2
Sonoma	167,497	440,337	162.9	255,273	52.4	670,765	300.5	240,426	43.5	782,643	367.3
Napa	69,378	118,582	70.9	162,268	133.9	155,177	123.7	95,049	37.0	237,311	242.1
Solano	149,386	223,074	49.3	351,312	135.2	223,454	49.6	231,381	54.9	378,531	153.4
Contra Costa	534,056	655,638	22.8	771,037	44.4	616,419	15.4	817,730	53.2	611,018	14.4
Santa Clara	922,357	1,412,834	53.2	1,348,370	46.2	1,408,025	52.7	1,275,817	38.3	1,260,181	36.6
Alameda	1,095,236	1,643,003	50.0	1,364,748	24.6	1,355,254	23.7	2,087,224	90.6	1,235,809	28.3
TOTAL	4,438,358	6,407,472	44.4	6,407,476	44.4	6,407,465	44.4	6,407,477	44.4	6,407,545	44.4

Round IIA*: Base Population, Projected Total Population for 1980, and Percent Change 1965-1980.

(*IIA Series, Full Loading Between IPLUM Runs)

Table 3



ROUND II.A - POPULATION PROJECTIONS BY COUNTY, 1980

FIGURE 5

was determined that virtually all of the oscillating zones were on the fringes and/or had high percentages of available vacant land in them.

Further Tests of the Model Package: Round IV, Network Measures

Several working memos written by Goldner as well as statements in his report to FHWA allude to the problems of largely vacant fringe zones.[12] In general, these problems stem from the form of the function which calculates the measure of attractiveness of zones for residential allocation. In the version of IPLUM used in the runs described above, this measure was essentially a function of the vacant land in the zone. Clearly, for large size, essentially undeveloped (i.e. largely vacant) zones, the level of attractiveness was quite high. At some point in the development of IPLUM, Goldner entered, and then deleted, a correction factor to deal with this problem. Based on the results obtained from Rounds IIA and IIB it was decided to reintroduce this correction factor.

The theoretical basis for this correction factor is that it is developable vacant land which determines the attractiveness of a zone rather than vacant land <u>per se</u>. This being the case, one measure of the developability of land would be in terms of the proportion of developable land in a zone which had actually been developed. Presumably this could be a surrogate measure of the available "infrastructure" e.g. sewer and water lines, roads, etc. in the zone. This correction factor was introduced into the IPLUM household allocation function.[13] Subsequently a new Round of the model system, called Round IV, was run; the results are shown in Table 4 and Figure 6. (Round III, a policy run, is described in Chapter V)

It can be seen in Figure 6 that a substantial reduction in the system's oscillations has been made, compared to Round II shown in Figure 5. Nonetheless, the

---

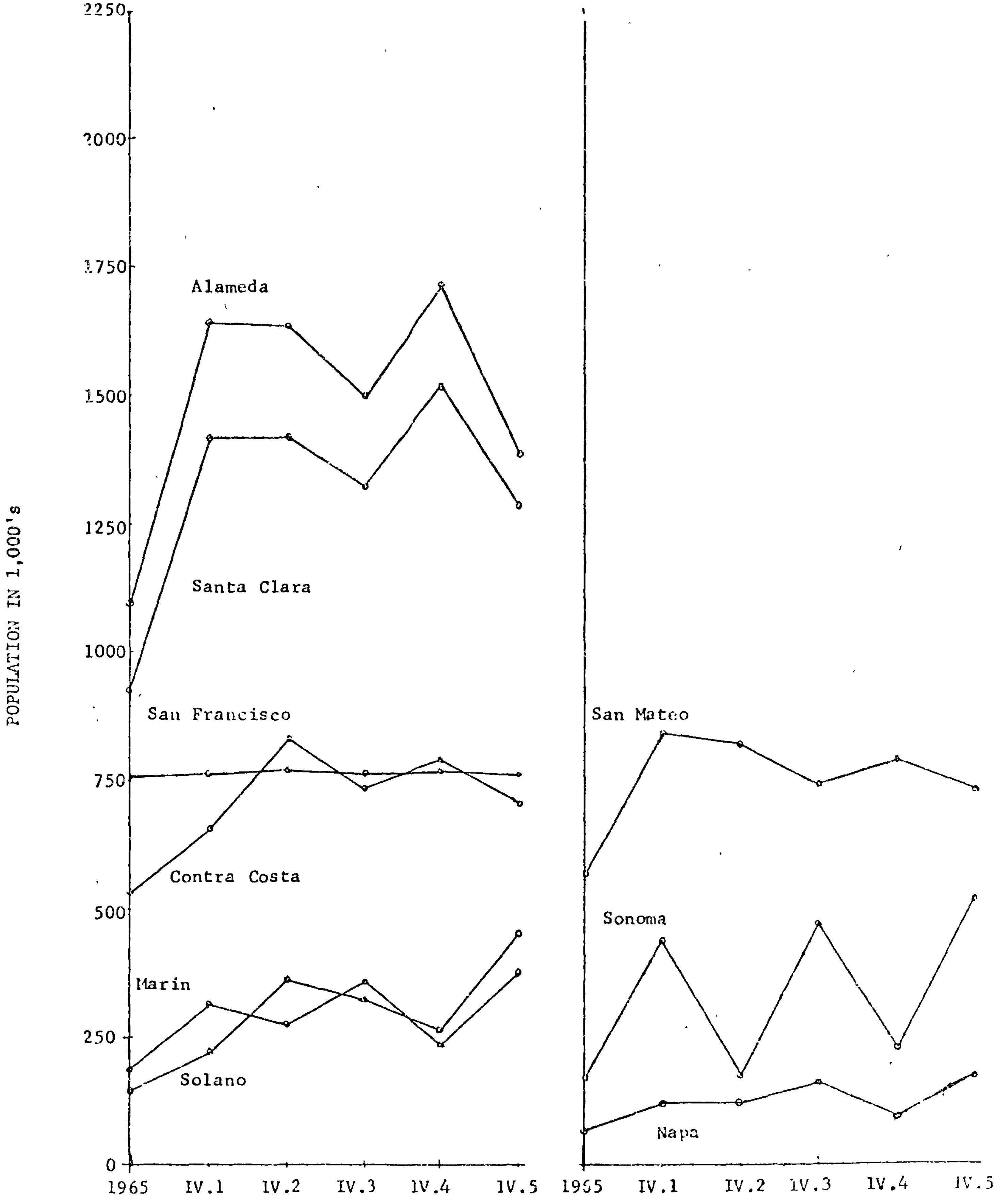
[12] See (3), Volume II, pp. 57-58, 94.

[13] Described in more detail in the following chapter.

		Run IV.1		Run IV.2		Run IV.3		Run IV.4		Run IV.5	
	1965	1980 - (1)		1980 - (2)		1980 - (3)		1980 - (4)		1980 - (5)	
San Francisco	754,754	759,821	.60%	767,173	1.6%	763,390	1.1%	766,364	1.5%	763,276	1.1%
Marin	192,523	313,982	63.0	275,484	43.0	363,621	88.8	238,932	24.1	380,640	97.7
San Mateo	553,271	840,265	51.8	762,003	37.7	746,089	34.8	788,324	42.4	735,260	32.8
Sonoma	167,497	440,337	162.8	242,874	45.0	474,468	183.2	230,152	37.4	520,828	210.9
Napa	69,378	118,582	70.9	118,737	71.1	163,118	135.1	95,836	38.1	179,825	159.1
Solano	149,386	223,074	49.3	363,296	143.1	328,929	120.1	262,109	75.4	450,116	201.3
Contra Costa	534,056	655,638	22.7	829,304	55.2	738,900	38.3	787,239	47.4	703,485	31.7
Santa Clara	922,357	1,412,834	53.1	1,412,478	53.1	1,325,489	43.7	1,520,842	64.8	1,287,851	39.6
Alameda	1,095,236	1,634,003	49.1	1,636,193	49.3	1,503,543	37.2	1,717,740	56.8	1,386,261	26.5
TOTAL	4,438,358	6,407,472									

Round IV: Base Population, Projected Total Population for 1980, and Percent Change 1965-1980

Table 4



ROUND IV - POPULATION PROJECTIONS BY COUNTY, 1980

FIGURE 6

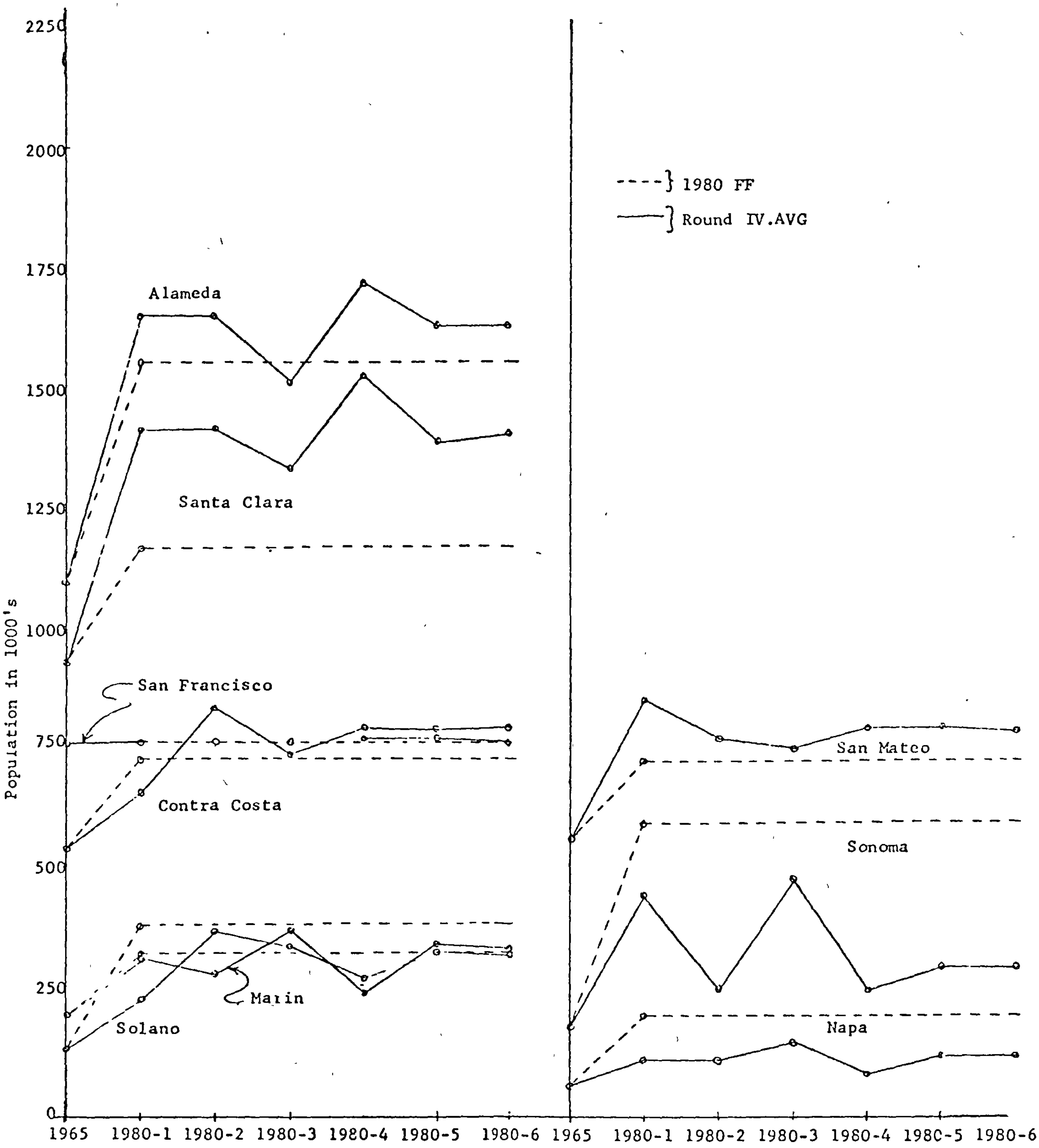
oscillations were still present and still significant. This oscillation was also visible in the network loadings, with network links connecting to zones of population oscillation showing correspondingly large volume oscillations. Looking back to the Beta-functions on which the population allocation in IPLUM is based, it was recalled that the mode of the β distribution was expected to be approximately one-third of the mean travel time of work trips. A look at the travel times on the loaded networks revealed that this was not the case. In the hope that changing the β values would improve the stability of the system, two runs, Run IV.3A and Run IV.4A, were made with increased values of β. These runs produced results very similar to those of Run IV.3 and IV.4, thus indicating that the model was not very sensitive to the values of β and that these values were not responsible for the oscillation problem.

Referring again to Figure 6, it appears that in going from Run IV.2 to Run IV.3 the system may have "overshot" a possible equilibrium point. Having done so, the subsequent attempt from Run IV.3 to IV.4 to return to the equilibrium point, overshoots in the opposite direction. Once this process has started, it appears as if it might well continue indefinitely. It seems that once having reached an "unstable" situation, the model system is unable to reach an intermediate and hopefully stable configuration. The next experiments were an attempt to help the model system to return to, or to achieve, such a stable situation. The trips generated from Run IV.3 and IV.4 were averaged together O.D. pair-by-O.D. pair. The resulting "averaged" trip matrix was loaded on the 1980-network and used as input to Run IV.5 AVG. The trips from Run IV.5 AVG were then averaged with those used as its inputs (i.e. the average of IV.3 and IV.4) and again loaded on the 1980-network. The resultant impedances were used as input to IPLUM Run IV.6 AVG. This procedure, as one would expect, did virtually eliminate the model system's oscillations. The results of these runs are shown in Table 5 and Figure 7. It

		Run IV.2		Run IV.3		Run IV.4		Run IV.5AVG		Run IV.6AVG	
	1965	1980 - 2		1980 - 3		1980 - 4		1980 - 5		1980 - 6	
San Francisco	754,754	767,173	1.6%	763,390	1.1%	766,364	1.5%	765,234	1.3%	764,950	1.4%
Marin	192,523	275,484	43.0	363,621	88.8	238,932	24.1	334,808	73.9	320,451	66.4
San Mateo	553,271	762,003	37.7	746,089	34.8	788,324	42.4	788,168	42.4	778,837	40.8
Sonoma	167,497	242,874	45.0	474,468	183.2	230,152	34.4	294,499	75.8	288,333	72.1
Napa	69,378	118,737	71.1	163,118	135.1	95,836	38.1	126,219	81.9	130,566	88.2
Solano	149,386	363,296	143.1	328,929	120.1	262,109	75.4	319,934	114.1	307,813	106.1
Contra Costa	534,056	829,304	55.2	738,900	38.3	787,239	47.4	780,736	46.1	788,455	47.6
Santa Clara	922,357	1,412,478	53.1	1,325,489	43.7	1,520,842	64.8	1,382,363	49.8	1,401,803	52.0
Alameda	1,095,236	1,636,193	49.3	1,503,543	37.2	1,717,740	56.8	1,621,579	48.0	1,626,328	48.5
TOTAL	4,438,358	6,407,542								6,407,556	

Round IV.AVG: Base Population, Projected Total Population for 1980, and Percent Change 1965-1980

Table 5



Round IV.AVG and Run 1980 FF - Population Projections by County, 1980

FIGURE 7

appears that if this procedure of trips averaging were imposed on the model system beginning with the 1980-3 run, i.e. with average of the 1980-1 and 1980-2 trips, the oscillations would be brought to within tolerable ranges. Consequently this trip averaging procedure, which is not a particularly "extreme" solution to the oscillation problem, was adopted as standard procedure in the model package.

As described earlier, two principal aims of this research effort were to attempt to wholly eliminate forecast errors due to the differences between free-flow and loaded network characteristics, and to partially eliminate the errors due to the incremental land development resulting from the changes in the transport networks. The first aim is achieved by using loaded networks as input to the land use model. The second aim is partly achieved by repeated iterations of the land use model and the network model. It should be noted that while the various intermediate steps in this iterative process do not purport to have any one-to-one correspondence to simulated points in time, the overall stream of iterations is probably similar to a simulated stream of time. (This notion is further explored in Chapter IV, p. 104).

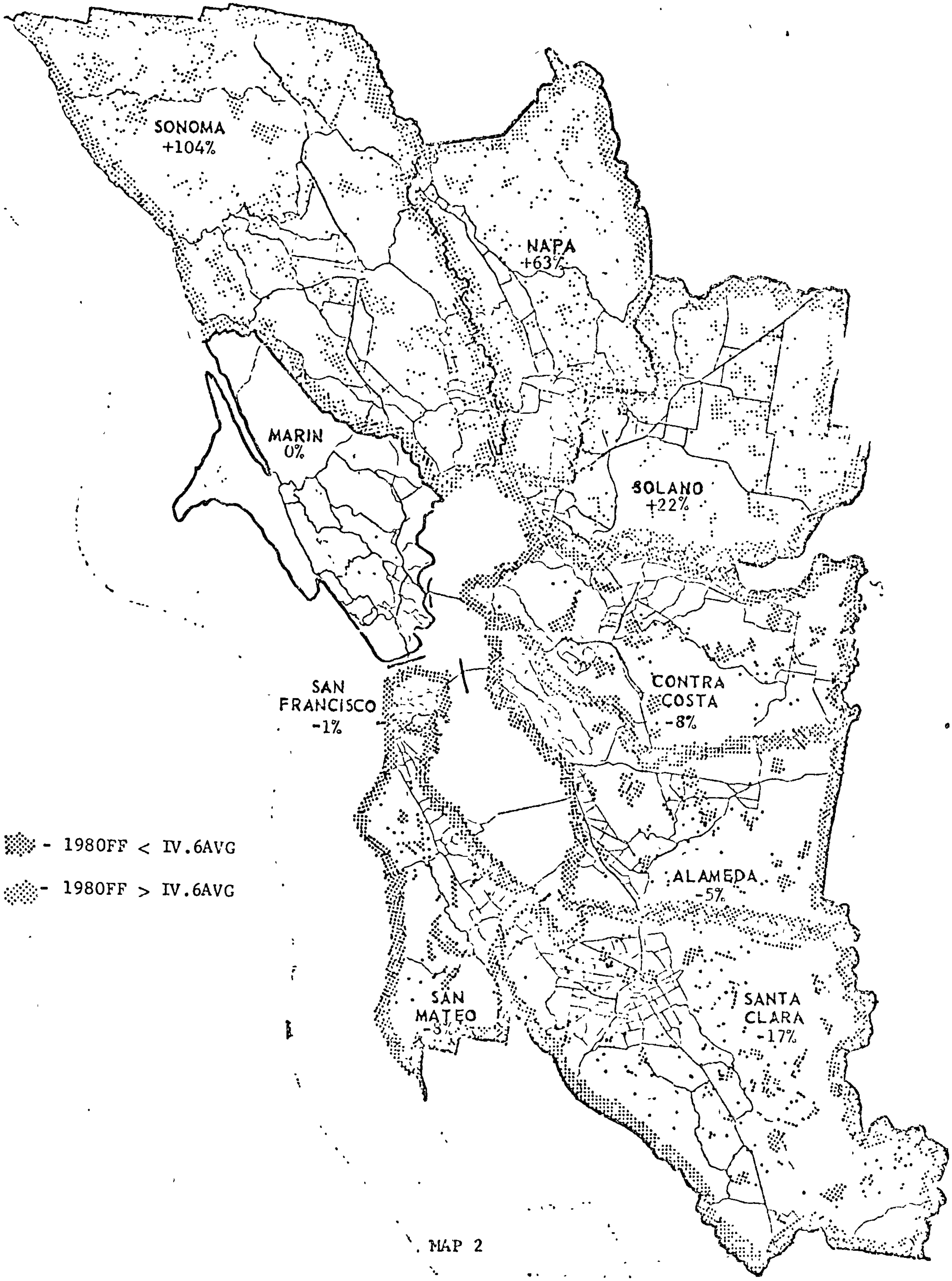
An obvious point of comparison for this integrated forecasting technique is with similar forecasts made in the traditional manner. The modified version of IPLUM was run with the "empty" (lacking any trip loadings) 1980 network impedances as input. This, of course, represents the standard procedure for unintegrated use of a land use model. This yielded a 1980 land use forecast based on the projected 1980 network without any consideration of the traffic which would develop as a result of the changed land use distribution. The county population forecasts from this run (called 1980-FF) are:

San Francisco	757,892	Solano	374,101
Marin	319,089	Contra Costa	725,874
San Mateo	716,536	Santa Clara	1,162,194
Sonoma	589,129	Alameda	1,548,474
Napa	213,256		

This set of forecasts does not reflect any of the feedback from land-use to transportation. If the actual links of the 1980 network had unlimited capacity, under which circumstance the travel characteristics of the links would not be debilitated as the volume on the links increased, then theoretically these forecasts would be acceptable though still not quite correct. The remaining theoretical error is due to the incremental development of changes in actual transport facilities which allows portions of them to come into public use and to have some effect on land use before the completion of other portions. Any model which makes its forecasts by discrete jumps through time will inevitably miss some of these phenomena.

The 1980-FF forecasts are plotted in dashed lines on Figure 7. It is quite clear from the figure that these results differ substantially from the results indicated by the stream of iterations. While it has not yet been possible to make a statistical comparison of these results, it appears that the results from the integrated model system are much more likely to be correct i.e. better forecasts. Finally, it should be noted that in those areas, shown on Map 2, where there has been substantial real growth in population (the southeastern counties), the integrated model system results are consistently greater than those of 1980-FF. This is an encouraging result and is suggestive of the potential advantages in forecasting accuracy to be had from the integrated model system.

Before concluding this chapter, some further discussion of the network modelling is in order. It is clear that it is of importance to have ways of assessing the "situation" on the transport networks. Since the origin-to-destination trip matrix will change, as described above, as the location of population and employment changes, the volume loadings on the network links will provide an accurate indication of the traffic conditions on the network. One possible measure is of total-trip-minutes". Unfortunately this measure does not accurately reflect the true



MAP 2

County Population Forecasts, 1980FF vs. Run IV.6AVG traffic situation on the network since there is no way to distinguish whether the total trip minutes come from large numbers of short trips or fewer numbers of long trips. After further consideration it was decided that the mean and standard deviation of; 1) volume/capacity ratio, 2) origin to destination impedance, and 3) trip length,would provide some useful information about the traffic situation on the network. These data are shown for Round IV.AVG in Table 6a.

While most of the oscillations in the county population forecasts can be seen to follow the oscillations of the network measures, these measures are not zone or county specific. Consequently it is not possible to compare zonal or county forecasts with the conditions on the relevant portions of the network. A crude zone-specific measure of congestion was therefore developed as follows. In each zone there is a network load-node. All traffic which either originates or terminates in the zone must pass through this load node. We look first at all of the network links which originate at the load node, the traffic on these links, and the volume/capacity ratio for the loaded links. It is then possible to compute an average volume/capacity ratio weighted by the traffic experiencing that level of congestion implied by the volume/capacity ratio. While this measure does give an indication of zone specific congestion, it has the fault of not distinguishing "through" traffic on the links, passing through the load-node (i.e. it neither originates nor terminates in that zone), from trips originating at the load node. Consequently, the particular placement of the load node in the zone affects the amount of through traffic and the results. Nevertheless, this zonal measure can be extended to the county level which averages the effect of particular load node placements, and thus comprises a county level measure of transport network congestion. These measures were calculated for the runs of Round IV.AVG and are shown in Table 6b.

If the behavior of the congestion measures is compared to the population projections, some interesting observations may be developed. There are two general

	IV.3	IV.4	IV.5AVG	IV.6AVG
Mean Vol./Cap.	0.798	0.883	0.827	0.811
S. Dev. Vol./Cap.	0.839	0.960	0.822	0.804
Mean Imped.	74.2	89.0	84.0	80.9
S. Dev. Imped.	38.3	52.0	44.3	42.5
Mean Trip Lgth.	32.0	43.5	36.2	34.1
S. Dev. Trip Lgth.	23.4	39.2	27.2	24.7

Network Characteristics on Input to IPLUM Run

Table 6a

	IV.3	IV.4	IV.5AVG	IV.6AVG
San Francisco	1.891	2.005	1.939	1.928
Marin	1.425	2.025	1.229	1.090
San Mateo	2.246	2.176	2.144	2.102
Sonoma	2.446	2.711	2.453	2.407
Napa	1.048	2.218	1.341	1.262
Solano	1.837	1.467	1.526	1.551
Contra Costa	1.558	2.311	1.567	1.566
Santa Clara	1.719	1.663	1.692	1.677
Alameda	1.810	1.922	1.777	1.766

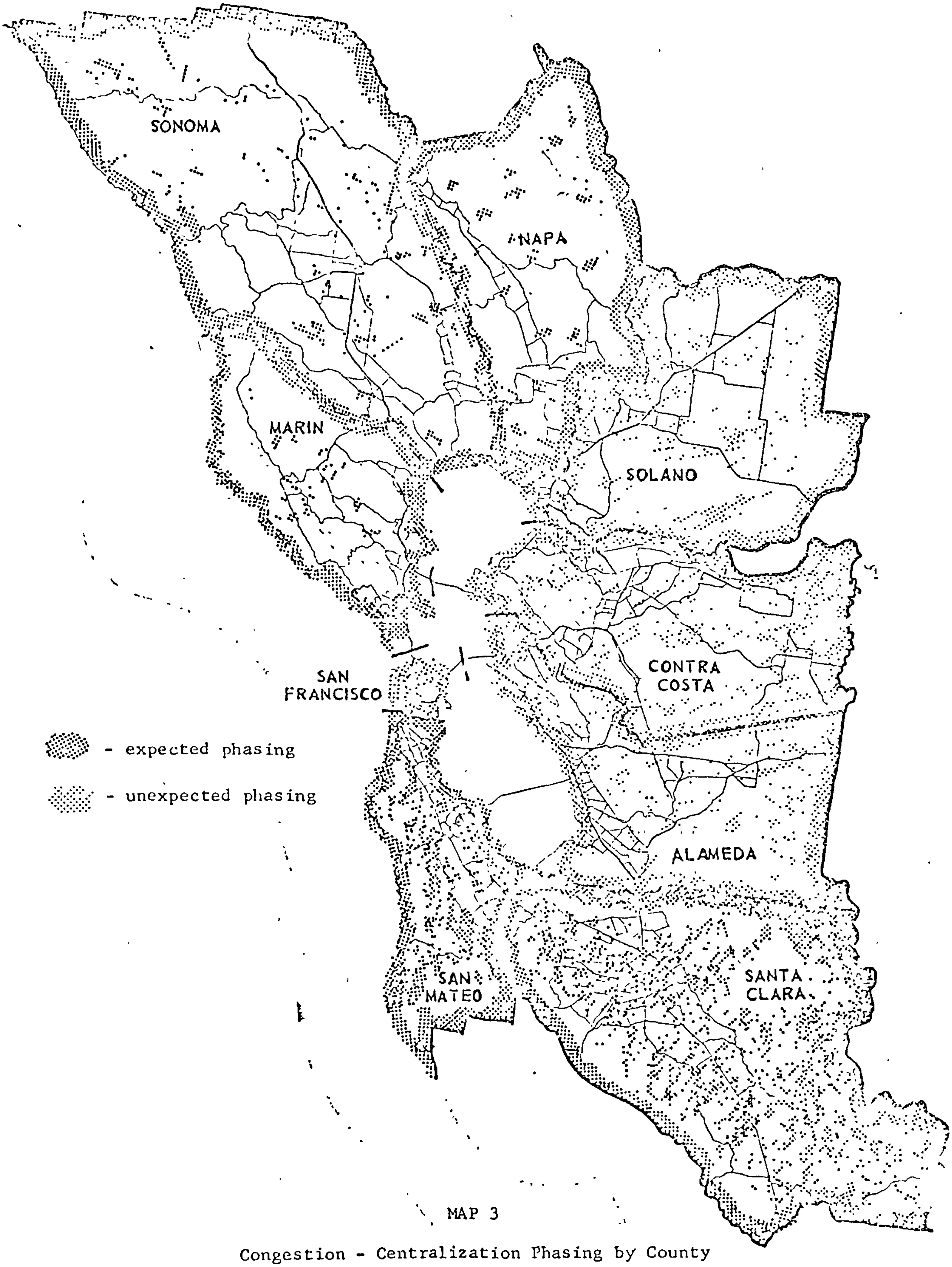
Congestion Measure on Input to IPLUM Run

Table 6B

classes of behavior to be observed. First, is where the population projections and congestion measures are one iteration "out-of-phase" with each other, e.g. IV.3 congestion is "high" and IV.3 population is "low" while IV.4 congestion is "low" and IV.4 population is "high". This situation occurs for five of the nine counties; Santa Clara, San Mateo, Sonoma, Napa, and Marin. This situation is what was expected of the model system, being of the general form: high population → high congestion → lowered population → lowered congestion → higher population → etc.

The remaining four counties; Alameda, Contra Costa, San Francisco, and Solano exhibit a different form of behavior. In these counties the population projection and congestion measures are "in-phase" with each other, e.g. IV.3 congestion is "high" and IV.3 population is "high" while IV.4 congestion is "low" and IV.4 population is "low". This situation is of the general form: high population → low congestion → lowered population → higher congestion → higher population → etc. This is a rather unexpected mode of behavior and is currently being investigated. There are several possible explanations to be tested. First, there is the possibility that the "through-traffic" ignored by the congestion measure is sufficiently important to be biasing the measure and making it an inaccurate indication to the true area specific levels of congestion. A second possibility is that congestion is not properly measured without considering numbers of originating trips and average trip lengths i.e. congestion in an area could increase if number of trips declined slightly and trip lengths decreased substantially. The fact that there appears to be a systematic spatial distribution of the counties displaying these modes of behavior, shown on Map 3, suggests that this problem can be resolved in a systematic fashion.

Two important positive points to be noted here will conclude the discussion of network measures. First, the overall network measures given in Table 6a do converge with the population projection convergence and are probably indicative



MAP 3

Congestion - Centralization Phasing by County of the desired equilibrium being reached. Second, the congestion measures shown in Table 6b also seem to converge with the converging population projections. The anomolous iteration-to-iteration behavior of the congestion measures is more likely to shed additional light on the nature of the process than it is to vitiate the results obtained. Finally, we note that in the development of a system of models and interrelationships as complex as this one, simplifying assumptions and certain gaps, both theoretical and operational, are inevitable. Here, however, and especially in view of the preliminary nature of these results, the question is not one of absolute accuracy of forecasts. The nature of the response of the whole system is the key result. The demonstration of the relative sensitivity of land use forecasts to transportation network conditions, and of traffic on the networks to land use forecasts, and that they reach equilibrium, is the principal output to date.

There is, of course, room for improvement to this model system. The following is but a sample of the questions to be explored in regard to the adequacy (or accuracy) of the system's forecasts

a) What are the consequences of redistribution of basic employment? The PLUM and IPLUM models are derivatives of the LOWRY model and as such the forecasts of basic employment are exogenous inputs to the models. It is clear that transportation networks do have an effect on the location of some basic employment.[14] Further, there are strong indications of decentralization of some basic employment in metropolitan areas. To the extent that this may represent a significant shift in workplace location, there are impacts to be felt throughout the area. How should the model system deal with this problem?

---

[14] See (6).

b) Both PLUM and IPLUM locate residences strictly as a function of work-trip distributions, available land, and subsequent, somewhat arbitrary, constraints. Work with other land use models indicates that locational amenities of potential residential locations have a significant impact. How can these models be modified to include these phenomena?

c) The networks now used in the system are highway only. Clearly one of the important questions for future metropolitan development will be that of utility of public (mass) transit systems. How can the model system be modified to include both highway and transit networks as well as a model split procedure?

d) The level of sectoral detail throughout the model system is too highly aggregated. Where and how can disaggregations be accomplished?

All of these questions, plus others, require further investigation.

References

1. Bay Area Transportation Study Commission, "Bay Area Transportation Report," BATSC, Berkeley, May, 1969.

2. Goldner, W. "Projective Land Use Model," Bay Area Transportation Study Commission, Technical Report 219, 1968.

3. Goldner, W. et al. "Projective Land Use Model-PLUM" Vol. I, II, III, Institute of Transportation and Traffic Engineering, University of California, Berkeley, March, 1972.

4. Harris, B. et al. "Research on an Equilibrium Model of Metropolitan Housing and Locational Choice: Interim Report." Planning Sciences Group, Institute for Environmental Studies, University of Pennsylvania, March, 1966.

5. Hill, D. M. "A Growth Allocation Model for the Boston Region," Journal of the American Institute of Planners, Vol. XXXI, No. 2, May, 1965.

6. Putman, S. H. "Intraurban Employment Forecasting Models: A Review and a Suggested New Model Construct," Journal of the American Institute of Planners, Vol. XXXVIII, No. 4, July, 1972.

7. Nathanson, J. "Basic Employment Model", Bay Area Transportation Study Commission, Technical Report 222, 1970 (Preliminary Draft)

8. U. S. Department of Transportation, Federal Highway Administration, "Urban Transportation Planning - General Information and Introduction to System 360" March, 1972.

9. U. S. Department of Transportation, Federal Highway Administration, Office of Highway Planning, Urban Planning Division. "GS-10, RFP-26", March 5, 1971.

IV. METHOD OF ANALYSIS: FURTHER ELUCIDATION

As a part of the development of the model package described in Chapter III, numerous detailed investigations were undertaken. In order to avoid further obfuscation of Chapter III's already complicated discussion, certain more detailed discussions of method were reserved for presentation here.

Before moving to the specific discussions, some mention of the data used in the project analyses is appropriate. First, the data set for the earliest prototype model runs was fabricated specifically for those test purposes and, the workability of the model package having been established, was simply stored for possible future use. The bulk of the project's analyses were conducted with data for the San Francisco Bay area, consisting of nine counties disaggregated to two-hundred ninety-one zones. This is the same data base as was used for the Bay Area Transportation Study and which, with some modification and up dating, is still being used by the Metropolitan Transportation Commission and the Association of Bay Area Governments. All the data used for this research project were obtained from the M.T.C. and, due to the limits of the project resources, had to be accepted as received, without further data collection or verification. A useful task to be undertaken as part of future work on this project would be to attempt to update this data base (from published sources) and subsequently do further testing of the model package. Many of the questions raised in the following discussion could at least begin to be answered if such updated data files were available.

Modification to IPLUM Allocation Functions

First, recall that IPLUM is the more recent, incremental version of PLUM. A copy of the computer program and documentation for the model were received by the project in early spring of 1972. In the process of adapting that program to our computer system several program errors were discovered and are noted in volume 2 of this report.

There are several constraints imposed on the IPLUM allocations during each program run. These are described by Goldner in Vol. II of his report.[1] The first question in regard to these constraints was one of the degree to which these constraints modified the model's "true" allocations. The second question was that of the theoretical and/or conceptual legitimacy of the individual constraints.

A set of experimental model runs were made investigate the operation of the constraint factors in IPLUM. $\dot{C}_1$ is the correction factor to neutralize the "distorting" effects of components of the allocative functions. It scales the allocated local serving employment increment (summed over all zones) to the area wide control total of local serving employment change (a calculated input to the model). $\dot{C}_2$ is the correction factor to neutralize the effect of the population-per-employed-resident for the base year, since IPLUM operates on increments rather than base year values. It scales the allocated residential non-working population increment summed over zones to a calculated input of area-wide residential non-working population change (control total for projected year). Four runs were made of IPLUM to test the effects of $\dot{C}_1$ and $\dot{C}_2$ on the allocation mechanism of the model. (1) Base run, (2) $\dot{C}_1$ not invoked, $\dot{C}_2$ remaining the same, (3) $\dot{C}_1$ invoked and $\dot{C}_2$ not invoked, and (4) both $\dot{C}_1$ and $\dot{C}_2$ not invoked. The essence of the results of these runs was that the model is critically sensitive to the value of $\dot{C}_1$ and $\dot{C}_2$. Further, there was some suspicion that the magnitude of C_1 (≈ 13) was indicative of potentially serious problems in the model's equation structure. There was clearly a need for further, and particularly thorough, investigation of these constraints.

---

[1] See (6) Vol. II pp. 78, 84-85, 92-93 and Chapter 6 in general.

The most critical components of the IPLUM model, indeed of most land use models, are the functions which allocate activities to geographic locations. Any thorough investigation of a land use model must begin with a detailed examination of these allocation procedures. It is appropriate here to try to present a prose description of the allocation procedures in IPLUM. In general the procedure has three steps. First, a function describing the probability (all other factors being constant) of travelling a given time-distance for a particular trip type e.g. work-to-home, work-to-shop, home-to-shop, is described and split into 31 intervals. Second, for each zone-to-zone pair, an algorithm uses the above mentioned probability function to produce a zone-to-zone matrix of trip probabilities. Finally, the elements of this matrix are "weighted" by indices of zonal "attractiveness". In the following discussion we deal only with the first two steps in this process.

We begin with a probability function as follows, divided into 6 intervals:

Travel Time = t	Probability of Trip of Length t
1 minute	.50
2 minute	.30
3 minute	.15
4 minute	.04
5 minute	.01
>5 minute	.00

Now consider that we shall investigate two zones of origin, Zone 1 and Zone 2. Suppose that around each of these zones we define all other zones to be in time rings. Thus around e.g. Zone 1 there is a ring of zones 1 minute away, a ring of zones 2 minutes away, etc. Similarly for Zone 2 there is also a set of such time rings. It is clear that for most regions, depending upon transport facilities, geography, etc., the numbers of destination zones in say, the 2 minute ring, for different zones of origin will be different. Consider the following table:

	No. of destination zones in ring t minutes from the origin zone				
Origin Zone	1 min.	2 min.	3 min.	4 min.	5 min.
Zone 1	5	0	3	0	2
Zone 2	2	2	3	1	2

Thus in the 2 minute ring for Zone 1 there are 0 destination zones while in the 5 minute ring for Zone 2 there are 2 destination zones. With this background we may consider the algorithms used in PLUM and IPLUM to produce the matrices of trip probabilities.

In the PLUM algorithm the essential notion seems to be that the travel probabilities are associated with travel to rings. Consequently the probability of a trip from a given origin zone to a given destination zone is equal to the probability of travelling to the destination zone's ring divided by the number of zones in the ring. Consequently the probability of a trip from Zone 1 to a zone in ring 1 is .5/5 = .1 while the probability of a trip from Zone 2 to a zone in its ring 1 is .5/2 = .25. While it could be argued that this procedure introduces the effects of zone-to-zone competition for locators, it is not clear that this is the most appropriate point at which to introduce these effects as opposed, for example, to doing so in the calculation of attractiveness indices. The issue is confounded by a treatment of the case of rings containing 0 zones which apparently stems from notions of intervening opportunities. When a destination ring contains 0 zones the entire probability of travelling to that ring is added to the next further ring. Looking again at our example, from Zone 1 there are 0 zones in ring 2. The zone 1 probability of travelling to a zone in ring 3 is (.30 + .15)/3 = .15. Summarizing the PLUM calculations for our example.

	Probability of trip to dest. zone in ring				
Origin Zone	1	2	3	4	5
Zone 1	.5/5	n.a.	(.3+.15)/3	n.a.	(.04+.01)/2
Zone 2	.5/2	.3/2	.15/3	.04/1	.01/2

In the IPLUM algorithm the emphasis seems to be on intervening opportunities and in general there is a higher level of consistency. The probability of travelling from a given origin zone to a given destination zone is equal to the probability of travelling to the destination zone's ring divided by the time to the ring. Consequently the probability of a trip from Zone 1 to a zone in ring 1 is .5/1 = .5 and the probability of a trip from Zone 2 to a zone in its ring 1 is also .5/1 = .5. Again, however, an inconsistency is introduced with the handling of rings containing 0 zones. This is done in a manner similar to that for PLUM, with the entire probability for the 0 ring being added to the next further ring. Thus, since there are 0 zones in Zone 1's ring 4, the probability of travelling from Zone 1 to a zone in ring 5 is (.04 + .01)/5 = .01. Summarizing the IPLUM calculations for our example.

	Probability of trip to dest. zone in ring				
Origin Zone	1	2	3	4	5
Zone 1	.5/1	n.a.	(.3+.15)/3	n.a.	(.04+01)/5
Zone 2	.5/1	.3/2	.15/3	.04/4	.01/5

Again, consider the "theory" implicit in the access-attractiveness types of allocation functions with which we are here concerned. In general one is attempting to describe the aggregate behavior of a group of tripmakers. Considering a particular origin, "other variables" being equal, the theory postulates that the trips to any given destination will be a function of the difficulty of reaching the destination and the degree to which that particular destination is capable of satisfying the trip purpose. The difficulty of reaching the destination is usually expressed in terms of travel time and/or cost. The degree to which the particular destination is capable of satisfying the trip purpose is usually expressed in terms of some measure of attractiveness or quantities of attractors located at the destination. Two of the "other variables" are particularly important and are therefore often included in these formulations. First is the possibility of intervening opportunities for satisfying the trip purpose before reaching the "intended" destination. And second, which is really just a different form of the first, there is the possibility of competition amongst alternative destinations. Both of these "other" variables have been used in various formulations of these allocation functions for the many land use models which have been developed in the past decade.

A convenient way to think about such access functions is to divide them into two components. The first component deals with the probability of a trip for a given trip purpose being of a certain length (in terms of time and/or cost). The second component deals with the attractiveness of the destination. This general notion seems to be the basis for the PLUM and IPLUM allocation functions as described by Goldner in Vol. II of his report, cited above.[2] Unfortunately it appears that the allocation functions as implemented in the models are inconsistent with this theoretical framework. In particular, the second step of the

---

[2] See (6) pp 53-58 and Chapter 4 in general.

allocation procedures, by virtue of the manner in which rings containing 0 zones are handled, produces an unacceptable inconsistency. With those procedures the probability of a trip of length t can, and does in practice, differ from origin to origin quite independently of the attractors. Looking at the numerical examples worked out above, in the IPLUM example, the probability of a 3 minute trip from Zone 1 is .15 while the probability of a 3 minute trip from Zone 2 is .05. While it might be argued that this difference stems from considerations of "intervening opportunities", close inspection of the algorithms will reveal that there is no basis for this contention.

In view of the above, it was decided to modify the first two steps of the IPLUM allocation procedure to make them more consistent with the general theory briefly mentioned above. The first modification made was simply to assure that the sum of the probabilities in the initial probability function was indeed equal to 1.00. This was, in part, the stated purpose of the C_1 correction factor, but which was discovered in our investigation not to be the case. This correction of the probability function ranged from 0% to 3% and produced a change in C_1 of 10% and changes in some sub-area allocations of up to 20%.

The second change involved asserting that the probability of a trip of length t is the same for all possible zones of origin. The algorithm was modified accordingly, with the problem of rings containing 0 zones being shifted to the attractor portion of the calculations. With this procedure, and the above numerical example the following table would be obtained:

	Probability of trip to dest. zone in ring				
Origin Zone	1	2	3	4	5
Zone 1	.50	.30	.15	.04	.01
Zone 2	.50	.30	.15	.04	.01

The calculation of attractors was left unaltered. The results of this modification were substantial, with the value of C_1 falling from 18.9 to 0.91 while changes in some sub-area allocations were as great as 50%.

The discussion thus far has described modifications and tests of the allocation function for households. In the IPLUM model a similar procedure is used for the allocation of "local-serving" employment. In this case, instead of work-to-home trips, the model is concerned with home-to-shop trips. The same modifications discussed above with reference to households were introduced into the employment allocation. This time a fourth IPLUM run was made. This, of course, resulted in significant changes in the employment estimates, and subsequent changes in the population estimates. The value of C_1 for this run was 1.000. The incremental "unconstrained" dwelling unit allocations from the above runs are tabulated below.

	Run-1	Run-2	Run-3	Run-4
San Francisco	65,531	66,097	64,085	48,989
Marin	40,261	40,350	31,490	32,448
San Mateo	76,361	76,566	102,725	110,302
Sonoma	102,954	103,370	48,652	68,910
Napa	26,000	26,057	12,709	14,607
Solano	59,493	59,521	32,564	34,996
Contra Costa	47,736	47,630	37,412	39,153
Santa Clara	123,035	122,391	223,091	191,712
Alemeda	97,129	96,750	70,318	81,100
TOTAL	638,493	638,727	623,042	622,213
C_1	19.204	18.980	0.912	1.000

The different runs were:

Run-1 = Base run

Run-2 = Sum of probabilities constrained to 1.00

Run-3 = Same, plus probability of a trip of a given length is the same for any origin - for households only.

Run-4 = Sane as Run-3 but extended to employment as well as households

The results of these runs in terms of incremental "unconstrained" local-serving employment are tabulated below.

	Run-1	Run-2	Run-3	Run-4
San Francisco	170,577	173,508	154,629	104,060
Marin	15,311	15,305	8,311	13,157
San Mateo	26,534	26,329	23,447	57,021
Sonoma	24,668	24,995	13,316	41,165
Napa	9,003	9,036	4,996	10,297
Solano	33,664	33,666	16,364	21,250
Contra Costa	39,470	39,218	34,151	39,163
Santa Clara	172,116	170,389	259,433	216,827
Alameda	112,431	111,327	89,124	100,830
TOTAL	603,767	603,766	603,765	603,766
C_2	0.596	0.596	0.603	0.591

With this form of the allocation function, several test runs of the model were made and runs of the model package were undertaken. After two rounds of the full model package were completed it was clear that there was excessive instability in the system which would have to be eliminated. In particular, about 10% of the zones showed substantial forecast oscillations in iterations 3, 4, and 5.

The first factor to be investigated in this regard was the manner by which the trips (both work and shop) were loaded on the networks between IPLUM runs. In the early version of the model package <u>all</u> of the trips generated in a run of IPLUM were subsequently loaded (by an incremental procedure) on the network. The travel time characteristics of this network were then used as input to the succeeding run of IPLUM. The incremental procedure used in the loading of networks loads 40%, 40%, and 20% of the projected trips onto the network in three passes.[3]

---

[3] See Appendix A-2

It was hypothesized that, in cases where successive iterations of IPLUM did not produce extremely different configurations of land use, the consequent trip matrices would not show extreme differences either. If this were indeed the case then it would be possible to effect a 30% reduction in computer usage for the network loadings by using a marginal loading procedure. This procedure could operate as follows:

1. When loading the network with trips generated by iteration (N) of IPLUM save the information for both the 80% loaded and the 100% (fully) loaded networks.
2) Use the information from the fully loaded network as input to the (N+1)th iteration of IPLUM and generate a new land use allocation and new trip matrices.
3) Calculate the difference, O.D. pair by O.D. pair, between the trip matrices generated in the (N+1)th IPLUM iteration and 80% of the trips in the trip matrices generated in the (N)th iteration of IPLUM.
4) This new matrix of trips (differences) is then added to the 80% loaded network (which was saved) from the (N)th iteration of IPLUM. The 80% loaded network is still saved and the new fully loaded network characteristics are used as input to the (N+2)th iteration of IPLUM.
5) This process is repeated in successive iterations. This new marginal loading procedure, after some modifications to the network programs package was adopted for future runs.

In Round II of the integrated package runs, the use of this marginal network loading procedure was initiated after the second iteration of IPLUM, for input to the third iteration. In the fourth iteration of IPLUM in this Round, certain zones displayed considerable instability with regard to their population allocations. Because of the extreme complexity of the complete model package, it was not possible to determine by inspection what the source of this instability was. For reasons which, at the time of this writing, are no longer apparent, the third and fourth iterations in this round were given the suffix B and are henceforth identified as II.3B and II.4B respectively.

An important procedure for investigating complex systems involves attempting to separate or isolate components of the system in order to localize sources of error or instability. In line with this notion it was decided to prepare an alternate version of Round II which did not use the marginal network loading procedure described above, but which used the straight-forward full loading of the network between each iteration of IPLUM. Since the marginal procedure had only been used to produce network inputs to runs II.3B and II.4B of IPLUM, it was only necessary to start this new version of Round II after run II.2 of IPLUM Three subsequent iterations were run, and are hereafter referred as II.3a, II.4A, and II.5A (again the reason for the choice of the suffix A has been lost in the mists of time). The land use allocations produced in runs II.3A and II.4A were only slightly different from those of II.3B and II.4B, and still demonstrated the instability which had been present before. Run II.5A was made in order to determine if the instability might have been a transient effect, but unfortunately it persisted. Tables 1, 2 and Figures 1, 2 show the outputs of these runs.

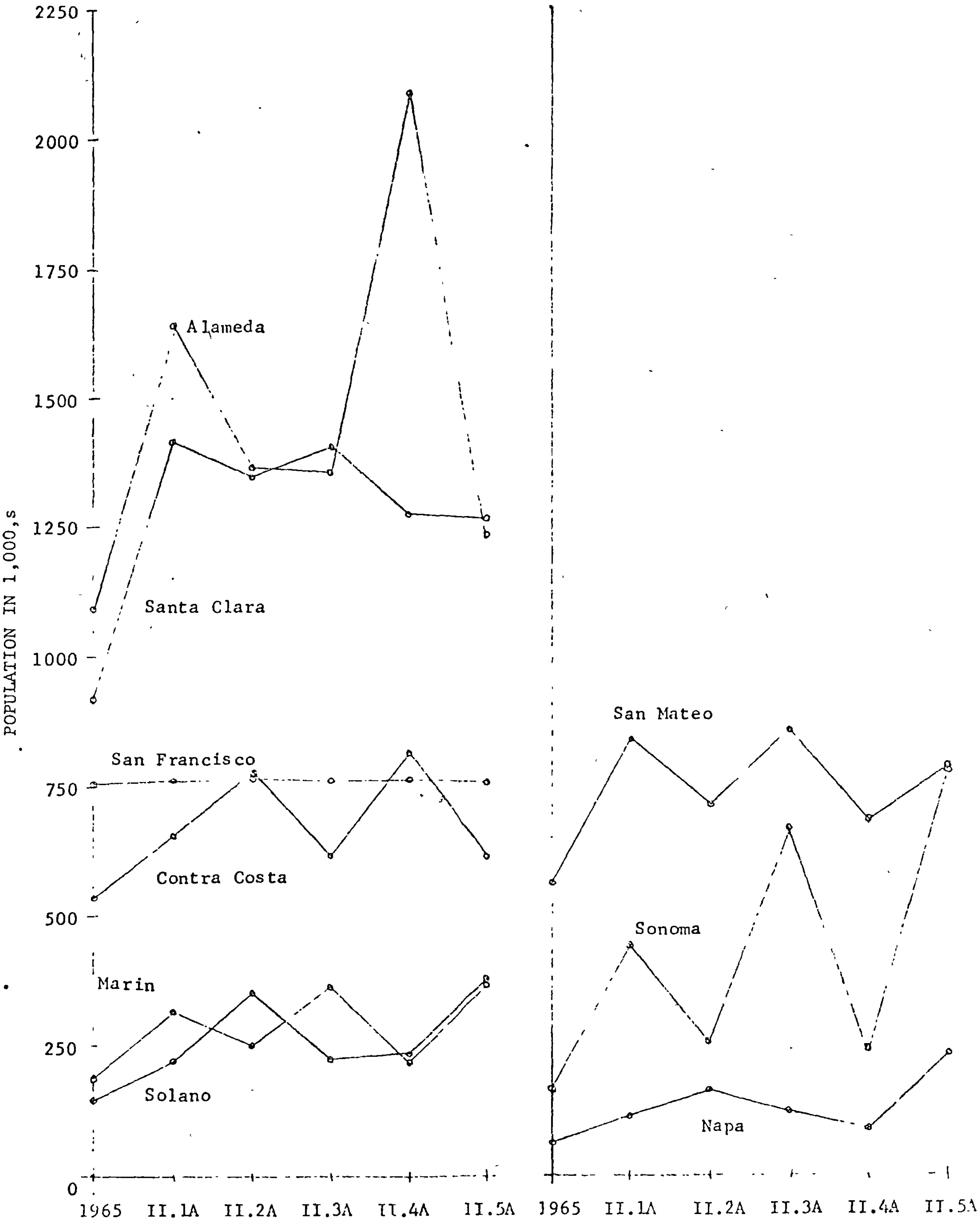
Closer inspection of the model runs indicated that the apparent instabilities were showing up in only a small number of the 291 zones in the San Francisco Bay Area (being used as the data base for the model runs). A telephone call to the Association of Bay Area Governments (ABAG) in San Francisco revealed that the two zones which were showing the largest oscillations were zones in which there actually was a great deal of development taking place in the Bay Area. Though this statement was reassuring, it did not solve the instability problem. Further examination of the changes of network impedances to the oscillating zones

		II.1A		II.2A		II.3A		II.4A		II.5A	
	1965	1980 - (1)		1980 - (2)		1980 - (3)		1980 - (4)		1980 - (5)	
San Francisco	754,754	759,821	0.67	762,455	1.02	759,217	0.59	760,607	0.78	758,369	0.48
Marin	192,523	313,982	63.1	251,919	30.9	360,995	87.5	220,628	14.6	362,523	88.3
San Mateo	553,271	840,265	51.9	719,364	30.0	858,236	55.1	678.678	22.7	781,160	41.2
Sonoma	167,497	440,337	162.9	255,273	52.4	670,765	300.5	240,426	43.5	782,643	367.3
Napa	69,378	118,582	70.9	162,268	133.9	155,177	123.7	95,049	37.0	237,311	242.1
Solano	149,386	223,074	49.3	351,312	135.2	223,454	49.6	231,381	54.9	378,531	153.4
Contra Costa	534,056	655,638	22.8	771,037	44.4	616,419	15.4	817,730	53.2	611,018	14.4
Santa Clara	922,357	1,412,834	53.2	1,348,370	46.2	1,408,025	52.7	1,275,817	38.3	1,260,181	36.6
Alemeda	1,095,236	1,643,003	50.0	1,364,748	24.6	1,355,254	23.7	2,087,224	90.6	1,235,809	28.3
TOTAL	4,438,358	6,407,472	44.4	6,407,476	44.4	6,407,465	44.4	6,407,477	44.4	6,407,545	44.4

Round IIA[*]: Base Population, Projected Total Population for 1980, and Percent Change 1965-1980.

(*IIA Series, Full Loading Between IPLUM Runs)

Table 1



ROUND II.A - POPULATION PROJECTIONS BY COUNTY, 1980

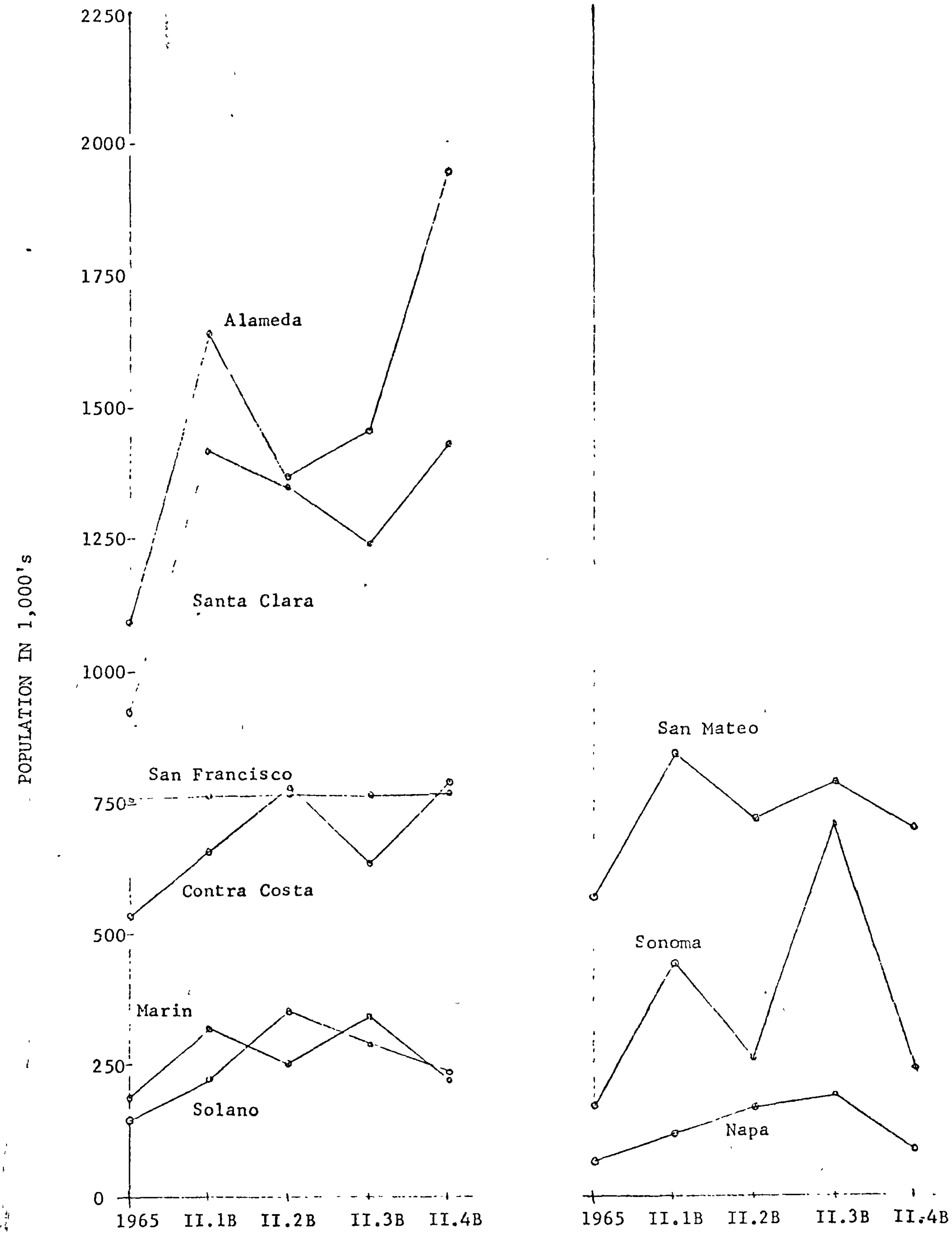
FIGURE 1

		II.1B		II.2B		II.3B		II.4B	
	1965	1980 - (1)		1980 - (2)		1980 - (3)		1980 - (4)	
San Francisco	754,754	759,821	0.671	762,455	1.02	758,553	0.503	760,438	0.753
Marin	192,523	313,982	63.08	251,919	30.85	341,856	77.57	223,377	16.03
San Mateo	553,271	840,265	51.9	719,364	30.0	787,382	42.3	700,500	26.6
Sonoma	167,497	440,337	162.9	255,273	52.4	708,799	323.2	244,780	46.1
Napa	69,378	118.582	70.9	162,268	133.9	190,562	174.7	93,982	35.5
Solano	149,386	223,074	49.3	351,312	135.2	288,630	93.2	229,551	53.7
Contra Costa	534,056	655,638	22.8	771,037	44.4	635,465	19.0	783,855	46.8
Santa Clara	922,357	1,412,839	53.2	1,348,370	46.2	1,239,020	34.3	1,429,129	54.9
Alemeda	1,095,236	1,643,003	50.0	1,364,748	24.6	1,457,279	33.1	1,941,722	77.3
TOTAL	4,438,458	6,407,472	44.4	6,407,476	44.4	6,407,468	44.4	6,407,474	44.4

Round IIB*: Base Population, Projected Total Population for 1980, and Percent Change 1965-1980

(*IIB Series, Partial Loading Between IPLUM Runs)

Table 2



ROUND II.B - POPULATION BY COUNTY, 1980

FIGURE 2

indicated that though there were significant changes in impedances, these alone would not account for the large changes in population in these zones. To confirm this observation, a special run of IPLUM was made. The impedances input to run II.3A and run II.4A were averaged together. The new impedances, the arithmetic average of the two separate matrices, were used as input to run II.AVG of IPLUM. The expectation was that if the oscillations were principally due to the impedance changes, despite the fact that the averaging procedure ignored the non-linearities of the network volume-delay function, the population allocations from II.AVG should be somewhere near the midpoint between the population allocations in II.3A and II.4A. The results are shown in Table 3 and are a rather emphatic denial of the hypothesis that in those areas where there are problems of oscillation, the impedances may be largely responsible. What was discovered, however, was that virtually all of the problem zones tended to be on the fringes of the Bay Area and were both large in total area as well as having a high percentage of vacant land. This bit of information led to further investigation of the IPLUM allocation function for residents.

The residential allocation function in IPLUM may be considered as having two components. The first component deals with the probability of a trip for a given trip purpose being of a certain length (in terms of travel time and/or cost) The second component deals with measures of attractiveness of possible trip destinations. The first component had already been thoroughly investigated and modified as described above. In this discussion we are concerned only with the second, or attractors, component of the residential allocation function.

In Goldner's published version of IPLUM a rather simple formulation is used for the measure of a zone's attractiveness. Called a measure of "opportunities", the variable is constructed as follows (using Goldner's notation):

	II.3A	II.AVG	II.4A
San Francisco	759,217	762,101	760,607
Marin	360,993	244,032	220,628
San Mateo	858,236	753,394	678,678
*Sonoma	670,765	247,984	240,426
Napa	155,177	102,127	95,049
Solano	223,454	227,754	231,381
Contra Costa	616,419	745,428	817,730
Santa Clara	1,408,025	1,400,946	1,275,817
*Alameda	1,355,254	1,923,773	2,087,224

Population Allocations for 1980, Three Runs of IPLUM
(*Counties with oscillation problems)

Table 3

$$O_h(i) = a_v(i) * \frac{h(i)}{a_r(i)}$$

where:

$O_h(i)$ = number of opportunities for new residential development in zone i

$a_v(i)$ = vacant acreage in zone i

$a_r(i)$ = residential acreage in zone i

$h(i)$ = number of housing units in zone i

This constructed variable is also referred to by Goldner as "the residential holding capacity of the zone".

While a number of objections might be raised to this measure of attractiveness, our principal concern is with only one detrimental characteristic. In a zone containing, say, a small town and a large quantity of vacant acreage, the measure of opportunities is likely to be quite substantially overstated. That this was indeed the case, became abundantly clear on inspection of the troublesome oscillating zones in the model runs described above. That Goldner was aware of this problem is clear from statements in his IPLUM report as well as in several of his project's working memoranda. In fact, he proposed a modification to the attractor formula which was to compensate for the problem. For reasons nowhere specified, he however moved this compensating factor to a part of the IPLUM program where it is rarely ever implemented. It is our hypothesis that due to the errors in the probability portion of his allocation function, this attractor correction was found to be ineffective or perhaps detremental to the model results and was therefore shifted elsewhere in the program.

The theoretical basis for this attractor correction seems to be based on the notion that the zones in which these problems appeared tended to have large quantities of vacant land, but lacked the "infrastructure" to serve the rapid development of this land. Goldner's proposed correction factor seems reasonable and is described below in his words:

> "The land development constraint is an approximation to the area in the zone served by adequate infrastructure, if it were known. Zones may have large quantities of vacant land but are still not developable due to the time-lag in supplying infrastructure. Utilities, water, sewers, schools, local government service, and other elements of the infrastructure cannot be supplied instantly; it takes years, sometimes decades for all these elements to be emplaced. For zones with relatively little development, the constraint (expressed as a proportion between zero and one) should be predominant, but for zones almost totally developed, the constraint should have virtually no effect.
>
> A function with this property is:
>
> $$G(x) = (1-e^{-3x})/(1-e^{-3})$$
>
> where x is the fraction of usable land developed. This function is invoked when zones have little development, but quickly increases so that it has a negligible effect with moderate increases in development. For example, at zero percent development the constraint is fully binding whereas at ten percent development the constraint is almost half gone."

This correction factor was then moved back into the allocation section of our version of IPLUM as a modification to the "attractors" component of the household allocation function.

It will be recalled that the product of the "trip probabilities" and the "attractors" is normalized before being used for the household allocation in IPLUM. This being the case, the introduction of the developability correction factor involves simply multiplying it times the "attractors" prior to the normalization. Thus the modified measure of "opportunities" or "attractors" is:

$$O_h(i) = a_v(i) * \frac{h(i)}{a_r(i)} * G(i)$$

where:

$$G(i) = \frac{1-e^{-3x(i)}}{1-e^{-3}}$$

and, where

$$x(i) = \frac{a_t(i)-a_u(i)-a_s(i)-a_v(i)-a_k(i)}{a_t(i)-a_u(i)-a_s(i)-a_k(i)}$$

and, where

$a_t(i)$ = total acreage in zone i

$a_u(i)$ = unusable acreage in zone i

$a_s(i)$ = streets/highways acreage in zone i

$a_v(i)$ = vacant (other) acreage in zone i

$a_k(i)$ = vacant (industrial) acreage in zone i

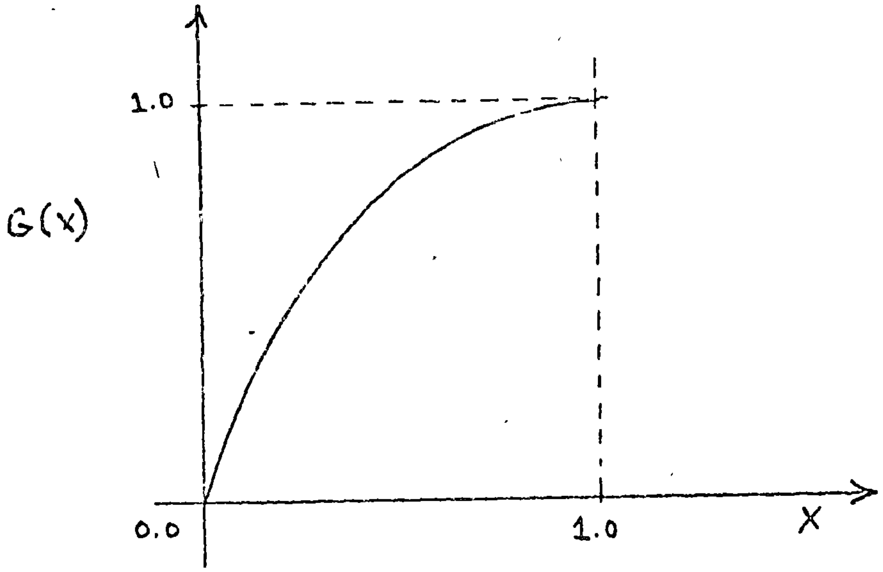
At this point two test runs of IPLUM with the modified "opportunities" measure were undertaken. Remembering that the purpose of this investigation was to discover the causes for oscillation in the population allocations, the inputs to runs II.4A and II.5A were used in two further test runs of the model. Before presenting the results of these runs, mention of the role of population control totals in the model is appropriate. In the IPLUM model, as is the case with virtually all urban land use models, there is a requirement for a regional control total for population. This value, exogenously supplied, is a constraint on the summation of the model's zonal allocations. An implication of this fact, with which we are concerned here, is that if from one run to another a particular zone shows a considerable decrease in population while the control total remains unchanged, then it is likely that other zones will have population increases to "absorb" the population from the decreased zone. The results of the two test runs are shown in Table 4 along with results of relevant other IPLUM runs. Note that the same impedance inputs were used in run II.4A and run II.4AG, which contained the new correction factor. Similarly the inputs to runs II.5A and II.5AG were identical. From these results it appears that oscillations in both the increasing and decreasing situations are somewhat damped by the addition of the correction factor.

	II.3A	II.4A	II.4AG	II.5A	II.5AG
San Francisco	759,217	760,607	766,434	758,369	762,909
Marin	360,995	220,628	225,962	362,523	406,388
San Mateo	858,236	678,678	739,142	781,160	751,859
Sonoma	670,765	240,426	243,444	782,643	536,659
Napa	155,177	95,049	89,339	237,311	183,622
Solano	223,454	231,381	240,993	378,531	457,115
Contra Costa	616,416	817,730	836,125	611,018	691,476
Santa Clara	1,408,025	1,275,817	1,398,746	1,260,181	1,347,906
*Alameda	1,355,254	2,087,224	1,867,354	1,235,809	1,273,610

Comparison of Round II Iterations 4 and 5 With and Without Land "Developability" Adjustment "G"--Projected Total Population for 1980.

Table 4

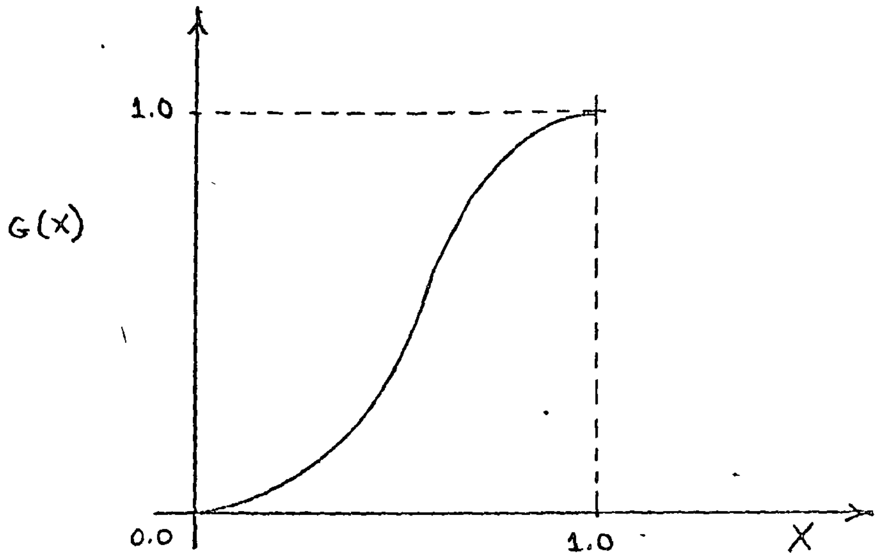
Further reflection on the problem of development in large vacant areas has suggested an alternative correction factor. Goldner's correction factor, described above, can be shown graphically as:



X = developed land/developable land

G(x) = developability correction factor

The implication of this form is that for largely undeveloped zones very small differences in X result in very substantial differences in developability. Once, however, substantial development has taken place in a zone, further development results in substantially smaller increases in developability. It is, however, probable that developability increases rather slowly as X increases until some threshold level is reached. At that point small additional increases in X result in large increases in developability until a substantial level of development is reached, at which point the curve would flatten out. This notion is shown graphically below.



This hypothesis was investigated as a student project with programming and computer time support from this project. The results are reported in Appendix C-2.

In summary, after further investigation of the IPLUM household allocation procedure and the results thereof, it was decided that further modification was necessary. Based on earlier, and subsequently deleted, work of Goldner; a correction factor, presumably a surrogate measure of infrastructure was introduced. This factor has the property of damping the attractiveness of fringe areas of the region which contain large quantities of vacant acreage; portions of which are not really ready for the development which, in the absence of the correction factor, the model would forecast. The introduction of this correction factor appears to have produced the desired effects in the model runs. A second investigation of a revised form of the correction factor was undertaken and seems to yield even better results This information became available too late to be included in the model runs for this project, but this new function should be used in future work with the model.

Possible Extensions to IPLUM

In the proposal prepared for this project, mention was made of a recognized need to improve the conceptual basis of the PLUM and IPLUM models. In particular, mention was made of attempting to use notions derived from the Herbert-stevens model, as modified at the University of Pennsylvania, for portions of the project analyses.[4] In fact, as one of the earlier project activities a prototype version of the Herbert-Stevens model was linked to the network model in the same manner as was first done for the IPLUM model. This specific work with the Herbert-Stevens model was not carried any further.

In line with this same work the question was raised as to how greater conceptual richness could be included in the IPLUM model. Attention was again directed to the allocation functions, this time with disaggregation of the household types and expansion of the structure of the attractiveness measures in mind. First a more detailed look was taken at the TOMM III model. The Time Oriented Metropolitan Model was one of the first Lowry derivative models, and one of the most substantively significant.[5] This model, developed by Crecine, was intended for use in Pittsburgh and was to have been a key component of a large, comprehensive, model system. Unfortunately, the contract was terminated before the model system was completed and, for that matter, before TOMM was calibrated. None the less, prototype versions of the model were tested and the substantive differences between this model and its progenitor are important and worth discussing.

First, and most important, was the fact that TOMM was an incremental model rather than an "instant-metropolis" model. In the TOMM model, which was intended for projection purposes, the base year distributions of <u>all</u> activities are included

---

[4] See (7)

[5] See (3), (5), and (9)

as determinants of the projection year's distributions of activities. The essential notion here is that, in contradistinction to the original Lowry model, not all households nor all "non-basic" employment is "free" to move in any given projection period, regardless of its length e.g. 5 or 10 years. Second, rather than treating population as homogeneous, the TOMM model disaggregates population into several types.

Referring first to the original TOMM model, the total household allocation function is rather simple, being strictly a function of access to employment. This allocation is embedded in a more complex process which should be described. First the land available for reallocation is calculated for each tract (zone). This is simply the total land minus, the sum of exogenously determined land use and the stable proportions of residential and commercial land use. As in the Lowry model, commercial land use is assumed to be preemptive vis-a-vis residential land use. Consequently, once the reallocable commercial land is calculated, the remaining reallocable land is for residential purposes.

The actual residential allocation in the model is done in terms of residential densities. That is, using Crecine's notation:

$$N^{H*}_{j,t}/A^{H*}_{j,t} = g \sum_i (E_{i,t}/Y_{i,j})$$

where

$N^{H*}_{j,t}$ = Total number of households in zone j at time t

$A^{H*}_{j,t}$ = Total residential land in zone j at time t

g = a scaling factor (constant)

$E_{i,t}$ = total employment in zone i at time t

$Y_{i,j}$ = trip index between zones i and j

Note that,

$$N^{H}_{j,t} = N^{H*}_{j,t} - N^{H*}_{j,t-1}$$

$$= \text{number of reallocated households in zone } j \text{ at time } t$$

Analogously,

$$A^{H}_{j,t} = A^{H*}_{j,t} - A^{H*}_{j,t-1}$$

There are several constraints which operate

1. A constraint on maximum density.
2. A minimum households constraint, corresponding to the stable households.
3. A constraint on the total number of households in the region, to which the sum of the zones is scaled by the factor g.

The model first calculates a residential density allocation for each zone and scales their sum to the regional total. Then the stable households (minimum) constraint is applied. Then a maximum residential density constraint is applied, with any excess being reallocated to all other tracts in proportion to their existing allocations.

The third important difference between TOMM and Lowry is that in addition to disaggregating households, TOMM uses a measure of amenities, albeit very much simplified, to determine the distribution of household types. In particular, given the total numbers of reallocated households per zone, the number of households of each type in each zone is postulated to be a function of the base year number of households of that type, and household specific work trip propensities. The equation is

$$N^{H1}_{j,t} = r_j \, [P_1 \, N^{H1*}_{j,t-1} + w_1 \sum_i (E_{i,t}/Y_{ij})]$$

where $N^{H1}_{j,t}$ = number of reallocated households of type 1 in area j at time t.

Where p_1 and w_1 are exogenously determined constants, and r_j is a scale factor such that the sum of all household types in a zone equals the total households in the zone (as previously estimated). The model then passes to the calculation of the "non-basic" employment allocations.

At a somewhat later date TOMM was slightly modified with respect to the allocation of household types within tracts.[6] The household specific work trip propensities were deleted. Again, given the total numbers of reallocated households per zone, the number of each household type which were in that zone in the base year. The equation is (using the above notation).

$$N^{N1}_{j,t} = r_j \sum_k (p_{1,k} \; N^{Hk*}_{j,t-1})$$

where $p_{1,k}$ is a matrix of household type-to-household type attractiveness indice The investigation of this household type specific calculation was carried further in another effort.[7] In this case the numbers of each household in each zone were postulated to be a function of: 1) the numbers of each household type which were in the zone in the base year, 2) the percent that each household type was of the total households in the zone in the base year, 3) population potential for each household type, 4) employment potential. These last two variables are not defined in the reference, but are probably some form of work-trip-accessibility measure. Attempts to calibrate these functions were not particularly successful.

Several years after the publication of TOMM, Crecine published a description of a more sophisticated version of the model which is referred to as TOMM-III.[8] The principal difference between this model and the earlier version was an enormous

---

6 See (2)

7 See (8)

8 See (4)

increase in the sophistication (and complexity) of the amenities measures used for allocation of households, by type, to zones. While there are other minor differences, they are of no substantive significance. The household types are allocated directly to zones, by type, rather than the two-step process of first allocating total households to zone and then disaggregating, which is used in the earlier versions of TOMM. A trial value of households, by type and zone, is calculated based on the "stable" households in order to start the model's iterative process. The actual household allocation is a function of ten independent variables.[9] The notation of this function is too complex to be worth repeating here. The variables used, in an additive linear function, to forecast location of type to households in zone I at time t, are

1. Accessibility to exogenous beaurocratic employment, from zone I
2. Accessibility to exogenous industrial employment, from zone I
3. Accessibility to endogenous commercial employment, from zone I
4. Housing prices in zone I
5. Percent change of household type L in zone I from time t-2 to time t-1
6. A measure of household demand potential in zone I
7. Index of deterioration of structure in zone I
8. Index of public school facilities in zone I
9. Index of other public facilities in zone I
10. Percentage of region's total households of type L which were in zone I at time t-1

This is an unusually long list of independent variables and raises strong doubts about the likelihood of ever achieving successful calibration. Crecine asserts that some attempt was made to calibrate portions of the model with data from Lansing, Michigan, but it is evident that much work remains before this model is really calibrated.

---

[9] See (4) pp. 38-39

Having reviewed the development and final form of the TOMM models, it was rather clear that the long list of independent variables made it very unlikely that the model, or an adaptation of its allocation function for use in IPLUM, could be calibrated for this project. Since it was also true that neither PLUM nor IPLUM had ever really been calibrated, the question became one of where the efforts might best be spent.

Explorations were made into the problems of calibrating PLUM or IPLUM. It was determined that most of the model's parameters are ratios of various items of base year data, one to another. These had been previously calculated by Goldner during the development of the model. The trip distributions did, however, seem to require more complete statistical estimation and provisions were made to obtain the necessary data to do this. Unfortunately the actual work was not undertaken as part of this project due to the press of other requirements.

Another investigation was made into the possibility of extending the formulation of the household allocation function to include measures of locational amenities along the lines attempted for TOMM III as described above. At the same time it was hoped to disaggregate the population into several classes. Consequently an effort was undertaken to perform an empirical analysis aimed at disaggregating and extending the IPLUM household allocation functions. A series of tests were done results of which are summarized below.

Households were disaggregated, by income class, into four types:

Household type 1 - Annual family income < $5,000
Household type 2 - Annual family income $5,001 → $10,000
Household type 3 - Annual family income $10,001 → $15,000
Household type 4 - Annual family income income > $15,000

This classification is somewhat arbitrary and, no doubt, other classifications could be selected which would probably yield similar results. The only readily available amenities data concerned housing and land use. Transportation data, as explicit variables, were excluded from the analysis pending the amenities results. The variables available were, by zone:

X_1 - % families with income \$10k → \$15k
X_2 - % families with income < \$5k
X_3 - % families with income \$5k → \$10k
X_4 - % families with income > \$15k
X_5 - % owner occupied residences
X_6 - % single family dwelling units
X_7 - % > 10 dwelling units per structure
X_8 - % structures built before 1939
X_9 - % structures in sound condition
X_{10} - median roons per dwelling unit
X_{11} - % land with slope > 15^o
X_{12} - average elevation of land
X_{13} - amount of water frontage
X_{14} - % gross residential land use
X_{15} - % basic employment land use
X_{16} - gross residential acres/basic acres
X_{17} - % vacant land
X_{18} - net residential density
X_{19} - median valuation
X_{20} - amount of unusable land
X_{21} - gross residential acres
X_{22} - unusable acres/gross residential acres
X_{23} - % unusable land

X_1, X_2, X_3, and X_4 were each used as a dependent variable and, though not in the same equation, as an independent variable. All other variables were used only as independent variables. Stepwise multiple regression procedures were used for this analysis, resulting in the following equations (variables with non-significant coefficients are deleted)

For families with income < \$5k

$$X_2 = -0.61\,X_1 + 0.25\,X_3 - 0.30\,X_5 - 0.58\,X_9 + 0.19\,X_{15} + 0.20\,X_{17} + 0.93 \qquad (r^2 = 0.64)$$

For families with income \$5k → \$10k

$$X_3 = 1.62\,X_1 + 0.18\,X_2 - 0.55\,X_4 - 0.17\,X_8 - 0.41\,X_9 + 0.001\,X_{18} - 0.00002\,X_{19} + 0.90 \qquad (r^2 = 0.68)$$

For families with income \$10k → %15k

$$X_1 = -0.06\,X_2 + 0.22\,X_3 + 0.30\,X_4 + 0.07\,X_5 + 0.16\,X_7 + 0.00007\,X_{12} + 0.04\,X_{14} - 0.0006\,X_{18} + 0.000001\,X_{19} - 0.09 \qquad (r^2 = 0.81)$$

For families with income > \$15k

$$X_4 = 0.68\,X_1 - 0.22\,X_3 - 0.14\,X_7 - 0.30\,X_9 - 0.05\,X_{11} + 0.00001\,X_{19} + 0.16 \qquad (r^2 = 0.67)$$

The key point here, is that these rather preliminary efforts indicate the feasibility of both disaggregating the population sector as well as expanding the number of explanatory variables in the household allocation function. More theoretical work is clearly desirable here, especially with regard to questions of how the transportation variables are to be included.

These empirical results may already subsume the effects of transportation facilities, in which case very little improvement in "explanation" of variance will be obtained by introducing say, access measures. If this were the case then it would be necessary to investigate what additional variables should be included and, in what way. All this remains to be undertaken in future research.

<u>Transportation Networks and Trip Generation</u>

The project's efforts to make the feedback connections which are so important a feature of the final model package involved quite detailed examination of the state-of-the-art of transportation network procedures and packages. In the course of these examinations it became quite obvious that there is a very great gap between what little micro-descriptive theory of traveller behavior exists, and the rather arbitrary procedures used by transportation packages. It was beyond the scope of this project to attempt to remedy this situation, but this is a set of problems which very clearly require further investigation. As further work such as this project's is undertaken, the gaps in theory of these network procedures will become increasingly critical. Several examples [illegible] the types of problems encountered are given below.

For the transportation networks the project received from M.T.C., the intrazonal and terminal impedances were apparently estimated by a combination of observation and educated guesswork--both completely outside the network. We were supplied with a vector and a vector of intrazonal times which are to replace the diagonal elements of the matrix. While this was perhaps suitable for the free flow network, it is not suitable for the loaded network derived from a capacity constrained loading procedure since both terminal and intra-zonal times would be invarient under different loading conditions. This in turn could cause a serious mis-estimation of the allocation functions in PLUM and IPLUM.

The network package could include a combined intra-zonal and terminal time, if the load node had only one link. The impedance of this link would be automatically added to every path connecting to the load node. This impedance then also serves as the intra-zonal time. To some extent this would still be an inaccurate representation, since some components of terminal times (e.g. parking fees) should be invariant under loading of the network while others should vary with varying loads. Unfortunately, we have no information on the determination of the terminal and intra-zonal times.

The Bay Area networks have load nodes imbedded in the network, with usually three or four links per node. This causes errors in the final impedance matrix insofar as the diagonal elements do not represent an intra-zonal time--merely an arbitrary link time--and the matrix as a whole does not include terminal times. This "error" however is corrected in a later program that inserts the appropriate values.

To utilize our network program in this way, dummy load links would have to be coded for each network. The link impedances would be some combination of the supplied intra-zonal and terminal times. Further, assumptions would have to be made with regard to the capacities of these links. If given a large capacity, there would never be any volume delays. If the capacity was too small, the volume delays would be excessive. In the wide middle ground, there is no criterion for choosing one capacity over another. Coding these links would have required substantial map work, rerunning of the network coding program, and retracing of the free flow minimum paths. It should be noted that the practice of imbedding load nodes into the network is described in the FHWA network package manual. It is probable that users of our package would have to modify their networks in a similar fashion. (The FHWA package per se has no provision for intra-zonal times. This may cause problems for FHWA users of PLUM and IPLUM.)

Two solutions outside the network package have arisen. One, suggested in the FHWA manual, would be to use 50% (or some other percentage) of the average impedance from each zone to all contiguous zones. Implementing this solution would have required the preparation of a zone contiguity table. The second solution was to weight the free flow intra-zonal impedances by the average percent change due to volume delays in the load links. This solution was implemented and is now an integral part of the model package.

Another question which was considered was that of procedures for model split calculation. There were three basic alternatives available. Firstly, it was possible to try to develop estimating equations based on data from the BATSC Household Survey. Secondly, the use of accessibility ratios could have been explored. Thirdly, a regional average for transit use could have been obtained, and then applied to every zone. These are now discussed in more detail.

First, regarding the development of estimating equations, the Household Survey had information on car ownership, income, and modal split on a zonal basis though the quality of the data was unknown, and there may have been a problem with relating survey zones to IPLUM zones. From this, it should have been possible to develop regression equations to estimate mode split. In order to use this equation for forecasts, it would have been necessary to forecast the independent variables. This would probably best be done by making a regional projection (for car ownership, income, etc.) and applying the same proportionate increment to all zones in the modes. This procedure would have had the advantage of providing some inter-zonal variation in transit usage, and the disadvantage that the forecasting procedures would tend to ossify that variation.

Second, regarding the use of accessibility ratios, the idea behind accessibility ratios is that they use an index of the relative attractiveness of transit and private auto as a major explanatory factor in modal split. The advantage of using it for this project was that it might have helped in assessing the effect of auto congestion on transit usage. The objections were twofold. Firstly, it would have required continual interaction between the traffic assignment model and mode split model before an estimate of equilibrium transit usage could have been obtained. Secondly, on the basis of NCHRP work, there did not appear to be a significant relationship between accessibility ratio and transit usage.

Finally, with regard to regional averages, the BATSC Survey provided data on the numbers of auto and transit travelers over the period 1940-65. It would have been possible from this to extrapolate in crude fashion the future transit usage, and apply this equally to all zones. This procedure had the advantage of simplicity and the disadvantage that variation between zones would be lost. As it was, the decision as to which of these methods would be used never needed to be made since, as will be described below, the transit network data were never a practical proposition for use on this project.

The following, quoted from a memo by P. Kuner of the project staff, describes the situation with regard to the transit network data.

> "The exact state of the transit networks is confusing at best. Basically, it will be difficult to make the networks usable, and once done, they will be of little value to the research. The reason for this is that they are not designed to respond to different levels of loading. The 1965 network is not designed to be loaded at all (only to produce interzonal impedances), and the 1980 and 1990 networks may be loaded only to produce utilization rates for selected routes, not to show changes under load. The networks would be valuable to us only if we wished to test modifications to them, and if we were able to divert trips from auto to transit via some function of their comparative impedances. Since at present we are not able to do this, the networks themselves are of little value. The impedance matrices produced by tree tracing on the networks, however, are of value. We presently possess the fare and time data necessary to construct them without using the networks.
>
> <u>1965 G Transit Network</u> - This network is presently coded as a CDC 3600 TRANPLAN highway network, and was last run (with unspecified difficulties) in 1968. It is in binary form, which is unreadable by

our computer. To obtain it, in readable form, we would have to get someone in California to write a program that would transform it to BCD or EBCDIC, and put it on tape.

We presently possess on tape, card images of several files that were used to make the final network. Processing these files involves changing all the node numbers to correspond to MAP zone numbers where possible, inserting links by hand where correspondance is not possible, and making miscellaneous corrections. The final product would have to be compared to a listing that we have on microfilm for verification. This processing involves several programs and a lot of hard work.

Given the final C network, however obtained, several problems still exist. First, the network as it stands can not be run on either our package or the BPR package. Recoding it for our programs requires use of X,Y coordinates for all the nodes. These exist only for the MAP nodes, which are only 273 of the 494 Transit nodes. I believe that the balance of the nodes could be handled mechanically by some sort of program, but I am not positive. At the worst, we would have to look up the nodes on a map and estimate the coordinates. Next, the network employs turn penalties to prevent trips from ilegally changing from one route to another. (Many routes share the same nodes). Our program can not handle turn penalties so we would have to insert dummy links and nodes for these cases.

An alternative to these procedures would be to utilize the BPR package (which can accomodate turn penalties). I have not investigated it in detail, but I believe that converting the network to BPR format and getting it operational on our computer would be as much work as converting it to our format. In addition, the BPR format is incompatible with the other program in the model system.

<u>1980 X and 1990 Y Networks</u> These networks are coded in an entirely different fashion than the G network. Like the G network, no coordinates exist for the nodes making it difficult to use with our program.

Instead of turn penalties, the networks are designed to be run on a program that can accommodate many different routes sharing the same links and nodes. We have a listing of this program, and probably we should use it rather than attempt to recode the networks. This would involve conversion from CDC FORTRAN to IBM FORTRAN and should be fairly easy, although the program does use certain language features specific to the CDC computer. The output of this program could be modified to be compatible with our other programs.

It should be noted that this summary of the network is quite simplified, e.g. there are two preprocessing programs that must be run before running the X or Y networks; somehow we will have to insert waiting times if there is a route change and these times are dependent upon headways which are dependent on the time of day which is being simulated; much data to correct the G network is available only on paper and will require keypunch and verification, and so forth."

Consequently, given the other tasks which needed to be done, and the project resources available, it was decided to omit transit networks from this version of the model package.

Various aspects of the trip generation procedures have been discussed in earlier sections of this chapter and in Chapter III. These procedures are also described, in somewhat more operational terms, in Volume 2 . None the less, a few comments on these parts of the package are appropriate here.

It will be recalled that in order to link the traffic network model and the land use model, it is necessary to obtain estimates of trip generation from the output of IPLUM. These trips are used to load the network and recalculate zone-to-zone travel times. The travel times obtained in this fashion are used in conjunction with estimates of transit times to give the zone to zone impedance for use in IPLUM.

The conceptual basis of residential models such as PLUM and IPLUM is that the locational behavior is governed largely (or entirely) by the journey to work time. This suggests that the interzonal travel times should be based on rush hour traffic volumes, when most work trips take place. This procedure also has the advantage that it deals with the traffic network at its most congested time, and so the effects of changes in land use controls or the road network or congestion can readily be monitored.

The network is more heavily loaded during the afternoon rush hour than in the morning. The B.A.T.S.C. Home Interview showed approximately 740,000 auto and transit trips starting in the period 7:00 a.m. to 8:00 a.m. and approximately 900,000 in the period 4:00 p.m. to 5:00 p.m.[10] Consequently, it was decided to

---

[10] See (1) pp. 25-30.

calculate impedances on the basis of afternoon (4:30 p.m. to 5:30 p.m.) traffic. Closer investigation of the Home Interview results shows that only about 35% of the total trip starts in the period 4:30 p.m. to 5:30 p.m. are home based work trips. After allowing for the number of transit and walking trips and for car occupancy rates, home based work trips account for 46% of all auto trips on the network, but this still leaves the majority of all trips unaccounted for.

The remaining trips are divided among the eight trip categories used by the B.A.T.S.C. Home Interview. The largest single category amongst these is non-home based trips. This causes considerable modelling problems, for while most of the other categories (given the knowledge that they either begin or end at home, and that they end in some central place, such as a shopping center or office complex) can be (with a little willing suspension of disbelief) fitted into a generalized retail category, neither the origins nor the destinations of the non-home based trips are known. In fact, it is likely that the origins and destinations could be obtained by consulting the full household interview tapes. However, for the purposes of obtaining rough estimates of trip generation, it has been assumed that these trips originate at work, and have a shopping center as an intermediate destination and home as a final destination. In later work it may be possible to refine this treatment.

The remaining trips can, with one or two qualifications be aggregated and treated in a fashion analogous to that used in retail trade gravity models. School trips are not susceptible to this treatment, since most schools draw their pupils from a distinct district near the school and not from a large region in competition with other schools as would be implied by a gravity model. However, school trips are relatively unimportant traffic generators during the evening rush hour, and can be ignored. It also appears that while social trips may be susceptible to gravity model treatment, since they could be expected to diffuse across a region, they are probably of a different nature from the other trips in that they are not made to specific centers. There are two options here: (1) design a gravity model specially for social trips with population as attractors (2) treat them as if they were "shopping" trips. The second option may not be as bad as it sounds since if zonal shopping floor space bears any reasonable porportion to zonal population the distribution of trips would probably work out to be very similar. The remaining five categories can reasonably be combined into a single category and distributed by use of a gravity model. The possible exception to this is Free Recreation which might reasonably be expected to include trips to the shore or to public parks, which might not correlate at all with the destinations of shopping trips or private business. However, this is a relatively small category and probably not too much is lost by treating it in this way.

Further problems arose in making the transition from the network model to the land use model. The impedance measure used in the land use model is taken to represent the difficulty of moving from one zone to another. This has to be a combination of auto travel time and transit. It also has to take into account the fact that only about 55% of work trips take place in the afternoon rush hour, and that nearly half the workers will face entirely different impedances. If the rush hour impedances are used, then the model would tend to locate that 45% of workers much closer to their work than would otherwise be the case if impedances on a less highly loaded network were used. A further, though possibly minor difficulty is that it cannot simply be assumed that a constant proportion of all work trips take place at rush hour regardless of location. Typically, work trips to manufacturing would take place at a different time fron work trips to offices or retail centers, and failure to take account of this could produce erroneous network loadings. The following paragraphs describe in more detail the current procedures for treatment of each trip type.

Regarding work trips, the location of basic employment is known, and the location of non-basic employment is generated in the land use model. Three adjustments are made to these employment figures before loading the transportation network:

i. Reduction by the ratio of work-to-home trips to all work based trips which take place at rush hour. For the purposes of rough estimates, the regional average is applied to each zone.

ii. Reduction to allow for transit usage and walking trips. For present purposes the regional average of 82% can be used. It is hoped that future research would be able to develop an estimating equation for transit use by <u>residential zone</u> based on B.A.T.S.C. data.

iii. Reduction by the car occupancy factor. For work trips this was 1.18 in 1965.

The combined effect of (i), (ii), (iii) based on regional averages is to give a conversion factor of:

$$\frac{28}{100} \times \frac{82}{100} \times \frac{1}{1.18} \times 100 = 19.7\%$$

i.e. car trips for period 4:30 p.m. to %:30 p.m. are equal to approximately 20% of total employment.

The home based non-work trips cover trip categories two through eight. The list is as follows.

2. Personal business, medical and dental
3. School
4. Visit friends and relatives
5. Eat meal, and commercial recreation
6. Convenience shopping
7. Comparison shopping
8. Free recreation accompany a person and other

In the present model package these are treated as one category. Note: there are so few school trips, they can be ignored. It was possible, to obtain estimates of trips-per-household in this conglomerate group from the B.A.T.S.C. regional data. Total person trips per day can then be estimated once zonal population is known. The trips are then distributed using a gravity model, using non-basic employment as an attractor. The total person trips per day distributed in this fashion are reduced to auto trips at rush by repeating steps (i) - (iii) above.

The non-home based trips (category 9) are considered as originating at places of employment. The person trips estimates are converted to auto trips by procedures (i) - (iii). They are then allocated to shopping centers using the PLUM and IPLUM work-to-shop allocation functions.

Finally, as one more example of the curious ways in which the state-of-the-art can affect attempts to join models into systems of models, the problem of negative trip volumes is described. It will be recalled from the earlier part of this chapter that an incremental trip loading procedure is used after the first few iterations of the model package.[11] With this procedure it is possible that the spatial redistribution of activities from one iteration to the next will result in a redistribution of trips which, on some links, will result in negative trips when one trip matrix is subtracted from another.

In other words, the volume of negative trips to be unloaded from a link can exceed the existing link volume because of the nature of the network assignment procedure. Generally speaking, in the assignment procedure some number of trips, originating at A and terminating at B, is to be loaded on the network. To do this, the shortest path through the network from A to B is found and the trips are added to each link in this path. Each time trips are loaded on a given link,

---

[11] See p. 71 in this report and Appendix A-2.

the time required to traverse that link increases. This is due to congestion of the link, and is expressed operationally in the workings of the volume-capacity ratio. Because of this effect, trips traversing a certain portion of the network will find that a link that was on the minimum (i.e. shortest time) path at one time, will not be on it at a later time.

When negative trips say, from A to B, are to be (un)loaded, the system is not able to deal with them appropriately. Any link on the minimum path from A to B may have gotten on the path by virtue of congestion on a "parallel" or alternative link and may thus have just a few trips which were loaded on it when the previous "best" link became overloaded. If the negative trips being (un)loaded are in excess of the existing trips on the link, there is no way for the program to determine what other link had been "best" prior to the loading of the link in question. Consequently there is no way to determine where the excess negative trips should be (un)loaded. These excess negative trips therefore wind up on the most recent "best" link which, in turn, is left with a negative trip volume. Operationally this is dealt with by restricting the volume/capacity ratio to be > 0, but the negative trip volume remains either to the end of the network run, or if positive trips (from another O.D. pair) are loaded on that link they reduce (or eliminate) the negative volume.

While operationally this does not occur frequently enough to cause any real difficulties, it does illustrate the kinds of problem which may arise with procedures which under normal circumstances are perfectly all right. Problems of this sort appeared regularly during this project, pointing up many of the areas where the paucity of integrative attempts has allowed important considerations to "fall between the cracks".

Development of the "New Process" Procedure

During the policy testing of the model package problems with the stability of the results continued to be encountered.[12] After much additional work on the problem it appeared that one of the fundamental difficulties was the size of the projection interval. In the existing model package a projection of 1980 is made, from a base of 1965, in one step. Successive iterations of the model package are simply attempting to find an equilibrium solution. It was thought that an alternative means of finding the equilibrium would be to approach it more gradually in a way analogous to making the projection from 1965 to 1980 in several steps rather than in one grand leap. It was hoped that such a procedure would both eliminate the remaining instabilities in the model system and be less expensive to operate by virtue of requiring fewer iterations of the model package. Work was undertaken to test this new process at a rather late date in the project. The evidence available at this time is that it will indeed do what we hoped.

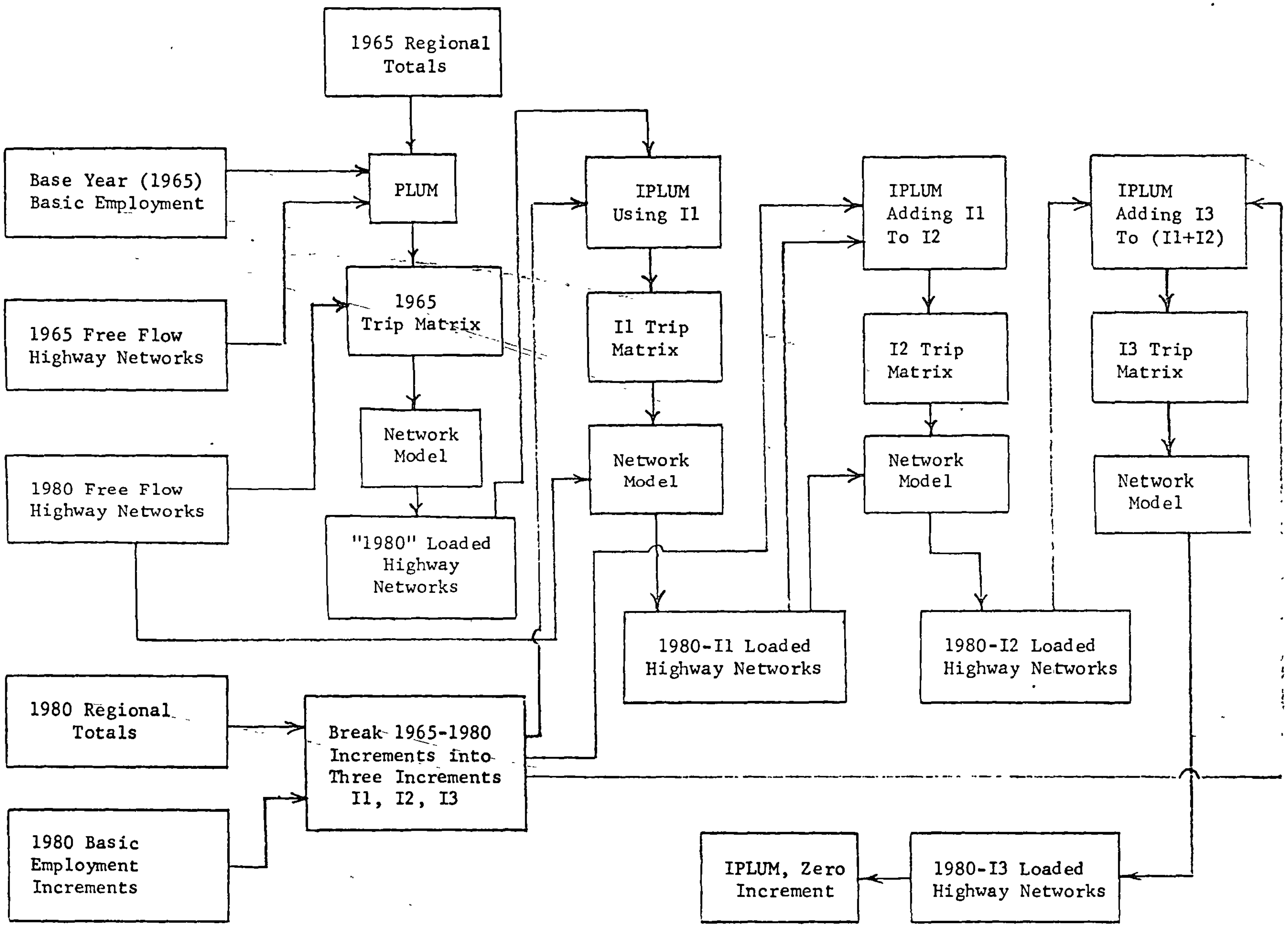
A verbal description of the process, shown in Figure 4, is as follows:[13]

1. Assume that the base year is 1965 and the projection year is 1980.
2. The base year (1965) trips are loaded on the 1980 network and zone-to-zone impedances are calculated.
3. With these impedances, one-third of the regional control totals, and one-third of the 1965→1980 basic employment increments as inputs, IPLUM is run, and new trip matrices generated.
4. The 1965 trips are subtracted from the new trip matrices and the remaining new trips are loaded on the already partly loaded 1980 network. New zone-to-zone impedances are calculated.
5. With these impedances, two thirds of the regional control totals, and an additional one third of the basic employment increments as inputs, IPLUM is run, and new trip matrices are generated.

---

[12] Described in the next chapter

[13] Compare with p. 26 of this report, which describes the original process.



SCHEMATIC OF NEW PROCESS FOR MODEL PACKAGE

re 4

6. The trip matrices from Step 4 are subtracted from these new trip matrices and the remaining new trips are loaded on the partially loaded 1980 network. New zone-to-zone impedances are calculated.
7. With these impedances, the full regional control totals and the final one third of the basic employment increments as inputs, IPLUM is run and new trip matrices are generated.
8. The trip matrices from Step 6 are subtracted from these new trip matrices and the remaining new trips are loaded on the partially loaded 1980 network. New zone-to-zone impedances are calculated.
9. With these impedances, the full regional control totals, and zero basic employment increments, IPLUM is run and should then produce an equilibrium solution.

This new process was run through step 8. The resulting population projections, where XA-1 = Step 3, XA-2 = Step 5, and XA-3 = Step 7, are shown in Figure 3. The dotted lines in Figure 3 represent the projections from Run IV.6AVG. It can be seen that the new process comes rather close after three recursions to where the old process was after six iterations. The fourth recursion of the new process was not run as described above, because of an unexpected operational deficiency in IPLUM. The model is theoretically capable of projecting the response, in terms of spatial redistribution of local-serving employment and population, to changes in basic employment and/or transportation facilities. However, it cannot operate with a zero change in basic employment. Since there were not sufficient resources (or time) on this project to correct this deficiency in IPLUM it was not possible to run the fourth recursion (Step 9) of the new process as originally intended.

As an alternative test, the impedances from Step 8 were used and IPLUM was rerun as in Step 7. This means that the last growth increment was added to the previous two-thirds, but using the impedances derived from the <u>prior</u> estimate of the last

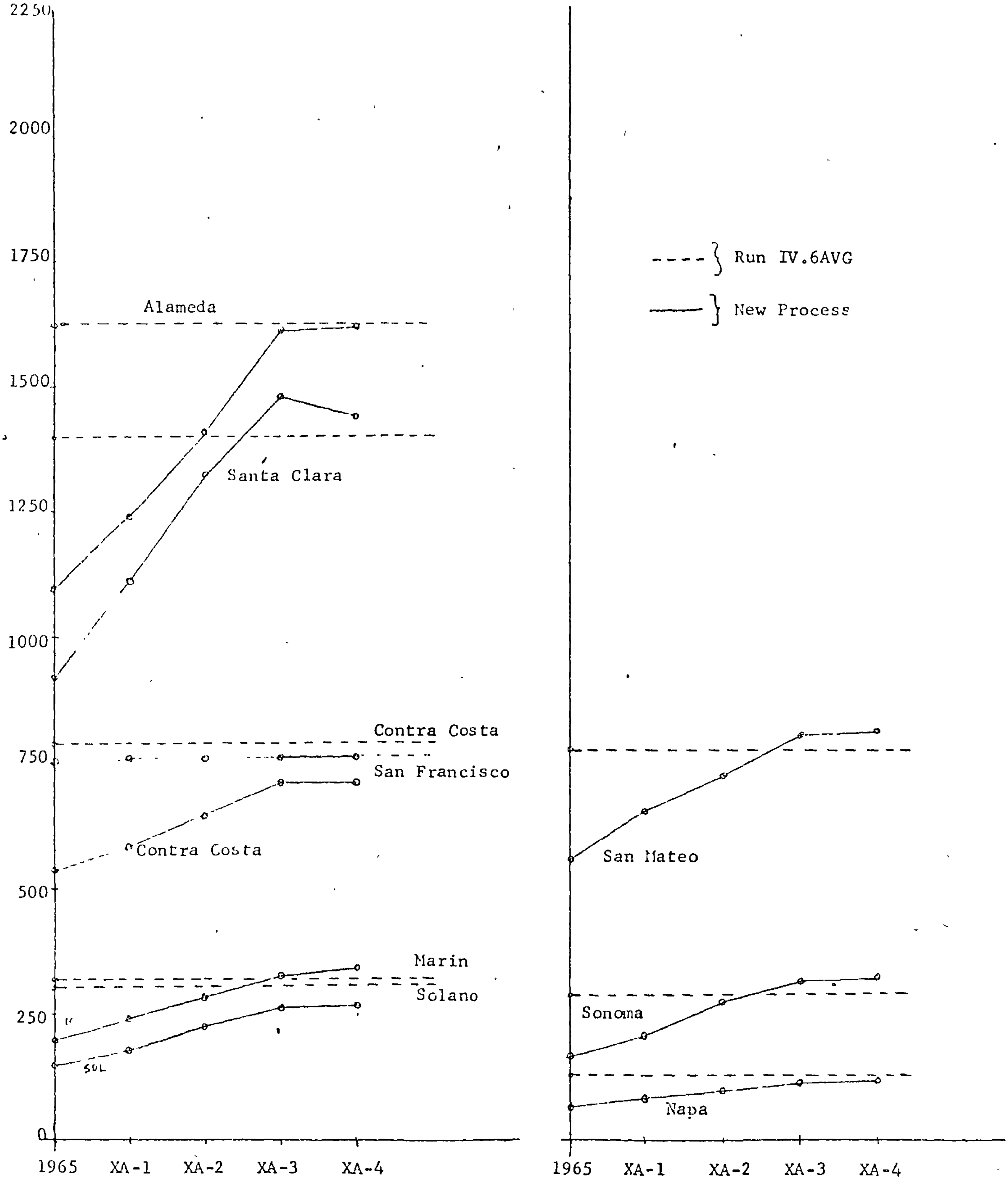


FIGURE 6

New Process, Population Projections for 1980 growth increment. The results of these two estimates of the spatial distribution of the last growth increment (one based on a two-thirds full network and the other on a three-thirds full network) should, if the new process had found an equilibrium in the first three steps, be nearly identical. In fact, the largest difference, at the county level, was less than five percent. These differences, shown as point XA-4 on Figure 3, are:

San Francisco	+0.08	Solano	+1.72%
Marin	+4.07%	Contra Costa	-0.12%
San Mateo	+0.77%	Santa Clara	-2.64%
Sonoma	+1.76%	Alameda	+0.31%
Napa	+2.29%		

While further testing is clearly desirable, the implication of these results is that by this new process the whole system reaches an equilibrium point in just three iterations as opposed to the five or six iterations required by the original procedure. Any further work with this model package should be done with this new process.

This concludes what turns out to have been the "grab-bag" chapter of this report, which attempted to document some of the more important perigrinations leading to the final results.

References

1. Bay Area Transportation Study Commission, "Bay Area Transportation Report" B.A.T.S.C., Berkeley, California, May 1969.
2. CONSAD Research Corporation, "Impact on Allegheny County Due to the Relocation of Residential and Commercial Activity in the East Street Valley" prepared for Department of City Planning, Pittsburgh, Pa., March 1967.
3. Crecine, J.P. "A Time-Oriented Metropolitan Model for Spatial Location" Technical Bulletin No. 6, Community Renewal Program, Department of City Planning, Pittsburgh, Pa., January 1964.
4. Crecine, J.P. "Spatial Location Decisions and Urban Structure: A Time Oriented Model" Discussion Paper No. 4, Institute of Public Policy Studies, University of Michigan, Ann Arbor, Michigan, March 1969.
5. Goldner, W. "The Lowry Model Heritage" Journal of the American Institute of Planners Vol. XXXVII, No. 2, March 1971.
6. Goldner, W. et. al. "Projective Land Use Model - PLUM" Vol. I, II, III. Institute of Transportation and Traffic Engineering, University of Calif., Berkeley, March 1972.
7. Harris, B. et. al "Research on an Equilibrium Model of Metropolitan Housing and Urban Choice: Interim Report" Planning Sciences Group, Institute for Environmental Studies, University of Pennsylvania, March 1966.
8. Lee, D.B. Jr. "Household Disaggregation in Urban Models" paper presented at Regional Science Association Meetings, Cambridge, Mass. Nov. 1968.
9. Lowry, I.S. "A Model of Metropolis" RM-4035 RC RAND Corporation, Santa Monica, California, 1964.

V. Policy Tests with the Model Package

Introduction

Having waded through more than one hundred pages of text, the reader may need to be reminded that the ultimate purpose of the development of the model package was to allow analysis of the possibilities of integrated land use control and transportation development policies. The aim of these policies was to achieve some balance between transportation development and land development. Since the data base for which most of the model work had been done was the San Francisco Bay area, this same base was used for the policy runs.

The project was able to examine copies of memos written in the Spring of 1971 by M.T.C. staff regarding tests of the PLUM. These tests were of the response of the PLUM model to changes in the transportation network and subsequent travel time and ranks data input. These tests concerned network changes only, and four kinds of test were carried out:

(1) Changing links (by addition or deletion)
(2) Changing link times
(3) Changing bridge penalties
(4) Incorporating transit links into the network (amodal network)

In all four cases the results showed that commercial employment tended to be slightly more sensitive than residential population in terms of their spatial redistributions. As might be expected, map zones showed greater variability than planning districts, which in turn showed greater variability than county totals.

All of these tests confirmed the notion that the PLUM model allocations would be sensitive to changes in the input impedances. This was not a surprising result, but some of the comments regarding the implications of the results were of considerable interest. For example:

> "It must also be kept in mind that the shifts in land use configuration is not the final statement of travel demand. In the system framework we still have to go through generation, distribution, modal split, and assignment before we can more adequately assess the travel demand let alone the level of evaluation."

> "By lengthening the times, the commute distance is shortened and the resultant allocations will be more compact. As would be expected the growth in the northern counties decreases appreciably while the more accessible areas closer to the employment centers increase greatly. ---However, even with changes in several counties by a magnitude of 25% - 30%, caution must be taken as to how this affects travel overloads.---(Only by following the process through to assignment could this question be answered.)"

These comments clearly indicate the desirability of making such tests with a model system incorporating the feedbacks which are a key ingredient of this project's package.

The inputs to the model package have been described in previous chapters of this report. The outputs of the model package are sets of rather highly aggregated predictions of equilibrium values; 1) on a zone-by-zone basis, of population, local serving employment, and land use, and 2) network link traffic volumes and times, the consequent interzonal travel times, and several ad hoc, spatially identified, network congestion measures. While this is a reasonably rich set of outputs it does, due to the levels of spatial and sectoral aggregation, restrict the kinds of policy which it is worth while testing. This restriction is due to both the outputs of the model package, by means of which the policy alternatives can be compared, and the nature of the inputs to the package, by means of which the policy is imposed on the model package's simulation of metropolitan dynamics.

This project is concerned with policies regarding land use and policies regarding transportation facilities, as well as with policies which integrate the two. Given the present version of the model package, land use policies must be constrained to considerations of:

(a) location of increments of basic employment

(b) land use restraints (open space policies)

(c) location of local serving employment.

Transportation policies must be constrained to considerations of:

(a) transportation facility improvements

(b) imposition of travel penalties

An interesting potential application of this model package would be in the area of policy stability testing. Once the model has been run to obtain predictions of land use patterns, on the assumption that the policy decisions are successfully carried out, the policy restrictions can be eased by allowing some proportion of the locators to move. If these free locators show a strong tendency to move into formerly restricted zones, it can be taken as an indication that the policy would be subject to strong external pressures, and would require strong enforcement.

A general strategy for testing policies with the model package, in accord with the above list could be as follows:

A. Land Use Policies

a) location of increments of basic employment - this would involve one or more model runs with patterns of basic industry different from the one currently predicted by BEMOD and used as input to the current model package. An important question exists regarding the centralization or dispersion of such employment. Current trends indicate strong tendencies toward dispersion. What are the consequences of attempting to alter this process? A policy such as this could be combined with biased changes in the transportation network which could improve intra-urban times while degrading urban → suburban times.

b) land use restraints - a particular question to be dealt with here has to do with the preservation of open space. A "green-belt" policy of restricting development in lands around the urban area, say between the urban and suburban areas and between the suburban and semi-rural areas, could be tested. An associated transportation facility might provide good access from one unrestrained area to another while restricting access to the "green-belts". Policies of "green-wedges" might also be tested, in an attempt to maintain alternate sections of open space between various built-up areas.

c) location of local serving employment - a question to be dealt with here is the continuing expansion of suburban shopping centers at the expense of the central business districts.

B. Transportation Policies

a) transportation facility improvements - many of these have already been tested by various model projects, and the more important questions here are regarding the relation of these policies to land use policies.

b) imposition of travel penalties - this often takes the form of bridge and/or tunnel tolls, but an interesting experiment would be to investigate a policy of allowing congestion to develop (i.e. a time penalty) in certain areas.

It should here be noted that there exist no reliable or acceptable indices to describe overall land use patterns. Consequently the description of the results of test runs will prove difficult. In particular it implies that most results will be presented at the county level, with detailed analysis only being done in areas of large estimated differences.

Two further simple but important points should be noted. First, the policy testing to be described below does not imply anything about feasibility of implementing the policy. As was mentioned both in Chapter 2 and in Appendix A-1, the potential effectiveness of a transportation or land use policy often has little to do with its social or political feasibility. Second, the following demonstrations of the models' sensitivity to policies and the reasonableness of its results do not constitute a proof of its forecasting accuracy. That the model package produces what seem to be sensible responses to changes in its inputs is certainly a desirable result and increases our confidence in its structure, but this does not offer further proof of its forecasting accuracy.

Strategy for Policy Tests

It is clear from even a cursory examination of the model system that despite its rather too aggregated levels of detail, there are an enormous number of alternatives which could be tested. Even for a particular type of policy there are usually many ways of expressing it as model inputs. When specifying a model run decisions must be made as to which sectors of activity in which zones will be changed by how much. Each difference in any of these numbers produces a "different" policy input. Until the overall sensitivities of the model system are known, it is difficult to make a priori decisions as to which differences are or are not important.

The model package, in its present form is both expensive and inexpensive to run. It is expensive because the cost of a run at the time of this writing, on an IBM 370-165 for the full Bay Area data set, is about $120 to $150 for a complete iteration of the network model, IPLUM, and the trip generation procedures. Consequently a full policy test round of five or six iterations of the system will cost about $600 to $900. It is inexpensive because this amount of money is often spent by transportation studies on one or two runs of only their network model package which yields much less information about the overall urban system, though it yields substantially more detail about the network.

The combination of the many possible alternative test runs plus the cost of running them,with the project resources remaining after the completion of the development of the model package,conspired to severely limit the number of test runs which could be undertaken.

It was initially decided to test three basic policy inputs:

1) a change in location of some of the basic employment increments.
2) an imposition of land use controls by altering land available for development.
3) specific improvements in the transportation network facilities.

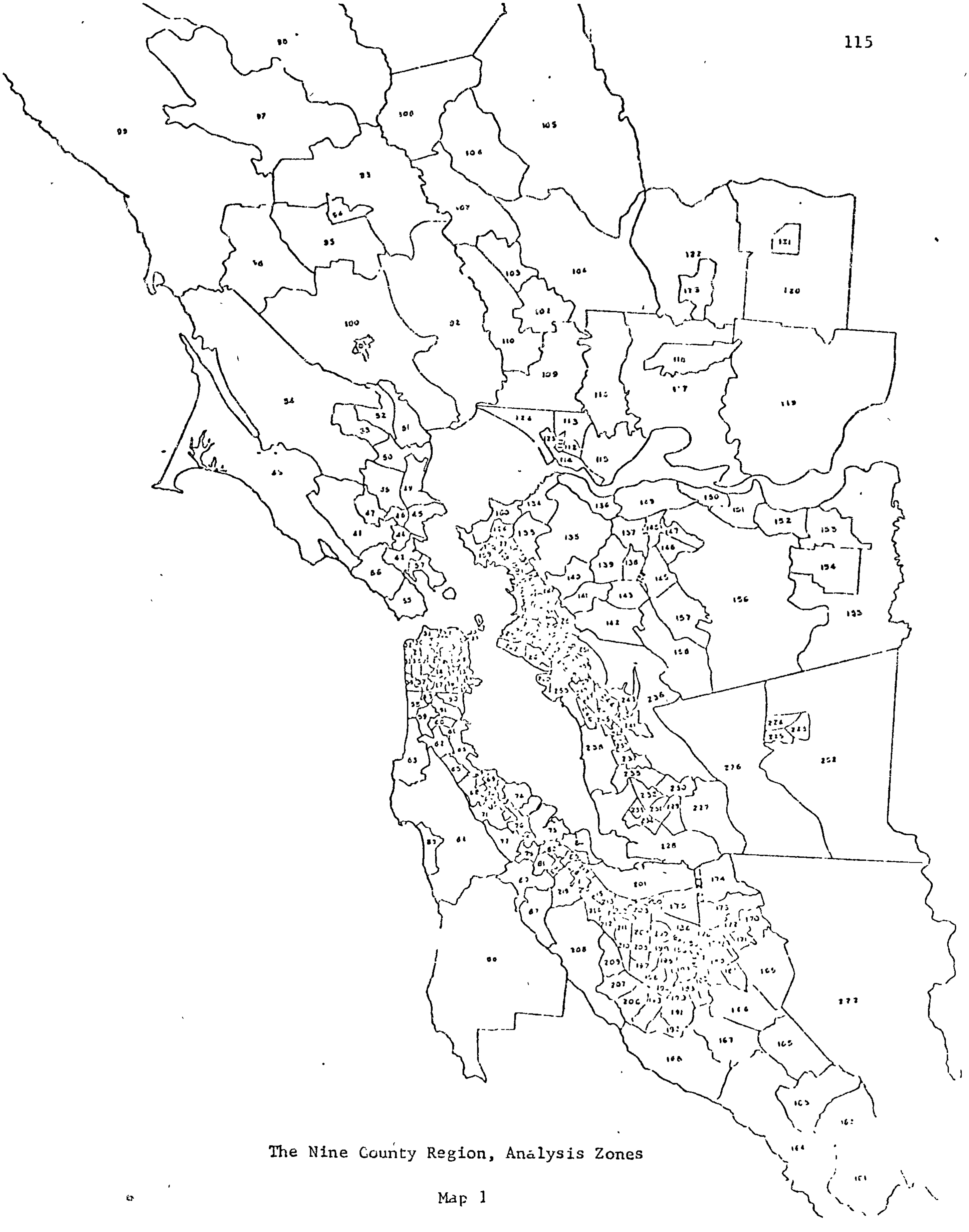
There were also the obvious questions of interactions between these and the consequent possible additional runs combining, say 1) and 3), or 2) and 3).

The first complete runs of the package suggested that the basic employment locations forecast exogenously by BEMOD lead to rapid urbanization of the eastern part of Alameda County, in particular the Amador Valley area (Pleasanton, zone 226) and the Crow Creek end of Castro Valley (zone 236).[1] This effect is so marked, and since, as mentioned in Chapter 3, this area is indeed experiencing problems of rapid growth, it was decided to concentrate attention on Alameda County (leaving the other BEMOD forecasts unaltered), and try to gauge the sensitivity of the model to policy input by examining the behavior of these areas for different patterns of industrial development. It was felt that in practice, existing industry would be hard to move, but that new development would be more susceptible to policy directives. With this in mind, the forecast 1980 distribution of basic industry increments was examined. Most of the increment fell in the area laying between San Leandro and Berkley in the northwest corner of the county, with small increments scattered elsewhere.

Two basic strategies seemed worth examining. One was to restrict all the increases to the present heavily urbanized areas of Berkley and Oakland. In this way employment would be directed away from easy access to the Dublin Canyon Road (leading to zones 236, 226) and so could be expected to lead to less urbanization in those areas. In addition, local congestion might be created, and this would

---

[1] See Maps 1 and 2 and note that both these zones are very large, though they contain a great deal of mountainous land, and see p. 29 of this report re BEMOD.



The Nine County Region, Analysis Zones

Map 1

The Nine County Region, Place Names and Highways

Map 2

lead to longer journey times and possibly to a more centralized urban pattern. The degree of congestion under this policy was to be examined, since there may well turn out to be a trade-off between dispersion of industry and population, and quality of life in central areas, and congestion measures might be useful in future evaluation efforts.

A second basic strategy would be to attempt to concentrate the growth of industry in several, scattered, main centers. With this type of policy however, there would probably be strong tendencies toward suburban sprawl, and complementary greenbelt policies would probably have to be introduced simultaneously and vigorously.

It was also desired to investigate the other potential policy inputs enumerated above. Consequently an overall strategy for policy testing was developed, consisting of three components.

1) Movement of basic employment in Alameda county to its northwestern corner, as described above.
2) Imposition of land use controls to restrict land development in the zones running along the mountains from Lafayette to Fremont.
3) Construction of two high speed, very limited access highways, one from Oakland to Livermore and the other from Fremont to Livermore. These roads would have no intermediate access points and would be superior to existing links.

The overall effect desired from this policy would be the establishment of a green belt along the area from Lafayette to Fremont. The centralization of fringe area development at Livermore with the hopeful elimination of sprawl development moving southeastward from the Berkeley to Fremont axis. This would also require very stringent land use control in the Livermore area in order to avoid sprawl emanating from that nucleus. And, the centralization of development along the eastern shore of the San Francisco Bay in the already developed Berkeley to San Leandro strip.

It should be stated at the outset that the full set of policy tests was not completed. This was solely a matter of inadequate project resources, particularly in terms of computer time.

More importantly, the runs that were done bore ample testimony to the power of this integrated model package to produce interesting and potentially useful estimates of the consequenc es of policy actions.

Policy Test 1: Round III Runs

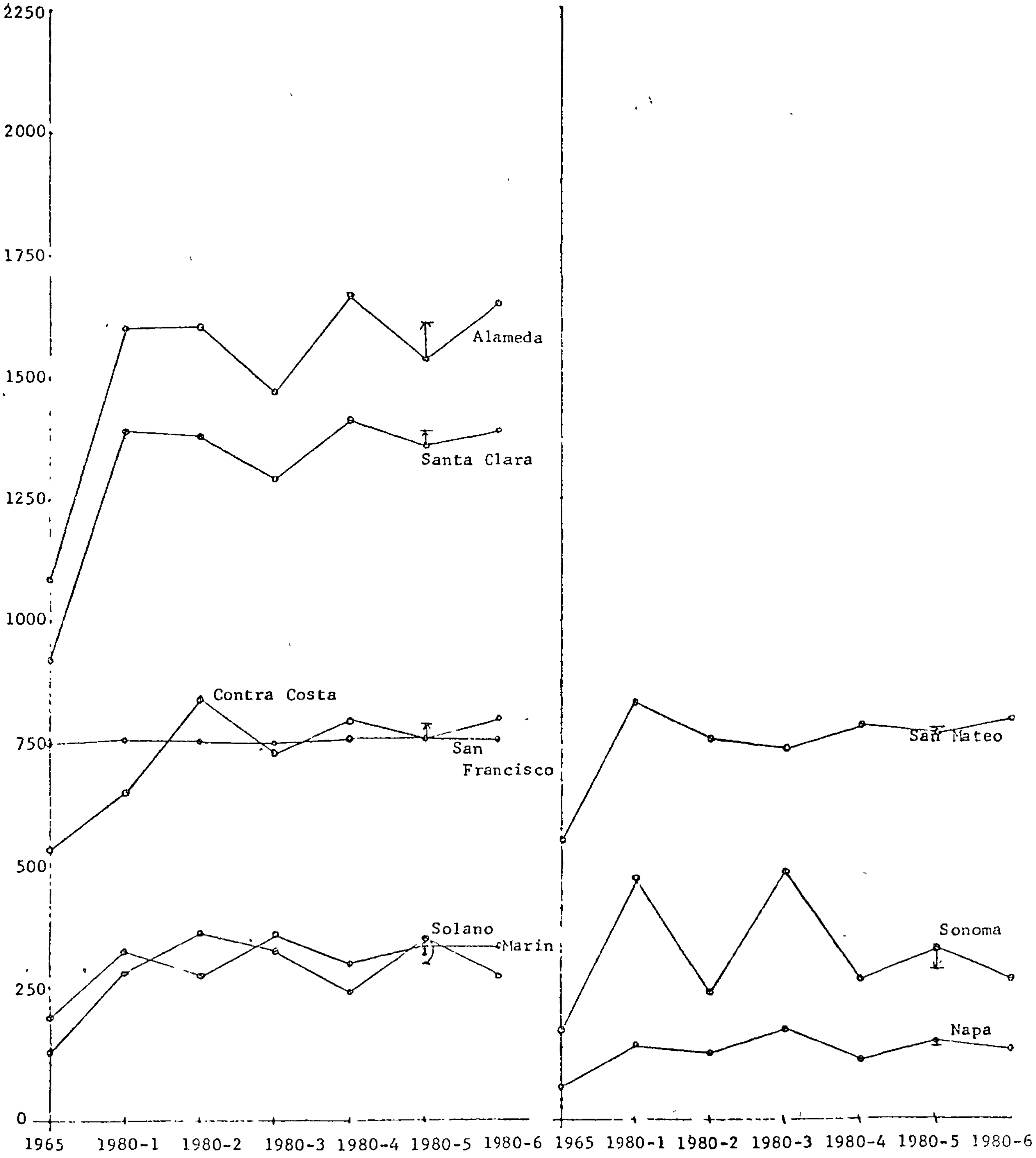
A full round of model runs was done to reflect the effects of the basic employment redistribution in Alameda county as described above. The results of this run are shown in Table 1 and Figure 1. Bearing in mind that the total basic employment in Alameda County was unchanged as a result of this policy, and that the 1965-80 change represented a little less than 30% of the 1980 total, the effects on the distribution of population were surprisingly large. Run IV.6 AVG which was tabulated on p. 55 in Chapter 3 is taken to be the base run. Comparisons are made between run III.5 and IV.6AVG and are given in Table 2. The general tendency of the result was that population moved out of Alameda, and to a lesser extent Contra Costa and Santa Clara, into the counties to the north of the region. Of these, Sonoma and Solano showed the greatest increases, followed by Marin and Napa. San Mateo and San Francisco showed only very small changes.

It would appear that the "natural" areas for workers in northern Alameda to live is in the northern counties, particularly since that section of Alameda, and closer sections of Contra Costa are already developed. The concentration of additional population in Sonoma and Solano suggests that commuters find most land available by using the Richmond-San Rafael, and the Carquinez bridges, both of which connect Contra Costa to the northern counties. On Map 3 the broad arrow represents the policy determined movement of basic employment while the county level shifts in population are shown as percentages below each county name.

		Run III.2		Run III.3		Run III.4		Run III.5		Run III.6	
	1965	1980-2		1980-3		1980-4		1980-5		1980-6	
San Francisco	754,754	762,464	1.02%	763,373	1.14%	765,262	1.39%	764,778	1.33%	765,224	1.39%
Marin	192,523	256,409	32.2	369,969	92.2	306,671	59.3	340,685	76.9	337,421	75.3
San Mateo	553,271	716,094	29.4	742,466	34.2	793,707	43.5	773,802	39.9	799,310	44.5
Sonoma	167,497	256,690	53.3	497,316	196.9	266,828	59.3	333,490	199.1	269,962	61.2
Napa	69,378	174,451	151.5	168,534	142.9	108,524	56.4	146,955	111.8	122,767	76.9
Solano	149,386	366,950	145.6	329,164	120.3	248,165	66.1	356,274	138.5	274,330	83.6
Contra Costa	534,056	778,539	45.8	733,951	37.4	794,572	48.8	760,834	42.5	796,591	49.2
Santa Clara	922,357	1,334,145	44.7	1,315,736	42.6	1,433,150	55.4	1,379,884	49.6	1,388,694	50.6
Alameda	1,095,236	1,761,796	60.9	1,487,035	35.8	1,690,663	54.4	1,558,833	42.3	1,653,243	50.9
Total	4,438,358	6,407,469	44.4	6,407,467	44.4	6,407,469	44.4	6,407,470	44.4	6,407,471	44.4

Round III: Base Population, Projected Population 1980, and Percent Change

Table 1

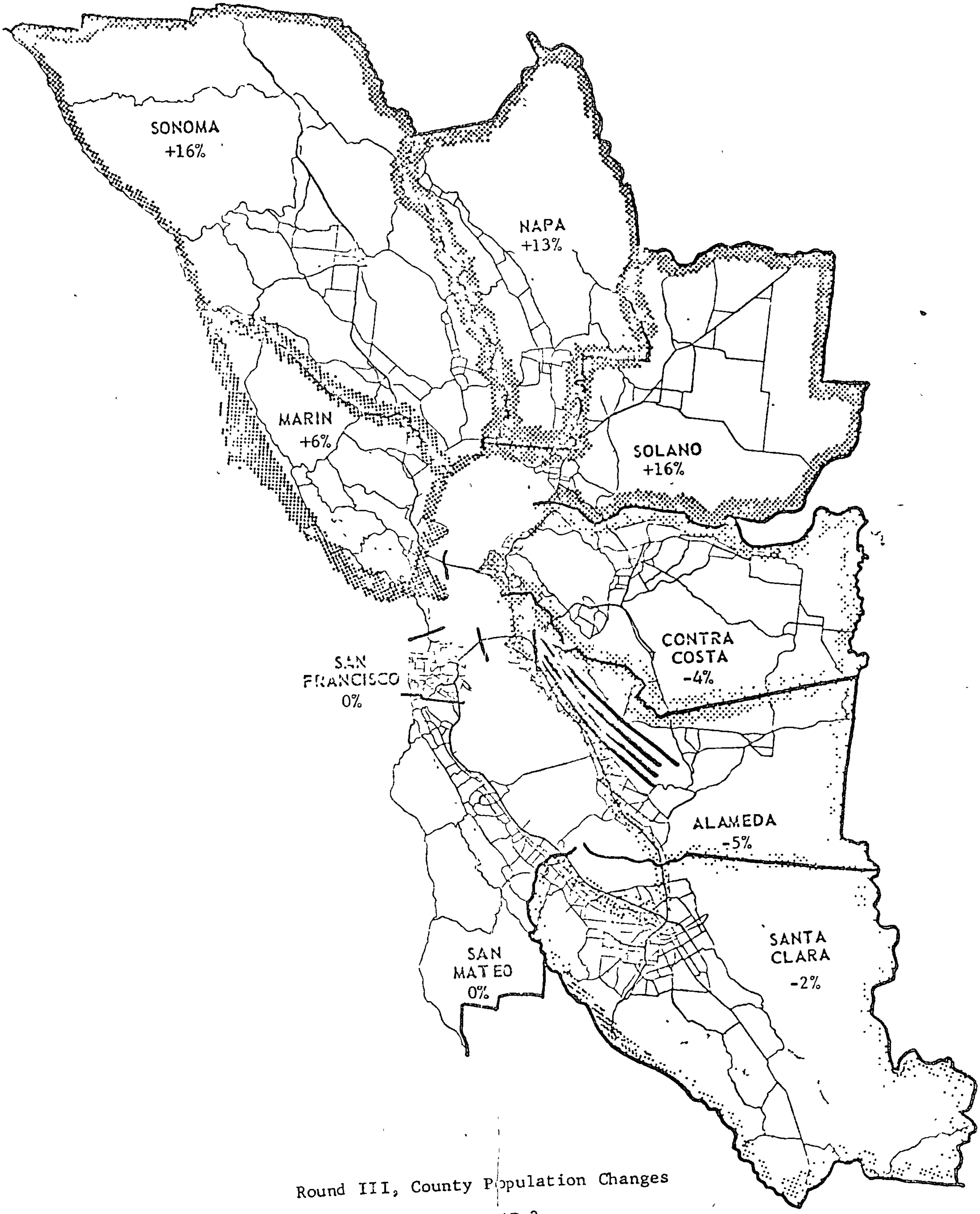


Round III - Compared to IV.6 Figure 1

	1965	IV.6AVG 1980	III.5 1980	% Diff.	Abs. Diff.
San Francisco	754,754	764,950	764,778	- 0.00	- 172
Marin	192,523	320,451	340,685	+ 6.31	+20,234
San Mateo	553,271	778,837	773,802	- 0.65	- 5,035
Sonoma	167,497	288,333	333,490	+15.66	+45,157
Napa	69,378	130,566	146,955	+12.55	+19,389
Solano	149,386	307,813	356,274	+15.74	+48,461
Contra Costa	534,056	788,455	760,834	- 3.51	-27,621
Santa Clara	922,357	1,401,803	1,379,884	- 1.37	-21,919
Alameda	1,095,236	1,626,328	1,558,833	- 4.16	-67,495

County Level Comparison of Run IV.6AVG and Run III.5, 1980 Population

Table 2



Round III, County Population Changes

MAP 3

A more detailed look at the zonal totals in Table 3 and Map 4 confirms the overall impression given by the county totals. In general, the zonal totals moved in the same way as the county totals. There was one major exception to this in Contra Costa where zone 135 showed an increase in population, while the other zones in the county that showed noticeable changes (i.e. above or around 10%) all showed decreases. However, this is in line with the overall pattern since this zone lies close to major roads running north and east, and contains substantial developable land.

In Marin, Napa and Sonoma the main zonal changes were found, as might be expected, in zones close to the Bay. There was one exception to this, in Sonoma, where the zone corresponding to the Santa Rosa area showed large increases. This is to be expected since examination of a map suggests that this is one of the more accessible areas in the county, and is already developing.

Almost all the zones in Solano showed large changes in resident population. The exceptions were zones 111, 112, and 125 which correspond roughly to the valley's area. Examination of the land available in those zones suggests that they are already well built-up. In contradistinction, Santa Clara, San Mateo, and San Francisco showed little change on a zone by zone basis.

The zonal changes in Contra Costa and Alameda were concentrated in the inland sections. This suggests that these areas lying to the east of the coastal hills are the "natural" commuting area for people working in southern Alameda. Looking at the map seems to confirm this, in that the main access routes from northern Alameda into the interland are through built up areas and thus likely to be congested. The routes from the south, while not such high capacity roads, are less likely to carry local traffic and are therefore less likely to be congested.

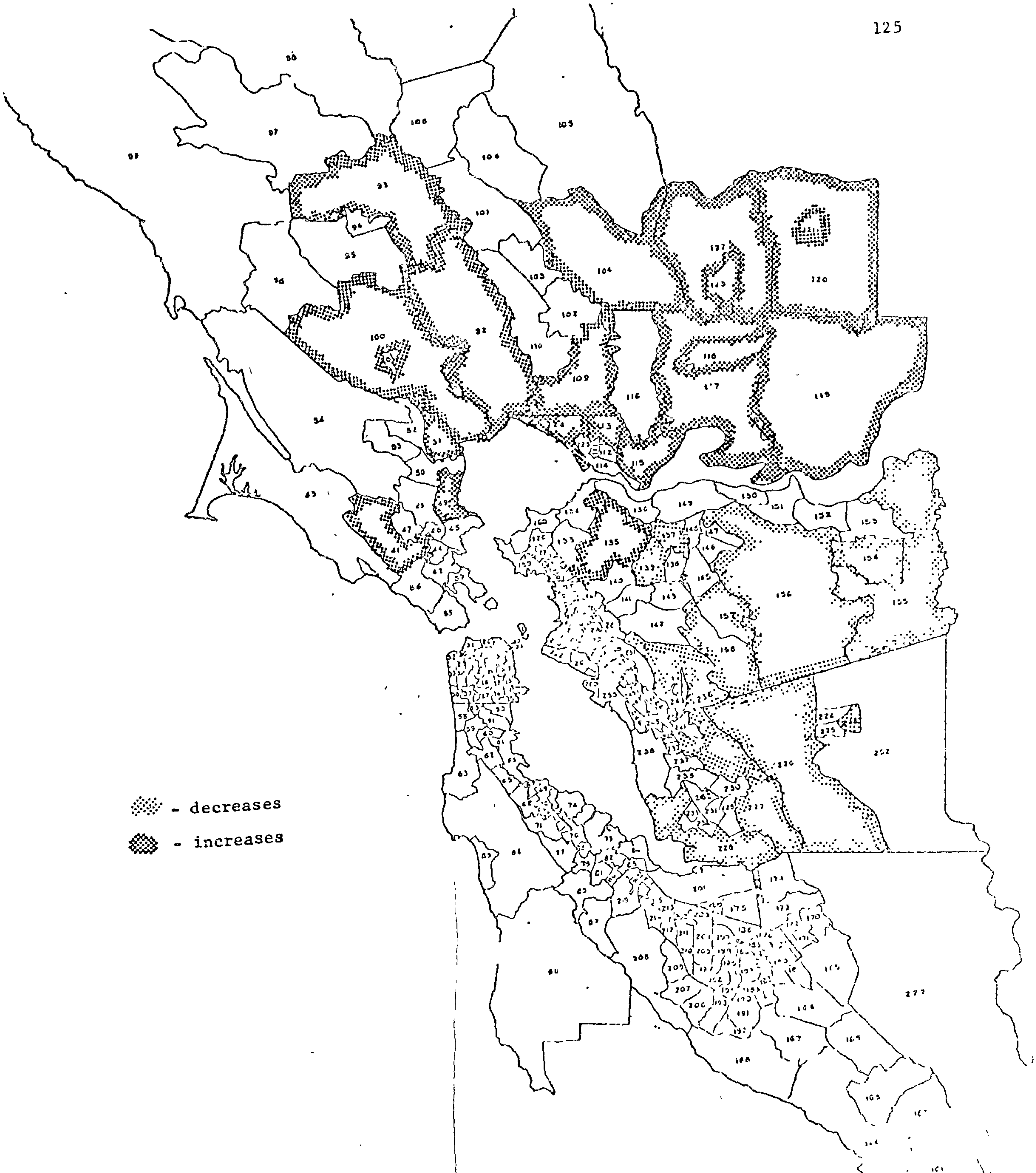
Table 3

Policy Test 1 Round III

Basic Employment Changes in Alameda County

This table shows zonal population changes of 10% (approximately) or more on a zone and county basis.

		Round III.5	Round IV.6AVG	
County	Zone	Population	Population	% difference
Marin	41)	8,106	7,262	11.6
	49)	32,909	29.274	12.6
	51)	17,235	14,317	20.4
San Mateo	No large changes			
Sonoma	92)	92,953	60,736	50.6
	93)	38,890	34,116	35.9
	100)	31,385	28,609	9.7
Napa	104)	7,043	5,606	25.6
	109)	37,542	31,924	17.6
	110)	7,325	6,638	10.3
Solano	113)	37,033	34,122	8.5
	115)	22,528	20,061	12.3
	117)	10,262	8,252	24.6
	118)	79,018	65,559	20.5
	119)	3,633	3,163	14.9
	120)	4,239	3,628	16.8
	121)	8,651	7,469	15.8
	122)	35,106	24,443	43.6
	123)	51,779	39,825	30.1
	124)	5,175	4,741	9.2
Contra Costa	135)	32,027	28,538	12.2
	137)	32,906	35,289	- 6.8
	139)	10,911	11,901	- 8.3
	150)	13,813	15,053	- 8.2
	151)	43,636	48,347	- 9.8
	154)	9,081	11,241	-19.2
	155)	2,963	4,514	-34.3
	156)	10,910	13,238	-17.6
	157)	18,184	21,192	-14.2
	158)	29,667	33,880	-12.4
Santa Clara	No major changes			
Alameda	223)	7,619	8,469	-10.0
	226)	115,658	137,291	-15.8
	227)	30,807	35,264	-12.6
	228)	53,607	63,047	-15.0
	236)	84,028	100,466	-16.4
	252)	10,637	12,649	-15.9



Round III, Zonal Population Changes

MAP 4

The congestion measures by county are shown in Table 4. It can be seen, as anticipated in the discussion above, that by comparison with Round IV there were significant increases in congestion in both Napa and Alameda counties as a result of this policy.

A measure of the weighted average trip length of trips originating in the counties gives the following results

	IV.6AVG	III.5
San Francisco	25.30	25.05
Marin	45.46	43.26
San Mateo	38.62	39.52
Sonoma	35.10	34.47
Napa	35.74	34.37
Solano	38.01	35.70
Contra Costa	44.43	44.40
Santa Clara	30.81	32.24
Alameda	34.91	36.57

Note that the trips (predominantly work trips) from Marin, Sonoma, Napa, and Solano are all somewhat less in run III.5, corresponding to the northward shift of the employment locations to which many of the residents of these counties commute. The trips from San Mateo, Santa Clara, and Alameda are, as one would hope, somewhat longer.

Finally, the total trip minutes spent by travellers on the network goes from 31,338,000 in run IV.6AVG to 33,523,220 in run III.5.

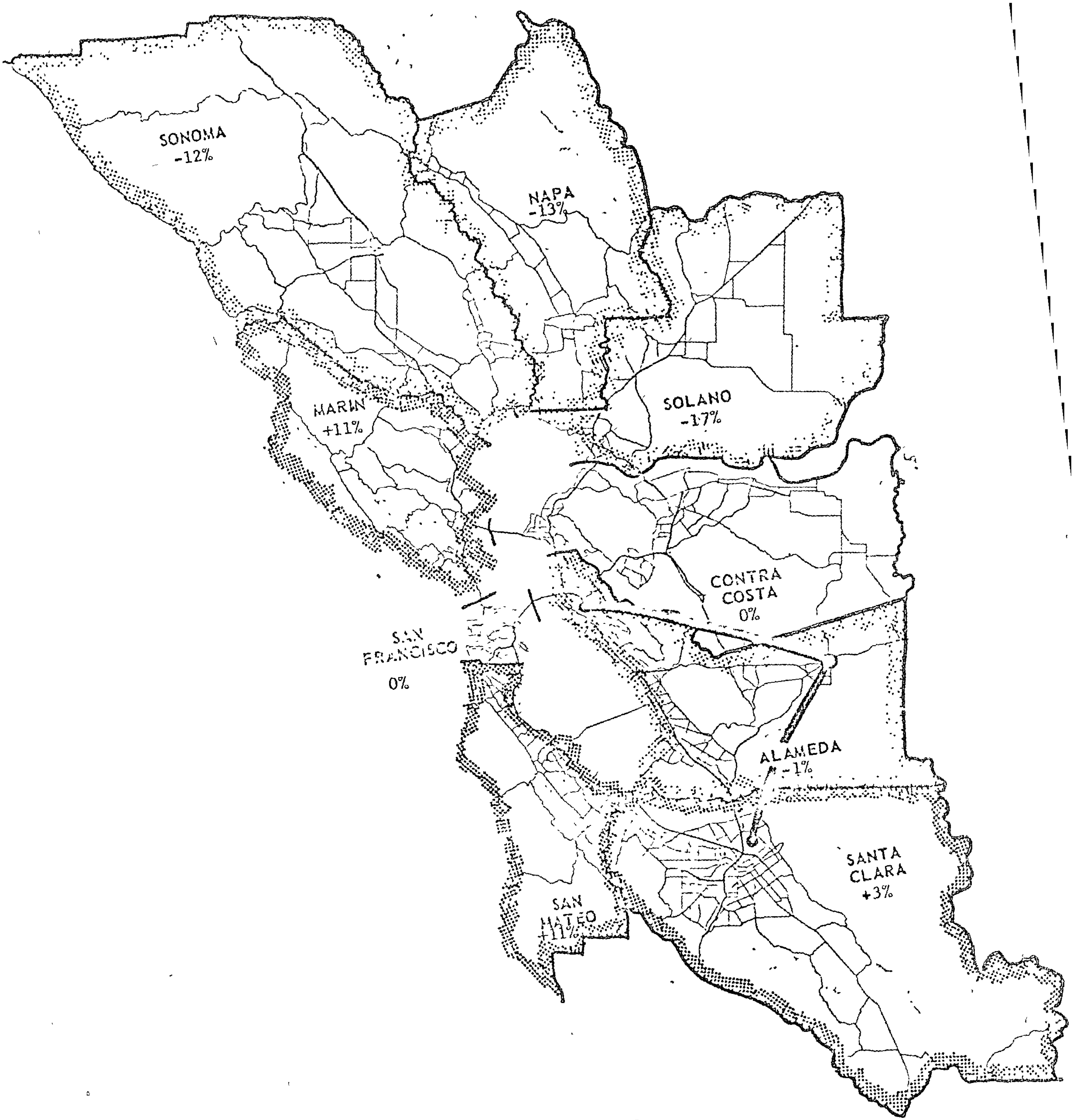
Policy Test 2: Round V Runs

The second full round of runs for policy testing, Round V, was an attempt to access the effect of highway network changes by installing high-speed, high capacity links (shown on Map 5) between Oakland (northern Alameda county) and Livermore Valley (central Alameda county), and between Livermore and an area just northwest of San Jose (northern Santa Clara county). The results of these runs

	III.3	III.4	III.5	III.6	IV.6AVG
San Francisco	1.890	1.935	1.925	1.926	1.928
Marin	1.398	1.307	1.075	1.067	1.090
San Mateo	2.243	2.166	2.133	2.115	2.102
Sonoma	2.44	2.503	2.413	2.419	2.407
Napa	1.059	1.591	1.216	1.422	1.262
Solano	1.859	1.584	1.514	1.556	1.551
Contra Costa	1.578	1.757	1.457	1.604	1.566
Santa Clara	1.716	1.675	1.683	1.674	1.677
Alameda	1.974	2.006	1.925	1.933	1.766

Round III Congestion Measure on input to IPLUM

Table 4



MAP 5

Round V, County Population Changes

are shown in Table 5 and Figure 2.[2] Again, using run IV.6AVG as a base, comparison is made with run V.5 and shown in Table 6. The overall regional pattern has some expected features, and some rather unexpected ones. Firstly, as might be expected, the three most northern counties lost population, some of which moved into the southern counties of Santa Clara and San Mateo.

The curious results came in Alameda, which showed little change in total population, and in Marin, which showed a large increase. The behavior of Alameda may be attributed to the redistribution of population between zones within the county, which will be investigated below. Marin is more difficult to explain. It is possible that Marin and the other northern counties are almost equally accessible with Marin just that much "closer in" that it only requires small changes in the network to tip the balance. In any event, it will be necessary to undertake a more detailed analysis to check this. On Map 5 the bold lines indicate the new network links while the county level shifts in population are shown as percentages below each county zone name.

At the zonal level, and beginning with Santa Clara, Table 7 and Map 6 show that at the zonal level the zones to the east and south of the county showed gains. The largest increases were in zones near the end of the new road. This is to be expected, since these are the areas rendered more accessible by the network changes.

San Mateo showed mixed changes. The largest change was an increase in zone 75 which is large, empty, and at the end of the Dumbarton Bridge and likely to benefit from the road. Large increases took place at the extreme northern end of the county, adjacent to San Francisco, while large negative changes occurred at the extreme south and west.

---

[2] Note that the basic employment relocation used in round III was not used in round V.

		Run V.2		Run V.3		Run V.4		Run V.5		Run V.6	
	1965	1980-2		1980-3		1980-4		1980-5		1980-6	
San Francisco	754,754	765,562	1.43%	767,672	1.71%	765,841	1.47%	766,898	1.61%	765,946	1.48%
Marin	192,523	265,593	37.9	384,865	99.9	327,091	69.9	357,235	85.6	399,042	76.1
San Mateo	553,271	741,227	33.9	852,372	54.1	777,238	40.5	826,875	49.5	789,550	42.7
Sonoma	167,497	322,346	92.4	240,455	43.6	298,937	78.5	250,551	49.6	332,285	98.4
Napa	69,378	166,038	139.3	111,759	61.1	130,706	88.4	114,863	65.6	136,468	96.7
Solano	149,386	433,627	190.2	230,153	54.1	312,623	109.3	256,759	71.9	317,144	112.3
Contra Costa	534,056	790,064	47.9	766,631	43.5	769,153	44.0	783,668	46.7	751,502	40.7
Santa Clara	922,357	1,318,723	42.9	1,443,655	56.5	1,331,590	44.4	1,438,544	55.9	1,366,135	48.1
Alameda	1,095,236	1,604,364	46.5	1,610,021	47.0	1,694,371	54.7	1,612,145	47.2	1,609,470	46.9
Total	4,438,358	6,407,477	44.4	6,407,470	44.4	6,407,469	44.4	6,407,472	44.4	6,407,475	44.4

Round V: Base Population, Projected Population 1980, and Percent Change

Table 5

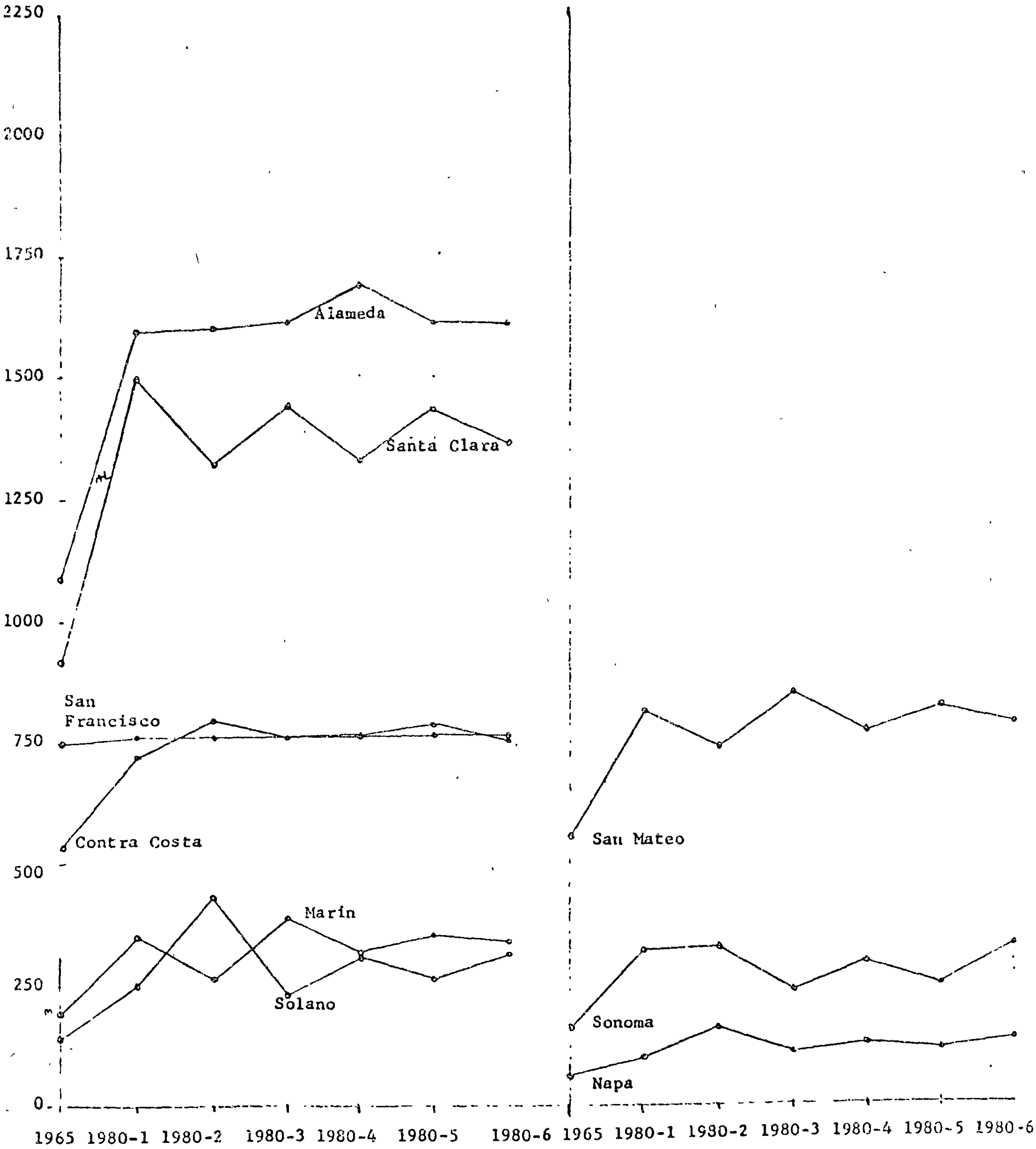


Figure 2 Round V

	1965	IV.6AVG 1980	V.5 1980	% Diff.	Abs.
San Francisco	754,754	764,950	766,898	+ 0.00	+ 1,948
Marin	192,523	320,451	357,235	+11.48	+36,784
San Mateo	553,271	778,837	826,875	+ 6.17	+48,038
Sonoma	167,497	288,333	250,551	-13.10	-37,782
Napa	69,378	130,566	114,863	-12.03	-15,703
Solano	149,386	307,813	256,759	-16.59	-51,054
Contra Costa	534,056	788,455	783,668	- 0.61	- 4,787
Santa Clara	922,357	1,401,803	1,438,544	+ 2.62	+36,741
Alameda	1,095,236	1,626,328	1,612,145	- 0.87	-14,183

County Level Comparison of Run IV.6AVG and Run V.5, 1980 Population

Table 6

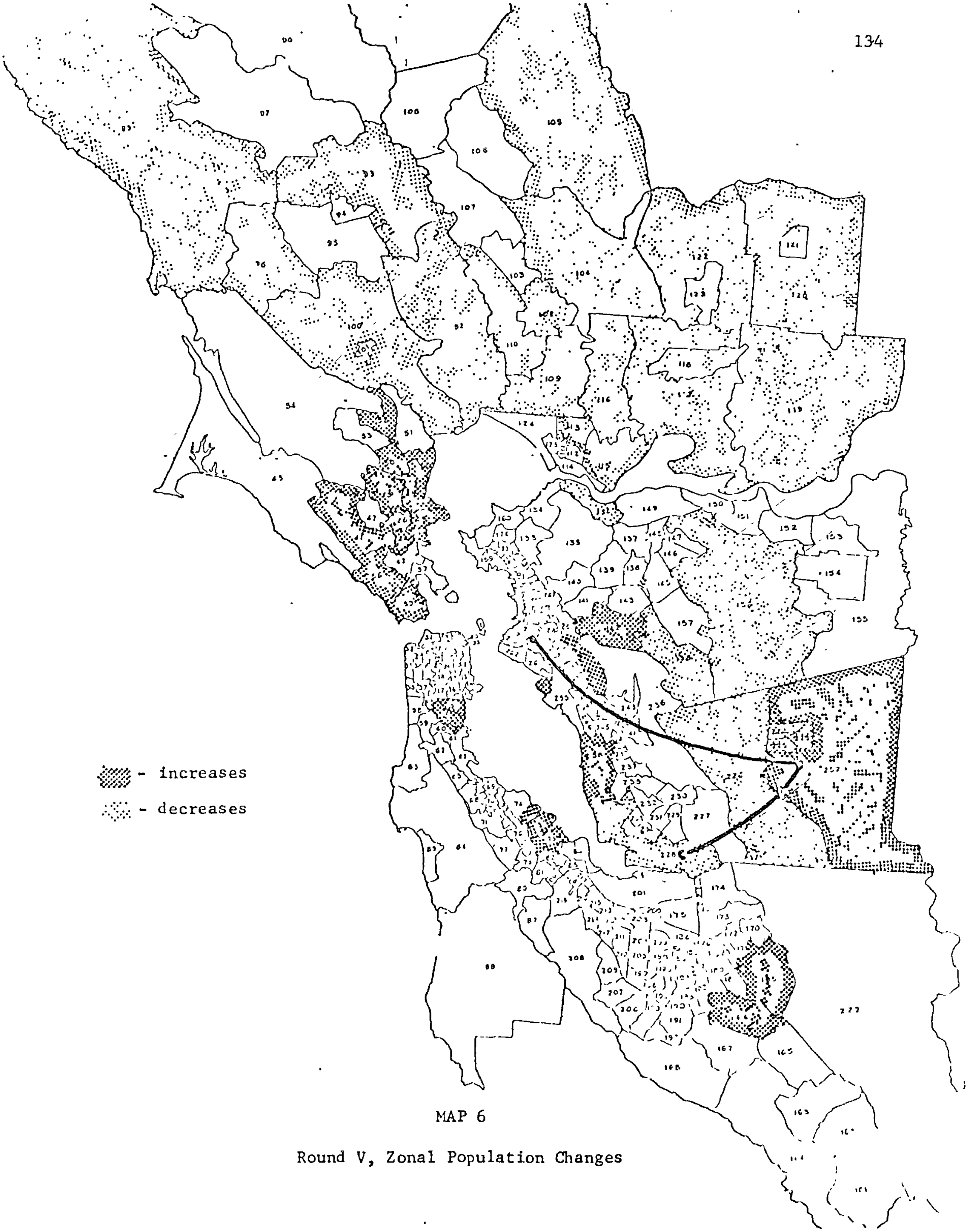
Table 7

Policy Test 2 Round V

New Highway Links in Alameda County

This table shows zonal population changes of 10% (approximately) or more on a zone and county basis.

County	Zone	Round III.5 Population	Round IV.6AVG Population	% difference
Marin	41)	10,506	7,262	44.67
	44)	23,845	20,199	18.05
	45)	21,126	18,574	13.74
	48)	47,381	42,148	12.41
	49)	33,675	29,274	15.03
	50)	26,270	21,854	20.20
	52)	14,428	12,157	18.68
	55)	18,136	15,383	17.90
	56)	19,874	16,289	22.00
San Mateo	75)	93,469	68,604	36.24
	88)	6,534	10,224	-36.09
	89)	6,930	8,136	-14.82
	90)	14,120	11,445	23.37
	91)	13,751	10,459	31.48
Sonoma	92)	47,813	60,736	-21.28
	93)	29,287	34,116	-14.15
	96)	23,288	28,988	-19.66
	99)	11,237	13,833	-18.77
	100)	18,773	28,609	-34.38
Napa	102)	51,998	57,740	- 9.94
	104)	4,437	5,606	-20.85
	105)	923	1,268	-27.21
	109)	25,106	31,924	-21.36
	110)	5,836	6,638	-12.08
Solano	113)	28,491	34,122	-16.50
	115)	17,618	20,061	-12.17
	116)	1,937	2,435	-20.
	117)	5,162	8,252	-37
	118)	50,345	65,559	-23.
	119)	2,385	3,163	-24.59
	120)	2,690	3,628	-25.85
	121)	5,996	7,469	-19.72
	122)	16,694	24,443	-31.70
	123)	25,258	39,825	-36.58
Contra Costa	136)	25,924	29,609	-12.44
	142)	27,184	23,157	+15.14
	156)	11,323	13,238	-14.47
	158)	30,264	33,880	-10.67
Santa Clara	166)	81,884	71,139	15.10
	169)	48,004	36,519	31.44
Alameda	223)	6,681	8,469	-21.08
	224)	11,291	15,128	-25.36
	225)	15,791	20,359	-22.44
	226)	89,023	137,291	-35.15
	228)	71,938	63,047	-16.26
	238)	107,725	85,910	25.39
	251)	18,846	16,958	11.13
	252)	20,429	12,649	61.15
	260)	20,021	16,894	18.51



MAP 6

Round V, Zonal Population Changes

Zones in Sonoma, Napa, and Solano showed fairly uniform decreases in almost all zones. Marin, however, showed uniform increases, though the largest was in the zone adjacent to the end of the bridge to Berkeley and Oakland.

Contra Costa showed, for the most part, only small changes. The northern zones followed Solano, and showed decreases. The southern and inland zones followed the inland areas of Alameda and also showed decreases. However, 142 in Contra Costa, which is rather near one end of the new road, increased, along with 251 in Alameda, which showed significant increases.

Alameda showed mixed behavior. As could be expected, zone 252 where the super highway ends, showed a substantial increase in population, though this only balanced the loss in population in zones 223-226 in the Livermore Valley. In part, this is a "hoped-for" result, since it suggests that such a limited access highway might save the areas it crosses but to which it does not give access.

A possible explanation for some of this unexpected behavior may be due to the fact that the allocation functions do not put people into nearest zone, but distribute travel behavior over many zones. Further, accessibility measures, via their normalization procedures, operate in terms of relative accessibility of one zone vis-a-vis all others. This is certainly true in the model system and probably true in reality. In any case, if the trip length measures, shown below, are examined, some surprising conclusions appear. These trip length measures are:

	IV.6AVG	V.5
San Francisco	25.30	26.01
Marin	45.46	45.02
San Mateo	38.62	39.61
Sonoma	35.10	42.69
Napa	35.74	41.65
Solano	38.01	46.27
Contra Costa	44.43	49.51
Santa Clara	30.81	34.78
Alameda	34.91	36.46

From these, plus the other projection results, it appears that the regional effect of the new road has been to shift the concentration of employment and population of the region towards the south. At the same time Marin county was "swamped" with new residents with San Mateo and Santa Clara taking the rest of the shift. Those residents remaining in the northern half of the region found themselves making much longer work trips. Even the work trips of residents in Santa Clara and Alameda increased, as they probably made use of the new road to commute to the Berkeley-Oakland area in northeastern Alameda and Contra Costa counties.

Total trip minutes went from 31,338,000 in run IV.6AVG to 32,494,593 in run V.5. The congestion measures for this round are shown in Table 8.

Policy Test 3: Round VII Run

Due to the press of time and resources a full round of runs was not done for a land use control test. The response of the model system to land use control was tested by re-running run IV.6AVG with the land use controls imposed by defining additional land in certain zones as being unavailable for development. The county level results of this run, called VII.1 are shown compared with the results of run IV.6AVG in Table 9. On Map 7 the county level shifts in population are shown as percentages below each county zone name, and the areas in which stringent land use controls were applied is shown shaded.

As can be seen on the map, the effect of these controls is to reduce population growth in the counties (Contra Costa and Alameda) where the controls were applied. Note that this is a reduction from the "no-controls" base run (run IV.AVG), but there is still substantial absolute growth in these counties over the period 1965 to 1980. The counties to the south and south-east (San Mateo and Santa Clara) also showed a relative decline in population. This presumably was due to declines

	V.3	V.4	V.5	V.6	IV.6AVG
San Francisco	1.713	1.925	1.908	1.916	1.928
Marin	0.724	0.935	0.902	0.920	1.090
San Mateo	1.861	2.066	2.056	2.052	2.102
Sonoma	1.856	1.856	1.910	1.863	2.407
Napa	1.368	1.306	1.308	1.268	1.262
Solano	1.500	1.468	1.428	1.367	1.551
Contra Costa	1.280	1.456	1.415	1.440	1.566
Santa Clara	1.368	1.590	1.587	1.588	1.677
Alameda	1.497	1.629	1.610	1.635	1.766

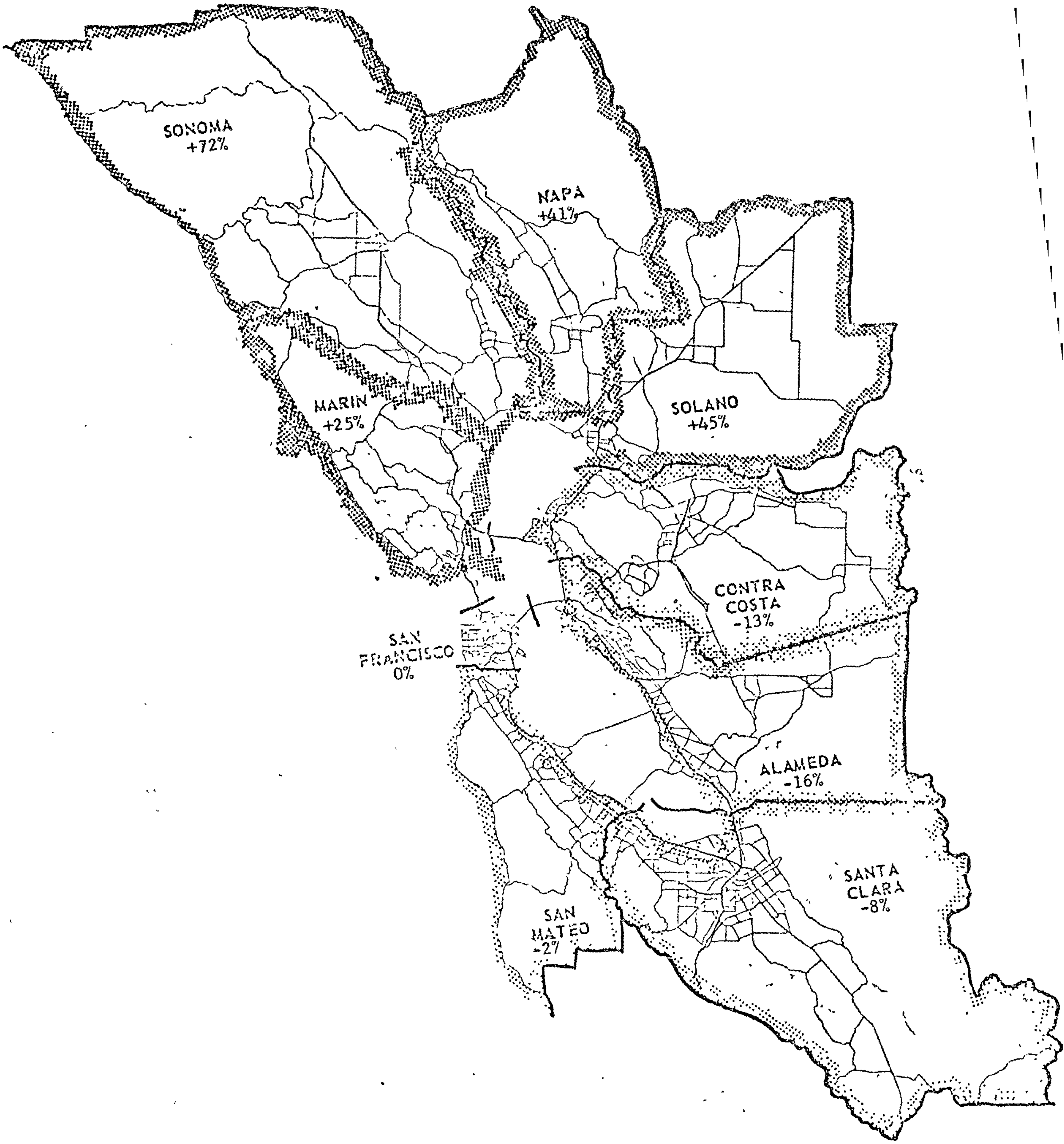
Round V. Congestion Measures on input to IPLUM

Table 8

	1965	IV.6AVG 1980	V.5 1980	% Diff.	Abs.
San Francisco	754,754	764,950	763,622	- 0.17	- 1,328
Marin	192,523	320,451	402,148	+25.49	+ 81,697
San Mateo	553,271	778,837	764,426	- 1.85	- 14,411
Sonoma	167,497	288,333	498,057	+72.74	+209,724
Napa	69,378	130,566	184,445	+41.27	+ 53,879
Solano	149,386	307,813	446,421	+45.03	+138,608
Contra Costa	534,056	788,455	682,986	-13.38	-105,469
Santa Clara	922,357	1,401,803	1,294,710	- 7.64	-107,093
Alameda	1,095,236	1,626,328	1,370,732	-15.72	-255,596

County Level Comparison of Run IV.6AVG and Run VII.1, 1980 Population

Table 9



MAP 7

Round VII, County Population Changes

in local serving employment in Contra Costa and Alameda to which employees residing in San Mateo and Santa Clara formerly commuted. Further indirectly induced declines in population and local serving employment took place throughout these four counties.

True to the notion that if a growing region is squeezed in one place it bulges out in another, there was substantial relative growth in the four northern counties; Marin, Sonoma, Napa, and Solano. This growth, for the most part, represents residential location of employees of San Francisco and Contra Costa (Berkeley, Oakland, etc.) who were unable to find residential space in Contra Costa and Alameda. Again, indirect effects via the local serving employment tend to amplify the situation.

The tabulation of zonal results is given in Table 10 and shown on Map 8. There are no results at the zonal level which are not expected from the county level results. It is, however, worth noting that virtually all the significant zonal changes take place on the urban fringe. This suggests that if appropriate policies were to be developed and implemented there would be a fair chance of having some effect on the shape of the future development of a region.

Summary

In a growing metropolitan region, the actual regional growth rate may be determined by factors over which the regional policy maker has little or no control. This is even more true for the local policy maker in sub-areas of the region. To the extent that the overall growth of the region is significant and beyond the control of the regional policy maker, he finds himself in a paradoxical position. A policy which seems perfectly reasonable for one part of the region may have serious, undesirable consequences elsewhere in the region. A salient feature of

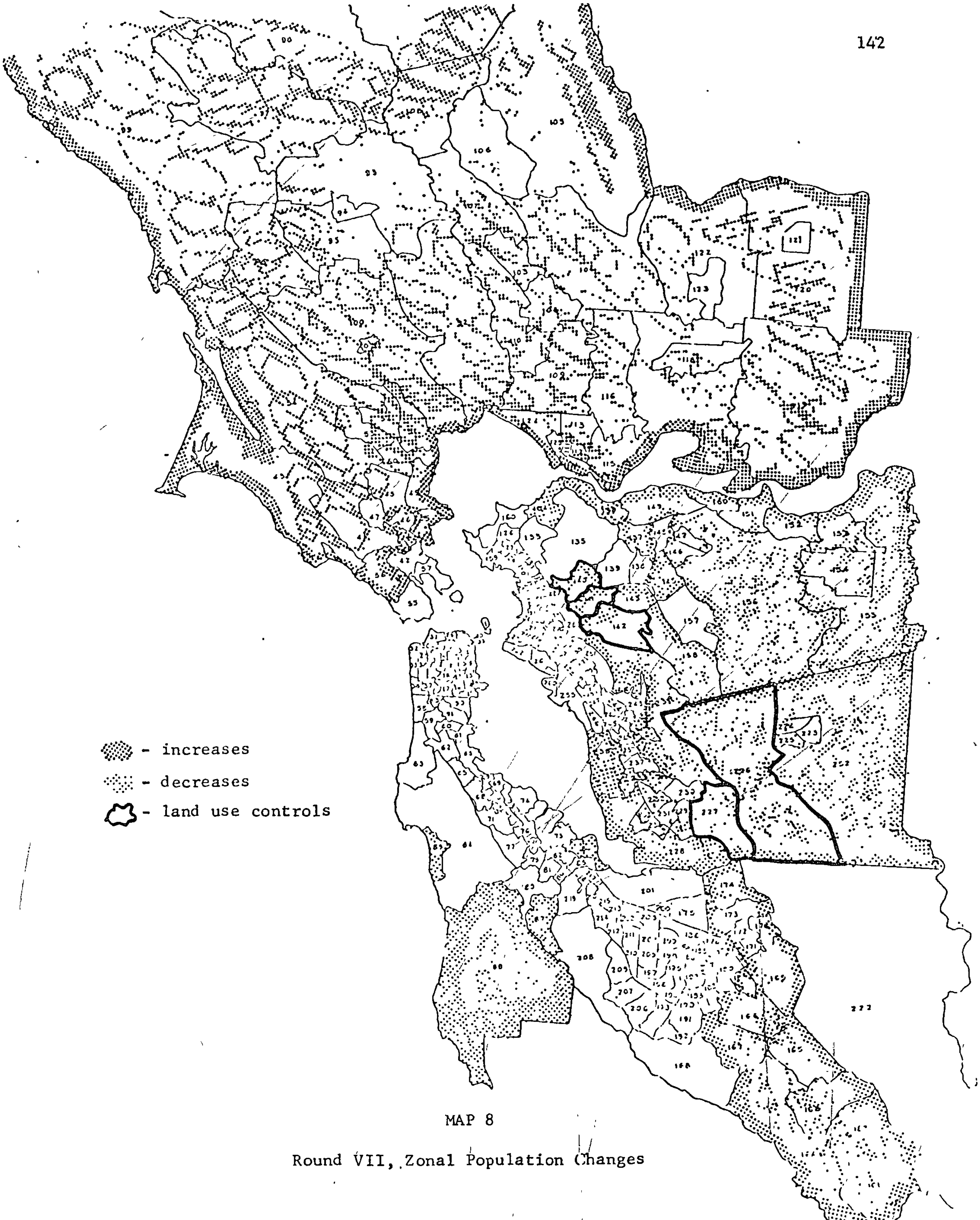
Table 10

Policy Test 3 Round VII

Land Use Control in Contra Costa and Alameda Counties

This table shows zonal population changes of 10% (approximately) or more on a zone and county basis

County	Zone	Round IV.6AVG. Population	Round VII.1 Population	% difference
Marin	41)	7,262	9,741	34.
	43)	5,415	6,434	12.
	44)	20,199	22,793	13.
	45)	18,574	23,006	24.
	47)	21,121	23,417	11.
	48)	42,148	53,482	27.
	49)	29,274	47,306	61.
	50)	21,854	31,656	45.
	51)	14,317	25,599	80.
	52)	12,157	17,923	47.
	53)	13,312	17,078	28.
	54)	8,599	12,344	42.
	56)	16,269	20,436	25.
San Mateo	84)	37,924	31,382	-17.
	87)	17,715	15,104	-15.
	88)	10,224	7,892	-23.
	89)	8,136	6,704	-18.
Sonoma	92)	60,736	160,479	164.
	93)	34,116	54,837	61.
	94)	29,299	38,191	30.
	95)	36,824	54,422	52.
	96)	28,988	53,638	86.
	97)	19,574	31,833	63.
	98)	10,575	12,025	14.
	99)	13,833	17,024	23.
	100)	28,609	46,551	63.
	101)	19,979	23,255	16.
Napa	102)	57,740	79,398	37.
	103)	2,448	3,006	23.
	104)	5,606	9,385	67.
	105)	1,268	2,250	78.
	106)	5,118	10,189	97.
	107)	9,109	13,062	43.
	108)	3,238	4,036	24.
	109)	31,924	46,941	47.
	110)	6,638	8,700	31.
Solano	113)	34,122	41,145	21.
	114)	29,033	36,352	25.
	115)	20.061	23,818	19.
	116)	2,435	3,103	27.
	117)	8,252	12,350	50.
	118)	65,559	102,619	56.
	119)	3,163	4,526	43.
	120)	3,628	7,109	96.
	121)	7,469	12,521	68.
	122)	24,443	54,748	124.
	123)	39,825	75,851	90.
	124)	4,741	5,813	23.
Contra Costa	134)	26,148	29,042	11.
	136)	29,609	32,999	11.
	137)	35,289	28,779	-19.
	140)	8,451	5,454	-35.
	141)	18,869	11,958	-37.
	142)	23,157	10,785	-53.
	143)	31,202	27,403	-12.
	145)	38,000	31,332	-17.
	146)	23,420	20,118	-14.
	149)	13,132	10,723	-18.
	150)	15,053	12,104	-20.
	151)	48,347	37,046	-23.
	152)	26,716	23,714	-11.
	153)	6,099	4,974	-18.
	154)	11,241	7,995	-29.
	155)	4,514	2,442	-46.
	156)	13,238	7,011	-47.
	158)	33,880	13,235	-61.
Santa Clara	161)	39,164	28,703	-27.
	162)	17,257	12,205	-30.
	163)	23,453	16,796	-28.
	164)	6,584	4,829	-27.
	165)	20,754	12,075	-42.
	166)	71,139	52,340	-27.
	167)	35,468	24,492	-31.
	169)	36,519	28,409	-20.
	170)	12,391	10,880	-12.
	172)	43,536	39,308	-10.
	173)	36,638	29,193	-20.
	174)	39,815	33,153	-17.
	181)	37,463	30,965	-17.
Alameda	224)	15,128	9,674	-36.
	225)	20,359	13,897	-38.
	226)	137,291	68,461	-50.
	227)	35,264	14,820	-58.
	228)	63,047	41,888	-34.
	229)	47,553	40,219	-16.
	230)	18.148	14,677	-19.
	232)	21.035	17,155	-19.
	234)	14,982	12,963	-18.
	235)	24,219	10,557	-19.
	236)	100,466	28,925	-79.
	238)	85,910	68,097	-21.
	241)	30,582	25,976	-15.
	252)	12,649	6,884	-46.



MAP 8

Round VII, Zonal Population Changes

these policy test runs is the fact of this regional interrelatedness. In each of these policy runs, a change in one portion of the region appears to produce significant changes in other parts of the region.

To Summarize these tests:

Test 1 - (Round III)

A policy to inhibit the spatial dispersion of basic employment in Alameda county by prohibiting new location in the south-central and eastern parts of the county, did reduce residential and local-serving employment sprawl in both Alameda and Contra Costa counties. In the absence of a comprehensive regional land use control policy, however, the desirable results in Alameda and Contra Costa were offset by the generation of considerable sprawl in Marin, Sonoma, Napa, and Solano counties.

Test 2 - (Round V)

A policy to centralize impending development in eastern Alameda county by the introduction of a high speed, very limited access highway did not seem to produce the desired results. There was a decline in population and local-serving employment in the area where it was to be centralized. The new highway seems to have been utilized for longer distance commuting and to have resulted in a general southerly shift in the location of the region's employment and population. This seems to have had a substantial effect only in Marin county, the southeastern portion of which was swamped with new residential, and local-serving employment, development.

Test 3 - (Round VII)

An attempt to limit development in western Contra Costa county and central Alameda county by the imposition of stringent land use controls was generally successful in those counties. Again, the lack of a regionwide policy allowed the development which was controlled in those counties to run rampant in the northern counties of Marin, Sonoma, Napa, and Solano.

Before continuing, it should again be noted that these tests are in no way representative of any real policy intentions known to the authors. They are presented here as examples of the use of the model package and, to a limited extent, as a demonstration of the types of phenomena which are likely to occur in reality.

The question may have been raised, in the mind of the perspicacious reader, as to why one particular model run in a round may have been chosen for analysis as opposed to another. This was simply a matter of selecting the run which seemed most likely to represent the "equilibrium" situation. The need for such selection, admittedly somewhat arbitrary, will be obviated in the future by use of the "new process" procedure, (described in the previous chapter) which seems to produce solutions closer to a true equilibrium. The merit of this assertion remains to be tested by further use of the model package.

Finally, with regard to the overall aims of this project, the conclusions, in one respect at least, are clear. There are strong interrelationships between transportation development and land development which are not adequately accounted for, even in complex computer simulation models, by traditional forecasting procedures. Further, it is clear that as more knowledge of these phenomena is developed, powerful instruments can be made available to policy makers for use in shaping the growth of the metropolis. Integrated, region-wide, planning of both transportation and land use is a necessary condition for the proper design and application of these policies. The integrated transportation and land use model package makes a significant step forward in accounting for these interrelationships. Hopefully it will be of use in future policy design and testing.

APPENDIX A-1

SURVEY OF LAND-USE AND DEVELOPMENT GUIDANCE TECHNIQUES

Richard B. Rubin

December 1972 (Rev. May 1973)

This work was performed as part of contract DOT-FH-11-7843 from the U.S. Dept. of Transportation, Federal Highway Administration, to the Institute for Environmental Studies, Department of City and Regional Planning, University of Pennsylvania under the direction of Dr. Stephen H. Putman, Principal Investigator

TABLE OF CONTENTS

I. INTRODUCTION 1

II. BACKGROUND AND CLASSIFICATION OF LAND GUIDANCE TECHNIQUES 1

III. REGULATION THROUGH THE POLICE POWER 3

- a. Legal Background 3
- b. Zoning 4
- c. Zoning Drawbacks 7
- d. Other Regulatory Devices 10

IV. OTHER LAND-USE AND DEVELOPMENT GUIDANCE TECHNIQUES 12

- a. Public Land Acquisition 12
- b. Public Investment Programming 13
- c. Taxation Policy and Incentives 13

V. PROBLEMS WITH OUR PRESENT "SYSTEM" 15

- a. Fragmentation and Parochialism in Use of Tools 15
- b. Problems in Planning and Administration 16
- c. Inability of Present Techniques to Meet Present needs 17
- d. Inequities of the System 18

VI. THE NEED FOR CHANGE 19

- a. Factors Promoting the Status Quo 19
- b. How to Proceed? 20

VII. PROPOSALS FOR IMPROVING THE PRESENT SYSTEM 22

- a. Administrative Recommendations 22
- b. Proposals for Reallocating Land Guidance Functions 25
- c. New or Underutilized Guidance Techniques 29
- d. Economic and Fiscal Guidance Tools 35

VIII. NEW PROPOSALS 39

- a. Activity and Promise at the Federal Level 41
- b. Advance Large-Scale Public Acquisition and Disposal of Land 42
- c. Proposed Operational Frameworks for Development Control 48

IX. CONCLUSION 51

BIBLIOGRAPHY

SELECTED BIBLIOGRAPHIC ANNOTATIONS

"SURVEY OF LAND-USE AND DEVELOPMENT GUIDANCE TECHNIQUES"

I. INTRODUCTION

This survey of "guidance techniques" is concerned with a wider spectrum of actions and policies for influencing and regulating land uses and land development, than is implied by the more traditional categorization, "land use controls." The literature in this field is quite extensive, especially when so broadly conceived as above. Thus, although this survey will touch upon most important areas, the emphasis will be upon proposals for the improvement of the present land-use and development guidance system (land guidance system for short). The development, present state, and problems of our present system are well explored and summarized in other sources, and will be dealt with more briefly here. The reader desirous of more detailed coverage of these areas will be directed to appropriate sources through footnotes.

II. BACKGROUND AND CLASSIFICATION OF LAND GUIDANCE TECHNIQUES

Our present land guidance system is quite naturally a product of American ideological development, as well as legal and historical accident. An appreciation of the social, economic, and psychological environment in which the American system of land-use controls developed is vital to a thorough understanding of its present structure and limitations. This country's vast endowment of land, the traditional American desire for growth, love of freedom, reverence of private property and free enterprise, distrust of government, and the large number of relatively independent local governments in this country are all elements of this environment.[1] A better grasp of our land guidance system can also be achieved with a knowledge of the more or less traditional classification of techniques and approaches that constitute the system. This

---

[1] See (24), pp. 1-15 for a more thorough description of this environment.

classification is grounded upon a primarily legal division of governmental powers and is basically as follows:[2]

a) Controls and regulations flowing from the government's "police power." The police power is the government's right to exercise "reasonable" control over persons and property within its jurisdiction for the protection of the security, health, safety, morals and general welfare of the community. Zoning is by far the best known and most widely used land-use control based on this power.

b) Power of eminent domain. This is the right of government to take private property for public use. In this country, this right may be exercised only if the owner receives "just compensation" for the taking.

c) Public spending policy. The government's ability to spend money and its policies with respect to the manner in which funds are to be spent represent very powerful influences on land use and development, even though these powers are rearely used effectively.

d) Taxation policy and incentives. Of course, the flip side of government spending is government taxation. Naturally, government policy with respect to taxation and related incentives, especially with respect to real property taxation, can be a powerful land guidance tool.

e) Government can also enter into contractural agreements and negotiations with private individuals. This ability can be used to do more efficiently that which might otherwise be done under one of the above powers, or to allow the government to achieve an objective whose attainment is otherwise blocked by legal difficulties.

---

[2] Refer to (30) for a comprehensive treatment of the structure and legal aspects of land-use planning and control. Also see (43), pp. 199-217 for a summary of the history and nature of traditional controls and the current administrative framework.

f) The government can also use its ability to disseminate information and mold public opinion to achieve some measure of influence over land-use and development.

III. REGULATION THROUGH THE POLICE POWER

A. LEGAL BACKGROUND

In this country, the state is the repository of the ultimate power to regulate non-federal land and its use within its boundaries. The Federal government does have some controls available to it, especially with respect to federally owned lands, but as the Constitution is now interpreted, the states retain the bulk of all direct regulatory power over non-federal land. Despite this legal supremacy of the states in police power regulation, it has been predominantly the local governments of this country that have actually wielded the power over land and land-use. The states have been satisfied to enact enabling legislation which delegates the power for regulation of land-use to one or more types of local government. Most of this enabling legislation was originally modeled after the Standard State Zoning Enabling Act (1926) and the Standard City Planning Enabling Act (1928) prepared by the U.S. Department of Commerce.[3] Once these acts had been passed and amended, which acts usually set forth the criteria upon which localities should base regulatory decisions only in very general terms, the states preferred to leave land-use regulation to the local governments. Although this is now changing, it has been true in the past that even the Federal government had more actual impact on land use, through indirect means and control of federal lands, than states had.

---

[3] See (2), p. xvii-xx or (43), p. 201, for example.

It must be noted that even the states themselves are limited in the use of police powers by certain constitutionality standards.[4] Regulations imposed under the police power:

a) must be construed to promote the health, safety, morals, or general welfare of the community. Often, this has been a strong restriction on the use of such regulations. However, the scope of the general welfare has been more liberally construed over the years. In 1954, in the case of Berman v. Parker, the U.S. Supreme Court found that: "The concept of the public welfare is broad and inclusive... It is within the powers of the Legislature to determine that the community should be beautiful as well as healthy, spacious as well as clean, well balanced as well as carefully patrolled."[5]

b) must apply equally to all, unless very good objective reasons exist for treating individuals differently.

c) must be "reasonable" and cannot be imposed if it can legally be construed to be an effective "taking" of property.

B. ZONING

Zoning is the most widely used land-use control. Basically, the zoning ordinance divides the local government's jurisdiction up into one or more sets of districts or zones. Within each zone uniform restrictions are placed on land use. Such restrictions typically include restrictions on the type of use permitted in the zone, maximum and minimum lot sizes, density restrictions, and bulk and height restrictions on buildings.[6] In general, zoning was basically a

---

[4] For detailed treatment of these and other legal aspects of land-use control, see (30) and the classic text on zoning, (7).

[5] (26), p. 124.

[6] See (24), pp. 41-53 or (7) for a more comprehensive treatment of zoning.

convenient extension of nuisance law, which made public in advance those types of land use which would be considered of nuisance value to the community.[7] The primary purposes of zoning when it was originally developed were: (a) to avoid conflicting land uses and (b) to stabilize and maintain property values in the community.[8] It received its first major push in 1916, when New York city adopted its first zoning ordinance.[9] In 1926, the U.S. Supreme Court settled the growing controversy over the new technique by upholding its constitutionality in the case of Village of Euclid v. Ambler Realty Co.[10] Zoning then caught on quite quickly. In 1968, 6,880 municipalities and 2,004 townships had zoning ordinances.[11] Today, Houston, Texas is well known as the only major American city not to have a zoning ordinance. The cit, relies on a program of careful enforcement of private deed restrictions, public land acquisition, and public facilities programming to guide land use and development.[12] In theory, if not in practice, the regulations imposed by zoning (and other land-use controls and guidance techniques for that matter) should be based upon a comprehensive or "master" plan setting forth the goals, policies, and general course of action of the community for the years to come.[13] Despite many legislative and judicial mandates to the contrary, this is almost never

---

[7] (50), p. 4.

[8] See (12), p. 316 and (41), p. 94.

[9] See (24), pp. 20-23. Also see (68) for an interesting and detailed historical account of the rise of zoning in America.

[10] See (68), pp. 213-253 for an extensive account of this case.

[11] (220, p. 5.

[12] See (74) and (49).

[13] See (13), pp. 112-115 for a brief discussion of the use of the general plan and the planning process as a technique for guiding urban growth.

the case. Even when a plan or planning process exists, all too often it isn't used in guiding the actual decisions made, or the plan itself is either too vague or too long-ranged to be of any practical assistance in determining policy or setting criteria.

To meet the challenge of accelerated urban growth, and the complex ramifications thereof, zoning has been made a more flexible and workable tool over the years through modifications and additional techniques:[14]

a) The variance. A simple escape valve, it allowed for exceptions to be made in cases of extreme hardship. As could be imagined, quite a bit of abuse and corruption was associated with the application of this amendment procedure.

b) The special permit: For some zones, a list of activities was written into the zoning ordinance. These were not allowed by right, but a land owner could apply for a "special permit" for such a use, issued at the discretion of the municipality. (Also called an "exception")

c) The "floating zone" or "unmapped zone" was a zone described in the text of the zoning ordinance but not specifically allocated a particular land area nor located on the zoning map. This zone would be created for uses that would be required, but whose desired location was in doubt or was decided to require individual discretionary treatment. A landowner, desiring to carry on such a use, would apply to the government to have the piece of land in question placed in this aspatial zone.

d) In "contract zoning" or "conditional zoning," the municipality merely agrees to a desired rezoning if the owner agrees to place certain conditions on the use of his land through a contract or deed restriction. This technique permits added discretion and flexibility on the

---

[14] (5), pp. 6-11

part of the government and has the added advantage of setting no troublesome legal precedents.

e) With changing technology, those uses which were once considered incompatible with a certain type of land use can now be made to be compatible and often desirable in a certain use area, like a factory that is neither noisy nor offensive, offering nearby jobs to a travel-weary community. Simple lists of permitted uses are no longer adequate. To cope with this change, a technique called "performance standard zoning" has been introduced. Instead of simply listing uses that are allowed or prohibited, performance standards are set that must be met by any use in the zone. Thus, if a research and development firm can conform to the standards of a residential zone, it would usually be permitted. However, great care must be taken in specifying proper standards.

C. ZONING DRAWBACKS

Over the years, quite an impressive list of criticisms of zoning has been compiled by planners and others who deal with land-use controls. Much of this criticism is justified, however not all of our land-use problems can be attributed to the failures of zoning. This can be better appreciated if we acknowledge a rather useful dichotomy in the drawbacks associated with zoning. While some of the criticism can justifiably be leveled at zoning and the manner in which it is presently being administrated, others are less fair in that they merely point out that zoning has been unable to cope with problems and needs for which its use was never intended.

Among the problems attributed to zoning which fall in this latter category are:

a) Zoning is a "negative" technique. It was intended to prevent undesirable uses. It cannot be successfully used by itself to insure that a given use will establish itself in a given place, since zoning a parcel for a certain use cannot insure that the owner or other developer will establish such a use.

b) Zoning was developed for regulating relatively small-scale development, as occurs on the small lots usually found in developed areas. It is too rigid, and inappropriate, for use on large scale developments.

c. Zoning was intended for the regulation of developed land. It is inappropriate to try to use it to regulate undeveloped land or as a tool to guide development in suburbanizing areas.[15] The development process is dynamic, complex, and founded upon private initiative for much of the land use that eventually occurs. Zoning is negative, primarily static, and flexible. It has been a dismal failure in effectively controlling development, especially the problem of urban sprawl. Zoning in undeveloped areas establishes specific use districts long before anyone could possibley determine the eventual development pattern and thus the wisdom of such a specification.[16] This premature restriction often leads to an undesirable development pattern and unnecessary and costly land speculation. On the other hand, zoning restrictions in such areas often have no effect, as they fall in a piecemeal manner under growing development pressures, stripping the community of any unified approach to the challenge of growth that might have been adopted if not for the false sense of protection offered by zoning.

---

[15] See (32), p. 13; (41), p. 98; and (46), pp. 15-16.

[16] See (46), p. 19.

d) Related to the above problems is that of "overzoning" and "underzoning."[17] Zoning for uses for which there is little or no demand, in the idle hope of attracting such a use, is counterproductive, especially when such actions only act to artificially increase land prices. The land goes idle or is developed in an inefficient and scattered manner. Underzoning for a use can also be counterproductive. The attempt to exclude a use often fails under strong development pressure, and where the weakest link breaks will often be far from the most desirable location for the use.

Of course, many of the criticisms of zoning are well founded, even when restricted to consideration of the regulatory role for which zoning was intended:

a) "Fiscal zoning" refers to the use of zoning and subdivision regulation "(1) to encourage forms of land development that minimize the need for services, and (2) to maximize the ratio between taxes reserved and service cost."[18] It is only natural for a community to try to exclude persons who require expensive municipal services and encourage the settlement of revenue-producing industry when local property taxes are the primary source of government revenue for providing required local facilities and services in a locality. However, on a metropolitan or regional scale, this results in a very inequitable and inefficient distribution of financial and public-service burdens.

b) Zoning provides a very piecemeal approach for accomodating changing conditions. The excessive and ad hoc manner in which variances are handed out violates the intent of the zoning ordinance and has an "eroding" effect on the community.[19]

---

[17] See (46), pp. 13-14.

[18] (22), p. 12. Also see (45), p. 28 and (32) pp. 18-19.

[19] See (32) pp. 13-14.

c) In many cases, zoning does not even fulfill its traditional objective of stabilizing land use and property values. Development pressures and corruption often result in rezoning not in the community's best interest. This is merely exacerbated by developers' tendency to choose lower-priced land zoned for less intensive use, fully confident in the ability to have the land successfully rezoned.[20]

d) Zoning is frequently constructed so as to limit and discourage even desirable discretionary flexibility.

e) The present system of appeals which relies heavily on the expensive, inefficient, and unstandardized judicial process results in a tremendous waste of private and public time and money.[21]

f) The present system of dealing with nonconforming uses is inadequate and fails to adequately distinguish between innocuous and undesirable cases of such uses.[22]

g) Zoning is all too often used by communities to exclude segments of society they find undesirable or financially burdensome.

D. OTHER REGULATORY DEVICES

Subdivision control is another land-use regulatory device. Under this form of regulation, the municipality demands that certain requirements and standards be met before a landowner can subdivide his land and sell or develop the resulting plots. This sort of zoning, in advance of development, is meant to control and provide standards for developing land.[23] For example, municipalities

---

[20] See (5), pp. 59-60.

[21] See (5), pp. 154-159.

[22] See (53), pp. 100-121 for a detailed discussion of this problem and suggestions.

[23] See (28), pp. 449-484 for a detailed treatment of subdivision regulation.

use an official map of public streets to provide for public streets and insure that existing and officially proposed public streets are not endangered by private construction. After land is designated to be used as a public street and this is officially adopted as part of the official map, a private landowner cannot build on land so designated without prior permission, and if he should subsequently build on this land without permission, he cannot collect any damages for the improvement if the land is then actually condemned for the road.

A more recent land-use regulatory concept, though still a kind of zoning, is the Planned Unit Development (P.U.D.).[24] This has been introduced as a more flexible and desirable alternative to traditional land-use controls when the development of a relatively large tract of land by a single owner is proposed. Since the requirements for the development and the eventual approval is based on the entire development in a case-by-case fashion, there is much more room for flexibility and discretionary control. Much more diversity and mixing of uses within the development is possible than would be possible under normal zoning and subdivision regulations.

Some states permit municipalities to plan for land outside their legal boundaries within certain limits. Some states have gone even further and have extended this extraterritoriality privilege to actual subdivision and official map control.[25] However, as Beverley Pooley points out:

> the usefulness of the device in metropolitan areas is very small; most metropolitan areas consist of one or more municipalities in a densely populated area. In most cases, the municipalities which form a part of such an area are islands - or at best, peninsulas - in a sea of incorporated territory... Therefore,... an adoption or extension of the extraterritorial zoning principle in unlikely to be the key to a solution of metropolitan planning problems.[26]

---

[24] For a detailed exploration of P.U.D.'s, see (40) and (72), the latter especially for legal aspects and model legislation.

[25] See (22), p. 6 and (67), p. 15 for details.

[26] (53), p. 23

IV. OTHER LAND-USE AND DEVELOPMENT GUIDANCE TECHNIQUES

A. PUBLIC LAND ACQUISITION

There are various ways in which government can acquire land: "land can be acquired not only be eminent domain but by excess condemnation, voluntary purchase, tax reversion, mortgage foreclosure, long-term lease, reclamation and gift and bequest."[27] Similarly, there are varied purposes for public land assembly. Among these are several mentioned by Fred Bosselman:[28] to help in the assembly of sites for large-scale developments, to resell or lease with restrictions for public objectives and to protect neighboring public facilities, for use in the future for public facilities, and to prevent immediate development and insure control over future development.

The use of the power of eminent domain, like that of police power, has legal restrictions, although not as severe. The primary restriction is that condemnation of land must be for a public purpose and must be statutorily authorized. However, this legal limit on the use of this power of compulsory purchase has become less restrictive over the years: "Indeed, concepts of public purpose will probably continue to change as the gorwing size and complexity of the modern metropolis requires actions which would not have been considered in the public interest before."[29] Of course, the legal restrictions are not the only problems associated with the exercise of the power of eminent domain. Because of the excessive cost and time loss that results when a landowner contests a condemnation action in court, it is often more expedient to negotiate for purchase or use some other

---

[27] (67), p. 11

[28] (11), p. 41

[29] (21), pp. 355-356

approach. However, the power of eminent domain is sometimes the only feasible approach, such as in the task of land assembly for urban renewal projects.[30] An approach which is between regulation and total fee acquisition is the acquisition of "development rights." In this approach, the government buys the owner's right to use or develop his land in a certain way, but the owner retains possession. This approach was at the heart of the British Town and Country Planning Act of 1947 and is the basis of scenic and conversation easements. A major drawback of this approach is that it is very difficult to make a fair estimate of the value of the development rights which are being condemned.

B. PUBLIC INVESTMENT PROGRAMMING

One very potent, although often underutilized, tool for guiding land development is government policy on public investment programming. A well formulated and coordinated program for the selective provision of public facilities and services can have a great, if not determinative, impact on land development in undeveloped areas. This is particularly true with respect to the placement of major highways and interchanges and the extension policy with respect to sewer lines. However, at present, because of jurisdictional and functional fragmentation of governmental authority and poor planning, few metropolitan areas have been able to afford themselves of this potentially effective land guidance tool.

C. TAXATION POLICY AND INCENTIVES

Another area of great potential for land-use and development guidance is taxation and incentives policy.[31] Of course, policy with respect to real property

---

[30] There is a tremendous literature dealing with urban renewal. For a brief introduction to this concept and redevelopment techniques see (32), pp. 84-87 and (28), pp. 485-519.

[31] See (9), pp. 565-577 for a discussion on the rise of taxation as a land-use control.

taxation is the most direct and forceful possibility. In evaluating policies with respect to property taxes, we will later find it very helpful to consider the real property tax as being made up of two components: a tax on the site and a tax on the buildings and other private improvements on the land. One must also be aware of another important point. The real property tax is a tax on value and not income. Although the tax rate may seem small, it is often equivalent to an "income tax" on the land of as much as 20-35%.[32] If the land being taxed is idle, the effective income tax rate is technically infinite. Such an important economic factor as this most surely would have a strong effect on the actions of private landowners with respect to landholding and development.[38]

The property tax is now used in most areas of the United States. It is used in many different locales with many different results; however, a prime effect is the discouraging of land improvement and building maintenance or modernization where it is most needed. In stagnant or blighted areas where tax assessment usually lags behind the decline in property values, the part of the property tax on buildings and private improvements discourages the needed new development and renovation.

Other taxes also have impacts on land-use and development. Three examples often cited are federal policies with respect to the capital gains treatment of land-based income, allowances on income taxes for certain classes of homeowners, and depreciation allowances for building owners.

The Federal government is also influential by way of its various incentives programs and grants-in-aid programs relating to home ownership, mortgage loans, urban renewal, and public housing, to mention a few.[34] Unfortunately, most of these incentive programs as well as tax policies are not coordinated into a single effective development policy tool.

---

[32] [21], p. 123.

[33] See [59], pp. 43-48 for a detailed evaluation of the effect of taxation on a landowner's ability to hold land.

[34] See [21], pp. 41, 64.

V. PROBLEMS WITH OUR PRESENT "SYSTEM"

A. FRAGMENTATION AND PAROCHIALISM IN USE OF TOOLS

Fred Bosselman characterized the present land guidance system as:

the feudal system under which the entire pattern of land development has been controlled by thousands of individual local governments, each seeking to maximize its tax base and minimize its social problems, and caring less what happens to all the others."[35]

Coke and Gargan convincingly demonstrate that fragmentation of land-use controls occurs in two fundamentally different ways: (1) areal or spatial jurisdictional fragmentation caused by the existence of thousands of relatively independent general purpose local governments and (2) functional fragmentation resulting from the existence of numerous levels of government, special districts and special purpose agencies with very little coordination of activities, even within the same local jurisdiction.[36] Coke and Gargan feel that the latter form of fragmentation is particularly damaging and hasn't received enough consideration. A case in point is the common lack of coordination and cooperation between the planning agency and the agencies which control highway and sewer construction. The result is that if anyone is actually doing any <u>effective</u> planning, it certainly isn't the planning agency. Of course, this is not to be interpreted to mean that "spatial" fragmentation is not a serious problem. Parochialism and self-centered myopia among local governments in their application of land-use and development controls is legend. Recently, exclusionary practices, only one part of the massive and intractible problem, has gained widespread interest.

---

[35] [12], p. 1

[36] See [22], especially pp. 1-2.

[37] [21], p. 6.

B. PROBLEMS IN PLANNING AND ADMINISTRATION

As was mentioned previously with respect to zoning in particular, there has been little serious effort to establish an effective, comprehensive planning process which would guide the establishment of land-use and development policy and set usable criteria for the administration of land guidance techniques. Only recently have federal grant-in-aid programs with comprehensive metropolitan and regional planning requirements started to change this situation. Another major problem is the failure of our present land-use planning and regulatory policies to reflect the complexity and pluralism of the land market and the development process. Marion Clawson observes: "Decision-making in the suburban land conversion process is highly diffused; there are many actors and many processes, complexly interrelated, with numerous feedbacks... No single person, group, or public agency is responsible for the kinds of suburbs that are built. . . ."[38]

Given this situation, governments having the responsibility of guiding land use and development to achieve accepted public goals should take these realities into consideration in the formulation of appropriate guidance policies. A policy which on the surface may appear to have rather obvious consequences might produce rather unexpected and undesirable results. A simple example is that of the locality that desires a shopping center in a particular location. The seemingly appropriate action of zoning this land for a shopping center may achieve just the opposite result by causing the price of the land to exceed that which would be economically workable.[39] In order to improve this situation

---

[38] [21] p. 5 See also pp. 58-140 for an extensive exploration of the suburban land market, the prime movers, and processes involved. [75], pp. 78-82 provides further discussion of this matter.

[39] See [46], pp. 12 - 13 and [14], pp. 62-68.

it is first necessary to improve the available data on land uses and substantially increase our empirical knowledge on the working of the development process and the factors which affect land development.[40]

Finally, among administrative problems of our present land guidance system, are the quite often repeated charges of ineptness, lack of professional staffing in capacities requiring them, serious and repeated instances of conflict of interest, and simple corruption.

C. INABILITY OF PRESENT TECHNIQUES TO MEET PRESENT NEEDS

As was especially true of zoning in particular, the total system of regulatory controls which now constitutes the bulk of our conscious efforts to control land-use and development are not capable of the positive, dynamic guidance required by the challenge of urban growth and related urban problems. These controls, especially as presently administered, are generally inflexible, overdetailed, and static. As F. Stuart Chapin points out, land development is a dynamic process and the sequence of events is critical.[41] The present inability of land-use controls to effectively deal with the timing of development is a critical shortcoming, urban sprawl being the major case in point.[42] "Simply stated, our methods of land-use regulation. . . were woefully inadequate to regulate undeveloped land in the urbanizing frontier in the face of the methods, scale, and speed of post-war growth."[43]

Another major shortcoming of our present system is its failure to sufficiently "take into account" the externalities and interdependencies involved in land use. As Clawson points out: "Externalities and interdependencies in land

---

[40] For a comprehensive survey of land-use information in the United States, see [19]. For recent empiracal work done on factors influencing urban development, see [16], (especially [15]), [17], [23] and [65].

[41] [17], p. 2

[42] See [11] for a discussion of sprawl and suggested preventive actions.

[43] [46], p. 116.

41 [17], p. 2

use affect the use and value of each tract of land more than do actions within each tract. Yet the mechanisms to reduce or eliminate negative externalities and/or to increase positive externalities, are weak or absent."[44] This, of course, is merely part of the more general failure to take into account more fully both total community and metropolitan or regional considerations in the establishing of land-use and development policy and the application of land guidance techniques.

D. INEQUITIES OF THE SYSTEM

Aside from such glaring examples as exclusionary practices and administrative corruption, there are other economic inequities built into our present system of land guidance techniques, which can be publically inefficient as well as privately injurious. A prime source of such inequities is the American "principle" that the owner of land is entitled to all increases in value of that land, even if such increases were prompted entirely by public investments and other actions. We thus find ourselves with a rather one-sided system under which we feel obligated to compensate a landowner for losses in value resulting from direct government regulations, while the losses suffered as a result of indirect public actions (such as building a highway interchange in the next town) and unearned gains resulting from various government actions are ignored. Not only is this "unfair," but the process by which the public hands out bonanzas to a lucky few is harmful to the general welfare because of the increased public cost and undesirable development that often results from the vigorous land speculation accompanying this real estate "jackpot" game.

Another major inequity, which also winds up decreasing general welfare, relates to developers or prospective developers. The complex, costly, and time-consuming character of the process by which a developer obtains the various permits and approvals necessary before construction can begin is well known.

---

[44] [21], p. 6.

Of course, when, on either valid or invalid grounds, a government agency finds a proposal objectionable, the costs of attempting to reach the construction stage are enormous. These costs, which are eventually passed on to the consumer in one form or another, result from an unnecessarily complicated and awkward administrative system, characteristic of most governments and the entire judicial appeal process which is involved. These costs are of course most objectionable when they are used to discourage or prohibit a development for what is actually exclusionary or otherwise invalid purpose.

VI. THE NEED FOR CHANGE

A. FACTORS PROMOTING THE STATUS QUO

One fact must be faced in exploring the need for change in our land guidance system. That is the resistance to change that will be forthcoming from those who are satisfied with the status quo. Of course there are the ideological barriers to change, the exclusionary and parochial interests less concerned with the more comprehensive welfare than their own self-centered desires, and the ubiquitous speculator, who relies on the present system to make a fast buck. However, there are also less glaring sources of such resistance. As Clawson points out, even our "urbanized regions" have sufficient land to absorb as much as two decades more of population growth within their present boundaries.[45] It is of paramount importance to realize that much of the population, especially the powerful and vocal segments, are satisfied with the present suburban growth process. Many suburbanites find their present existence desirable, despite the planners' criticisms of sprawl, monotony, and exclusionary practices. They enjoy the temporary insulation from city problems their traditional controls

---

[45] [21], p. 8

afford them. The problems that do exist in the development process don't appear to be critical[46] and so many people have no great incentive to push for change. This widespread apathy toward the need for improvement on the part of the population is only worsened by the wide variation in development pressures throughout the country.[47] In many areas where development pressures are not strong, the fear of radical, new governmental techniques dominate any weak concern over poor development control.

Despite this public resistance and apathy, there is no doubt about the need for change. There are serious land-use and development problems ranging from costly and inefficient spread of sprawl to the more socially oriented problems of racial exclusion and lack of sufficient low-income housing. We must improve our land guidance system if we are to solve these pressing problems. The need for sweeping changes has been pushed home by a recent rash of major reports recommending with a remarkable degree of unanimity a vast array of actions and new approaches, some of them very radical by all but very recent standards.[48] This, as well as the recent federal legislation that it has spawned, provides very strong support to the cries for reform which planners have been uttering for years.

B. HOW TO PROCEED?

The need to reform is clear. The question remains as to how we can best improve the system. In considering the various alternatives and mixes of approaches, it is illuminating to conceive of the space of possible actions as being defined by five "axes of variation", so to speak, which define dimensions of

---

[46] See [21], pp. 3-4.

[47] For example see [27], pp. ii-iii, 22

[48] See [32] for an excellent summary and analysis of five of these major reports which include [43], [1], and [2].

variability. Mentioned by Dennis O'Harrow,[49] these axes may be expressed through the extremes of the dimension of variability they represent:

1. Government level: Federal → local
2. Incentives → Prohibition
3. Regulation → Acquisition
4. Development agent: Public → Private
5. Complete control → Free Market

Each policy and technique we consider would be a point in an "action space" defined by the appropriate position along each of these axes.

Another important point should be made concerning proposed improvements. Significant differences exist among the tasks of guiding land-use and development in developed areas, undeveloped areas, and in developed areas that are to undergo redevelopment. It is essential that we appreciate this three-way distinction and adopt separate policies and techniques for each of these three different tasks. As we have pointed out above, we cannot expect zoning, a device intended to regulate land-use in developed areas, to be an effective tool in guiding development[50] unless it is integrated with other planning policy.

Finally, any search for new and improved approaches to land guidance must consider the matters of effectiveness and economic, legal and political feasibility. Of all these, the one that will ultimately prove least consequential, despite previous and present difficulties, is the question of legal feasibility. Fred Bosselman is not alone in observing: "The courts have adapted constitutional standards to meet the needs of a changing society. The increasingly grave problems caused by uncontrolled development on the fringes of metropolitan areas create both the need and the constitutional justification for new techniques."[51]

---

[49] [48], pp. 140-141.

[50] See [32], pp. 13-14; [46] pp. 23-25; and [67], p. 69.

[51] [11], p. 69.

VII. PROPOSALS FOR IMPROVING THE PRESENT SYSTEM

Whatever the actual controls and techniques employed, a professional and well-coordinated planning process should be the foundation of any land guidance system. There is no doubt that we must improve the quality of and strengthen the role played by planning in such a system if we are to achieve satisfactory results. However, we must be careful not to interpret this as a mandate for the use of the most large-scale, comprehensive, long-range planning possible. There is a substantial need for local level planning as well as higher-level planning efforts, and truly comprehensive long-range planning on even the metropolitan scale is either impossible or too vague to be useful. The most important objective is the achievement of the proper mix and coordination of short-, mid-, and long-range planning at all levels of government. Even in theory this is a big order, and the political realities of pluralism in this country are not to be ignored in trying to achieve this goal.

A. ADMINISTRATIVE RECOMMENDATIONS

Recommendations for improvements in administration of land-use and development regulations have been numerous and are among the most traditional suggestions forwarded by critics of the present implementation of land-use controls.[52]

One major administrative suggestion is that the state set standards of size and administrative and planning competence on units of local government that could exercise land-use and development controls.[53] Governmental units not meeting these standards would forfeit these powers to a higher level of government. Along with such recommendations come suggestions that states set forth

---

[52] For some examples of recent suggestions in this area see [32], pp. 25-39.

[53] See [32], p. 16 and [21], pp. 352-353 for examples of such suggestions.

in their enabling legislation more explicit standards and criteria for administrative procedure and the grounds upon which discretionary decisions should be made. Hopefully, such action would permit greater uniformity and fairness in the application of controls, make local procedures more efficient, and expedite the appeals process in both the courts and superior review agencies.

Prominent among administrative recommendations is the suggestion to replace the various zoning, building codes, planning, and appeals agencies and officials in a municipality by a single administrative agency.[54] Such an agency, which Dennis O'Harrow has referred to as a "development control agency,"[55] would be responsible to either the local legislature or chief executive and would perform all functions associated with the planning and regulation of land-use and development, except for "adopting and amending plans and ordinances and adopting and making major amendments to the zoning map, which will remain the prerogative of the legislature."[56] A complementary measure would replace the various ordinances and rules dealing with land-use and development with a single "development ordinance."[57] These measures would insure a simpler, more consistent, and efficient regulatory process. The process might be made even quicker and less costly for both the community and prospective applicant through an extension of the "comprehensive permit process" utilized in a limited fashion under the Massachusetts Zoning Appeals Law.[58] The American Law Institute is presently completing a Model Land Development Code, which would serve as a comprehensive model enabling act for the states to use, and which incorporates

---

[54] See [32], pp. 16-17.

[55] [48], p. 153.

[56] [32], p. 17

[57] See [67], p. 20.

[58] See [12], pp. 170-172.

many suggested administrative improvements within its provisions.[59]

Naturally, administrative recommendations are not limited only to changes within a municipality. Many proposals have been made to establish higher level (usually state) agencies to hear appeals from local regulatory agencies. Such special-purpose, quasi-judicial agencies would relieve the court system of a great burden for which it is not well suited anyway. Appeals from such an agency could go directly to the state's highest court.[60] Such an agency could also serve to develop and police uniform administrative standards for local governments, oversee the regulation and development of land and projects of statewide concern, and even exercise determinative authority over local controls in certain instances.[61] A state appeals process has been established by the Massachusetts Zoning Appeals Law of 1969 for the limited purpose of providing for needed low and moderate-income housing by overruling exclusionary zoning decisions on the local level.[62]

Marion Clawson made a novel suggestion to decrease the monetary temptations for corruption in zoning and recoup for government some of the unearned value landowners acquire through government action.[63] He proposed that zoning and rezoning be sold through public auction to the highest bidder among persons who owned or had an option on land in an area deemed appropriate for the new use. So far, this proposal does not seem to have been taken very seriously.

---

[59] See [2] for a first tentative draft of the model act.

[60] See [21], pp. 345-346.

[61] See [32], pp. 17-18 and [5], pp. 166-167 for examples of such proposals.

[62] See [12], pp. 164-180 for a discussion of this law and an evaluation of its success so far.

[63] See [20], for the details of this proposal.

B. PROPOSALS FOR REALLOCATING LAND GUIDANCE FUNCTIONS

Many proposals for the improvement of our land-use and development guidance system have involved the creation of more rational governmental units and the vesting of powers and functions traditionally the domain of local governments in such units and pre-existing, higher levels of government. Although such proposals have met with much resistance and have struck directly at the heart of the difficult problem of irreconcilable differences of interest among localities, they have been among the more favored approaches forwarded by planners.

Among the least controversial proposals in this area is the formation of so-called Councils of Governments (C.O.G.'s) for metropolitan areas and other natural planning regions. A C.O.G. is "a voluntary association of governments designed to provide an area-wide mechanism for key officials to study, discuss, and determine the means - - cooperative, if possible - - of solving common problems."[64] C.O.G.'s do not generate the opposition that more radical proposals (such as the creation of a metropolitan government) encounter, because they are usually only voluntary associations with an advisory capacity, meant to foster cooperation, but not involving any real loss of autonomy on the part of member government units.[65] However, one must trade the relative lack of opposition to such a proposal against the typical impotence of such governmental bodies in effecting meaningful improvements in area-wide planning and land-use and development guidance. The C.O.G.'s lack of any real power, the zeal with which member governments guard their prerogatives, and the usual one-vote-per-government unit rule all tend to severely limit the efficacy of this approach.

---

[64] [22], p. 32. Quoted from: John C. Bollens and Henry J. Schmandt, The Metropolis, Harper and Row, New York, 1965, p. 392.

[65] See [22], pp. 32-38 for a discussion of C.O.G.'s and an illustrative example of the Metropolitan Council of the Twin Cities area, which, the reader is cautioned, has more power than the typical C.O.G.

A much more radical and demanding proposal calls for the creation of metropolitan governments to assume control over certain governmental functions in a metropolitan area (which could include hundreds of local government units). It would be unnecessary and undesirable for such a government to assume responsibility for all governmental functions. If a metropolitan government were to be established it would concern itself only with such area-wide problems as the location, timing, and control of major commercial and industrial centers and major public facilities such as utility, sewer, and transportation systems.[66] The provision for the necessary amount and variety of housing for the region would be an example of another responsibility preferably entrusted to a metropolitan government.

There are few examples of what could really be considered metropolitan governments in this country. Nashville, Tennessee, and Dade County, Florida, are two oft-mentioned examples.[67] A successful use of metropolitan government could certainly help remedy many of the problems associated with the geographic fragmentation of land-use and development controls. However, the political feasibility of creating such governments in most metropolitan areas is small and the intractable problem of fundamentally conflicting groups of interests within a metropolitan area do not disappear under the magic wand of metropolitan government.[68] An often discussed variant of the metropolitan government proposal is the utilization of the county in a planning and even governing capacity, far more intensive and urban-oriented than this basically rural unit of government has traditionally assumed.[69] The use of the pre-existing county

---

[66] See [67], p. 14.

[67] See [22], pp. 39-46 for a discussion of the use of metropolitan government and an evaluation of the experience in Nashville, Tennessee.

[68] See [27], p. 28 and [22], p. 20.

[69] See [22], pp. 52-57 for a more thorough discussion of this approach and the illustrative example of Marion County, Indiana.

units has both advantages and disadvantages when compared to the establishment of a new metropolitan unit, which must be evaluated on a case-by-case basis.

Recently, there has been an upsurge in suggestions calling for increased state planning, review, and even selective determinative control in the field of land guidance. States do have the ultimate authority in these matters and are generally large enough to insure a reduction in the parochialism characteristic of most lower units of government. In recent years, several states have enacted legislation placing varying degrees of authority in a state agency to regulate or review local regulation of land-use and development, usually limited to areas of specific statewide concern. Much of this legislation is so new that there has been insufficient time to evaluate conclusively its effectiveness or impact on the likelihood of further state involvement. However, a brief mention of some of the major examples of such state action would demonstrate the scope and general direction of these pioneering efforts.[70]

One of the first and probably deepest intrusions of a state into what is usually the traditional function of a local government occurred in Hawaii in 1961 with the passage by that state of the Land Use Law:

> The Land Use Law gave state agencies a degree of control over the use of the state's land resources that was far in excess of that enjoyed by other states. It created a state Land Use Commission and directed it to divide the entire state into four districts: Conservation, agriculture, rural, and urban. The Land Use Law authorized land in the urban district to be used for whatever purpose is permitted under the local zoning regulations. Lands in agricultural and rural districts were to be used only in compliance with regulations of the state Land Use Commission, and lands in the conservation district were to comply with regulations of the State Department of Land and Natural Resources.[71]

---

[70] [12] provides a rather in-depth description and evaluation of most of the recent examples of major state inroads in the land-use and development control domain.

[71] [12], p. 5.

However, as is pointed out, there were atypical factors involved in the case of Hawaii and it is generally felt that any general state-wide control of land-use elsewhere in the United States is unlikely at this time.[72]

The Maine Site Location Law has given the Maine Environmental Improvement Commission the authority to regulate the location of major industrial and commercial activities and impose conditions upon them, or even totally prohibit them under certain circumstances, to insure the protection of the environment.[73] Despite a small budget and staff, the Commission has proved to be rather effective. However, the law's concentration on environmental considerations is likely to be a limiting factor on the ability of the Commission to exercise regulatory authority in a more general context in the future.

In 1970, the Vermont Legislature passed the Environmental Control Law in response to threats on the environment from rapid unregulated development, promoted by skiing and second homes.[74] The law gives the State Environmental Board two areas of responsibility:

a) "a judgment function exercised in issuing development and subdivision permits through seven district commissions."[75]

b) The preparation and adoption of a statewide comprehensive land use plan in three stages.

Although the Board does not have comprehensive powers, it has been rather effective, especially in the area of inter-agency coordination.

---

[72] See [12], pp. 5-6 for these factors and pp. 7-34 for a discussion of the operation and success of this unique system.

[73] See [12], pp. 187-204 for details and an evaluation of this program.

[74] See [12], pp. 54-107 for a detailed exploration of the law, its workings, and effect.

[75] [12], pp. 56-57.

Quite a bit of other legislation extending state activity in the land guidance process, too extensive and varied to adequately treat here, has recently been enacted. Most of this legislation can be classified broadly as dealing with "critical areas," land-use study commissions, or wetlands and shoreland.[76]

Extension of state involvement in land-use and development planning and control, especially in areas of undeniable state-wide concern, has great potential. However, the states have shown a great aversion to taking over what traditionally have been local powers and using them in the innovative manner called for. Should this aversion be overcome, which pending federal legislation is intended in part to accomplish, it will be necessary to guard against extremes in the other direction, in which the state involves itself in detailed matters of only local concern.

C. NEW OR UNDERUTILIZED GUIDANCE TECHNIQUES

An unorthodox and innovative approach to the rational timing of development on the local level has been tried through an extension and modification of the usual land-use and development controls. In 1969, the town of Ramapo (population of approximately 76,000) in Rockland County, New York attempted to meet the problem of explosive urban development.[77] The zoning ordinance of the town was amended so as to require individuals wishing to develop their land for residential purposes to obtain a "Residential Development Use" special permit. The permit would be granted only if the land parcel in question met certain minimum standards with respect to the availability of certain facilities and public services on or near the parcel.[78]

---

[76] See [12] for detailed examples of other such state programs.

[77] For a more detailed treatment of the "Ramapo story," see [64], pp. 1-2 and [50].

[78] This rather unorthodox zoning technique has been upheld by the State's highest court in the case of Golden v. Planning oard of Town of Ramapo, 30 N.Y. 2d. 359, 285 N.E. 2d 291 (1972).

By utilizing a "development point" system to "score" a parcel with respect to available facilities and services, such as sewerage, drainage, roads, schools, fire protection and recreational facilities, a fairly objective criterion for granting special permits has been established. The parcel in question must have at least 15 such development points before a permit can be issued. The town has not introduced this concept as a trick to prevent further development, but merely to insure that the town can adequately accommodate the growth. In meeting its acknowledged obligation to provide for continued, although controlled, urban growth, the town has proceeded with an overall development plan for the provision of the expanded public facilities and services that will be required by further development. This plan is in turn used to guide the preparation of an 18-year capital improvement program and, finally, a 6-year operating capital budget.

Although this novel approach to timed development appears to have great promise, one must be cautious in ones expectations. The sophistication and relative absence of parochial self-interest which characterize this town's intentions and planning operation were not only critical factors in the court's decision to permit this unorthodox approach, but are also likely to be determinative of the success the town may achieve in the use of this approach. It is unlikely that more than a small percentage of other local governments across the country are similarly capable and motivated. Still, this new approach is worthy of consideration, especially if it could be implemented through agencies at higher levels of government.

Earlier in the paper, the concept of "performance standard zoning" was mentioned as a more flexible variant of traditional use district zoning. It is possible to expand on this idea to create techniques which take environmental factors into account and can better guide urban development. A prime example of such an approach, described and advocated by Jacob Kaminsky, is based on the concept of the Environmental Characteristic Type or E.C.T.[79] Under this approach, the E.C.T. would replace the use type as the basis of land-use controls and would be used for the classification by development characteristics of planning sub-areas. Areas would receive E.C.T. classifications that would depend on:

a) the size of the area,

b) physiographic characteristics,

c) community-activity range, and

d) the ability of planners to envision development patterns.

Land would be categorized into development types based on such area characteristics as building density, ratio of man-made to natural surface in the outdoors, and the type and intensity of activities currently prevailing. As was done in the Suggested General Development Plan for the Baltimore region,[80] land would be classified into eight development types (instead of the usual use-types): open space, rural, conservation, sub-urban, moderate-urban, urban, town centers, and industrial. Within each type, any use conforming with the nature and goals of that type would be permitted. Kaminsky also discusses various sets of "environmental" ratios and standards, such as the F.H.A. Land Use Intensity scale ratios, which could be employed in a more flexible and effective land-use and development regulatory process.[81]

One of the most widely discussed techniques of land-use and development regulation involves an ingenious marriage of the police power to eminent domain. Labelled compensative or compensable regulations or compensatory payments, this concept was first introduced by Jan Krasnowiecki and Ann L. Strong.[82] Proposed

---

[79] See [36] for a thorough treatment of this approach.

[80] See [36], pp. 10-11.

[81] See [36], pp. 25-37.

[82] See [37]

as a method for preserving open space, the form originally suggested by Jan Krasnowiecki and Anne Louise Strong is one in which the government agrees to compensate the present and future owners of the property, until the full assessed value of the unregulated property has been cummulatively exhausted, for the difference between actual sale price and this guaranteed assessed value.[83] It has been suggested as a more widely applicable tool in the attempt to guide land-use and development. "Compensatory payments are a means of legitimizing regulations which would otherwise be unconstitutional."[84] Because of the legal restrictions on the use of police power, some regulations imposed on the use of land which were deemed desirable have been voided by the courts, usually on the grounds that they were construed to be "unreasonable" or an effective taking of property without compensation. Often, just the expectation of such a decision has prompted governments to reject as infeasible certain regulations they had been considering. By guaranteeing the owner just compensation for any damages he has suffered as a result of what would otherwise be an invalid use of police power, a government using this technique can often be confident that the desired regulation will remain in effect.

Several variations of this technique, each with its own advantages and disadvantages, have been proposed. David Heeter discusses two possible forms:[85]

a) Damages equivalent to the difference between the value of the land with and without the regulation are paid to the landowner in some manner. Usually, it is suggested that the compensation reflect only the "invalid portion" of the regulation and be made only when the owner demonstrates his intent to develop in violation of the pertinent regulation.

---

[83] [37], pp. 91-92, 95-96.

[84] [32], p. 11.

[85] [32], pp. 11, 61-62.

b) The landowner would be forced to choose between uncompensated regulation and sale of the total fee interest in the land at unregulated value through suit in an "inverse condemnation" proceeding in which the government would be forced to purchase the land. This variant is often preferred to the first because it would discourage all those not seriously damaged by the regulation from seeking compensation and would eliminate the very difficult task of trying to justly determine the actual loss in value resulting from the regulation.

In all cases, compensable regulation offers the advantages of uninterrupted regulation without the costly need to condemn all property that cannot be adequately controlled through the exercise of police power alone. However, compensable regulations do have their drawbacks. At the present time, the constitutionality of such a technique is much in question, although the present trend in legal decisions does seem encouraging.[86] In addition to administrative difficulties, one must also be aware that in some cases there are other land guidance techniques which can achieve the same results more cheaply and effectively. Still, as one of several tools in a total land-use and development control system, compensable regulations promise to have very useful applications.

In its report, Urban and Rural America: <u>Policies for Future Growth</u>, the Advisory Commission on Intergovernmental Relations suggests the use of a technique it refers to as "reserved lands," which can serve three basic purposes in the development process:[87]

1) Allow for recreational and other open space needs
2) Allow for flexibility in the development process to provide for possible changes and unforseen land needs of various uses.
3) Provide greenbelts to serve as perimeters of communities or towns, affording a sense of self-containment.

---

[86] For example, refer to [11], p. 27.

[87] See [1], pp. 117-118 or [32], pp. 72-73, 62-63, 12.

The term "reserved lands" actually refers to four techniques which can be used singly or in combination to achieve the above goals:

1) "Reserved development unit" (RDU) is a planned development which has tentative approval for certain uses, but has no detailed plan at the time it is initially formed. Development would normally occur in other parts of the community before the start of construction in the RDU, the timing and details of which would depend on market conditions and capital availability.
2) "Holding zones" are zones in which all uses except open space, agriculture, forestry, and rural residential are prohibited for a certain time period or indefinitely. Compensation would be required in order to exercise such control. These zones could be used by local government if it felt that development was inappropriate at that time, adequate urban facilities for the area could not be provided at the time, or if the eventual use of the land had not yet been decided upon and possibly inappropriate uses were to be prevented.
3) "Expansion sites" are oversized parcels on which certain types of centralized uses would be located. The parcels would be oversized to provide room for expansion of the activity expected with community growth. This technique could prevent the extreme cost associated with expansion of facilities in developed areas and the alternative possibility of scatteration of activities that may best be centralized. Civic centers and shopping centers are two such activities.
4) Public policy with respect to "land bank withdrawals" is a fourth technique that could be used to guide development. Land kept in a land bank would be developed for urban use only when deemed advisable. This banking of land could be used to prevent sprawl and promote an orderly and economical development process.

A hybrid classification that has recently been added to the growing list of suggested development techniques is the "Planned Development Zone."[88] In such a zone only three types of land-use would be permitted:

a) singly-family homes on lots of at least ten acres.

b) agriculture

c) planned developments. These would be allowed by special permit only, and the development would have to be rather sizeable and would have to adhere to standards set down by the municipality (similar to P.U.D. requirements).

Such zones would have to be renewed at least once every five years to determine if the classification was still appropriate for all property in the district. Their use would be in areas that are as yet undeveloped but in which the onset of urban development is imminent. Once development progresses past a certain point, the classification would be replaced by an appropriate normal zoning classification.

D. ECONOMIC AND FISCAL GUIDANCE TOOLS

It is ironic that some o the potentially most potent tools available to government to influence and guide land-use and development are not new proposals at all, but have always been part of the traditional government function at all levels of government. These are of course the governments' powers to spend money and tax. What is new in the many proposals made relating to the use of these powers is the call for a more vigorous, comprehensive, and, above all, coordinated application of new and old forms of these powers in such a way that specific public land-use and development goals are promoted at the same time that needed public services are being supplied and revenue is being raised. As has been mentioned above, the manner in which government spends its money has

---

[88] See [11], pp. 11-21.

a substantial, almost determinative, impact on the path of development. This power could be used to effectively promote public development goals. The numerous suggestions for a more intelligent use of the public capital programming function in guiding land-use and development for the general public good are all the more convincing when one realizes the absence of many of the legal and political restrictions, which confront other approaches, in the exercise of this government power.[89]

The use of government taxation and incentive policies as a tool to guide land-use and development is another area of great potential, with many opprotunities for improvement. Naturally, most recommendations concern themselves with the local real property tax and property value assessment policy and procedure. The majority of these recommendations are founded on the assertions that present assessment and property tax polices:

1) are prime factors in the perpetuation of fiscal zoning and other related "parochialisms";
2) are basically inequitable and do not discourage excessive land speculation, as they could be made to do; and
3) discourage development and desirable maintenance and improvement efforts on the part of landowners.

As was suggested in the discussion of "fiscal zoning", the reliance of local government on property taxes has been a major factor in sustaining exclusionary practices.[90] Proposals for the correction of this difficulty usually involve the introduction of a statewide or regional tax and the provision of certain services on a similar basis, instead of the usual community-by-community basis.

---

[89] See [6], pp. 186-197 for an examination of the need and possibilities and some major working examples.

[90] See [43], pp. 211-213.

In some cases, it is recommended that the property tax be partially or totally replaced by other taxes, although such a drastic change is unlikely in the near future.

From other sources comes the suggestion that the private improvements part of the property tax (which taxes both land and improvements as one) be removed from the property tax altogether or at least be taxed to a lesser extent than the site itself.[91] Such recommendations are intended to deal with the second and third undesirable effects of present policies listed above. By assessing at true market value and placing a high tax on the land and public improvements (a land value or site value tax), the government would be able to recoup much of the unearned gains landowners enjoy as a result of public investment. However, equity is not the only aim of such a move. By drastically cutting down on the possibility of quick, unearned profits from land ownership, speculation and associated ills of urban sprawl and regulatory corruption would be correspondingly lessened. In addition, by removing or substantially reducing the tax on private improvements (i.e. buildings), a great discouragement to private development and improvement is eliminated. On the positive side, a high, market-value land tax would encourage the development of idle land when the opportunities for development are ripe.

One major drawback of an exclusive land or site value tax is the possibility that even a 100% tax rate on the land value alone would not raise as much revenue as the present property tax.[92] In such a case, the best that could be done would be the substitution of a levy which taxes land at a much higher rate than other real property. This so-called "graded tax" has been used in Pittsburgh, Hawaii, and western Canada, but has the administrative disadvantage

---

[91] See for example: [60], pp. 23-26 and [45], pp. 41-42. Even back in 1939, this was strongly recommended: [44], p. 216.

[92] See [45], p. 42

of requiring two separate value assessments - one on the site and another on the private improvements.

Where the annual land valuation or graded tax is found to be inappropriate or infeasible, an alternative recommendation is an additional tax on land value increments.[93] The rationale for such a tax rests mainly on the argument stating the inequity of permitting "unearned" land income. Related both to public facilities programming and local tax policies is the suggestion that residents of a newly developed area be made to pay for the costs of providing the necessary public facilities for the community.[94] This policy would make more explicit to new residents the cost of their new homes and might tend to discourage the publicly expensive practice of leapfrogging in which many developers participate in the search for cheaper land.

Most other suggestions relating to taxation and incentives jump directly from the local to the federal level. One common recommendation is that federal tax depreciation allowances on buildings be limited to 100% of the initial value for all owners of the property, instead of the present policy by which successive owners can each "start anew" after a certain time period.[95] Such a change would tend not to discourage improvement and renewal as do the present rules. Another suggestion is that income from increases in land value would not be permitted to be treated as capital gains for federal income tax purposes.[96] This would have an effect similar to the suggestion that a land value increment tax be imposed.

---

[93] See [45], p. 43.

[94] See [21], pp. 160-161, 348.

[95] See [21], pp. 350-351.

[96] See [21], p. 350.

Of course, we must not forget the well-known tangled maze of federal grants-in-aid and incentive programs.[97] By doing a better job of rationalizing and coordinating these numerous programs, the Federal government could have an even greater impact on the quality of land-use planning and the general path of development in this country.[98] And, finally, we mention one further recommendation, which, although not relating to incentives in the usual sense of the word, is nevertheless appropriately considered here. Much of the difficulties in the area of planning for and guiding land-use and development stem from a lack of appropriate information and a generally poor public appreciation of the problems urban America faces and the measures required to alleviate these problems. Although it would by no means eliminate all of our thorny problems, the adoption of a more accurate and complete land-use and development data system for our planners and decision-makers and an improved program of "civic education" is clearly indicated.[99]

VIII. NEW PROPOSALS

In this section, some of the recent proposals for major changes in the present approach to the planning and control of land-use and development are briefly explored. Because of the boldness of these proposals and the radically different government role they envision, it was decided best to treat them in a separate section. These proposals involve greater federal participation and/or the use of large-scale advance public acquisition and disposal of land as the

---

[97] See [9], pp. 563-564 and [21], p. 374 for brief discussion of these federal incentives and their application in land-use and development matters.

[98] See below for further discussion of the potential federal role.

[99] See [75], pp. 90-91 and [13], pp. 119-120, respectively, for discussions of these points.

foundation of a more positive role of government in the development process and, as such, represented major shifts in the approach to land guidance. There are substantial grounds for seriously considering these proposals. Premier among these is the very near unanimous call for the implementation of such sweeping changes by various major commissions and study efforts. The five well-known reports summarized and evaluated by David Heeter[100] are the primary examples.

Although Heeter finds variations in the recommendations and emphases of these reports, he concludes that "the reports are largely in agreement in both their findings and their recommendations" and they all "recommend a sweeping restructuring of the entire land-use regulatory system."[101] The reports recommend an expanded role for both the Federal government (including the establishment of a national urban growth policy) and the use of advance public acquisition of land. The advisory Commission on Intergovernmental Relations states in their report: "On balance...the Commission concludes...that there is a specific need for immediate establishment of a national policy for guiding the location and character of future urbanization, involving Federal, State, and local governments in collaboration with the private sector of the national economy."[102] Speaking of the five reports he reviewed, Heeter stated: "They recommended that local governments be empowered to acquire land through negotiation or condemnation and to sell or lease it subject to conditions which will promote a wide variety of public objectives."[103]

---

[100] The five reports are: [32]; [43]; [2]; Report of the Task Force on Housing and Urban Development, The Canadian Federal Task Force on Housing and Urban Development, 1969; and New Directions in Connecticut Planning Legislaction, American Society of Planning Officials, 1967. They are summarized in [32], as noted previously.

[101] [32], p. 2.

[102] [1], p. 129

[103] [32], pp. 11-12.

A. ACTIVITY AND PROMISE AT THE FEDERAL LEVEL

It is to be noted that these reports are not merely representative of the hopes of a group of idealistic but unrealistic thinkers. Recent action and statements by both the executive and legislative branches of the Federal government suggest that these recommendations are being taken very seriously. "In his January, 1970, State of the Union message, President Nixon proposed that the United States 'develop a national growth policy' and that the 'Federal government... assist in the building of new cities'."[104] In Environmental Quality, (Washington, D.C., 1970, p. xiii), he further stated: "I believe we must work toward development of a National Land Use Policy."[105]

In the legislative branch there has been even more encouragement. The Housing and Urban Development Act of 1970 "had as a stated public objective 'the development of a national urban growth policy'."[106] Senator Henry M. Jackson, Chairman of the Committee on Interior and Insular Affairs, has well expressed the Congressional concern in these matters:

> The rapid and continued growth of the Nation's population, expanding urban development, proliferating transportation systems, large-scale industrial and economic growth, conflicts in patterns of land use, fragmentation of governmental entities exercising land use planning powers, and the increased size, scale and impact of private actions, have today created what many Americans perceive to be a national land use crisis.
>
> That such a perception is widely shared in the Federal government - in both the legislative and executive branches - is demonstrated by the more than two hundred land use policy measures before Congress.[107]

Two of the most important bills pertinent to our present discussion, now being considered by Congress, are the President's proposed "National Land Use Policy

---

[104] [66], p. xxi

[105] [66], p. xxi

[106] [8], p. 231. This article provides a comprehensive review of the federal legislation of 1971 dealing with the development of a national urban growth policy.

[107] [70], p. 1.

Act of 1971" (S. 992, H.R. 4332 in the 92nd Congress)[108] and the "Land Use Policy and Planning Assistance Act of 1972" (S. 632).[109] The latter bill has been amended and just recently has been passed by the Senate by the Committee on Interior and Insular Affairs.[110] If enacted, it would "provide Federal technical assistance and a granr-in-aid program to the individual states to assist them in developing and improving their capacity for land use planning and management. The Act also provides important new authority designed to improve coordination between the planning efforts of the Federal and state governments."[111] a new Office of Land Use Policy Administration would be created under the Secretary of the Interior to aid in the administration of the act and $100 million a year would be authorized for eight years for the grant-in-aid program.

B. ADVANCE LARGE-SCALE PUBLIC ACQUISITION AND DISPOSAL OF LAND

At least as far back as 1939, governmental reports have discussed and demonstrated the great benefits of governmental ownership of land in excess of current needs.[112] The concept of using large-scale advance public acquisition and disposal of land as a major method to guide land-use and development has received rather strong new support from the five reports mentioned above, as well as many other sources. Despite the support for such an approach, a 1966

---

[108] [12], Appendex: pp. 2-13.

[109] See [71].

[110] See [71]. Also refer to [70] for a very thorough review and comparison of these bills and other recent proposed and enacted legislation in this area.

[111] [71], p. 21

[112] [44], pp. 326-328. See 444, pp. 11-16 for an excellent annotated bibliography of the selected American and Canadian works dealing with public land acquisition and dating from the 1930's.

questionnaire survey, yielding responses from 144 cities, suggested that less than 30% of cities with populations over 50,000 in the United States carry on any type of advance land acquisition.[113] In those cities which have such programs, they are limited, and "no large-scale plans for influencing orderly land development were reported."[114]

In addition to the five major reports mentioned previously, the idea of using advance public land acquisition as a land guidance technique has been discussed and supported by many others.[115] In a widely quoted article, John Reps summarized his recommendation for improving our land-use and development system as follows:

> I propose that land at the urban fringe which is to be developed for urban uses should be acquired by a public agency. Acquisition, in fact, should run well ahead of anticipated need and include the purchase or condemnation of idle or agricultural land well beyond the present urban limits...
>
> Land scheduled for early development should be designed in detail, conforming to a general, comprehensive, and long-range metropolitan growth plan. The public agency, directly or indirectly, should install all street and utility improvements and should identify and retain all sites needed for such public facilities as parks, schools, and other neighborhood and community needs. The remaining land then be disposed of to private builders by sale or lease, the aggregate price to reflect full acquisition and improvement costs but not profit. The terms of the sale or lease should include adequate safeguards to insure development only in conformity to the detailed plans prepared for the area.[116]

Despite the extensive change called for by such a proposal, it is often offered as the only means for achieving some of the major goals this country has in the land development and housing fields. Marion Clawson comments: "If we wish

---

[113] [62], pp. 4, 16-17. This is an excellent source of information on the general area of advance public land acquisition.

[114] [62], p. 4.

[115] For example, see [62]; [57]; and [6], pp. 164-185.

[116] [57], pp. 2-3.

really to cure the major deficiencies in the suburban land conversion process, public land purchase must be initiated and carried out on an adequate scale."[117] The lack of adequate low-and moderate-income housing is the deficiency cited most often.

A rather long list of advantages and benefits associated with the adoption of this tool of land guidance has been amassed.[118] Among the more prominent are:

a) It would decrease considerably the cost and difficulty government presently incurs in obtaining sites for public facilities.

b) It would go a long way to eliminate excessive speculation and unearned value increments on the part of the landowners.

c) Better planning and assurance of the implementation of these plans, unimpeded by previous development, private landowners, and an unrestricted development process, would afford more efficient and desirable development and lower-cost housing for the public. Urban sprawl could be easily eliminated.

d) Achieving goals, such as the provision of adequate new low-income housing, which is practically impossible with our present system, would be within reach.

e) The private developer and builder, especially the smaller one, would also be greatly aided by this new approach to development control. By curbing land speculation, performing most of the planning, providing most of the land improvements necessary prior to construction, and substituting a simple restricted sale or lease of land for the long

---

117 [21], p. 355. Also see pp. 9, 361.

118 See for example: [57], pp. 4-5 and [62], pp. 4-6, 100-102.

and complicated regulatory obstacle course now dominant, the government could greatly reduce the costs and risks that developers presently experience.

Regardless od these advantages, we must establish whether or not the use of large scale advance public land acquisition as a development control is legally and politically feasible in the United States. As is becoming more and more the case, the legal prospects are not at all grim. Reps observes: "A persuasive case can be presented that an ample public purpose exists in the need to provide well-planned additions to urban areas through large-scale public land acquisition and management... I suggest, therefore, that while many fascinating and complex legal problems would be involved, we need not fear any fundamental reasons in law standing in the way of such a program."[119] More concrete evidence of the legal feasibility of such an approach is mentioned by Ann Louise Strong: "Public power to condemn land for new development has been affirmed. The U.S. Supreme Court, by denying certiorari in the Rosso* case, upheld the right of a public body... to condemn undeveloped land for resale or lease for development in accordance with public plans."[120]

---

[119] [57], p. 6. See also [62], p. 18-22 for a discussion of the legal questions involved.

[120] [66], p. xxii. *Commonwealth of Puerto Rico V. Jorge I. Rosso and wife Carmen Descartes, opinion 67-172, El Tribunal Supremo de Puerto Rico, December 7, 1967; cert. denied, 393 U.S. 14 (Sup. Ct. 1969)

The question of political feasibility is of course a different matter. Here, as in most instances concerning land-use and development controls, the political opposition is likely to be much more effective than possible legal considerations in blocking the use of this approach. The greatest opposition is likely to come from rural and suburban areas for the obvious reasons. However, one source of concentrated opposition, that many would feel to be most natural, is likely not to arise. As mentioned above, private developers, who might be expected to vehemently oppose such a "socialistic" incursion into the private land market, will benefit greatly from such an approach and are thus unlikely to oppose it. James Rouse, himself a well-known private developer, testified before Congress in support of this form of government control of development and stated: "This is no threat to the private homebuilding industry, but an asset."[121] Another major consideration is that of financial feasibility. Local and state governments are already at or past their financial breaking-points and a large-scale program of public land acquisition would require vast sums of money. Thus, such programs would have to be limited at first. Federal support such as that available under the Housing Act of 1965 could be expanded to help finance initial investments. However, after a while, the program could be economically self-supporting, as the acquired land is sold or leased for development and relatively inexpensive undeveloped land at the urban fringe is acquired for future growth.[122]

An additional boost to the case of those supporting advance public land acquisition is the fact that it not only looks good on paper, but has substantial real-world success. The European experience with this approach, especially the examples of Stockholm and Amsterdam, is extensively cited by American

---

121 [6], p. 109

122 See [21], p. 139; [44], pp. 325-326 and [57], p. 3.

supporters of this public development role.[123] Ann Louise Strong has gone so far as to state: "Public landownership has been a crucial element of almost all successful European planning."[124]

Although far from as extensive as the European experience, the United States has had experience with public development and obviously, with public ownership of land. As a matter of fact, one could use the fact that the federal government at one time owned 3/4 of the land in the country, to dispute ideological arguments against an "unprecedented" degree of government control of land. In addition to cases of cities whose original development was planned and controlled by the government (such as Washington, D.C.), there has been limited experience with restricted application of advance land acquisition techniques in this country.[125]

Naturally, like most proposals, the use of large-scale public land acquisition and disposal has its disadvantages and drawbacks, although they do not seem overwhelming.[126]

a) There is a social cost in tying up public funds in land purchases.

b) There is further loss in revenue as the tax receipts on land withdrawn from private ownership stop flowing.

c) There will be some difficulty in deciding how "large-scale" the acquisition should be. What percentage of the land to be developed in an urban area should be acquired by government in advance? This is not an easy question to answer.[127]

---

[123] See, for example, [44], pp. 312-317; [62], pp. 106-108; and [57], pp. 8-11.

[124] [66], p. xxxi.

[125] See [62], pp. 4, 65-97 for discussion of the experiences of Montgomery County, Maryland and Richmond, Virginia with limited advance land acquisition.

[126] See [62], p. 6

[127] See [21], pp. 358-359 and [57], pp. 6-7 for comments on this question.

d) There will be substantial costs involved in the planning and administration of such a program. Of course, these costs must be balanced against the savings that such an approach is expected to permit.

C. PROPOSED OPERATIONAL FRAMEWORKS FOR DEVELOPMENT CONTROL

The bulk of the above discussion dealt with the general concept of advance public land acquisition as a land-use and development control. Next are discussed some of the major operational formats proposed for use of this technique.

One of the most popular proposals is the construction of "new towns" to absorb the urban growth of our nation. Although new towns can be built by private developers without the use of public land acquisition, past experience has demonstrated that the power of public land acquisition is essential if this is to become a desirable and significant form of urban development.[128] "In May of 1969, the National Commission on Urban Growth recommended that federal assistance be provided for the construction, by the year 2000, of 100 new towns with populations of 100,000 each, and 10 new cities of one million people each."[129] A related concept involves the construction of "new communities." These differ from new towns in that they are not self contained economic entities, but smaller developments that become part of a larger urban area and consist of a population of 15,000-20,000 with a diverse social, economic, and ethnic background.[130] The Advisory Commission on Intergovernmental Relations recommends the use of such new communities as a major approach to the urban development problem.[131]

---

[128] See [63] for a more thorough discussion of the use of the new towns approach.

[129] [66] pp. xxi-xxii.

[130] See [32], pp. 66-74.

[131] See [1], chapters IV-VI for the details of their recommendations and findings. Also see [6], pp. 92-95 for a discussion of an A.I.P. policy statement recommending government involvement in the construction of new communities.

Another major operational construct for the guidance of land development through the use of advance public acquisition and disposal of land is the "development district." Marion Clawson introduced this concept in an article advocating the creation of what he called "suburban development districts."[132] Various other authors have recommended the same type of approach, although varying in administrative and structural details. One of the most comprehensively developed proposals based on the development district construct was the heart of a report by Henry Bain.[133] Bain's development district concept can be briefly characterized as follows: The district would provide a means of concentrating a process of staged development in a limited number of disignated areas, leaving the remaining unbuilt land to be set aside for later development or open space purposes. Each district would consist of an area of a size sufficient for a new community, yet capable of being planned for in detail as a unit (thus avoiding the insurmountable problems one encounters in idealistically trying to plan in detail for an entire metropolitan area). Development would be spread over several years.

To implement this process, an agency would be established by the county, preferably as a new governmental department or a public-private corporation, and would be charged with the administration of all development districts in the county. The agency would base its detailed plans for a district on the general plan for the area or region. "Owners of sizable tracts of land designated for development at lower and medium densities would be encouraged to submit applications for planned-unit zoning, and all remaining land in this category would be rezoned by sectional map amendment. Land designated for intensive development, and other land not developed within a reasonable time, would be acquired

---

[132] See [18].

[133] [6]. Note that the development district proposal presented in this paper was expressly tailored for the Montgomery-Prince George bi-county region of Maryland and thus may contain specifications not appropriate for general application.

by development district for sale or lease to developers subject to conditions as to the character and time of development."[134] Further, the agency administering the districts would coordinate the provision of public services with this development process and work with other agencies to insure the timely completion of public improvements. Each district would be a financially separate unit, financed through county bond issues which would be repayed through assessments and taxes collected within the district. The district would be dissolved as a functional unit once development in the district was substantially complete.[135] A major criticism of the "development district" is that it merely adds another layer of confusion and fragmentation to the already fractured and uncoordinated system of governments and special districts.[136] Hopefully, such a district could be structured and empowered so as to improve instead of worsen the present state of intergovernmental cooperation in an area.

Despite the many proposals for the use of public agencies and corporations to carry out large-scale urban development programs, it was perhaps not until the New York State Urban Development Corporation was established in 1968 that such an entity was actually created and given the necessary wide-ranging powers to accomplish its mission.[137] The U.D.C. is empowered to acquire real property, by condemnation if necessary, to build, sell or lease property, to negotiate with government agencies for the provision of public utilities and services, and borrow money. The Corporation is to provide only low-income housing and industrial development in "substandard" areas or those areas in danger of becoming substandard. The Corporation also has the rare power to exempt itself from obeying

---

134 [6], p. iv.

135 See [6] for elaboration. Also see [46], pp. 26-39 for a similar concept called the "development sector".

136 See [6], p. 116

137 See Laws of New York, 1968, Chapter 174 or [6], pp. 130-134 for details on the purpose and operation of this public corporation.

local zoning ordinances and building codes if it finds this necessary. Despite the tremendous power the Urban Development Corporation has on paper, experience has demonstrated that it has been prevented from exercising its powers in the face of intense political pressures and opposition from localities. In fact, only gubernatorial vetos have saved the Corporation from losing on paper the powers it has been afraid to use in practice.[138] The future efficacy of the U.D.C. and prospects for agencies like it is very unsure at this time.

IX. CONCLUSION

Many of the pieces of legislation and proposals mentioned herein are still too recent to permit a valid evaluation of their absolute and relative merits at this time. Compounding this is the custom-made nature of many of the programs and suggestions, which were structured with the character and needs of a particular state or region in mind. Although one may despair at this yet further example of fragmentation and nonuniformity in the application of land-use and development controls, it is perhaps more productive to rejoice at a pluralistic system in which thousands of separate "laboratories" can experiment and hopefully produce a few Ramapo's. The efficacy and feasibility of the various proposals, especially in combination, and with respect to various areas of different character and needs, is not yet well established. However, it is clear that there is a great need for improvement in our present approach. Though it is fairly certain that no single approach will alone be sufficient to centend with the complex urban problems that confront this country, some of the proposals discussed above will certainly prove to be useful in an improved land guidance system.

---

[138] See [61].

BIBLIOGRAPHY

\+ Indicates a work which has been annotated. Annotations follow this list in a separate section.

(*) Indicates a work which has been annotated in work [58] listed below. The number given in parentheses is that used in work [58] to identify the appropriate annotation.

[1] Advisory Commission on Intergovernmental Relations. Urban and Rural America: Policies for Future Growth. Washington, D.C., April, 1968.

[2] The American Law Institute, A Model Land Development Code (Tentative Draft No. 1) Philadelphia, Pa., 1968.

[3] American Society of Planning Officials. Regulatory Devices (ASPO National Planning Conference). Chicago, Illinois, 1969.

+[4] American Society of Planning Officials, Land-Use Policies (Papers presented at the Land-Use Policies Short Course held at the 1970 ASPO National Planning Conference). Chicago, Illinois, 1970.

*(92) +[5] Babcock, Richard F. The Zoning Game: Municipal Practices and Policies. Madison, Wisconsin: The University of Wisconsin Press, 1966.

+[6] Bain, Henry. The Development District: A Governmental Institution for the Better Organization of the Urban Development Process in the Bi-County Region (Prepared by the Washington Center for Metropolitan Studies for the Maryland-National Capital Park and Planning Commission), 1968.

[7] Bassett, Edward M. Zoning: The Laws, Administration, and Court Decisions During the First Twenty Years. New York: Russell Sage Foundation, 1940.

[8] Beckman, Norman. "Toward a Development of a National Urban Growth Policy: Legislative Review 1971", Journal of the American Institute of Planners, Vol. XXXVII, No. 4 (July, 1972), pp. 231-249.

[9] Benscher, J. H. Land Use Controls: Cases and Materials (3rd ed.). Madison, Wisconsin: The College Printing and Typing Company, 1964.

[10] Boezi, Reynold A. "Strategies for Metropolitan Reorganization," in work [4], supra, pp. 39-49.

+[11] Bosselman, Fred P. Alternatives to Urban Sprawl: Legal Guidelines for Governmental Action (Research Report No. 15 prepared for the National Commission on Urban Problems). Washington, D.C., 1968.

+[12] Bosselman, Fred and David Callies. The Quiet Revolution in Land Use Control Washington, D.C.: Council on Environmental Quality, 1971.

+[13] Chapin, F. Stuart, Jr. "Existing Techniques for Shaping Urban Growth," in work [35], infra, pp. 108-130.

*(83)[14] Chapin, F. Stuart, Jr. Urban Land Use Planning (2nd ed.). Urban, Illinois: University of Illinois Press, 1965.

[15] Chapin, F. Stuart, Jr. and Shirley F. Weiss. "Land Development Patterns and Growth Alternatives," in work [16], infra, Chapter 13.

*(27)[16] Chapin, F. Stuart, Jr. and Shirley F. Weiss (editors). Urban Growth Dynamics in a Regional Cluster of Cities. New York: John Wiley and Sons, Inc., 1962.

*(26)[17] Chapin, F. Stuart, Jr. and Shirley F. Weiss. Factors Influencing Land Development - Evaluation of Inputs for a Forecast Model. Chapel Hill: Institute for Research in Social Science, University of North Carolina, 1962.

[18] Clawson, Marion. "Suburban Development Districts - A Proposal for Better Urban Growth," Journal of the American Institute of Planners, Vol. XXVI, No. 2 (May, 1960), pp. 69-83.

[19] Clawson, Marion with Charles L. Stewart. Land Use Information: A Critical Survey of U.S. Statistics Including Possibilities for Greater Uniformity (Resources for the Future). Baltimore, Maryland: The Johns Hopkins Press, 1965.

[20] Clawson, Marion. "Why Not Sell Zoning and Rezoning? (Legally, That Is)," Cry California, Vol. II, No. 1 (Winter 1966-67), pp. 9,37.

+[21] Clawson, Marion. Suburban Land Conversion in the United States: An Economic and Governmental Process. Baltimore, Maryland: The John Hopkins Press, 1971.

+[22] Coke, James G. and John J. Gargan. Fragmentation in Land-Use Planning and Control (Research Report No. 18 of the National Commission on Urban Problems). Washington, D.C., 1969.

[23] Czamanski, Stanislaw. "Effects of Public Investments on Urban Land Values," Journal of the American Institute of Planners, Vol. XXXII, No. 4 (July, 1966), pp. 204-217.

*(84)[24] Delafons, John. Land-Use Controls in the United States (2nd ed.). Cambridge, Massachusetts: The M.I.T. Press, 1969.

[25] Dyckman, John W. "Socio-Economic and Technological Forces in Urban Expansion," in work [35], infra, pp. 1-27.

*(206)[26] Feiss, Carl. "Planning Absorbs Zoning," Journal of the American Institute of Planners, Vol. XXVII, No. 2 (May, 1961), pp. 121-126.

+[27] Garrison, William L., Edgar M. Horwood, and Duane F. Marble. Progress Report on a Study of Land Development Problems at Freeway Interchanges (Research Report No. 20, Highway Economic Studies). University of Washington, March, 1960.

[28] Goodman, William I. and Eric C. Freund (editors). Principles and Practice of Urban Planning. Washington, D.C.: International City Managers' Association, 1968.

[29] Haar, Charles M. (editor). Law and Land: Anglo-American Planning Practice. Cambridge, Massachusetts: Harvard University Press and M.I.T. Press 1964.

(*69)[30] Haar, Charles M. Land-Use Planning: A Casebook on the Use, Misuse, and Re-Use of Urban Land (2nd ed.). Boston, Massachusetts: Little, Brown and Company, 1971.

[31] Hagman, Donald G. "Public Acquisition and Disposal of Lands," in work [4], supra, pp. 5-17.

+[32] Heeter, David. Toward a More Effective Land-Use Guidance System: A Summary and Analysis of Five Major Reports, Chicago, Illinois: American Society of Planning Officials, 1969.

[33] Hill, Carroll V. "Planning, Property Taxation, and Zoning: A Trilogy in Urban Development," Planning 1967. Chicago, Illinois: American Society of Planning Officials, 1967, pp. 282-290.

[34] Horack, Frank E., Jr. and Val Nolan, Jr. Land Use Controls. St. Paul, Minnesota: West Publishing Co., 1955.

(*20)[35] Housing and Home Finance Agency. Urban Expansion - Problems and Needs: Papers Presented at Administrator's Spring Conference. Washington, D.C., 1963.

+[36] Kaminsky, Jacob. Environmental Characteristics Planning: A Concept for Land Use Planning and Land Development Standards. Baltimore, Maryland: Regional Planning Council, July, 1968.

[37] Krasnowiecki, Jan and Ann Louise Strong. "Compensable Regulations for Open Space - A Means of Controlling Urban Growth," Journal of the American Institute of Planners, Vol. XXIX, No. 2 (May, 1963), pp. 87-101 (incl. comments).

[38] McGrath, Dorn C., Jr. "Implementing National Policies: Bigger Carrots, Bigger Sticks," in work [4] supra, pp. 29-37.

(*160)[39] Makielski, S.J., Jr. The Politics of Zoning - The New York Experience, New York, New York: Columbia University Press, 1966.

[40] Mandelker, Daniel R. Controlling Planned Residential Developments. Chicago, Illinois: American Society of Planning Officials, 1966.

[41] Marks, Harold. Protection of Highway Utility (National Cooperative Highway Research Program Report 121). Washington, D.C.: Highway Research Board, 1971.

[42] Maryland Planning and Zoning Law Study Commission. Final Report. Legislative Recommendations, Baltimore, Maryland: Colony Press, Inc., 1969

[43] National Commission on Urban Problems. Building the American City (Report of the National Commission on Urban Problems to the Congress and to the President of the United States). Washington, D.C., 1968.

*0)+[44] National Resources Committee. Urban Planning and Land Policies (Volume II of the Supplementary Report of the Urbanism Committee to the National Resources Committee). Washington, D.C., 1939.

[45] Netzer, Dick. *Impact of the Property Tax--Effect on Housing, Urban Land Use, Local Government Finance* (Prepared for the Consideration of the National Commission on Urban Problems). Washington, D.C., 1968.

*(55)+[46] Noble, Jack. *A Proposed System for Regulating Land Use in Urbanizing Counties*. Chicago, Illinois: American Society of Planning Officials, 1967.

[47] Noble, Jack. "The Zone Busters Are Coming," in work [4], supra, pp. 55-59.

+[48] O'Harrow, Dennis. "Proposals for New Techniques for Shaping Urban Expansion," in work [35], supra, pp. 137-159.

[49] Olson, William A. "City Participation in the Enforcement of Private Deed Restrictions," *Planning 1967*. Chicago, Illinois: American Society of Planning Officials, 1967, pp. 266-270.

[50] Oser, Alan S. "Innovator in Suburbs Under Fire," *The New York Times*, Section 8, March 28, 1971, pp. 1,6.

[51] Pickard, Jerome P. *Taxation and Land Use in Metropolitan and Urban America* (Urban Land Institute Research Monograph 12). Washington, D.C.: Urban Land Institute, 1966.

[52] Plager, Sheldon J. "The XYZ's of Zoning," *Planning 1967*. Chicago, Illinois: American Society of Planning Officials, 1967, pp. 271-277.

*(89)[53] Pooley, Beverley J. *Planning and Zoning in the United States*. Ann Arbor, Michigan: Legislative Research Center, The University of Michigan Law School, 1961.

[54] Rothkopf, Arden H. *The Law of Zoning and Planning* (3rd ed., based on the original text by Charles A. Rothkopf). New York, New York: Clark Boardman Co., Ltd., 1969.

[55] Rawson, Mary. "Property Taxation and Zoning Bylaws - The Positive and the Negative in City Planning," *Planning 1967*. Chicago, Illinois: American Society of Planning Officials, 1967, pp. 278-282.

*(211)[56] Reps, John W. "Requiem for Zoning," *Planning 1964*. Chicago, Illinois: American Society of Planning Officials, 1964, pp. 56-67.

+[57] Reps, John W. "The Future of American Planning: Requiem or Renascence?," *Land-Use Controls*: *A Quarterly Review*, Vol. 1, No. 2 (1967), pp. 1-16.

[58] Schiffman, Irving. *The Politics of Land-Use Planning and Zoning: An Annotated Bibliography*. Davis, California: Institute of Governmental Affairs, University of California, 1970. [* notes refer to annotations in this volume]

[59] Schmid, A. Allan. *Converting Land From Rural to Urban Uses* (Resources for the Future, Inc.). Baltimore, Maryland: The Johns Hopkins Press, 1968.

[60] Schmid, A. Allan. "Tax Policy and Land-Use Guidance Techniques," in work [4], supra, pp. 19-26.

[61] Shipler, David K. "U.D.C. Housing Agency: Paper Tiger," The New York Times, August 5, 1972, p. 38.

+[62] Shoup, Donald C. and Ruth P. Mack. Advance Land Acquisition by Local Governments: Benefit-Cost Analysis as an Aid to Policy (Prepared by the Institute of Public Administration for the United States Department of Housing and Urban Development). Washington, D.C., 1968.

[63] Slayton, William L. "New Cities: Policies and Legislation," Planning 1967. Chicago, Illinois: American Society of Planning Officials, 1967, pp. 197-174.

[64] Stollman, Israel. "Editorial: Ramapo," Planning, Vol. 38, No. 6 (July, 1972), pp. 108-109.

[65] Strong, Ann Louise. Factors Affecting Land Tenure on the Urban Fringe. Philadelphia, Pennsylvania: Institute for Environmental Studies, 1968.

[66] Strong, Ann Louise. Planned Urban Environments. Baltimore, Maryland: The Johns Hopkins Press, 1971.

[67] Sussna, Stephen. Land Use Control: More Effective Approaches (Urban Land Institute Research Monograph 17). Washington, D.C.: Urban Land Institute, 1970.

[68] Toll, Seymour I. Zoned America. New York, New York: Grossman Publishers, 1969.

[69] United States Senate Committee on Banking, Housing and Urban Affairs. Land-Use, Planning, and Management Programs: Hearings Before the Subcommittee on Housing and Urban Affairs on S.J. Res. 52 and Title II of S. 1618. Washington, D.C., 1971.

+[70] United States Senate Committee on Interior and Insular Affairs. National Land Use Policy: Background Papers on Past and Pending Legislation and the Roles of the Executive Branch, Congress, and the States in Land Use Policy and Planning (Committee Print, 92nd Congress, 2nd Session). Washington, D.C., April, 1972.

[71] United States Senate Committee on Interior and Insular Affairs. Land Use Policy and Planning Assistance Act of 1972 (Report No. 92-869 of the Committee to Accompany S. 632). Washington, D.C., June 19, 1972.

[72] Urban Land Institute. Legal Aspects of Planned Unit Residential Development with Suggested Legislation (Urban Land Institute Technical Bulletin 52). Washington, D.C., 1965.

[73] Walker, David B. "Needed Components of a National Urban General Policy," in work [4], supra, pp. 65-74.

[74] Welch, Louie. "The Strength of Planning in Houston," Planning 1967. Chicago, Illinois: American Society of Planning Officials, 1967, pp. 253-259.

+[75] Wheaton, William L. C. "Public and Private Agents of Change in Urban Expansion," in work [35], supra, pp. 64-94.

(*9)[76] Willhelm, Sidney M. *Urban Zoning and Land-Use Theory*, New York, New York: The Free Press of Glencoe, 1962.

(*72)[77] Williams, Norman, Jr. *The Structure of Urban Zoning--And Its Dynamics in Urban Planning and Development*. New York, New York: Buttenheim Publishing Corporation, 1966.

(*35)[78] Wingo, Lowdon, Jr. (editor) *Cities and Space*: *The Future Use of Urban Land* (Resources for the Future). Baltimore, Maryland: The Johns Hopkins Press, 1963.

SELECTED BIBLIOGRAPHIC ANNOTATIONS

Note: Only author and short title used to identify work. See bibliographic reference above for more complete information. See work [58], supra, for a far more extensive annotated listing.

[4] ASPO, Land-Use Policies, 74 pp.

This ASPO publication consists of seven papers presented at the land-use policies short course held at the 1970 ASPC National Planning Conference. The inspirations for this set of papers were major recent reports that have dealt with the problem of land-use and development and have suggested sweeping changes in our present regulatory system. These papers (five of which were summarized and analyzed in detail by David G. Heeter in work [32], supra) represent a new interest in controlling and guiding land development and a growing willingness to take rather bold steps in this area of public concern. Although much of the material summarizes some of the findings of these reports and some more recent ones, there is also considerable critical comment and discussion in such areas as tax policy, public acquisition and disposal of land, "zone busting," and the general path the state of New Jersey has been following in land-use and development controls.

[5] Richard Babcock, The Zoning Game, 202 pp.

Definitely among the more readable and interesting books to deal with the field of zoning. The author uses his personal professional experience with zoning litigation to describe the present state of zoning, dispels what he feels to be misconceptions on "both" sides, and make suggestions for the improvement of the present system. Illustrative cases and examples are used liberally. The author concludes that despite its bad reputation among professionals, zoning is worth "saving" through improvement.

[6] Henry Bain, The Development District, 204 pp.

Although written for the bi-county Montgomery - Prince George region of Maryland, this report is an excellent general work on the concept of the development district in particular and, more broadly, that of public land acquisition and disposal in the management and guidance of urban land development. The main body of the report describes the basis for the use of development districts in the Bi-County Region, introduces the concept of this approach, explores its workings in detail, and ends with recommendations for its implementation. In a section of the report labelled "appendix," although constituting the major part of the text, Mr. Bain does an excellent job of summarizing recent proposals generally supporting the use of a development district type of approach and exploring several related areas.

[9] J.H. Beuscher, Land Use Controls, 577 pp.

A rather thorough treatment of the legal aspects of land-use controls. The author makes extensive use of case material and editorial comments to deal with most of the traditional areas of land-use controls. Some less traditional areas, such as the use of taxation and conditional grants-in-aid programs, are dealt with briefly.

[11] Fred P. Bosselman, *Alternatives to Urban Sprawl*, 69 pp.

This report concerns itself with the question of the constitutionality of three major techniques that have been proposed to stop urban sprawl and promote a desirable urban growth process: (1) Planned Development Zoning; (2) Compensative Regulations; and (3) Public Land Assembly. The approach is legally oriented. These three techniques are described and various possible constitutional objections that have been or may be made against each are discussed, following a section demonstrating the need to prevent urban sprawl. The author acknowledges the legal stumbling blocks involved in implementing these techniques, but basically remains optimistic about the upholding of their constitutionality. He concludes that such techniques must be used and that the changing needs of our society will insure their judicial acceptance.

[12] Fred Bosselman and David Callies, *The Quiet Revolution in Land Use Control*, 327 pp. + appendix.

This book reviews the substance of a number of innovative land-use control programs, mostly at the state level, which have recently been introduced, and then proceeds to explore their operation and "success/failure" records. An appendix is included which contains the text of much of the relevant legislation. A great deal of interview work lies behind the comments made with respect to the actual operation and success of these innovations, which are for the most part too recent to be adequately evaluated at the present time.

[13] F. Stuart Chapin, "Existing Techniques for Shaping Urban Growth"

In this article, the author explores the present techniques used to shape urban growth and discusses their deficiencies. He groups the existing techniques into five broad classes: (1) the general plan and planning process; (2) public policy as a means for shaping urban growth; (3) public works as a mechanism for stearing urban expansion; (4) regulatory devices for guiding urban expansion; and (5) civic education as a factor in urban expansion. The second half of the paper suggests a more effective system of techniques under five headings analogous to the above, emphasizing coordination and reorganization. The stress on coordination is formalized in a concluding section on construction of an Urban Development Guidance System through careful coordination of techniques and governments.

[21] Marion Clawson, *Suburban Land Conversion in the United States*, 406 pp.

This book deals with land use and the urban land development process. It focuses on the suburbs and the process by which land is converted to urban use. The author states an intention to treat this subject comprehensively and to carefully evaluate the meaning of the data presented or discussed. No single program of action is advocated, although a number of suggestions for improvement are discussed, and the importance of their interrelationship in application is stressed. The author explores the process of suburban development since World War II and the "actors" involved in the process in detail; examines the process and its results in the Northeast with the help of several metropolitan case studies; makes some projections on results of a continuation of this process over coming decades; and ends with some suggestion for improvement.

[22] James C. Coke and John J. Gargan, Fragmentation in Land-Use Planning and Control, 91 pp.

This exploration of fragmentation in land-use planning and control has as its main theme the importance of distinguishing two forms of fragmentation: the geographic fragmentation resulting from the existance of large numbers of local and other governments, each exercising land-use controls in space; and the functional fragmentation caused by the lack of coordination among governments, within and between levels of government, in the provision of various public facilities and services. The authors conclude that not enough attention has been given the latter form of fragmentation and that this functional fragmentation must be arrested before any significant improvement can be achieved.

[27] William L. Garrison, Edgar M. Horwood, and Diane F. Marble, Progress Report on a Study of Land Development Problems at Freeway Interchanges

This is a progress report on a study that was incomplete at the time of writing. The study includes four major subsections: (1) a general review of the literature in the area of land use control and interchange problems; (2) a look at the magnitude and character of development around freeway approaches and interchanges; (3) a critical exploration of land use controls and their impact; and (4) some suggestions for the improvement of development controls and policy with respect to highways. This progress report deals primarily with projects two and three above. Preliminary results and suggestions are discussed. Extensive appendices emphasize projection of the supply and demand for land for various uses.

[32] David Heeter, Toward a More Effective Land-Use Guidance System, 120 pp.

This is a summary and analysis of five reports, with an emphasis on the development of an effective system for the regulation and manipulation of land-use, development, and redevelopment. The author not only summarizes and compares the pertinent proposals of these reports in detail, but he tries to do so in a way that constructs a fairly unified, "conglomerate" proposed system from the individual sets of proposals. This approach is founded upon the author's claim that all of the reports fundamentally agree in suggesting sweeping changes, although there are admittedly considerable variations in the techniques and extent of change recommended among the five. The need for approaches that transcend purely local considerations and the contention that land use and development must be considered and approached separately in three steps - development of rural land, regulation of developed land, and redevelopment of developed land - to achieve an effective land-use guidance system, are major themes in this book.

[36] Jacob Kaminsky, Environmental Characteristics Planning, 45 pp.

This report, an outgrowth of work done on the Suggested General Development Plan for the Baltimore Region, proposes the use of "Environmental Characteristics" districts as the foundation of land-use planning and regulation in place of the less flexible use districts. Along a continuum from open space to "town center," a set of eight development types, based on intensity of activity and metropolitan function, is introduced. An Illustrative series of development standards, to be applied to such district types, are discussed. The use of this more flexible approach to development planning on the regional and local levels is touched upon briefly.

[44] National Resources Committee, Urban Planning and Land Policies, 366 pp.

This very interesting and well-done report covers three major areas: (1) planned communities; (2) urban living conditions; and (3) urban land policies. Despite the 1939 date of the report, it seems surprisingly recent in its coverage, concerns, outlook, and conclusions. Especially in the area of land policies, the suggestions made were obviously ahead of their time and are often identical to the "bold new" proposals being made today. The report made extensive use of the European experience to support its findings in some areas.

[46] Jack Noble, A Proposed System for Regulating Land Use in Urbanizing Counties, 49 pp.

This report results from a study based on a suggested land development regulation program for Anne Arundel County, Maryland. The suggested regulatory program revolves around the use of a "development sector" concept in dealing with land not yet built up. The important difference that must exist between the regulation of developed and undeveloped land; the respective appropriate places for discretion and for precise rule-making in the regulatory process; and the facts that, (a) development cannot be stopped (nor would this be desirable) and (b) the "ultimate" use of presently undeveloped land can never be foreseen with certainty, are dominant themes of the report.

[48] Dennis O'Harrow, "Proposals for New Techniques for Shaping Urban Expansion".

In this paper, Dennis O'Harrow begins with a discussion of the failure of present techniques to adequately deal with present urban expansion, especially the failure of land-use controls to handle development on the scale at which it now occurs. He briefly outlines the various dimensions along which government approaches can be varied. The bulk of the paper then discusses a large number of new techniques and administrative formats for dealing with urban development.

[57] John W. Reps, "The Future of American Planning: Requiem or Renascence?"

Reps presents a rather convincing argument for the adoption of the technique of advance land acquisition and management by states and their political subdivisions to meet the need for a rational, equitable, and socially efficient public approach to guiding urban growth. He finds that present land-use and development controls are totally inadequate. Successful European and American examples are cited and a suggested system of advance land acquisition is outlined and discussed.

[62] Donald C. Shoup and Ruth P. Mack, Advance Land Acquisition by Local Governments, 127 pp.

The primary purpose of this report is to help local governments in making rational and socially desirable decisions with respect to advance public acquisition of land for sites of future public facilities and other land development objectives. This is achieved through the development of a set of guidelines based on benefit-cost analysis. However, the authors are cognizant of and have great respect for the advantages and disadvantages associated with advance acquisition that are not easily translated into dollars and cents. Detailed case studies, two questionnaire surveys, and a look at the European experience are used to obtain more concrete results and illustrate the methods developed. In addition, the authors do a fine job of summarizing the theoretical and legal aspects of advance land acquisition and discussing its application as a land-use control technique.

[70] U.S. Senate Committee on Interior and Insular Affairs, National Land Use Policy, 212 pp.

This Committee Print was intended for use by members of the Senate Committee on Interior and Insular Affairs. It contains a great wealth of information on past and proposed federal legislation dealing with or related to land-use policies, on federal programs having impacts on land use, and on recent and innovative state legislation in the field. Included are copies of major bills being considered, comparisons and evaluations of the prominent pending legislation, and bibliographic material on recent writings in land-use and development.

[75] William L.C. Wheaton, "Public and Private Agents of Change in Urban Expansion."

This is an excellent discussion of decision-making in the urban development process. Wheaton's basic premise is that the development pattern that ultimately emerges is the result of many diffused and independent decisions made over time, which are subject to various market considerations, but are not under the control of any single group, including government. He concludes that direct application of government power is impossible and the best approach would include the development of an improved, comprehensive system of market data for use by both private investment and public decision-makers; development of metropolitan master plans; and institution of better professional and bureaucratic standards of practice in dealing with land and land development.

APPENDIX A-2

Investigation of Traffic Assignment Algorithms

Peter Kuner, Research Fellow

Department of City and Regional Planning,

University of Pennsylvania

April 1973

The research described in this paper is a portion of the work done under contract DOT-FH-11-7843 issued by the Department of Transportation, Federal Highway Administration, to the Department of City and Regional Planning, University of Pennsylvania.

I. INTRODUCTION

This paper briefly discusses the history of efforts to simulate network traffic and the problems of simulating congestion. Four algorithms with varying approaches to these problems are explored with respect to their theoretical and practical advantages. Statistics for comparing simulations using different algorithms are discussed and, finally, the results of experimental simulations are analyzed.

II. SIMULATING TRAFFIC AND CONGESTION

Historically, the simulation of mono-modal traffic in a network has been done in two separate stages. First the routing between origins and destinations is determined using the minimum impedance criterion, then trips are assigned to paths through the network. The matching of trips to a path is usually done by an "all or nothing" algorithm, which is to say that there is only one best path from an origin to a destination and all trips from that origin to that destination take that route. This is a sort of "damn the torpedoes" approach, and in recent years considerable attention has been paid to ways of handling the sunken ships of overloaded links, through the application of capacity restraint and volume delay procedures.

In a simplified way, the transportation simulation cycle has become minimum path selection, or tree tracing; loading the network with trips; recomputing the network impedance characteristics under load; and iterating the cycle in some fashion.

These somewhat cumbersome procedures are the result of trying to analyze an analogue problem, one of flows, in a digital fashion. In addition to abstracting the highway into a network, vehicle flows are also abstracted into large discrete pieces of time, vehicles per hour or vehicles per day. Finally the continuous process of movement is separated into route selection on an empty network, and route usage. This last abstraction in a sense is responsible for the inaccuracies in the simulation of congestion. Too often the process unrealistically overloads the "best path" links, while unrealistically leaving many links completely unloaded.

One series of attempts at solving these latter problems involved initially determining several "best" paths. Then "capacity restraint" procedures load trips onto a path until the volume reaches the capacity. Once this happens, successive trips are loaded onto the next best path and so forth. "Diversion curve" procedures allocate traffic among several best paths according to a function of their relative impedances and/or capacities.

These procedures have not become standard practice for several reasons. First, there is a theoretical problem in determining more than one best path. If the next best and the third best path must be entirely different from each other, as well as different from the best paths, the "diversion" process will force traffic onto highly unlikely routes. On the other hand, if each path may be the same but for a few links the difference may be meaningless. Finally, given the present and short run foreseeable state of computer technology, determining one best path will remain an expensive process and determing more than one best path outrageously expensive.

In "all or nothing" algorithms, congestion is simulated by applying a volume delay function to the free flow, or base path impedances, and iterating the cycle of tree tracing and loading. The disadvantage of this procedure is in its "all or nothing" characteristics. It forces all the traffic going from one origin to a destination to take the same route. However, it is an improvement over ignoring the effects of volume delays. Our research involved formulation of four different algorithms for applying a volume delay function, two of which allowed traffic to take several routes from one origin to another. Two algorithms were rejected on theoretical grounds and therefore were partially tested, and the remaining two were extensively tested.

III. FOUR ALGORITHMS

1. Whole Network Loading

This procedure consists of tree tracing, loading all trips, and applying the volume delay function to adjust the link impedances. These adjusted link impedances are then treated as free flow impedances, the link volumes are zeroed, and the process repeated. Since this process "in the raw" causes considerable oscillation of link volumes and impedances from one iteration to the next, a limitation is placed on the adjustment of the link impedances by weighting the adjusted impedances heavily with the free flow, base impedance. The experience of other experimenters has shown that four to nine iterations are necessary for convergence, a fairly expensive proposition! (1), (3).

While this procedure adjusts link volumes such that the final impedances matrix reflects congestion, it does not address the problem of "all or nothing" loading.

2. Incremental Whole Network Loading

This procedure attempts to avoid overloading links beyond their capacity by loading a proportion of the total trips onto free flow minimum paths, adjusting the link impedances (without any limiting factor), recomputing the minimum pahts, and then repeating the process for succeeding proportions. If, for example, trips are assigned to the network in five, twenty percent increments, the first twenty percent are assigned to minimum free flow paths. The next twenty percent are assigned to minimum paths that are computed on a twenty percent loaded network. The next twenty percent are assigned to paths computed on a forty percent loaded network, and so on. Another tactic is to use unequal increments, for example 50%, 30%, and 20%.

The disadvantage of this procedure is the cost in repeating the tree tracing and loading many times. The advantages are that excessive loadings of links of a given minimum path tend to be minimized, and that the trees are determined subject to the previous loading of the network.

3. Incremental Whole Network Reloading

Instead of accumulating the trips as in the previous method, the increments are accumulated and the same trips successively reloaded. Thus with five twenty percent increments, the first increment would be loaded onto free flow paths. The link volumes are then set to zero, and the minimum paths determined. Then the first plus the second (i.e. 40%) increment of trips are reloaded. This procedure reiterates until the last iteration when all trips are reloaded onto the minimum paths computed on a network whose impedances reflect almost complete loading conditions. As in the previous method, unequal increments may be used.

The concept behind reloading accumulating increments is to develop the impedance characteristics of a loaded network in a gradual fashion, thus avoiding the oscillations that result from repetitive loading of all trips. This method is, of course, an "all or nothing algorithm.

4. Incremental Tree by Tree Loading

A method developed by the pioneering CATS study was to partially integrate route selection and route usage. One tree was traced, that is the minimum paths were determined from one origin to all destinations, and then all trips from that origin were loaded. The link impedances of the tree were adjusted, and the process repeated for the next tree. Thus the trips from the first tree "encounter" a free flow network, while the trips from the last tree encounter an almost fully loaded network.

One problem is that the order in which trees are selected affects individual link volumes. One attempt at dealing with this bias involved selecting trees randomly. This does not seem to be an ideal solution since it only randomizes the bias, and does not reduce its magnitude.

The advantage of the CATS method is that the assignment process is done once, and the tree tracing done once, or at most a second time (to obtain interzonal impedances of the final loaded network). The disadvantages are in the biases caused by the order of tree selection, and the "all or nothing" route selection.

To attack those problems, the original CATS method was incrementalized. After each tree is traced, it is loaded with a portion of the trips rather than all trips. The magnitude of the bias caused by the order in which trees are selected is reduced according to the proportion of trips loaded in each iteration. To illustrate, consider five, twenty percent increments. On the first iteration, the paths from the first origin are selected on the basis of an empty network, and the paths from the last origin are selected on the basis of almost a twenty percent loaded network. On the second iteration, the first origin paths are selected on almost the same basis as the first iteration, last tree origin's paths, and so forth. Five equal increments divide the bias of tree origin order by approximately five. Similarly, the trips from an origin to a destination may be assigned to five different paths.

IV. TESTS AND STATISTICS FOR COMPARING ALGORITHMS

Two sets of data were used in testing these methods. One was the 1965 San Francisco Bay Area Highway Network and the other was a small test network of 114 links. Using our coding methods, the San Francisco Network contains 5107 links and 2022 nodes, 291 of which are load nodes. The volume delay function used was taken from the Bureau of Public Roads: (1)

$$\underline{I' = I^{o} \times (1.0 + 0.15 \times (v/c)^{4})}$$

I' = Modified link impedance

I^{o} = Free flow link impedance

V = Link volume

C = Link capacity

Early runs ran into difficulties from overloaded links, partly due to large traffic generators connected to the network by a single link and partly due to an overly generous trip generation model. As a remedial step, the volume delay formula was modified by placing an upper bound of 2.50 on the volume capacity ratio. The maximum delay factor was $(1.0 + 0.15 \times 2.5^4)$ which is approximately equal to 39. Thus a link with an impedance of one minute could increase to an impedance of thirty-nine minutes, but no more. Links whose delay factors were "restrained" in this way were flagged for further investigation.

The main criteria used in comparing methods was total travel time, which is of course, an extension of the minimization of disutility that is expressed in the "best path" criteria. Secondary criteria were the number of links loaded beyond 2.5 times capacity, the mean and standard deviation of actual link volume to capacity ratios, restrained volume to capacity ratios, and the number of unloaded links on the network. The last criteria was developed in a traffic assignment study by Huber, Boutwell, and Witheford. (2)

The clearest indication of network congestion is obtained from the mean (and standard deviation) of the volume to capacity ratio. The other statistics reflect the design of the network and the quality of the assignment technique as well as the network state. A large number of restrained links, for example, can indicate poor network design, over estimated traffic generation, or simply an overloaded network. The effects of restraining links are shown by comparing the means of the actual volume capacity ratio to the restrained volume capacity ratio.

Runs on the small test network confirmed the intuitive idea that all the statistics moved together, within limits. For example, the best possible loading will leave some links with zero volume. However, up to this point, the number of links with zero volume will decline as "better", i.e., lower total network travel time, solutions are found. Similarly, "better" solutions always have lower mean volume to capacity ratios, and fewer restrained links.

Extensive testing was not done on the first method for a number of reasons. First, there was an aversion to the theoretical inelegance of attempting to force convergence. It is clear that the paths from one iteration to the next vary a great deal, and that the decision of when to stop iterating is fairly arbitrary. (3)

Finally, it appears that successive iterations would not produce more routes, i.e. that the number of zero loaded links would remain fairly constant.

V. RESULTS OF SIMULATION RUNS

1. Whole Network Loading and Incremental Whole Network Loading

One iteration of the Whole Network Method is identical to a single, one hundred percent increment of the Incremental Method, i.e. all trips are loaded on paths derived from the unloaded network. Table 1 compares two single increment runs with five, twenty percent increments.

The results of these runs show that incremental loading is superior to non-incremental loading, and also illustrates a difficulty in interpreting the statistics.

Run 1 loaded forty percent of the traffic in one increment, an identical load to the second iteration of Run 3. All the statistics for Run 3 are superior, indicating that for a light traffic load, two increments are superior to one.

Run 2 loaded one hundred percent of the traffic in one increment, an identical load to the last iteration of Run 3. With this heavier load, Run 2 has a total travel time that is significantly less than Run 3, but the number of unloaded links is three times greater.

The difficulty with the total travel time statistic is that it is the summed product of final link impedances multiplied by link volumes. On links with heaviest volumes, the impedances are computed on the basis of restrained volume to capacity ratios. Where there are relatively few restrained links, the total travel time statistic is more accurate, e.g. at the forty percent level.

The single increment statistics appear better because fewer links are used overall, even though they are restrained more. This is indicated by the total number of unloaded links, and the mean and standard deviation of the volume to capacity ratio.

2. Incremental Whole Network Reloading

This method was not extensively tested because it contains the same disadvantage as a single increment - the final allocation is "all or nothing". While the link impedances might be as good as previously discussed methods, fewer links would be utilized which would be reflected in larger total travel time, more unloaded links, and a higher volume to capacity ratio.

3. Incremental Tree by Tree Method

It is clear that the finer the increment of trips loaded at one time, the better the network will be utilized. The last method, loading increments on a tree by tree basis rather than a network basis, gives much finer increments than the other methods. In effect, a single iteration of this method loads increments of the total number of trips divided by the average number of trips generated from each load node.

Table 2 compares the two incremental methods, whole network and tree by tree on a small test network. The total travel time for the fully loaded network via the tree by tree method is less than the whole network method at the sixty percent load level. The number of restrained links and volume/capacity statistics of the tree-by tree method are roughly equivalent to the eighty percent level of the whole network method.

These preliminary conclusions from the small test network were confirmed, albeit less dramatically, by the results of runs on the San Francisco Bay Area Highway Network, about fifty times larger than the test network. Table 3 shows the results of three runs using the tree by tree method, and may be compared with the whole network runs in Table 1.

One run uses a single increment of 100%, similar to the CATS Method. A second run uses two increments of 50% each, and a third, three increments of 40%, 40%, and 20%. The worst of these three incremental tree by tree runs, the single increment, is comparable to the best incremental whole network run. The second tree by tree run has about 7% less total travel time, and the third run using three increments has about 8% less total travel time. Greater improvements are indicated by the other statistics, for example, the best tree by tree run has over twenty percent fewer unloaded links, and over thirty percent fewer restrained links than the best whole network run.

As the costs of computing three tree by tree increments are about the same as five whole network increments, it is clear that the tree by tree method is superior to the other methods. To this optimistic conclusion should be appended several notes of caution. I believe that the incremental tree by tree algorithm is indeed superior, but under many circumstances it may be only marginally superior - and may not be cost justifiable. Preliminary investigations indicate that the total travel time statistic is subject to the particular volume delay function used, and the general under or overloading of the network as well as being subject to restraints on the volume to capacity ratio. As we were interested in studying congestion, our trip generation model was designed to be generous, and this perhaps increased the differences between methods.

A more interesting statistic of the quality of the algorithm is the number of unloaded links. The significance of this statistic, however, remains to be determined. It may be said that an algorithm that loads ninety percent of the links utilizes the network more than an algorithm that loads eighty percent of the links, but we cannot say it is necessarily better, let alone ten percent better.

Finally, there are considerations of the purpose of the traffic assignment. If one is interested in examining traffic loads on specific links, e.g. over loaded links or under utilized links, many iterations of the tree by tree method may be justifiable. If one is interested in obtaining an accessibility pattern, e.g. an inter-zonal impedance matrix, the differences between few and many interations are negligible, and not cost justifiable.

In conclusion, it is hoped that more work will be done with these algorithms to more precisely identify their characteristics with different size networks, different volume delay functions, and varying load patterns.

	Run 1	Run 2	Run 3				
Increments	40%	100%	20%, 20%, 20%, 20%, 20%				
Level	40%	100%	20%	40%	60%	80%	100%
Total Time (000's)	11,322	58,246	2,959	10,767	23,188	41,070	63,319
Unloaded Links	974	974	983	885	675	471	363
Restrained Links	65	622	2	38	187	401	686
Mean v/c	.412	1.018	.206	.410	.729	1.047	1.363
Stan. Dev. v/c	.601	1.508	.301	.542	.919	1.131	1.290
Mean v/c Restrained	.401	.800	.206	.408	.627	.943	1.191
Stan. Dev. v/c Restrained	.548	.893	.300	.531	.710	.810	.861
Total Links	5104	5104	5104				
Total Trips	214,065	547,066	107,033	214,065	332,994	428,136	547,066

Table 1: Incremental Whole Network Method:

40% Increment, 100% Increment, and Five 20% Increments

	Whole Network					Tree by Tree	
Increments	20%	20%	20%	20%	20%	40%, 40%, 20%	50%, 50%
Level	20%	40%	60%	80%	100%	100%	100%
Total Time (000's)	105	372	1.662	8,876	50,816	1,237	1,252
Unloaded	20	8	6	5	3	5	5
Restrained	3	10	16	22	36	29	27
Mean v/c	.316	.742	1.177	1.603	2.076	1.996	1.883
Stan. Dev. v/c	.641	.941	1.445	1.989	2.382	2.367	2.207
Mean v/c Restrained	.340	.670	.959	1.169	1.360	1.342	1.273
Stan. Dev. v/c Restrained	.548	.737	.814	.841	.825	.892	.888
Total Links					114	114	114
Total Trips					15,203	15,203	15,203

Table 2: Comparison of Incremental Methods: Whole Network and Tree by Tree

Increments	100%	50%, 50%	40%, 40%, 20%
Level	100%	100%	100%
Total Time (000's)	59,555	54,468	53.197
Unloaded	374	325	287
Restrained	574	456	422
Mean v/c	1.328	1.276	1.259
Stan. Dev. v/c	1.257	1.111	1.074
Mean v/c Restrained	1.193	1.178	1.169
Stan. Dev. v/c Restrained	.839	.807	.796
Total Links	5,104		
Total Trips	547,066		

Table 3: Incremental Tree by Tree Method:

Comparison of Three Runs on Bay Area Highway Network

BIBLIOGRAPHY

1. "Urban Planning System 360 - General Information", U.S. Dept. of Transportation, Federal Highway Administration, Bureau of Public Roads, June 1970.

2. "Comparative Analysis of Traffic Assignment Techniques With Actual Highway Use", Matthew J. Huber, Harvey Boutwell, David K. Witheford. National Cooperative Highway Research Program Report 58, Highway Research Board 1968.

3. Op. Cit. 1, p. 20. Also see Research Study of Intercity-Intra Urban Interfaces (12) Metropolitan Implementation: User's Guide, pp. 50-54. CONSAD Corporation, Pittsburgh, Pa.

4. Protective Land Use Model - PLUM, three volumes, U.S. Dept. of Transportation, Federal Highway Administration, February 1972.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1976-211-173/781

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE

APENDICE BIBLIOGRAFICO

ARQ. ALEJANDRO VILLANUEVA EGAN

ENERO, 1978.

Palacio de Minería Calle de Tacuba 5, primer piso. México 1, D. F.

[illegible] Sec., Vol. [illegible] pp. 31–48. Pergamon Press, 1975. Printed in Northern Ireland

A COMPUTER-ORIENTED LAND USE FORECASTING MODEL WITH MAPPING CAPABILITY†

CARL F. MEYER
University of Missouri, St. Louis

ANDRE B. CORBEAU
St. Mary's University, Halifax

and

HAROLD L. MACK
University of Missouri, St. Louis

(Received 6 March 1974)

Abstract—This paper describes the philosophy, algorithms, and implementation of a computer simulation model for land use forecasting, and the capability of generating maps of these forecasts. The model is applicable to SMSA's organized on a census tract basis by counties. The forecasts are macro to the census tract level for industrial, residential, commercial and public land uses. Modeling is based on an economic data base of the region, and is extremely flexible for the user. An example set of forecasts and maps is included for illustrative purposes.

is paper describes the philosophy and implementation of a computer-oriented land use ecasting model, currently operational for the St. Louis, Missouri, SMSA.‡ Its features lude: (a) macro forecasting at the census tract level; (b) capability to provide decade- e regional maps of the forecasts; (c) substantial flexibility for the user; (d) capability of ng either large or small economic data base; (e) easily modified to adapt to other ons, (f) can be altered to accommodate different numbers of uses, options, and land units.

Throughout this description of the model, we will use the specific parameters of the St. uis region; it should be noted, however, that any or all of these items may be easily anged. The two major properties of the model are its requirement for minimal input ojections and its large degree of user-optional flexibility.

The model is designed to forecast industrial, residential, commercial, and public land e by census tract in ten year increments to the year 2030. These land uses are ranked and ocessed in the above order so that the same parcel of land cannot be allocated to more in one use. The rationale for the ordering itself is that industry, in general, will acquire atever new land it needs first, having the economic power to do so. Residential was nked second primarily because commercial establishments generally are attracted to ate near residential areas. Public use consists of government facilities, streets, right-of- y, and the like, and as such was considered to represent a residual use. New land is quired by each of the uses from vacant and agricultural land. The model assumes that the id units to be forecasted constitute counties in the region, and that land unit numbering is contiguous within each county. This feature enables minimum input projections: only county-wide population and density forecasts by decade by use are required.

Modeling begins with current-decade (1970) land use information as well as othe pertinent economic data. This is provided by a data base of relevant variables for eac tract, stored on a direct-access device for use by the model. The forecasts by decade, by tract, for each use are generated in four phases: desirability ranking on a county initial allocation to existing land (intensive land use), allocation to vacant and agricultural land (extensive land use), and finally a combination of intensive-extensive use on a count wide basis. Input to the model consists of population and density projections, ranking criteria, and any optional control parameters the user specifies. Output consists of summary reports from each phase, a decade summary by use for each county by land unit, and ample but optional diagnostic information. In addition to printed reports, the forecasted variables are stored by decade on a direct-access device for use by the mapping featu Maps are produced by desired decade and use through stand-alone programs which retrieve the required forecasts from these files. Currently, the variables capable of being mapp are population, acreage, or density (people/acre), although any variable could be mapped

Description of data base

The model accesses an economic data base of 580 variables for each land unit to com the land use forecasts. The existing data base was provided by the East–West Gateway Coordinating Council, a regional planning commission. The required information for each land unit consists of current-decade land use data (acreage and population), centroid distances to all other land units, and a vector containing the identifier number of those tracts bordering the land unit in its county. The remaining variables constitute information on the soil type in the land unit, availability of water, utilities and sewerage, representation of clusters of development, and access to transportation characteristics. These economic variables are used by the model to rank the land units for each land use in the allocation procedure, as well as to provide user-specified overrides, changes, and controls of the forecasting.

Model input

In addition to the data base, certain other information is required for modeling, entered via parameter and control cards. The parameter input and description is shown in Exhibit A; Exhibit B contains the analogous description for the control input. The required input parameters are exogenously determined macro forecasts for the metropolitan area. The St. Louis SMSA is composed of seven countries in Missouri and Illinois and the City of St. Louis. The macro forecasts are county (and city) predictions of population, employment, and density, by land use, at ten year intervals through the year 2030. The function of the model is to allocate the population forecasts to the land units in the region within the specified density constraints. The application of the optional dispersion parameters will be discussed later. The various explanations for the control information will be presented within the context of their specific use. In general, they are designed to allow the user to manipulate the forecasts at the census tract level: the required data and control parameters apply to individual land units only insofar as they are members of a county. The optional control data may be divided into two types: (a) to limit the degree of land utilization, (b) to provide rules which specify a particular land unit or a class of land units [illegible] to be altered.

---

† This research was supported by grants from the Office of Water Resources Research, Department of Interior, and the St. Louis Missouri District, U.S. Army Corps of Engineers.

‡ User information and program (source code, object code, or executable module) is available at cost m the Computer Center, University of Missouri—St. Louis.

Input to the mapping program specifies for each map desired the decade, use, variable, and optional counties. If only decade, use, and variable are given, the entire region is mapped.

Model philosophy

Prior to a description of how the forecasts are produced, we briefly examine the real world process reflected by the model. Observed at a single point in time, land in a large region will be allocated to a variety of uses, and within each category will be also at various stages of development. Through the passage of time both use and development patterns will change gradually for the region as a whole. The use pattern changes in response to economic growth as unused (vacant/agricultural) land is assimilated into the four use categories The development pattern is altered through more intensive utilization of land.

Land in a large region is hierarchal in the sense that some units possess attributes which make them more economically desirable for development than others. The first function of the model is to rank the land units in order of desirability for development in a particular use according to user-specified criteria. The second function is the allocation of forecast population to presently used land, leading to more intensive utilization. The third function is to alter the use pattern of land through time and is accomplished by allocating vacant and agricultural land to the various uses. The final function is to simultaneously apply both intensive and extensive modes on a county-wide basis to smooth the forecasts.

Model operation: ranking procedure

The initial step in the formulation of a forecast is the determination of the desirability of the land units in a county for each land use. This determination is achieved with the use of a set of criteria specified by the user (FILL control input). Each criterion assigns a numerical value to the individual land units with the provision that the larger the value, the more desirable the land unit, most desirable assigned a rank of 1. The values themselves are either values of requested variables in the data base, or user-specified functions of data base variables. The model incorporates a user-written subroutine to perform any such required calculations

Subsequent to ranking the land units for each use for each criterion, the ranks are summed to provide a single ordering on the basis of desirability. From this ordering, an initial estimate of the percentage of the county population to be assigned to each land unit is devised, by use, by decade. This percentage is computed by inverting and normalizing the above sum of the ranks. In particular, let X_j be the sum of the ranks for land unit j. Then p_j, the initial percentage, is obtained from:

$$P_j = \frac{1}{X_j \sum_j \frac{1}{X_j}} (100) \text{ for all } j.$$

The rationale leading to the formulation of this equation is that summing the individual ranks for a land unit results in a single order for desirability. The initial percentage then should be proportional to the inverse of this order, since the most desirable land unit for any criterion has rank 1. This initial estimate, being a percentage of the next decade population, becomes a first approximation to the ultimate forecast for the land units. This approximation is then compa o the previous-decade population percentage and smoothed with a random number generator to obtain a final percentage allocation for each land unit by use.

The methodology employed in determining these final percentages is designed to maint the thrust of the ranking procedure without extreme deviations from previous-dec allocations. The smoothing process is based upon randomizing about one-half of the la unit's half-difference between initial and previous-decade percentages. This routine, sho in Chart 1, applies the simulation effect to the percentage calculation in such a way as

Chart 1

Simulation routine for final percentage allocation

```
Let $N$ be the number of land units in a county.
Let $e_j$ = half-difference for land unit $j$
      $= \dfrac{\text{initial percentage}_j - \text{previous-decade percentage}_j}{2}$
Compute modified half-differences, $e_j'$, as follows.
   For $|e_j| < 0.0002$, $e_j' = e_j$
   For $|e_j| \geq 0.0002$, compute $\delta_j$, the increment to $e_j$, in order of rank of the $e_j$, smallest first.
Set $\text{Bal}_j = 0$, the residual from randomizing to be distributed among land units with larger $|e_j|$.
Select $R$, a random number between $\pm 0.5$.
Set $\delta_j = R|e_j|$.
Set $\text{Bal}_j = -1(\delta_j)$.
BEGIN: Set $e_j' = e_j + \delta_j$
            $j = j + 1$ (in rank order)
          $\text{Bal}_j = \text{Bal}_{j-1}$
          Select $R$
   For $R$ and $\text{Bal}_j$ of opposite sign, set $\delta_j = \text{Bal}_j + R|e_j|$
                                                  $\text{Bal}_j = \delta_j$
Begin:
   Otherwise, set $\delta_j = \text{Bal}_j$
              $\text{Bal}_j = 0$
              Begin:
   When all $N$ land units have been processed,
      Set: final percentage = initial percentage $- e_j'$.
```

guarantee that the final percentages sum to one. The calculation proceeds after the lar units in each county have been ordered on the basis of smallest (absolute value) half-diffe ence ranked first. Processing begins with the first land unit encountered in order who half-difference exceeds 0·02 per cent. The land units are then ranked, most desirable firs on the basis of these final per cents for subsequent processing.

The ranking criteria provided by the FILL control cards apply to land units by u only as members of a county In order to allow the ranks of specific land units or class of land units to be altered, the RANK control input is provided This one of four gener types of control information capable of modifying results of the model: their comm structure is shown in Chart. 2. The RANK control type provides the rules by which t final ranking of land units in a county may be changed.

Chart 2

Structure of control rules

The four control rule types, RANK, XDEN, EXCP, and STAT have the following common stru whose form is that of a logical expression. A rule is executed if its logical expression is true. The lo expression can consist of at most nine items.

A	#	*B*	And/or	*C*	#	*D*	Action	Amount
1	2	3	4	5	6	7	8	9

B, *C*, *D* represent numeric constants, variables in the data base, or change in [illegible] bles being fore ie symbols # represent the logical operators ≤, =, <, >, ≥, ≠. The Action t [illegible] f the logical e sion is true may be one of three types. increase by Amount, decrease by Amount, or insert Am

Model operation: initial allocation

In this phase, population is allocated to previous-decade land already assigned to the corresponding use, subject to the user-supplied density levels. This density, in terms of people/acre, is the maximum to which land can be developed.

Under the assumption that more desirable land units will attract more extensive development, the forecasted county population is divided into two segments. The first is allocated to the land units on the basis of the final percentage described above. The second is allocated solely on the basis of desirability. As a result, one or a few units, depending on capacity, may receive all the population from this second segment. Specification of the amount of county population to be thus dispersed on the basis of desirability is achieved through the DISP input parameter percentages, by decade by use.

For each decade to be modeled, processing begins for each county, each use, by calculating the two segments of the population to be allocated. Land units are allocated in rank order first by the percentage computed by the randomizing routine. If the number of people to be so allocated will violate the density constraint, then the maximum that can be allocated is assigned, with the remainder placed in a residual to be used in the next phase. If no residual results, then as much of the dispersion segment, if any, is allocated to the land unit consistent with the density level All land units are thus forecasted in county rank order If any of the dispersion segment remains, it is placed in the residual of the most desirable land unit in each county.

Two control types are available for use in conjunction with the initial allocation phase, XDEN prior to allocation, EXCP afterwards: both apply to densities. The XDEN control allows user-specified modifications to be made to the DEN input parameters by land unit or classes of land units. The EXCP control allows for the changing of densities by land unit or classes of land units after initial allocation, and the subsequent alteration of residuals. The reason for providing density modifications both before and after initial allocation is due to the effect of the DISP parameter. If there is no population dispersed through the county, then there would be no need for the EXCP control.

Model operation: vacant-agricultural allocation

The function of this phase of the forecasting procedure is to allocate any residual population generated in the previous step to vacant and agricultural lands. Since no development of a specific use has occurred on these two types of land, densities must be determined at which development will proceed. It is unrealistic to allow new land to be used at the same intensity as currently-used land. Consequently, a range of densities based upon the desirability of the land units is computed in the following manner. The most desirable unit is allowed to be developed to the density of current land, by use, specified by the DEN parameters. The remaining units in the county are assigned initial density estimates by selecting a random decimal percentage from specially constructed intervals, which will be applied to the input threshold density. The width of an interval is dependent upon the proportion of currently-used land to the total acreage of the unit, while its midpoint decreases uniformly with desirability. When all land units in a county have been assigned an initial density estimate, final densities are computed by smoothing the initial estimates of contiguous units The algorithm for these computations is shown in Chart. 3. Note that three inter related elements comprise the density at which vacant and agricultural land is

developed: the desirability of the unit, the amount of current land use acreage, and development of surrounding units in the county. This provides for reasonable, smo development throughout the county. Nevertheless, changes to the final density for a spec land unit or classes of land units may be effected through use of the STAT control typ

Chart 3

Vacant-agricultural density algorithm

The land units in a county are processed in order of desirability. Let I_j be the initial density estimat land unit ranked jth.

Then I_1 = density level for this county, this decade this use, from the DEN parameter.
For each of the remaining units, I_j will be drawn randomly from an interval whose center is a unifo decreasing constant, and whose width is the proportion of this land use to total acreage.

The mid-point of the interval $\text{MD} = \dfrac{N - j + 1}{N}$

Half the width of the interval $\text{HW} = \dfrac{\text{acres this use}}{2(\text{total acres})}$

Upper bound of the interval $\text{UP} = \min(1, \text{MD}(1 + \text{HW}))$
Lower bound of the interval $\text{LO} = \max(0, \text{MD}(1 - \text{HW}))$

$$I_j = \left[\left(\frac{\text{UP} - \text{MD}}{2}\right)\text{R} + \text{LO}\right][\text{DEN parameter}]$$

where R is a uniform (0, 1) random number.

The final densities are computed by taking a weighted average for each land unit, of the initial dens of contiguous (in its county) land units.

During the initial allocation phase of the modeling, unallocated population by use accumulated for each unit as a result of the density specifications (DEN). This popula is now allocated to the vacant and agricultural land in those units with non-zero residu The following procedure is applied first to vacant land, then to agricultural. The acre required to fully allocate the residual is calculated using the computed density descri above. If this is less than or equal to the acreage available, the population is assign current-use acreage is increased, and the donating-type acreage is decreased If the acre required exceeds available acreage, all available is used, and the residual decremented by amount of people actually allocated. This procedure is applied to all land units in a cou in county rank order.

The control information applicable to this phase of forecasting is of two types: th pertaining to restrictions on the use of vacant and agricultural land (HOLD), and th pertaining to restrictions on the use of forecasted land (ZONE) These two types of cont input differ in structure from the others in that each entry applies to a specific named la unit, and the values represent acreage per decade. The ZONE information restricts acre for any of the four uses to the maximum specified by the user The HOLD control in allows the specificaton of an amount of either vacant or agricultural (or both) acreage to withheld from development for the decades desired. These features allow for the capabi to accommodate specific projections for a land unit, such as recreation areas, zoning la and depopulation/land restoration.

Model operation: county-wide allocation

On a theoretical basis, this phase should not be invoked: if the input parameters a control information have been properly specified, the two prior phases should have fo casted both intensive and extensive land uses with no residuals remaining Theoretic however, there is no way to precisely forecast the information required by the model, co sequently some method is needed to attempt to allocate any residuals remaining from vacant-agricultural phase. The function of this last phase of forecasting is to evenly allocate the county residual, that is, the sum of all residuals attached to the land units in the county, on the basis of use mode and desirability. County residuals, by land use, are distributed in three steps: intensive, extensive, and intensive. Following the general modeling philosophy outlined above. In any step, processing is in l[illegible]d unit order, the most desirable receiving population first. The first step attempts to allo[illegible] the county residual to developed land, within the density specifications (DEN). As ma[illegible]eople as can be allocated are assigned to the land units in order. Any residual remaining [illegible]assed to the second step for allocation to vacant and then agricultural land. This step [illegible]ates in the same fashion as phase three, and presumably generates more developed la[illegible]r each use. Since this second step can allocate new acreage, and does so at lower de[illegible]s than those specified through the DEN parameter, any residual still remaining afte[illegible] step should be able to be resolved by intensification of use. The third step, then [illegible]ame in function and operation as step one, [illegible]oring developed land to the maximum [illegible] This allocation is the final attempt by model to eliminate residual population. A [illegible] three, a non-zero residual generates an error message which states that the mod[illegible] the county, use, and decade was unsuccessful.

Model operation: mapping

The mapping feature, while not an inherent function of the modeling itself, should logically be described at this point. The seven counties and city which constitute the St. Louis SMSA have been divided into a total of 452 land units, with the specific breakdown listed in Table 1. In order to provide maps of the region, a grid system was superimposed on a census tract map. The grids were defined to be 5000 ft², resulting in a total of 5300 cells necessary to define the region. Next, a direct-access file was built containing each cell in the region (row-column designation) and its corresponding land unit number. Maps are produced by decade by use either for the entire region or for specific counties of population, acreage, or density from the forecast files generated by the model. The value of a forecasted variable for each land unit is assigned to all the grid cells which are assigned to it. This method does introduce a certain arbitrary element in the maps since not every land unit has linear boundaries, but the degree of distortion is minimal.

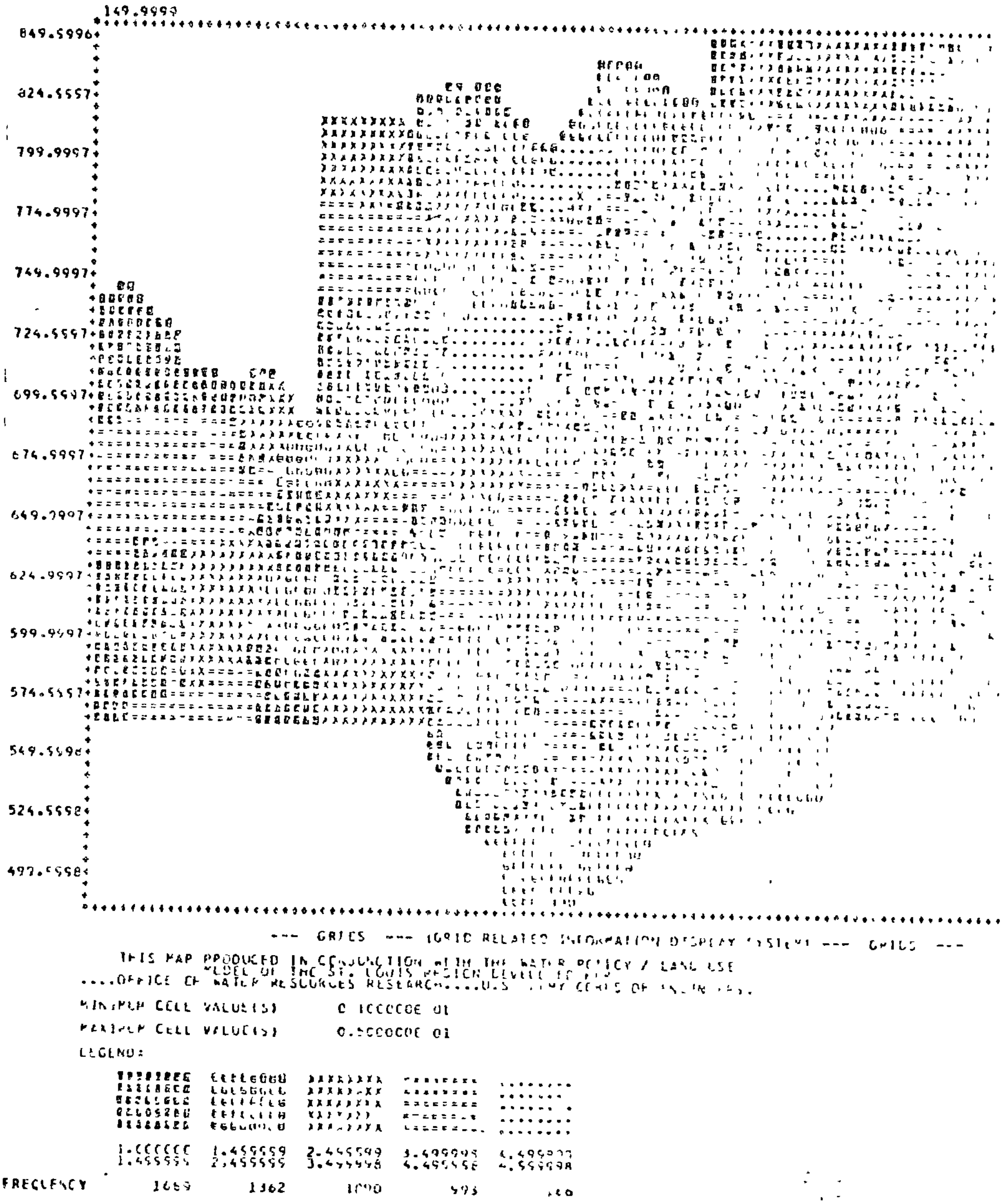
A land unit overlay, shown in Fig. 1, was produced by creating a pseudo forecast file whose variables were essentially the land unit colors of the census tract map.

Model output

Output from the modeling effort may be divided into two categories: summary reports and diagnostic information. There are four summary reports for each decade, one

Table 1. Land unit components

County	Land units
City of St. Louis	1–126
St. Louis	127–273
St. Charles	274–288
Franklin	289–303
Jefferson	304–323
Monroe	324–330
St. Clair	331–394
Madison	395–452



```
--- GRIDS --- (GRID RELATED INFORMATION DISPLAY SYSTEM) --- GRIDS ---
THIS MAP PRODUCED IN CONJUNCTION WITH THE WATER POLICY / LAND USE
MODEL OF THE ST. LOUIS REGION DEVELOPED FOR
....OFFICE OF WATER RESOURCES RESEARCH....U.S. ARMY CORPS OF ENGINEERS.
MINIMUM CELL VALUE(S)      0.1000000E 01
MAXIMUM CELL VALUE(S)      0.5000000E 01
LEGEND:

          1.000000  1.499999  2.499999  3.499998  4.499997
          1.499999  2.499999  3.499998  4.499998  4.999998
FREQUENCY    1669      1362      1090       593   [illegible]
```

Fig. 1. Land unit overlay.

Table 2

```
                    ***** OFFICE OF WATER RESOURCES RESEARCH (U.S.D.I.) *****

                      WATER POLICY / LAND USE MODEL OF  THE ST. LOUIS REGION

                                   LISTING OF CONTROL CARDS
                                   ------- -- ------- -----

1....             TO    2030
2.....PRRO        LF
3....       IND FILL          V135
4....       IND FILL          V130
5....       IND FILL CALL6    V14
6....       IND FILL CALL6    V10
7....       RES FILL          V138
8....       RES FILL          V14
9....       RES FILL CALL6    DV130
10....      RES FILL CALL6    V141
11....      COM FILL          V130
12....      COM FILL          V131
13....      COM FILL          DV14
14....      COM FILL          V141
15....      PUB FILL          DV131
16....      PUB FILL          V37
17....      PUB FILL          DV14
18....      PUB FILL          V10
19....      PUB FILL          V141
20.....PRRO IND POP           640       730       800       880       920       1020
21.....PRRO IND DEN           17        17        18        19        21        25
22.....PRRO IND DISP          8         8         6         5         5         5
23.....PRRO IND FILL CALL2    V72
24.....PRRO IND FILL CALL9    V594
25.....PRRO IND XCEN          V1        EQ 328                                  +112
26.....PRRO IND EXCP          V48       EQ 1        OR  V48       EQ 2          +10
27.....PRRO IND STAT          DV130     LT 50                                   -20
28.....PRRO RES POP           30500     49900     70100     96100     105300    127600
29.....PRRO RES DEN           9         14        18        22        25        25
30.....PRRO RES DISP          5         8         7         6         5         5
31.....PRRO RES FILL CALL8    V79
32.....PRRO RES XCEN          V40       EQ 1        OR  V48       EQ 2          +5
33.....PRRO RES EXCP          DV130     GT 100                                  -5
34.....PRRO RES STAT          DV14      LT 10                                   -5
35.....PRRO COM POP           1170      1360      1560      1730      1880      7060
36.....PRRO COM DEN           20        24        28        29        32        32
37.....PRRO COM DISP          10        10        8         8         7         5
38.....PRRO COM FILL          DV14
39.....PRRO COM FILL CALL8    V79
40.....PRRO COM FILL          DV130
41.....PRRO COM STAT          DV14      LT 10                                   -5
42.....PRRO COM STAT          DV131     LT  5                                   -10
43.....PRRO PUB POP           590       570       610       600       630       620
44.....PRRO PUB DEN           10        11        13        17        18        19
45.....PRRO PUB DISP          3         3         3         3         3         3
46.....PRRO PUB FILL          DV14



                    ***** OFFICE OF WATER RESOURCES RESEARCH (U.S.D.I.) *****

                      WATER POLICY / LAND USE MODEL OF  THE ST. LOUIS REGION

                                   LISTING OF CONTROL CARDS
                                   ------- -- ------- -----

47.....PRRO PUB FILL          DV131
48.....PRRO PUB FILL          DV131
49.....PRRO PUB STAT          DV131     LT 10                                   -5
```

corresponding to each of the last three phases, plus a decade summary (which includ information from phase one). Also produced is a listing of the input information alon with any logically-detectable errors. The decade summary lists each land unit in sequenc order, interrupted after the last land unit in each county by the county totals. The information displayed, by land use, consists of: previous and current decade populations, chang and per cent change, previous and current decade acreages, change, and per cent chang the percentage of population allocated, the county rank, and previous and curren decade densities.

The summary report generated by the initial allocation phase displays by county, b use: people allocated and percentage, residual and percentage, the number of land uni with non-zero residuals, the units with maximum and minimum densities and these value and the most desirable land unit.

The vacant-agricultural phase summary report lists by county, by use: the populatio and acreage allocated using agricultural and vacant land, the residual remaining, and th amount of and unit with the maximum residual, the units with maximum and minimu resultant densities and these values, and the units with maximum acreage acquisition an this value.

The summary report of the county-wide allocation phase displays by county, by us the population, acreage where applicable, and the number of land units effected by each c the three steps. Diagnostic information consists of the intermediate calculations of th vacant-agricultural densities, as well as the listings of the effects of the four types of contr rules.

Table 3

```
                  ***** OFFICE OF WATER RESOURCES RESEARCH (U.S D.I.) *****                       PAG

                     WATER POLICY / LAND USE MODEL OF  THE ST. LOUIS REGION

                                 DECADE SUMMARY FOR 1970-1980

                           POPULATION                               ACREAGE                             % CNTY(CNTY)   1970
                   1970    1980  CHANGE(% OF 1970)         1970    1980  CHANGE(% OF 1970)          POP    (RANK)    DEN

324 INDUSTRIAL   -    0      51       51(*****%)     *        5       5       0(    0%)     *       8.10(  6)       0
    RESIDENTIAL    1257    2418     1161(   92%)     *      386     386       0(    0%)     *       7.96(  6)       3
    COMMERCIAL        9      55       46(  511%)     *        9       9       0(    0%)     *       4.76(  7)       1
    PUBLIC            5      36       31(  620%)     *       14      14       0(    0%)     *       6.  (  7)       0

325 INDUSTRIAL       25      54       29(  116%)     *        5       5       0(    0%)     *       8.52(  5)       5
    RESIDENTIAL    1820    3216     1396(   76%)     *      364     364       0(    0%)     *      10.56(  5)       5
    COMMERCIAL       22      81       59(  268%)     *        5       5       0(    0%)     *       6.95(  6)       4
    PUBLIC            9      80       71(  788%)     *        8       8       0(    0%)     *      11.57(  3)       1

326 INDUSTRIAL      100     151       51(   51%)     *       10      10       0(    0%)     *      27.71(  2)      10
    RESIDENTIAL    6015    7900     1885(   31%)     *      792    1469     677(   85%)     *      45.73(  1)       7
    COMMERCIAL      599     380     -219(  -36%)     *       19      19       0(    0%)     *      36.37(  1)      31
    PUBLIC           30     120       90(  300%)     *       12      12       0(    0%)     *      15.06(  3)       2

327 INDUSTRIAL       75      81        6(    8%)     *        3       3       0(    0%)     *      16.54(  3)      25
    RESIDENTIAL    5734    7303     1569(   27%)     *      433     847     414(   95%)     *      24.00(  2)      13
    COMMERCIAL       37      96       57(  154%)     *        9       9       0(    0%)     *       8.24(  5)       4
    PUBLIC            5      64       59( 1180%)     *        8       8       0(    0%)     *       9.26(  6)       0

328 INDUSTRIAL      129     206       77(   59%)     *        1       2       1(  100%)     *      27.71(  1)     129
    RESIDENTIAL     700    1820     1120(  160%)     *      111     377     266(  239%)     *       6.  (  7)       6
    COMMERCIAL      393      80     -313(  -79%)     *        4       4       0(    0%)     *      23.19(  2)      98
    PUBLIC           65      10      -55(  -84%)     *        1       1       0(    0%)     *      25.65(  1)      65

329 INDUSTRIAL       50      77       27(   54%)     *        4       4       0(    0%)     *      12.16(  4)      12
    RESIDENTIAL    2033    4017     1984(   97%)     *      706     706       0(    0%)     *      13.22(  3)       2
    COMMERCIAL       61     377      296(  485%)     *       25      25       0(    0%)     *      11.74(  3)       1
    PUBLIC            5     190      185( 3700%)     *       19      19       0(    0%)     *      11.97(  4)       0

330 INDUSTRIAL        0      20       20(*****%)     *        1       1       0(    0%)     *       3.26(  7)       0
    RESIDENTIAL    1272    3726     2454(  193%)     *      220     366     146(   66%)     *      12.24(  4)       5
    COMMERCIAL        9     101       92( 1022%)     *        8       8       0(    0%)     *       5.71(  4)       1
    PUBLIC           25      90       65(  260%)     *        9       9       0(    0%)     *        .  (  2)       2
```

Table 4

```
***** OFFICE OF WATER RESOURCES RESEARCH (U.S.D.I.) *****                                        PAGE   15

              WATER POLICY / LAND USE MODEL OF  THE ST. LOUIS REGION

                          DECADE SUMMARY FOR 1980-1990

                             POPULATION                          ACREAGE                        % CNTY(CNTY)   1980    1990
                    1980    1990  CHANGE(% OF 1980)      1980   1990  CHANGE(% OF 1980)    POP    (RANK)    DEN     DEN

324 INDUSTRIAL        51      75     24(  47%)    *        5      5      0(   0%)    *    10.40(   6)      10      15
    RESIDENTIAL     2418    4282   1864(  77%)    *      386    386      0(   0%)    *     8.58(   7)       6      11
    COMMERCIAL        55     108     53(  96%)    *        9      9      0(   0%)    *     7.84(   7)       6      12
    PUBLIC            36      49     13(  36%)    *       14     14      0(   0%)    *     8.64(   6)       2       3

325 INDUSTRIAL        54      83     29(  53%)    *        5      5      0(   0%)    *    11.43(   5)      10      16
    RESIDENTIAL     3216    4930   1714(  53%)    *      364    364      0(   0%)    *     9.88(   5)       8      13
    COMMERCIAL        81     111     30(  37%)    *        5      5      0(   0%)    *     8.11(   6)      16      22
    PUBLIC            80      73     -7(  -8%)    *        8      8      0(   0%)    *    12.86(   5)      10       9

326 INDUSTRIAL       151     149     -2(  -1%)    *       10     10      0(   0%)    *    20.54(   2)      15      14
    RESIDENTIAL     7920   10785   2865(  36%)    *     1489   1489      0(   0%)    *    21.61(   2)       5       7
    COMMERCIAL       380     390     10(   2%)    *       19     19      0(   0%)    *    26.63(   1)      20      20
    PUBLIC           120      99    -21( -17%)    *       12     12      0(   0%)    *    17.40(   3)      10       8

327 INDUSTRIAL        81      81      0(   0%)    *        3      3      0(   0%)    *    12.70(   3)      27      27
    RESIDENTIAL     7303   11454   4151(  56%)    *      847    847      0(   0%)    *    22.95(   1)       8      13
    COMMERCIAL        76     182     86(  89%)    *        9      9      0(   0%)    *    13.21(   4)      10      20
    PUBLIC            64      83     19(  29%)    *        8      8      0(   0%)    *    14.69(   4)       8      10

328 INDUSTRIAL       206     228     22(  10%)    *        2      2      0(   0%)    *    26.92(   1)     103     114
    RESIDENTIAL     1820    4698   2878( 158%)    *      377    377      0(   0%)    *     9.42(   6)       4      12
    COMMERCIAL        80      96     16(  20%)    *        4      4      0(   0%)    *     8.38(   5)      20      24
    PUBLIC            10      11      1(  10%)    *        1      1      0(   0%)    *     6.71(   7)      10      11

329 INDUSTRIAL        77      87     10(  12%)    *        4      4      0(   0%)    *    12.03(   4)      19      21
    RESIDENTIAL     4017    7015   2998(  74%)    *      706    706      0(   0%)    *    14.06(   3)       5       9
    COMMERCIAL       377     305    -72( -19%)    *       25     25      0(   0%)    *    22.15(   2)      15      12
    PUBLIC           190     155    -35( -18%)    *       19     19      0(   0%)    *    22.00(   1)      10       8

330 INDUSTRIAL        20      27      7(  35%)    *        1      1      0(   0%)    *     5.96(   7)      20      27
    RESIDENTIAL     3726    6736   3010(  80%)    *      866    866      0(   0%)    *    13.50(   4)       4       7
    COMMERCIAL       101     188     87(  86%)    *        8      8      0(   0%)    *    11.68(   3)      12      23
    PUBLIC            90     100     10(  11%)    *        9     10      1(  11%)    *    17.70(   2)      10      10
```

Table 5

```
***** OFFICE OF WATER RESOURCES RESEARCH (U.S.D.I.) *****                                        PAGE   2

              WATER POLICY / LAND USE MODEL OF  THE ST. LOUIS REGION

                          DECADE SUMMARY FOR 1990-2000

                             POPULATION                          ACREAGE                        % CNTY(CNTY)   1990    2000
                    1990    2000  CHANGE(% OF 1990)      1990   2000  CHANGE(% OF 1990)    POP    (RANK)    DEN     DEN

324 INDUSTRIAL        75      95     20(  26%)    *        5      5      0(   0%)    *    11.99(   5)      15      19
    RESIDENTIAL     4282    5347   1065(  24%)    *      386    386      0(   0%)    *     7.63(   7)      11      13
    COMMERCIAL       108     128     20(  18%)    *        9      9      0(   0%)    *     8.25(   7)      12      14
    PUBLIC            49      52      3(   6%)    *       14     14      0(   0%)    *     8.67(   6)       3       3

325 INDUSTRIAL        83     109     26(  31%)    *        5      5      0(   0%)    *    13.63(   3)      16      21
    RESIDENTIAL     4930    6483   1553(  31%)    *      364    364      0(   0%)    *     9.25(   6)      13      17
    COMMERCIAL       111     140     29(  26%)    *        5      5      0(   0%)    *     9.07(   5)      22      28
    PUBLIC            73      74      1(   1%)    *        8      8      0(   0%)    *    12.25(   5)       9       9

326 INDUSTRIAL       149     152      3(   2%)    *       10     10      0(   0%)    *    19.09(   2)      14      15
    RESIDENTIAL    10785   15004   4219(  39%)    *     1489   1489      0(   0%)    *    21.40(   2)       7      10
    COMMERCIAL       390     432     42(  10%)    *       19     19      0(   0%)    *    25.69(   1)      20      22
    PUBLIC            99     148     49(  49%)    *       12     12      0(   0%)    *    19.51(   1)       8      12

327 INDUSTRIAL        81      84      3(   3%)    *        3      3      0(   0%)    *    10.83(   6)      27      28
    RESIDENTIAL    11454   15399   3945(  34%)    *      847    847      0(   0%)    *    21.96(   1)      13      18
    COMMERCIAL       182     244     62(  34%)    *        9      9      0(   0%)    *    15.66(   3)      20      27
    PUBLIC            83     105     22(  26%)    *        8      9      1(  12%)    *    17.29(   4)      10      11

328 INDUSTRIAL       228     234      6(   2%)    *        2      2      0(   0%)    *    26.84(   1)     114     117
    RESIDENTIAL     4698    7835   3137(  66%)    *      377    377      0(   0%)    *    11.18(   5)      12      20
    COMMERCIAL        96     112     16(  16%)    *        4      4      0(   0%)    *     8.94(   6)      24      28
    PUBLIC            11      13      2(  18%)    *        1      1      0(   0%)    *     6.34(   7)      11      13

329 INDUSTRIAL        87      98     11(  12%)    *        4      4      0(   0%)    *    12.33(   4)      21      24
    RESIDENTIAL     7015    9869   2854(  40%)    *      706    706      0(   0%)    *    14.08(   4)       9      13
    COMMERCIAL       305     285    -20(  -6%)    *       25     25      0(   0%)    *    18.30(   2)      12      11
    PUBLIC           155     112    -43( -27%)    *       19     19      0(   0%)    *    18.47(   2)       8       5

330 INDUSTRIAL        27      28      1(   3%)    *        1      1      0(   0%)    *     5.29(   7)      27      28
    RESIDENTIAL     6736   10163   3427(  50%)    *      866    866      0(   0%)    *    14.50(   3)       7      11
    COMMERCIAL       188     219     31(  16%)    *        8      8      0(   0%)    *    14.07(   4)      23      27
    PUBLIC           100     106      6(   6%)    *       10     10      0(   0%)    *    17.47(   3)      10      10
```

Table 6

***** OFFICE OF WATER RESOURCES RESEARCH (U.S.D.I.) *****   PAGE [illegible]

WATER POLICY / LAND USE MODEL OF THE ST. LOUIS REGION

DECADE SUMMARY FOR 2000-2010

			POPULATION				ACREAGE			% CNTY(CNTY)	2000	2010
		2000	2010	CHANGE(% OF 2000)		2000	2010	CHANGE(% OF 2000)		POP (RANK)	DEN	DEN
324	INDUSTRIAL	95	114	19(20%)	*	5	5	0(0%)	*	12.94(4)	19	[illegible]
	RESIDENTIAL	5347	7152	1805(33%)	*	386	386	0(0%)	*	7.44(7)	13	[illegible]
	COMMERCIAL	128	154	26(20%)	*	9	9	0(0%)	*	6.91(6)	14	17
	PUBLIC	52	60	8(15%)	*	14	14	0(0%)	*	10.14(6)	3	4
325	INDUSTRIAL	109	143	34(31%)	*	5	5	0(0%)	*	15.35(3)	21	[illegible]
	RESIDENTIAL	6483	7525	1042(16%)	*	364	364	0(0%)	*	7.63(6)	17	[illegible]
	COMMERCIAL	140	162	22(15%)	*	5	22	17(340%)	*	9.37(5)	28	7
	PUBLIC	74	70	-4(-5%)	*	8	8	0(0%)	*	11.67(5)	9	8
326	INDUSTRIAL	152	163	11(7%)	*	10	10	0(0%)	*	16.61(2)	15	16
	RESIDENTIAL	15004	20233	5229(34%)	*	1489	1489	0(0%)	*	20.71(2)	10	13
	COMMERCIAL	432	470	38(8%)	*	19	19	0(0%)	*	24.27(1)	22	[illegible]
	PUBLIC	148	145	-3(-2%)	*	12	12	0(0%)	*	22.66(1)	12	12
327	INDUSTRIAL	84	87	3(3%)	*	3	3	0(0%)	*	10.25(7)	28	29
	RESIDENTIAL	15399	22859	7460(48%)	*	847	847	0(0%)	*	21.02(1)	18	26
	COMMERCIAL	244	301	57(23%)	*	9	22	13(144%)	*	17.41(2)	27	13
	PUBLIC	105	122	17(16%)	*	9	9	0(0%)	*	20.37(2)	11	13
328	INDUSTRIAL	234	235	1(0%)	*	2	2	0(0%)	*	24.16(1)	117	117
	RESIDENTIAL	7835	10162	2327(29%)	*	377	377	0(0%)	*	13.52(5)	20	[illegible]
	COMMERCIAL	112	116	4(3%)	*	4	4	0(0%)	*	6.47(7)	28	[illegible]
	PUBLIC	13	17	4(30%)	*	1	1	0(0%)	*	4.01(7)	13	17
329	INDUSTRIAL	98	109	11(11%)	*	4	4	0(0%)	*	12.41(5)	24	27
	RESIDENTIAL	9869	13239	3370(34%)	*	706	706	0(0%)	*	13.73(4)	13	18
	COMMERCIAL	285	295	10(3%)	*	25	25	0(0%)	*	17.08(3)	11	11
	PUBLIC	112	92	-20(-17%)	*	19	19	0(0%)	*	15.39(4)	5	4
330	INDUSTRIAL	28	29	1(3%)	*	1	1	0(0%)	*	5.19(7)	28	29
	RESIDENTIAL	10163	14910	4747(46%)	*	866	866	0(0%)	*	15.31(3)	11	17
	COMMERCIAL	219	232	13(5%)	*	8	8	0(0%)	*	14.07(4)	27	[illegible]
	PUBLIC	106	94	-12(-11%)	*	10	10	0(0%)	*	15.75(3)	10	9

Table 7

***** OFFICE OF WATER RESOURCES RESEARCH (U.S.D.I.) *****   PAGE 34

WATER POLICY / LAND USE MODEL OF THE ST. LOUIS REGION

DECADE SUMMARY FOR 2010-2020

			POPULATION				ACREAGE			% CNTY(CNTY)	2010	2020
		2010	2020	CHANGE(% OF 2010)		2010	2020	CHANGE(% OF 2010)		POP (RANK)	DEN	DEN
324	INDUSTRIAL	114	131	17(14%)	*	5	5	0(0%)	*	14.33(4)	22	26
	RESIDENTIAL	7152	7632	480(6%)	*	386	386	0(0%)	*	7.25(7)	18	19
	COMMERCIAL	154	184	30(19%)	*	9	9	0(0%)	*	9.81(6)	17	20
	PUBLIC	60	71	11(18%)	*	14	14	0(0%)	*	11.76(6)	4	5
325	INDUSTRIAL	143	155	12(8%)	*	5	5	0(0%)	*	16.93(3)	28	31
	RESIDENTIAL	7525	8934	1409(18%)	*	364	364	0(0%)	*	6.45(6)	20	24
	COMMERCIAL	162	248	86(53%)	*	22	22	0(0%)	*	13.19(5)	7	11
	PUBLIC	70	105	35(50%)	*	8	8	0(0%)	*	16.69(3)	8	13
326	INDUSTRIAL	163	167	4(2%)	*	10	10	0(0%)	*	18.18(2)	16	16
	RESIDENTIAL	20233	21617	1384(6%)	*	1489	1489	0(0%)	*	20.53(2)	13	14
	COMMERCIAL	470	455	-15(-3%)	*	19	19	0(0%)	*	22.94(1)	24	23
	PUBLIC	145	111	-34(-23%)	*	12	12	0(0%)	*	17.64(2)	12	9
327	INDUSTRIAL	87	90	3(3%)	*	3	3	0(0%)	*	9.91(6)	29	30
	RESIDENTIAL	22859	23056	197(0%)	*	847	847	0(0%)	*	21.89(1)	26	27
	COMMERCIAL	301	253	-48(-15%)	*	22	22	0(0%)	*	13.50(4)	13	11
	PUBLIC	122	91	-31(-25%)	*	9	9	0(0%)	*	14.53(5)	13	10
328	INDUSTRIAL	235	234	-1(0%)	*	2	2	0(0%)	*	23.11(1)	117	117
	RESIDENTIAL	10162	11133	971(9%)	*	377	377	0(0%)	*	10.59(5)	27	29
	COMMERCIAL	116	128	12(10%)	*	4	4	0(0%)	*	7.55(7)	29	32
	PUBLIC	17	18	1(5%)	*	1	1	0(0%)	*	5.85(7)	17	18
329	INDUSTRIAL	109	112	3(2%)	*	4	4	0(0%)	*	12.23(5)	27	28
	RESIDENTIAL	13239	15566	2327(17%)	*	706	706	0(0%)	*	14.78(4)	18	22
	COMMERCIAL	295	323	28(9%)	*	25	25	0(0%)	*	17.20(2)	11	12
	PUBLIC	92	99	7(7%)	*	19	19	0(0%)	*	15.74(4)	4	5
330	INDUSTRIAL	29	31	2(6%)	*	1	1	0(0%)	*	5.44(7)	29	31
	RESIDENTIAL	14910	17342	2432(16%)	*	866	866	0(0%)	*	15.46(3)	17	20
	COMMERCIAL	232	289	57(24%)	*	8	41	33(412%)	*	13.37(3)	29	7
	PUBLIC	94	135	41(43%)	*	10	10	0(0%)	*	15.77(1)	9	13

Table 8

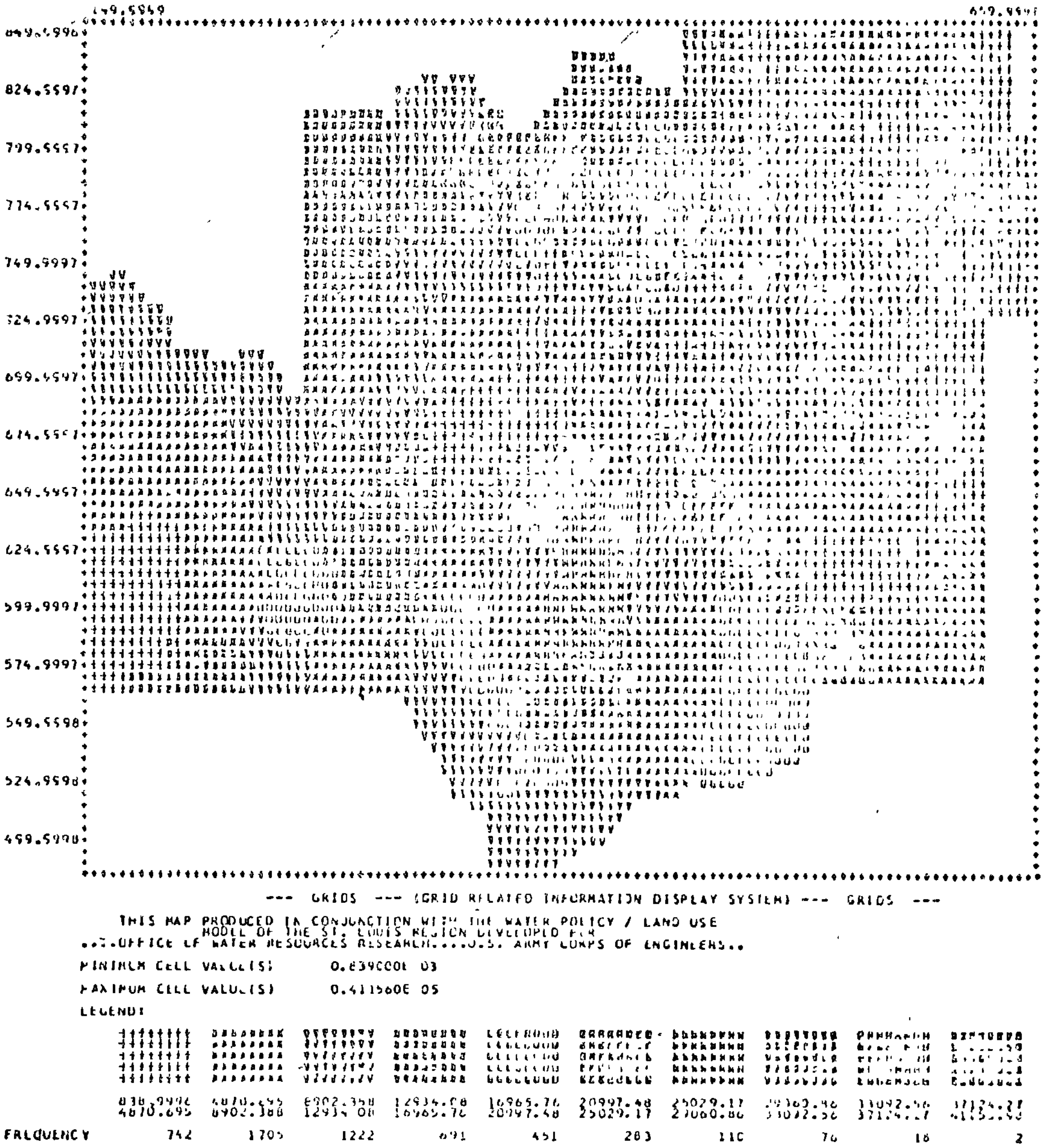
```
                        ***** OFFICE OF WATER RESOURCES RESEARCH (U.S.D.I.) *****                          PAGE  41

                            WATER POLICY / LAND USE MODEL OF  THE ST. LOUIS REGION

                                        DECADE SUMMARY FOR 2020-2030

                           POPULATION                                    ACREAGE                           % CNTY(CNTY)   2020    2030
                  2020    2030  CHANGE(% OF 2020)        2020    2030  CHANGE(% OF 2020)          POP    (RANK)     DEN     DEN

324  INDUSTRIAL    131     149     18(  13%)     *        5       5      0(   0%)     *     14.70(  4)      26      29
     RESIDENTIAL  7632   10344   2712(  35%)     *      386     386      0(   0%)     *      8.11(  7)      19      26
     COMMERCIAL    184     211     27(  14%)     *        9       9      0(   0%)     *     10.28(  6)      20      23
     PUBLIC         71      72      1(   1%)     *       14      14      0(   0%)     *     11.66(  6)       5       5

325  INDUSTRIAL    155     173     18(  11%)     *        5       5      0(   0%)     *     17.01(  3)      31      34
     RESIDENTIAL  8934   11223   2289(  25%)     *      364     402     38(  10%)     *      8.00(  6)      24      27
     COMMERCIAL    248     235    -13(  -5%)     *       22      22      0(   0%)     *     11.43(  5)      11      13
     PUBLIC        105      77    -28( -26%)     *        8       8      0(   0%)     *     12.44(  5)      13       9

326  INDUSTRIAL    167     186     19(  11%)     *       10      10      0(   0%)     *     18.12(  2)      16      18
     RESIDENTIAL 21617   25575   3958(  18%)     *     1469    1489      0(   0%)     *     19.34(  2)      14      17
     COMMERCIAL    455     511     56(  12%)     *       19      19      0(   0%)     *     22.78(  1)      23      26
     PUBLIC        111     117      6(   5%)     *       12      12      0(   0%)     *     18.94(  2)       9       9

327  INDUSTRIAL     90      99      9(  10%)     *        3       3      0(   0%)     *      9.78(  6)      30      33
     RESIDENTIAL 23056   28367   5313(  23%)     *      847     946     99(  11%)     *     21.07(  1)      27      29
     COMMERCIAL    253     322     69(  27%)     *       22      22      0(   0%)     *     15.66(  3)      11      14
     PUBLIC         91     116     25(  27%)     *        9       9      0(   0%)     *     18.83(  3)      10      12

328  INDUSTRIAL    234     255     21(   8%)     *        2       2      0(   0%)     *     23.32(  1)     117     127
     RESIDENTIAL 11153   11316    163(   1%)     *      377     377      0(   0%)     *     10.73(  5)      29      30
     COMMERCIAL    128     128      0(   0%)     *        4       4      0(   0%)     *      8.08(  7)      12      32
     PUBLIC         18      19      1(   5%)     *        1       1      0(   0%)     *      4.48(  7)      18      19

329  INDUSTRIAL    112     123     11(   9%)     *        4       4      0(   0%)     *     12.08(  5)      28      30
     RESIDENTIAL 15566   18633   3067(  19%)     *      706     706      0(   0%)     *     14.60(  4)      22      26
     COMMERCIAL    323     334     11(   3%)     *       25      25      0(   0%)     *     16.26(  2)      12      13
     PUBLIC         99      87    -12( -12%)     *       19      19      0(   0%)     *     14.05(  4)       5       4

330  INDUSTRIAL     31      35      4(  12%)     *        1       1      0(   0%)     *      4.99(  7)      31      35
     RESIDENTIAL 17342   22140   4798(  27%)     *      866     866      0(   0%)     *     17.35(  3)      20      25
     COMMERCIAL    289     319     30(  10%)     *       41      41      0(   0%)     *     15.51(  4)       7       7
     PUBLIC        135     132     -3(  -2%)     *       10      10      0(   0%)     *     19.60(  1)      13      13
```

Model example

An example of the application of the forecasting model was prepared for this article for illustrative purposes. While the entire region was modeled, only Monroe County, Illinois, was selected for specific description here because it has the smallest number of census tracts of any county in the SMSA. Monroe is a rural downstate Illinois county just Southeast of the City of St. Louis; all four land uses were forecasted.

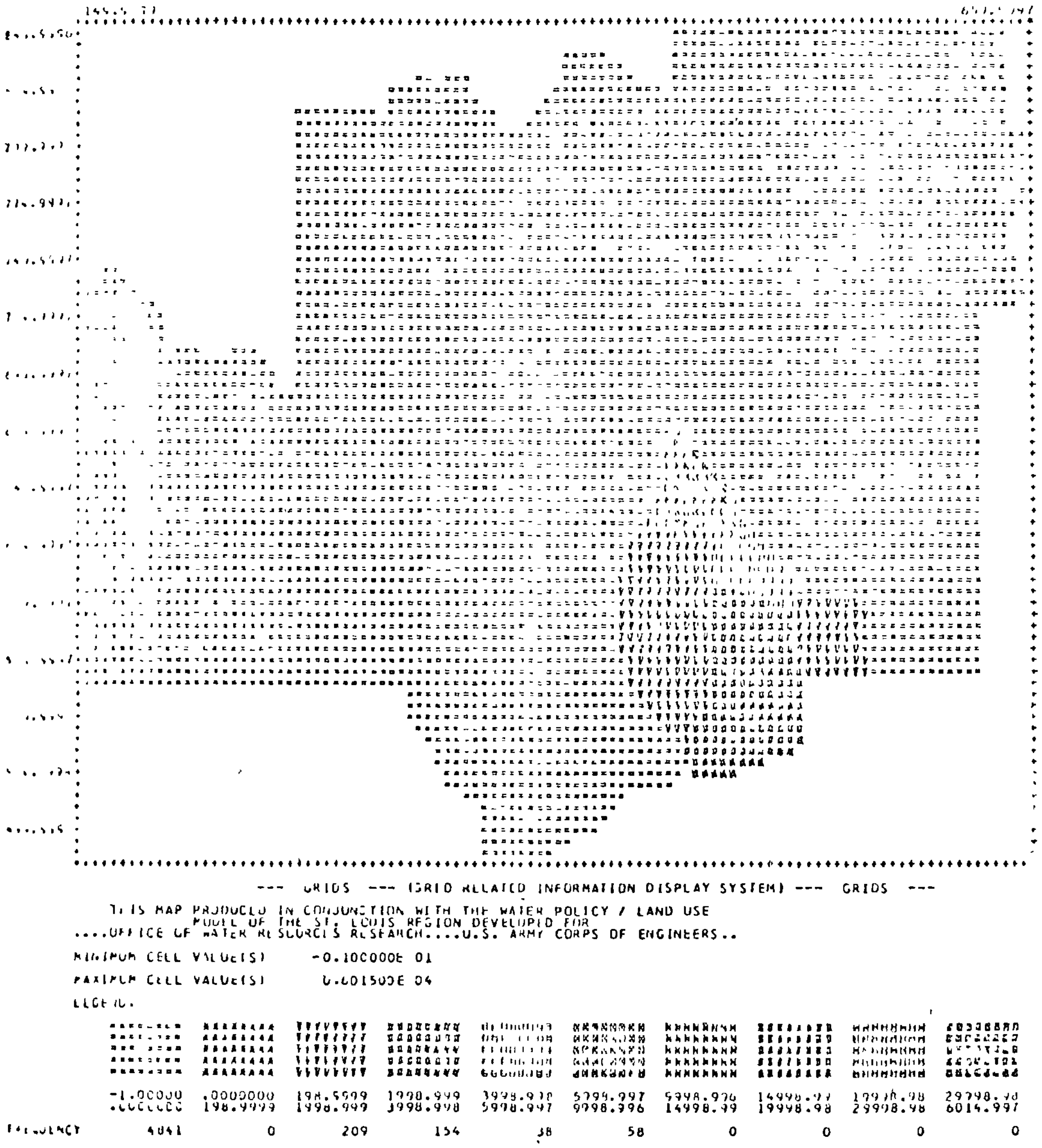
Table 2 lists the control cards specifying the various parameters used in forecasting; the parameter input was supplied by the Regional Industrial Development Corporation, a St. Louis planning agency. Exhibit C lists, by control card number, the explanation of each item of information supplied to the model The field MNRO identifies the input as pertaining to Monroe. Cards 3–19 were common to all counties. The specification of 'CALL 8' on the FILL cards invokes a portion of a user-written subroutine to reverse the general ranking criterion. Recall that the larger the value of a ranking variable, the lower its associated land unit's rank number (more desirable). In order to obtain the opposite effect, i.e : the larger the value of a ranking variable, the higher its associated land unit's rank number (less desirable), use of the subroutine is required.

Tables 3–8 contain the summary reports for each of the decades modeled, 1980–2030. Aggregating the acreage information in these tables, it is possible to construct the incremental acreage allocation by decade. The results are shown in Table 9, which may be used, for example, to see that the large jumps in new acreage for residential use occur during 1980–1990, and then during 2020–2030. The reason for the increase in 2020–2030 can be explained as follows. The initial large acreage allocation during 1980–1990 resulted in considerable residential land available at low densities. This enabled intensive residential



```
          --- GRIDS --- (GRID RELATED INFORMATION DISPLAY SYSTEM) --- GRIDS ---

    THIS MAP PRODUCED IN CONJUNCTION WITH THE WATER POLICY / LAND USE
             MODEL OF THE ST. LOUIS REGION DEVELOPED FOR
...OFFICE OF WATER RESOURCES RESEARCH....U.S. ARMY CORPS OF ENGINEERS..

MINIMUM CELL VALUE(S)       0.839000E 03

MAXIMUM CELL VALUE(S)       0.411560E 05

LEGEND:

           838.9996  4870.695  6902.358  12934.08  16965.76  20997.48  25029.17  29060.86  33092.56  37124.27
           4870.695  6902.388  12934.08  16965.76  20997.48  25029.17  29060.86  33092.56  37124.27  41155.96
FREQUENCY       742      1705      1222       691       451       283       110        76        18         2
```

Fig. 2. 2030 projected population.



```
---  GRIDS  --- (GRID RELATED INFORMATION DISPLAY SYSTEM) ---  GRIDS  ---

    THIS MAP PRODUCED IN CONJUNCTION WITH THE WATER POLICY / LAND USE
          MODEL OF THE ST. LOUIS REGION DEVELOPED FOR
....OFFICE OF WATER RESOURCES RESEARCH....U.S. ARMY CORPS OF ENGINEERS..

MINIMUM CELL VALUE(S)     -0.100000E 01

MAXIMUM CELL VALUE(S)      0.601500E 04

LEGEND.

          -1.00000  .0000000  198.9999  1998.999  3998.998  5998.997  9998.996  14998.99  19998.98  29998.98
          .0000000  198.9999  1998.999  3998.998  5998.997  9998.996  14998.99  19998.98  29998.98  6014.997

FREQUENCY     4041         0       209       154        38        58         0         0         0         0
```

Fig. 3. 1970 Monroe County residential population.

Table 9. Acreage increment summary—Monroe County

	Current use	Increments during					
	1970	1970–1980	1980–1990	1990–2000	2000–2010	2010–2020	2020–2030
Industrial	29	1	0	0	0	0	0
Residential	3012	2023	0	0	0	0	137
Commercial	79	0	0	0	30	35	0
Public	71	0	1	1	0	0	0

development on these lands for approximately the next 40 years. By the year 2020, however, the densities in all land units approached the maximum specified, so that extensive use prevails during this period.

The map of the forecasted residential population for 2030 for the entire region is shown in Fig. 2, while Fig. 3 displays the 1970 residential population for Monroe County only.

REFERENCES

1. Aleshin S. and Mednitzki V. Algorithm problems in an interregional inter-industry model, *Regional Science Association Papers*. Vol. XXIV (1970).
2. Anderson Theodore R. Social and economic factors affecting the location of residential neighborhoods: *Papers and Proc. of the Regional Sci. Assoc.* 9 (1962).
3. Beckmann Martin J. The analysis of spatial diffusion processes, *The Regional Science Assoc. Papers*. 25 (1970).
4. Berry B , Simmons James W. and Tennant Robert J. Urban population densities structure and change, *The geogr. Rev.*
5. Boyce David E. Communication in the field of urban development models, Urban development models. Special Report 97. Highway Research Board.
6. Burns Leland S. A programming model for urban development, *The Regional Science Assoc Papers*. 11 (1963)
7. Clark Colin. The location of industries and population, *The Town Planning Rev*. 35, 281–298 (1964)
8. Colenutt R. J. Building linear predictive models for urban planning, *Regional Studies* 2, 139–143 (1968).
9. Cripps E. L and Foot D. H. S. A land-use model for sub-regional planning, *Regional Studies* 3, 243–268 (1968).
10. Friedmann John. Regional economic policy for developing areas, *The Regional Science Assoc. Papers*. 11, 1963.
11. Goldberg, Michael A. An economic model of intrametropolitan industrial location, *J. Regional Sci.* 10, 75–79 (1970).
12. Hamburg John R., Creighton Roger L and Scott Robert S Evaluation of land use patterns, *Urban Development Models* Special Report 97, Highway Research Board.
13 Herbert John D. and Stevens Benjamin H. A model for the distribution of residential activity in urban areas, *J. Regional Sci*. 2, 25–41 (1960)
14. Jandy Geza. Problems of location in operations research models, *Regional Science Assoc. Papers*. 22, 195–202 (1969).
15. Lee Douglass B. Jr. Models and techniques for urban planning, Cornell University. Buffalo, New York, Sept. (1968).
16. McLaughlin J. B. Simulation for beginners: the planning of a sub-regional model system, *Regional Studies*. 3, 313–323 (1969)
17. Niedercorn John H. and Kain John F. An econometric model of metropolitan development, *Papers and Proc. of the Regional Science Assoc* 11, 1128–1276 (1963).
18. Pitts Forrest R. Problems in computer simulation of diffusion, *The Regional Science Assoc. Papers*. 11, 941–1015 (1963).
19. Rhodes T. Computer models in sub-regional planning, *Regional Studies* 3, 331–336 (1969)
20. Shoup Donal C The optimal timing of urban land development, *The Regional Science Assoc. Papers* 25, 231–257 (1970).
21. Tolley George S. Population adjustment and economic activity. three studies, *The Regional Science Assoc. Papers*. 11, 561–573 (1963).

22 Wilson A G Developments of some elementary residential locations models J. *Regional Sci* 9, 377–285 (1969)

23. Wilson A G Models in urban planning. a synoptic review of recent literature, *Urban Studies* (1968)

24 Wilson A G. Research for regional planning, *Regional Studies* 3, 3–14 (1969).

25 Wilson Alan. Disaggregating elementary residential location models, *Regional Science Assoc Papers*. XXIV, 1219–1243 (1970).

Exhibit A. Parameter input

Type	Description	Default
TO	Specifies the last decade to be modeled	2030
DO	Specifies the counties and land uses to be modeled	all
POP	Specifies the county population by decade, by use	required
DEN	Specifies the county density (people/acre) by decade, by use	required
DISP	Specifies the county dispersion percentage by decade, by use	zero

Exhibit B Control input

Type	Description	Default
FILL	Specifies the variables to be used in ranking the land units, by county, by use	required
ZONE	Specifies the maximum acreage in the named land unit to be assigned to a particular use, by decade	ignored
HOLD	Specifies the amount of acreage of vacant and/or agricultural land to be withheld in the named land unit, by decade	ignored
RANK	Specifies a rule used to alter the land unit ranking, by county, use, and decade	NA
XDEN	Specifies a rule used to alter the input density of named or classed land units, by county, use, and decade	NA
EXCP	Specifies a rule used to alter the results of the initial allocation phase of named or classed land units, by county, use, and decade.	NA
STAT	Specifies a rule used to alter the computed density at which vacant and agricultural land is developed for named or classed land units, by county, use, and decade.	NA

Exhibit C Example control information

Industrial (IND) Card number†	Description
3	Rank desirability directly with the access to transportation code—industrial (interstates, airport, truck, rail, barge)
4	Rank desirability directly, with previous-decade employment
5	Rank desirability inversely to previous-decade residential population
6	Rank desirability inversely to recreational acreage
20	County forecast population by decade, 1980 through 2030
21	County forecast density by decade, 1980 through 2030
22	County forecast dispersion percentages by decade, 1980 through 2030.
23	Rank desirability inversely with distance to Interstate 255 (this is a growing residential area)
24	Rank desirability inversely to distance from the southern most part of St. Louis
25	Increase by 112 people/acre the input density of land unit 328
26	Increase by 10 people/acre the density after initial allocation of any land unit which is either a major or minor non-contiguous community
27	Decrease by 20 people/acre the vacant-agricultural density computed for any land unit whose change in industrial employment after initial allocation is less th[illegible]
Residential (RES)	
7	Rank desirability directly with the access to transportation code—residentia (interstates only)
8	Rank desirability directly with previous-decade population
9	Rank desirability inversely to the current change in industrial employment
10	Rank desirability inversely to the existence of office centers
28	County forecast population by decade, 1980 through 2030
29	County forecast density by decade, 1980 through 2030
30	County forecast dispersion percentages by decade, 1980 through 2030
31	Rank desirability directly with distance to Interstate 255
32	Increase by 5 people/acre the input density of any land unit which is either a ma minor non-contiguous community
33	Decrease by 5 people/acre the density after initial allocation of those land whose change in industrial employment exceeds 100
34	Decrease by 5 people/acre the vacant-agricultural density computed for those units whose change in residential population after initial allocation is less than
Commercial (COM)	
11	Rank desirability directly with the access to transportation code—residential
12	Rank desirability directly with previous-decade commercial employment
13	Rank desirability directly with the current change in residential population
14	Rank desirability directly with the existence of office centers
35	County forecast population by decade, 1980 through 2030
36	County forecast density by decade, 1980 through 2030
37	County forecast dispersion percentages by decade, 1980 through 2030
38	Double the importance of the rank element associated with the current chan residential population
39	Rank desirability directly with the distance to Interstate 255
40	Rank desirability directly with the current change in industrial employment
41	Decrease by 5 people/acre the vacant-agricultural density computed for any unit whose change in residential population is less than 10
42	Decrease by 10 people/acre the vacant-agricultural density computed for any unit whose change in commercial employment after initial allocation is less than
Public (PUB)	
15	Rank desirability directly with the current change in commercial employment
16	Rank desirability directly with previous-decade public population
17	Rank desirability directly with the current change in residential population
18	Rank desirability directly with recreational acreage
19	Rank desirability directly with the existence of office centers
43	County forecast population by decade, 1980 through 2030
44	County forecast density by decade, 1980 through 2030
45	County forecast dispersion percentages by decade, 1980 through 2030
46	Double the importance of the rank element associated with the current chan residential population
48–48	Triple the importance of the rank element associated with the current chan commercial employment
49	Decrease by 5 people/acre the vacant-agricultural density computed for any land whose change in commercial employment is less than 10

† From Table 2.

Comput. & Urban Soc., Vol. 1, pp. 159 to 168 Pergamon Press, 1975. Printed in Northern Ireland

A COMPUTER SIMULATION OF URBAN GROWTH

JARIR S. DAJANI
Assistant Professor of Civil Engineering and Policy Sciences, Duke University, Durham, N. C., U.S.A.

and

MARK W. REINHARDT
Graduate Student Division of Engineering and Applied Physics, Harvard University, Cambridge, MA., U.S.A.

(*Received* 26 *August* 1974)

Abstract—This paper presents a computer model for simulating urban growth, given a set of economic and transportation parameters. The model is designed for both forecasting and pedagogic purposes. It allows the planner or the student to produce a spatial forecast of land uses in a given geographic region, on the basis of exogenously determined growth parameters, employment patterns and transportation alternatives. The program calculates future levels of activity and land areas required for each activity. It also allocates land uses in space. These land-use allocations are then printed out for each future year, and are determined on the basis of minimizing the time spent in transportation between each new activity and all existing activities in the region.

INTRODUCTION

FORECASTING future urban development patterns which might result from alternative policy options is an integral part of any planning effort. This, however, is not an easy undertaking. A number of complex simulation models have been developed for this purpose during the past two decades. These include such models as Chapin's probabilistic model of urban growth [1], Hills' Empiric model [2], Lowry's Model of Metropolis [3], Herbert and Stevens' Regional growth model [4], Schlager's land use design model [5], and others. Comprehensive reviews of the state-of-the-art in urban and regional modeling are given elsewhere [6-8]. Recent attention has been paid to the utility of these models in exploring the complex interaction between transportation and land use and providing a useful input into the formulation of urban and regional growth policies [9]. Parallel recent interest has focused on the development of computer programs and algorithms for the generation and evaluation and selection of architectural and industrial layouts and space allocation schemes [10, 11]. The computerized model which is proposed in this paper combines some of the relevent characteristics and approaches which are used by urban modelers and industrial engineers. Its purpose is to allow the planner to explore the interface between transportation system design and the form that city growth will take. The computer calculates the growth of the city on a yearly basis and shows this growth in the form of changed land uses. Least travel time between complimentary land uses is used as the determinating factor in locating land uses in the city.

The purpose of the model is to allow the planner to produce a spatial forecast of land uses for a given urban settlement, under a predetermined set of conditions. These conditions include growth parameters, employment patterns, and transportation alternatives. It is intended that the model will evolve into a tool which is useful in evaluating alternative transportation and growth policies, especially as they might affect future land use patterns.

THE MODEL

The urban region is represented by a square matrix X of variable size $w \times w$. I element in the matrix represents a city block or zone. These zones can be defined by su imposing a square grid pattern on the city, but they can also take any irregular sl provided that each zone is not contiguous to more than four other zones. The transfc ation from irregular shapes to the square matrix is made by defining those contigu zones, and by describing the connectivity of these zones by a transportation netw Each zone can be assigned any one of the following land uses:

(1) Primary Industrial (I)—Areas of manufacturing and finance industries dealin a scale much larger than the city being modeled. The primary function of t industries in the city is that of an employer.
(2) Residential (R)—Areas of housing for the city's inhabitants.
(3) Service Industrial (S)—Areas of service industries (stores, schools, churches, ernment, etc.) catering to the needs of the city's inhabitants.
(4) Parks (P)—Open land specifically allocated to recreational use.
(5) Open Land ( )—Land available for development into one of the above uses.

In addition, other land uses can be entered into the matrix by assigning characters c than the code letters given to the corresponding matrix points. Since these characters be overlooked by the computer in assigning land uses, the areas represented mus developed to their ultimate use or be areas where development is not feasible suc rivers, mountains, floodplains, etc.

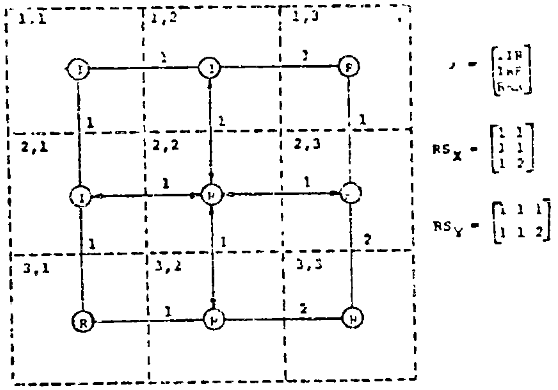


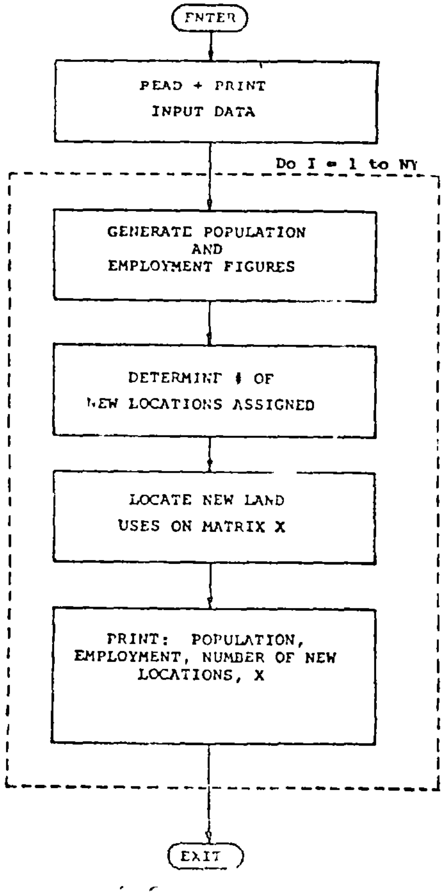
Fig. 1. Land use and road speed matrix.

The region's transportation system is represented by two other matrices. These ma represent connectivity and speed along any two major axis (say, north–south and east–w and are denoted by RS_X and RS_Y. Each element in these two matrices represents a between two zones in the X matrix. It should be noted again, that the use of the two m axis is only conceptual and represents a general direction of movement rather than a str orthogonal transportation network. An idealized 3 × 3 land use matrix representi primary industrial and residential region is shown in Fig. 1. Superimposed on it transportation network whose connectivity and speeds are shown by the two ma described above. The sizes of the matrices RS_X and RS_Y will depend on the size of the land-use matrix X. For a $w \times w$ land-use matrix, RS_X will be $(w) \times (w-1)$ and RS_Y will be $(w-1) \times (w)$. The values of the different elements of these two matrices represent the relative speeds along the respective links. If, for example, all entries representing local arterial streets are given a value of 1, then freeways may be represented by an entry of 3, indicating that they run at three times the speed. Values of 0 can also be entered to indicate the absence of a connection or a transport link between two contiguous land use zones.

Once an urban region is described in terms of land use and road speed matrices, the model proceeds to perform the following four functions:

(1) It generates aggregate population and employment growth figures for any number of predetermined sequence of future years.
(2) It calculates the corresponding changes in land use which are required to accommodate the projected growth.
(3) It locates new projected land uses on the land use matrix, using a maximum accessibility or minimum total travel time criterion, as described below.
(4) It produces an updated land use matrix and a set of updated statistics for the urban region at the desired future year.

These four steps are shown in Fig. 2 and discussed below.



Population, employment, and land use projections

Population and employment projections are made by applying the economic method. This method is dependent on a number of ratios that represent quantit relationships between different parts of an urban economic system [12]. On the bas ratios representing the proportion of employment in primary or export industries to employment and to employment in local service industry and the ratio between employ in either of these types of industries to total urban population, the planner can esti future employment in primary and service industries, as well as expected future populat This, of course, assumes that the system of ratios remains applicable through the projec period. This assumption is not very critical in this model, because projections car made on the basis of yearly increments, and thus the ratios can be changed at any fu date, if so desired.

Once figures are obtained for future primary industry employment, (PIE_K), ser industry employment (SIE_K) and total population (POP_K) in a given year K, the prog proceeds to compute the number of zones in matrix X to be allocated to each of the land uses described above. This is done by applying scale and density parameters wh are predetermined and entered with the input data. The scale parameter establishes size of the land area or zone to be represented by an entry in the land-use matrix X. given as the number of zones per square mile. The density parameter represents the nun of employees (or residents) to be accommodated per square mile, in a zone of a given l use classification. Three such density parameters are given, corresponding to prim industrial, service and residential classifications. The number of zones needed to acc modate recreational park demands is obtained by using a planning standard expressec the number of residents served by one square mile of parkland.

Land use allocation

In the version of the model described in this paper, the land-use allocation proces based on the premise that all newly generated activities locate so as to minimize to travel time to some set of other existing activities in the urban region. Primary industr for example, are concerned with minimizing their total travel to all other primary indust already located in previous rounds. Service industries are concerned with accessibi to their patrons, the residents of the region, as well as to their suppliers in both the prim and service sectors. They will thus tend to locate so as to minimize their total travel these three groups. Residents will need to travel to all four land-uses and will thus atten to minimize total travel time to all existing facilities in the region. Parks would be ration located so as to minimize total travel time to these parks from all residential areas.

Total travel time, tt_{ij}, between a given zone i and all other zones of land-use classificati j is defined by number of trips expected to originate in zone i and terminate in a zone classification j (n_{ij}), multiplied by the average travel time (t_{ij}) between zone i and all zo of classification j, actually in existence. In other words

$$tt_{ij} = n_{ij} t_{ij} .$$

The average travel time is calculated from the road speed matrix, by adding the distar between the zone in question and all zones of classification j, and dividing by the to umber of zones in that classification. The total travel time for en zone which available for development is the summation of equation (1) above over all classifica which are relevant under the travel time minimization criteria describ

The number [illegible] ip interchanges (n_{ij}) which can be expected to occur per day between two zones of a given classification are obtained by applying the density and scale parameters to trip interchange values which can be estimated in trips/resident (or employee) per day. These latter values are entered in a matrix (N). There could conceivably be a number of zero entries in the matrix denoting improbable interchanges, such as between a pair of zones classified as recreational parks. The present version of the model assumes symmetry in movement, and thus that the number of trips per zone per day between land uses i and j is the same as the number between j and i. A typical interchange matrix is given below, where the x's represent a possible interchange, and thus an entry in the matrix:

$$N = \begin{array}{c} \\ I \\ R \\ S \\ P \end{array} \begin{array}{c} \begin{array}{cccc} I & R & S & P \end{array} \\ \left[\begin{array}{cccc} X & X & X & 0 \\ 0 & X & X & X \\ 0 & X & X & 0 \\ 0 & X & 0 & 0 \end{array}\right] \end{array}.$$

An implicit assumption in the use of this matrix and equation (1) above, is that the number of trips between two zones is solely dependent on land use, and is distance independent. This, of course, is an oversimplification. It has long been established that the number of trips is also a function of distance and trip purpose. An exponential decay relationship exists between the number of trips undertaken and trip length, as hypothesized by the gravity model. The exponent of decay is a function of trip purpose [13]. Both of these factors could be incorporated into the calculation of the total travel time (tt_{ij}) by discounting the number of trips with larger lengths and increasing the number of shorter trips, by factors appropriate to each type of trip interchange.

In order to perform the allocation process described above, the model defines five travel time matrices: one for each of the four land uses considered in the model and a fifth to be used for intermediate travel time calculations. This last one will be designated (*TA*). The other four will be designated *TI*, *TR*, *TS* and *TP*, corresponding to primary industry, residential zones, service industries and parks, respectively. All of these matrices will have dimensions (w) × (w) and each cell in them will thus represent a corresponding zone in the land-use matrix (X). An entry in each of the first four matrices represents the summation of all travel times from all existing zones occupied by the land use represented by the matrix, to the zone under consideration. In other words, an entry in a given zone in matrix (*TI*) represents the total travel time from that zone to all developed industrial zones. By dividing this entry by the number of developed industrial zones, the average travel time t_{II}, which is used in equation (1) above, could be obtained. On the other hand, an entry in matrix (*TA*) represents the least travel time between that zone and the zone under consideration. Whenever a new activity is located in a given zone, the value of the entry corresponding to that zone in the (*TA*) matrix is set to zero, and a shortest route algorithm is used to compute minimum travel times between that zone and every other zone in the matrix [14]. The travel times are determined in proportion to distance and in nverse proportion to the speeds given in the two road speed matrices.

Thus, whenever a new activity is to be assigned a geographic location, the model first earches for such a location, by considering each available, or undeveloped, zone in the X) matrix. For each such zone the total travel time is found, by adding the corresponding ntries in the matrices which are relevant for the activity in question, after multiplying them by the appropriate number of trips. If, for example, it is desired to locate a reside zone, then the total travel time (tt_R) for each available zone is given by

$$tt_R = n_{RI}\,t_{RI} + n_{RS}\,t_{RS} + n_{RP}\,t_{RP}$$

where n_{RI}, n_{RS} and n_{RP} are the number of trips between pairs of land uses obtaine applying the scale and density factors to the appropriate entries in matrix N, and t_{RI}, and t_{RP} are the entries corresponding to the location under consideration in the *TI*, and *TP* matrices, divided by the number of zones developed for industrial, service and uses, respectively. The zone having the minimum value of total travel time tt_R ident the location to be used for the residential activity being considered for location. Once location has been determined, it is located on the *TA* matrix, the shortest paths to it f all other zones are identified, calculated and entered in the matrix. The total (*TA*) m is then added to the matrix of residential travel times, and an updated matrix $TR^* = T$ *TA* is obtained, which will be used in calculating the minimum travel time location of next activity.

In any given year, the model must allocate growth in each of the four land-use catego The allocation of activities to available zones is done sequentially. The first activit be allocated is that of primary industries, since it is assumed that this is the basic act that will create new employment opportunities and attract an increase in populat which in turn will generate an increase in the demand for service industries and recreati facilities. Once primary industries are located, residential zones, service industries parks are located, in that order. This location sequence can be easily varied in the mo r order to satisfy a different set of assumptions.

Figures 3–5 show the computer output for a sample run. In Figure 3 the input dat printed out. This includes the Road Speed Matrices and the matrix (X) which include number of existing land uses. Figures 4 and 5 show the printout describing the spa patterns of growth over the 5 yr period between 1979 and 1984.

SUMMARY AND CONCLUSION

The computer simulation model presented herein is an attempt to present a techni which has potential usefulness in simulating the spatial aspects of urban and regio growth. Upon making projections of a region's future population and employment, the two categories of basic and service industries, the model proceeds to estimate l demands in four general land use categories, which are required to accommodate the p jected growth. It then allocates these activities on the basis of some set of criteria pertain to the desirability of activities to be accessible to other existing activities in the regi The spatial location is determined, to a large extent, by the quality of the transportat system: a main variable which can be changed and its influence on the emerging land- pattern tested. Factors other than transportation, such as land value, for example, have been built into the model, but their incorporation is feasible. Other refinements suc the use of different residential densities which are tied in to different types of employm the use of peak hour transportation interchanges, and the incorporation of any form trip-distribution model are also feasible.

The model has the capability of being extended and calibrated for a given metropol area and thus used for predicting expected growth patterns in a region, under differ economic and transportation policy considerations. It also presents an approach which pedagogic value for analyzing the interaction between transportation, economic growth land-use.

```
N=        18          RO=          15790        G= 2.49999E-02
DR= 4.39500E+03       DS= 2.53500E+03           T=           1974
KP= 1.75000E+04       TP= 3.00000E+00           PIPI= 9.99999E-02
BP= 1.59999E-01       SIPI= 1.99999E-01         SIST= 1.99999E-01



         KC= 1.44999E-01          DP= 6.10000E+03
         NY=              20      SCL= 4.00000E+00
         [illegible]= 2.19999E-01 RSI= 4.49999E-01
         W= 3.00000E+00           PER=          5:
```

```
[illegible]             [illegible]              --P----------------
[illegible]             111111111112111111       | R      P        |
111111111121111111      111111111112111111       |  P  PP P        |
31111111112111113       113111111112111131       |   PPPPPR   R    |
11333211111111131       111131113112111311       | S PPSSIP P  P   |
11111111133111111       111131111132113111       |  RSPPPSSRR      |
11111111113323111       121211111111213111       |P R  PRRRPS  R   |
11112111111233111       112311113111213111       |PPPPIIII RP S    |
11221111111121111       111321113111123111       |PPPPIIIRISPR     |
11111111331112211       111312111131122211       |PPPPIISPISPPSRRR |
11111111111112122       111312111112112111       |PRPPPPRRRRSRRR   |
11111221112311211       111312211121131211       |     R PRRISS   R|
11333111121111111       113111211121131211       |  SS R  RIP IR   |
13111111111113121       111111211121113121       |SS  BSR PRP   R  |
31111211112111111       111112111112113121       |       I         |
11111111111111112       111112111112113112       |  RSBB  R        |
11112111111111111       111121111112113111       |  BS      R      |
11111111111111111                                ---RR-------SIII--S

   RSX                    RSY                           X
```

Fig 3. Sample computer output: Initial conditions

```
YEAR:                                 1979



TOTAL POPULATION=                     118842         INCREASE=     13780
PRIMARY INDUSTRIAL EMPLOYMENT=         22308         INCREASE=      2598
SERVICE INDUSTRY EMPLOYMENT=           17226         INCREASE=      1998

RESIDENTIAL SQUARES=                      77         INCREASE=         8
PRIMARY INDUSTRIAL SQUARES=               20         INCREASE=         2
SERVICE INDUSTRY SQUARES=                 27         INCREASE=         3
PARK SQUARES=                             27         INCREASE=         3
```

```
--P----------------
| R      P        |
|  P  PP P        |
|    PPPPPP   F   |
| SPPPSSIP P  P   |
|  PSPPSSRP       |
|P PR PPPPRSSPR   |
|PPPPIIIIRPPPSP   |
|PPPPIIIFISPP S   |
|PPPPIISRISPPSRRB |
|PPPPPPPPPPSRPR   |
|  R RPPRFFISSS R |
|  SSPR PPIP IR   |
|SS  PSF PRP IIR  |
|       P         |
|  RSPP  P        |
|  RS       P     |
---RF-------SIII--S
```

Fig [illegible] imple computer output: 1979 projections,

```
YEAR:                                 1984



TOTAL POPULATION=                     134405         INCREASE=     15603
PRIMARY INDUSTRIAL EMPLOYMENT=         25326         INCREASE=      2938
SERVICE INDUSTRY EMPLOYMENT=           19489         INCREASE=      2263

RESIDENTIAL SQUARES=                      89         INCREASE=        11
PRIMARY INDUSTRIAL SQUARES=               23         INCREASE=         3
SERVICE INDUSTRY SQUARES=                 30         INCREASE=         3
PARK SQUARES=                             30         INCREASE=         3
```

```
--P----------------
| R       P       |
|  R  FP R        |
| R PPPPPP    B   |
| SRPPSSIBSP  P   |
|  RSPPRSSRP  P   |
|PPPPPPRPRISSPP   |
|PPPPIIIIRFPPSP   |
|PPPPIIIPISRPPS   |
|PPPPIISPISPRSFRR |
|PRPPPPPPPRSPRFP  |
|  PPRPPPPPISSSRR |
| RSSRRPPPIPIIRP  |
|SSPRISPSPPP II   |
|       SR P   I  |
|  RSRP  P    I   |
|  RS      R      |
---RF-------SIII--S
```

Fig. 5. Sample computer output: 1984 projections.

REFERENCES

1. Chapin F. S. A model for simulating residential development, *J Am. Inst. Planners* 31 (2), 120–12 (1965).
2. Hill D. M. A growth allocation model for the Boston region, *J Am. Inst. Planners* 31 (2), 111–12 (1965).
3. Lowry I. S. *A Model of Metropolis*. Rand Corporation, Santa Monica, California (1964).
4. Herbert J. D. and Stevens B H. A model for the distribution of residential activities in urban area *J Reg. Sci.* **2** (2), 21–36 (1960).
5. Schlager K. J. A land use plan design model, *J Am Inst Planners* 31 (2), 103–111 (1965)
6. Reif B. *Models in Urban and Regional Planning*. Intext Educational Publishers, New York (1973
7. Kilbridge M., O'block R and Teplitz P V. A conceptual framework for urban models, *Manageme Sci.* **15** (6), B246–266 (1969)
8. Hemmens G. C. (Editor) *Urban Development Models* Highway Research Board, Washington, D C (1968).
9. Khan A M Land use and transport interaction: policy and planning implications, *J. High Spee Ground Transportation* 8 (1), 125–152 (1974).
10 Francis R L and White J. A *Facility Layout and Location: An Analytical Approach*. Prentice-Hal Englewood Cliffs, N.J. (1974)
11. Muther R and McPherson K., Four approaches to computerized layout planning, *Indl Engng* 2 (1, 39–42 (1970)
12 Goodman W. I. and Freund E C (Editors) *Principles and Practice of Urban Planning* Internation City Managers' Association, Washington, D C., (1968).
13. Catanese A J. (Editor) *New Perspectives in Urban Transportation Research*. Lexington-Heath, Lexin ton, Mass (1972)
14 Rogers A *Matrix Methods in Urban and Regional Analysis* Holden-Day, San Francisco (1971

APPENDIX 1

THE COMPUTER SYSTEM

The program was written in the PL/1 language. This language was chosen for its efficient charac representation. The computer time used by the model is, of course, highly dependent on the input dat [illegible]cularly important is the matrix size (w) [illegible] (w) of between 15 and 20 was chos[illegible] much of the [illegible] with the model A matrix size in this range allows a reasonable number of [illegible]ations while usi little computer time The program was compiled on the PL/1 optimizing compiler and stored on dis [illegible]

The model consists of a main program and four subroutines, DATA, RNEW, TWO and RAND. Each of these subroutines performs a key role in the model and a discussion of the function of each is given below.

DATA. The function of this subroutine is to read and print the input data, including the Road Speed Matrices RS_X and RS_Y. The Trip Generation Variables are also calculated in this subroutine. DATA is called at the beginning of the program and at any point where a data change is called for.

RNEW. This subroutine establishes the Travel Time Matrices *TI*, *TR*, *TS*, and *TP* from the present Road Speed Matrices and previously located land uses. This subroutine is called in two situations:

1. RNEW is called after DATA at the beginning of the program. It searches the matrix *X*. If initial land uses (*I*, *R*, *S*, and *P*) are assigned in the input data, RNEW will find them and update the Travel Time Matrices to include them.
2. If a change is made in the Road Speed Matrices at some point during the execution of the program, the Travel Time Matrices are set to zero and RNEW is called. The subroutine will recalculate the Travel Time Matrices for all existing land uses and the new Road Speed Matrices.

TWO. Once a land use element is located the shortest travel time from that location to all other points on *X* is represented as the matrix *TA*. This matrix is computed through successive callings of the subroutine TWO. When the element is located (say at point $X(L, M)$), *TA* is initialized at 1000. $TA(L, M)$ is then given the value of 0. The subroutine TWO computes a new point on TA each time it is called. It does this by checking each free point in *TA* that is within 1 link (along the *X* and *Y* axis) of an already assigned point and computing the travel time from the corresponding point on *X* to $X(L, M)$. After all free points are checked in this way the travel times are compared and the lowest is assigned to its corresponding location on *TA*. TWO is called until all locations on *TA* are assigned a travel time.

RAND. This subroutine is a random number generator and takes on a random value between 0 and 99. RAND is called when it is necessary to choose between two or more locations on *X* of equal merit for a particular land use element. Many random number generators have been written for PL/1 which can be used here.

APPENDIX II

The inputs to the model are given below in the order in which they are entered in the data deck:

Overall factors

(1) Matrix Size—(w)—establishes the city matrix *X* as being of size ($w \times w$).
(2) Initial Primary Industrial Employment-(PIE_0)—# persons employed by Primary Industries at the start of the model.
(3) Growth Rate—(G)—the growth rate (%) of the Primary Industries in the city (expressed as a decimal).
(4) Service Constant—(K_s)—percentage of population employed by Service Industries (expressed as a decimal).
(5) Residential Density—(D_R)—average residential density in the city (persons/square mile).
(6) Primary Industrial Density—(D_I)—average primary industrial density (employees/square mile).
(7) Service Industry Density—(D_s)—average service industry density (employees/square mile).
(8) Year at the Start of the Model-(Y).
(9) Number of Years for the Projection—(NY).
(10) Scale—(SCL)—scale of the matrix *X* in squares (matrix locations) per square mile.
(11) Park Constant—(K_p)—population/square mile of parks needed.
(12) Employment Factor—(f_e)—total population/total number of workers.
(13) Period—(PER)—number of years between printouts. (This value does not affect the functioning of the model but controls the printout only.)

Trip generation variables (one way trips/person (or employee)/day)

(14) Primary Industry to Primary Industry—(II)
(15) Residence to Residence—(RR).
(16) Residence to Service Industry—(RS).
(17) Residence to Parks—(RP)
(18) Service Industry to Primary Industry—(SI).
(19) Service Industry to Service Industry—(SS)
(20) Work Trips—(WT).

After these values, predetermined land uses on the matrix *X* are entered. These may be "dead spots" in the matrix which are entered as a character not used by the model but of significance to the programmer. They may also be existing locations of Primary Industries, Service Industries, Residence, and Parks. These

will be considered by the computer and are entered as *I*, *S*, *R* and *P* respectively. The procedure for enter these land uses is as follows:

(21) Number of Predetermined Land Use Elements entered (0 if none)
(22) Coordinates of *X* and character code for each element

The Road Speed Matrices (RS_X and RS_Y) are entered next. All locations on these matrices are given a value of 1 but any or all can be changed here to any whole number between 0 and 9 inclusive.

Horizontal roads

(23) For each set of roads of a given speed enter:
  (i) Number of roads (locations on Road Speed Matrix).
  (ii) Road speed.
  (iii) Coordinates of each road on RS_X
(24) Enter 0 at the end of horizontal road speed assignments

Vertical roads

(25) For each set of roads of a given speed enter:
  (i) Number of roads (locations on Road Speed Matrix).
  (ii) Road speed.
  (iii) Coordinates of each road on RS_X.
(26) Enter 0 0 at the end of vertical road speed assignments.

The data entered into the computer may be changed at any year during the run of the model.

(27) If no data changes are desired, enter 0 0. If data changes are desired, do the following for e change:
  (i) Enter year of change.
  (ii) Enter 1 if road changes are to be made;
  0 if no road changes are to be made.
  (iii) Enter changed data set in the order above. Enter 0 0 after the last changed data set

Comput. & Urban Soc., Vol 1, pp 195–206 Pergamon Press, 1975. Printed in Northern Ireland

A COMPUTER ROLE IN MONITORING URBAN NETWORKS

R. BAXTER
Department of Architecture, University of Cambridge, Cambridge, England

and

R. BRADY
Applied Research of Cambridge Ltd., 4 Jesus Lane, Cambridge, England

(*Accepted* 19 *March* 1975)

Abstract—Service networks of various types have as a common denominator a data structure which describes their geometry and topology. To this data structure can be attached information pertaining to non-spatial attributes of the network. If the data is computerised it becomes attractive to conceive a system which interactively monitors changes in the network. This paper describes the structure of a service network information processing (SNIP) system emphasising the hardware configuration necessary, the software support and the organisation of the data files.

INTRODUCTION

THE INCREASING use of analytical and simulation techniques in the study of service networks is bound to involve added use of the computer. Most of these mathematical exercises demand information about the networks and once this is computerised in a systematic manner it opens a whole new field of applications. One of these is monitoring.

Monitoring implies the continuous examination and assessment of two types of event. Firstly, changes occurring in reality within the system over which an Authority has responsibility and secondly the impact of possible decisions on the system performance. Under the first heading there is an evident need for an up-to-date record of the existing system and means of interrogating and updating this record. For an underground service network this implies a data base containing topological and geometrical information about the network located in a three dimensional space.

The maintenance of such an up-to-date data base is only possible given an efficient and accurate updating system. In order to meet these specifications it can be argued that some form of interactive communication with the data base is sought. Furthermore, the spatial relationships intrinsic to the network suggest a graphical capability when retrieving and modifying user defined parts of the network.

In this paper a computer system, SNIP (Service Network Information Processing), is described which allows for the interactive retrieval, manipulation and modification of service networks. By allowing for the maintenance of an up-to-date data base it satisfies the first of the enunciated requirements of a monitoring system. Through the creation and maintenance of data in logically and consistently structured files it ensures an easy transition to the second, prognostic and diagnostic, role of a monitoring process. Design applications such as pipe-sizing, flow routing and congestion locating become far more tangible exercises when used in conjunction with up-to-date data bases. The establishment of multi-service data bases would permit the immediate identification of conflicts when new services are laid.

The system is intended to be used for any type of network. Typical applications for underground services would be water supply, sewerage pipes, electrical cables, telephone cables, district heating systems and gas mains. For purely descriptive purposes a sewer network is chosen for this application. The system calls for computerised files describing the network topology and geometry, attributes associated with the links and Ordnanc
digital map overlays.

At the outset it is useful to explain briefly the hardware environment in which t
system operates The operator works from a storage tube connected to a com
might be a multi-access system or a dedicated mini). In the example being descr
storage tube is a Tektronix 4010 with a linked hard copy unit. The user is able to cc
cate with the program in the computer using both the keyboard and the cursor
controlled manually by two thumb screws (one for vertical alignment, the other for
tal alignment). The physical link to the computer is achieved using a modem and
Office Datel service. The ability to receive 1200 baud is made possible by a Modu
processor acting as an interface to the main Atlas computer. Atlas has both r
tapes and disc as backing store.

The various operations that can be performed using the SNIP system are outli
lowed by a more detailed description of the graphical display used and the struct
the data files.

THE SNIP SYSTEM IN OPERATION

The SNIP system operates in a multi-access time sharing environment to wh
attached interactive graphical devices (storage tubes in this example). The main p
is written in Fortran which calls routines to generate displays and handle interactic
calls file handlers to organise the transfer of data from the three files to the main p
The program is compiled and dumped on magnetic disc. This the user loads and
Working interactively it is possible to interrogate and modify the network.

Apart from the main program which controls the interactive tasks, there are the im
subsystems concerned with display and file handling for the data files. These are de
later.

Three different levels of operation can be defined. Operations on the whole n
operations on a particular link, and the editing of the data pertaining to a link
structure is reflected by a hierarchical command structure within the SNIP prograr
user can either be in COMMAND status, LINK status or EDIT status.

Whilst in COMMAND status (see Fig. 1) the user can interact with the system b
tioning the cursor over any of the boxes labelled:

DRAW MAP
LINK?
RESTART
ESCAPE

If the cursor is located on DRAW MAP an overlay may is added to the screen; if l
on LINK? the user enters LINK status; if located on RESTART the screen is clear
the network redrawn; and if located on ESCAPE the session is terminated, the ne
data base held on magnetic disc updated, and the final display drawn on a high resc
flatbed plotter.

When in LINK status the user has to define a link using the cursor, and is then a
operate on that link by positioning the cursor over any of the boxes labelled.

EDIT
STOP EDIT
DISPLAY
REMOVE

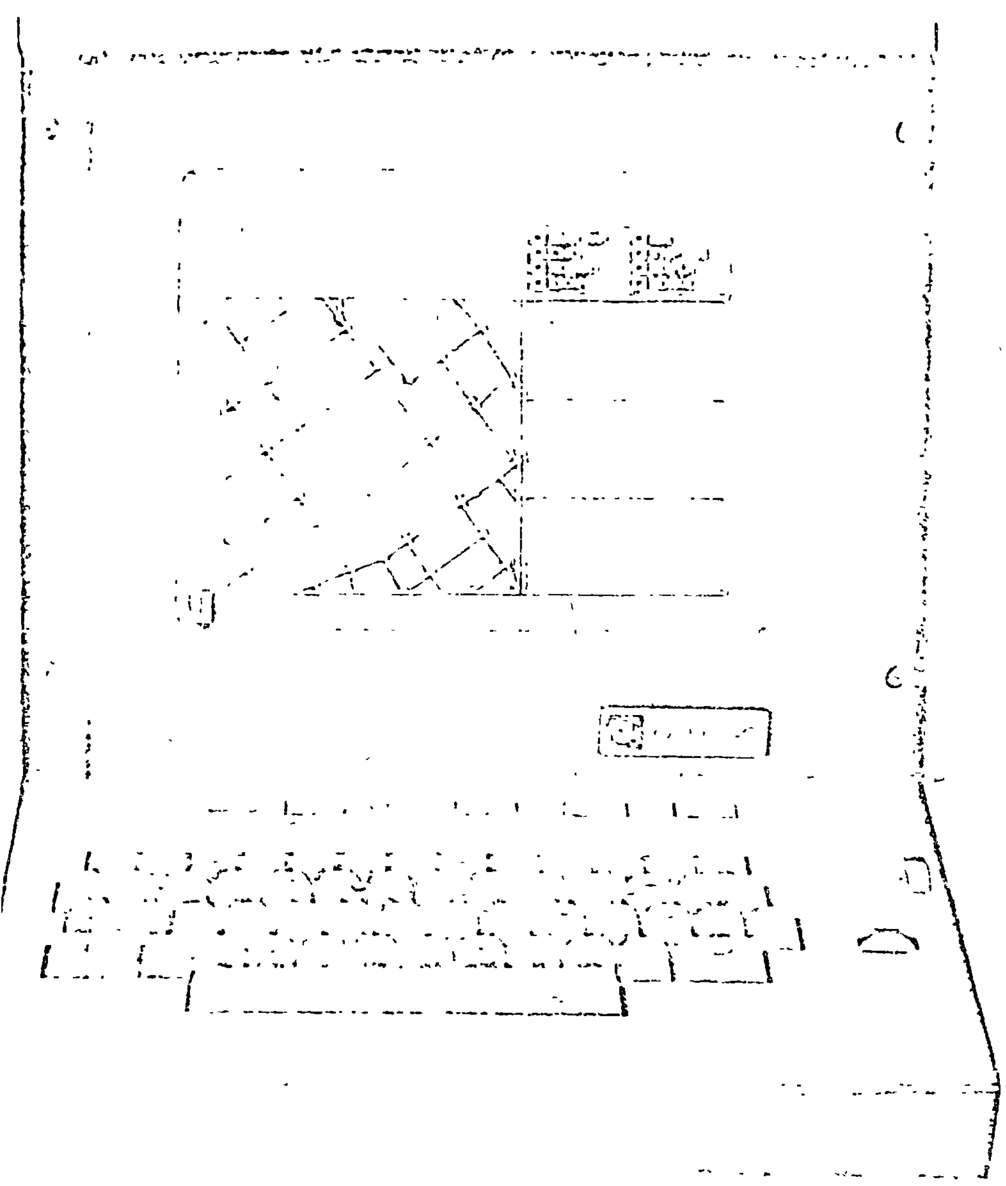


Fig. 1. SNIP system entered

If the cursor is located on EDIT, the user enters EDIT status; if located on STOP EDIT when in EDIT status the user is returned to LINK status; if located on DISPLAY, details of the current link are displayed in one of the right-hand boxes; if located on REMOVE the current link is deleted from the network data base.

Figures 2 and 3 show the various operations that are possible when in COMMAND status. In Fig. 3 LINK status has been entered and link 1–2 set up as the current link, (i.e. the data pertaining to link 1–2 as marked on the screen, is held as a current document for editing, display, or removal.) The coordinates of the end nodes of the link are identified, interactively using the cursor. The data file which holds the information describing the current displayed network is then searched. The input coordinates are matched (to a given level of tolerance) with the coordinates of a link in the data file. This link is set up as the current link. If no match occurs the system responds with an error message.

Figures 4–6 show the operations possible when in LINK status. When details about a link are displayed (as in Fig. 4) the length of the link is calculated from the coordinates of the nodes, and is displayed to scale together with the invert levels and the diameter of the link. The diameter is obtained from a component file which also contains further details. After displaying the information the system responds with another LINK 1–2, reminding the user of his current status.

When EDIT status is entered the system responds by filling in the first two sections of the next information box, with a column of data identifiers, and the values held for current link:

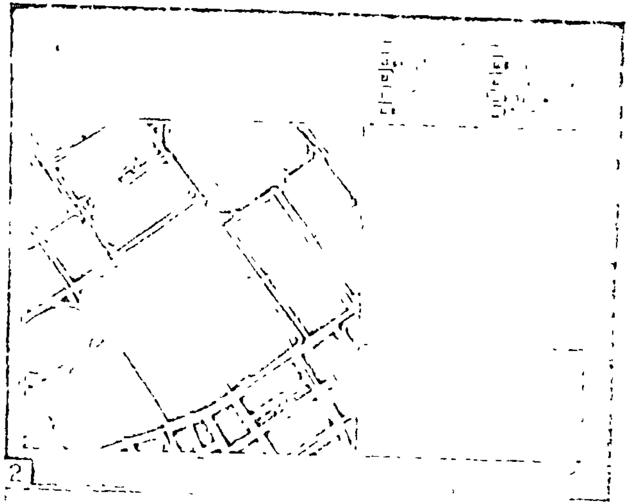
INV1	2434	(downstream invert level of node)
INV2	2485	(upstream invert level of node)
INST	300357	(installation date)
MAINT	300357	(last maintenance date)
FUNCTN	3	(link function)
CODE	333	(component code)

The user can then point to a particular data item using the cursor and type in a changed data value from the console keyboard after the system responds with an '=' sign. Figures 7 and 8 show examples of editing.

It is therefore possible, using the system, to display information about selected parts of a network (the user selects a particular area when entering the SNIP program), spatially reference the network diagram (by overlaying an O.S. digital map of the selected area), and display information about selected links. As well as displaying information the system also allows the user to update the data base by:

(a) editing data held about a link (Fig. 9)
(b) removing a link (Fig. 10)
(c) adding a link (Fig. 11)

Addition of a link is accomplished by default if only one of the two coordinate pairs entered via the cursor can be identified in the data base. A new link is set up with all data values set to zero. Subsequent editing then enters the relevant data.


2


3

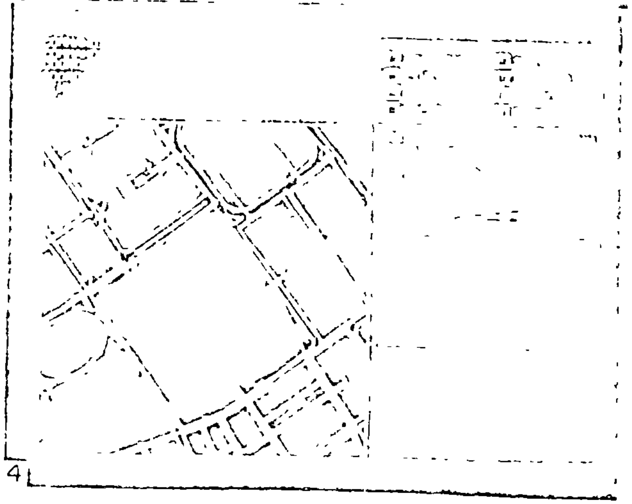
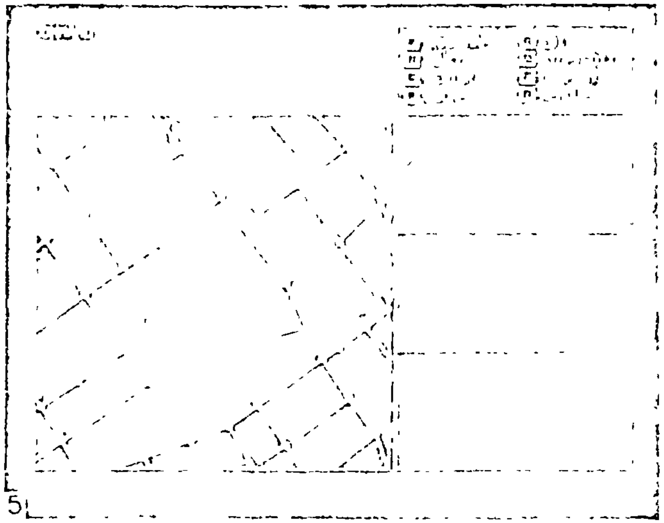

4

Fig 2 Addition of map to the screen
Fig. 3. Link located


5

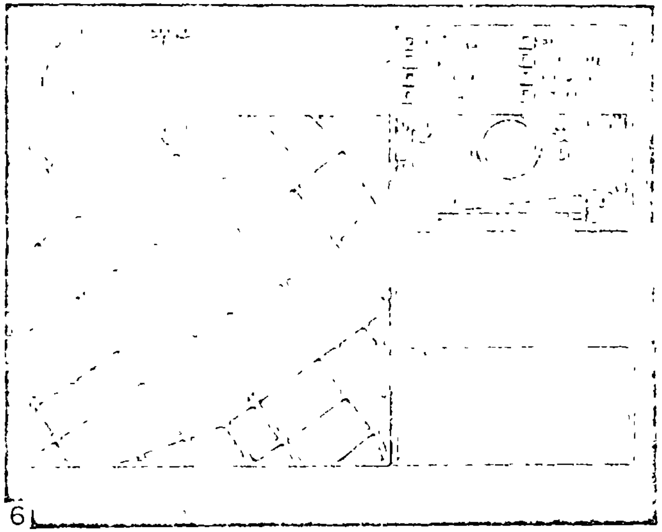
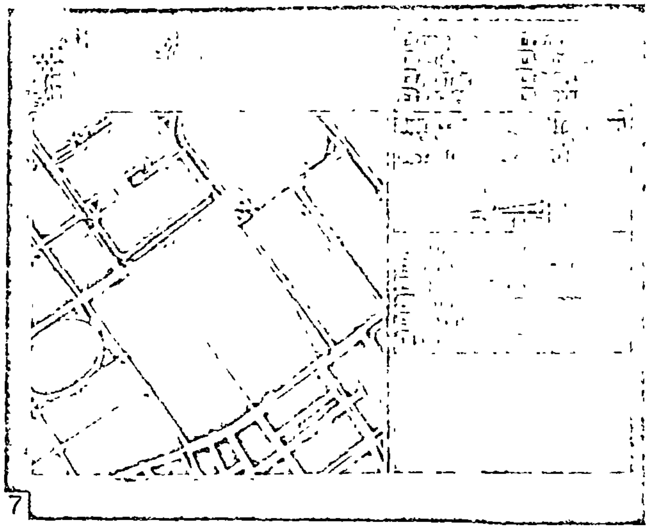

6

Fig. 5. Network with link removed.
Fig 6. New link defined

DISPLAY

The whole system is based upon visual representation of data and the ability to interrogate the resulting picture interactively. This requires suitable software facilities. The software adopted here is GINO-F (although the availability of GINO-F is not a prerequisite of the SNIP system) which is a Fortran derivation of GINO (Graphical Input and Output) as designed and implemented by the University of Cambridge Computer Aided Design Group [1]. The GINO-F system takes the form of a library of Fortran subroutines enabling the user to generate displays A feature of the system is its transportability between mainframe computers (already extant on ICL 1900 series, IBM 360/70. Univac 110 DC6600, PDP11, PRIME, NOVA) and independence of any particular graphical device to which the display is destined. An initialisation routine is provided for each device



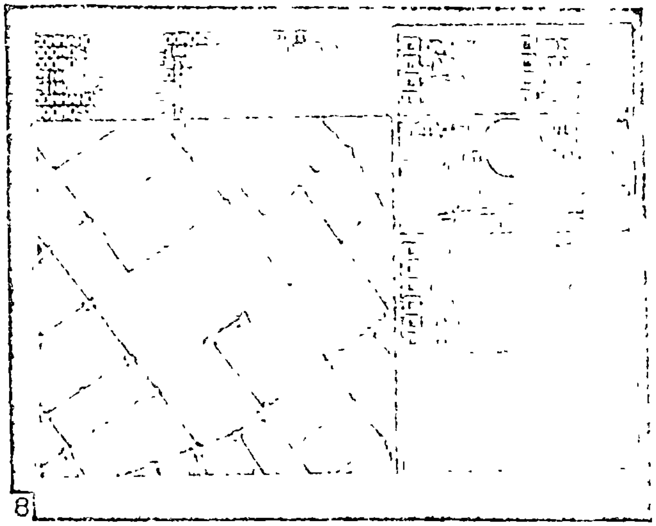


Fig 7. Changes to link information.
Fig. 8. Information added about newly defined link.

and this alone requires modification to redirect the display to an alternative device. This is made possible by a series of post-processors, one for each device type, which are normally provided by an interested party. New devices can be speedily incorporated into the package and used. Apart from transportability and device independence, a graphics language must be able to create pictures.

In the GINO-F system, the unit in terms of which pictures are generated is the picture segment—the smallest unit that may be independently displayed or plotted. Picture segments are generated sequentially by the user, who controls the size and content of the segment. Any number of segments can be viewed together. A picture segment consists of a sequence of picture parts, each added by a call to a GINO-F routine. Routines are available for a group of built-in picture parts which include points, lines, characters and circular arcs. This is supplemented by a facility to create user defined picture parts; the

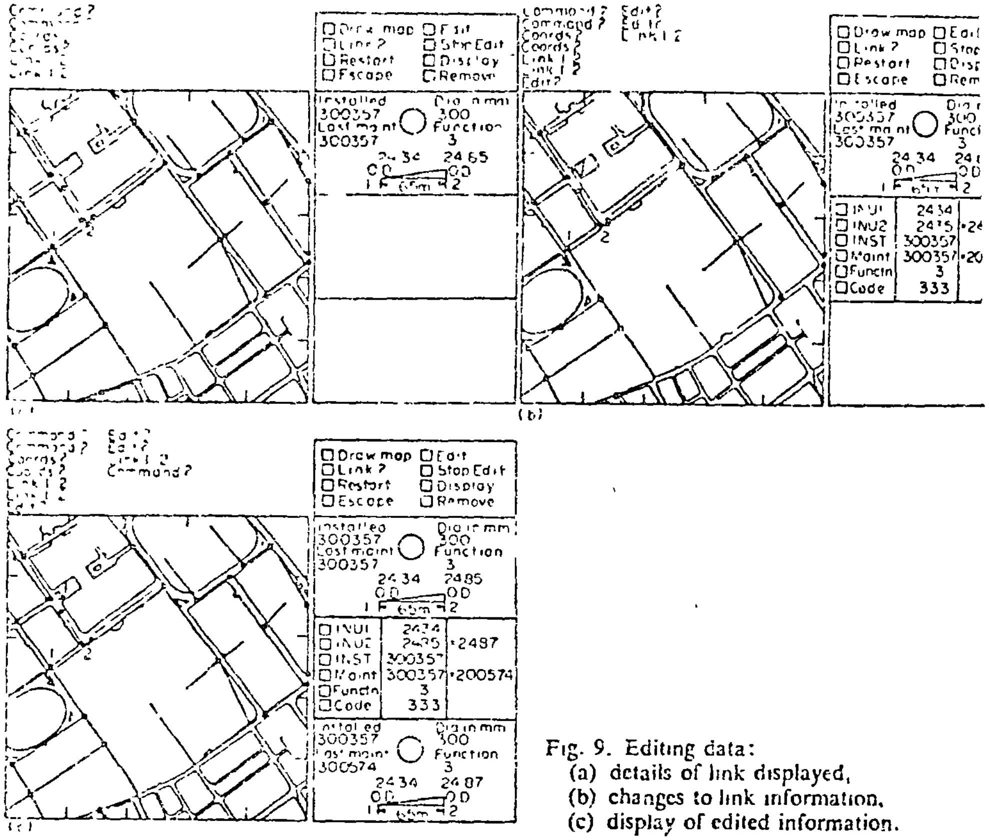


Fig. 9. Editing data:
(a) details of link displayed,
(b) changes to link information,
(c) display of edited information.

user-defined object thus created can then be called as many times as required. Pictu segments can be generated in two or three dimensions the latter being displayed as tw dimensional projections. Windowing or clipping routines exist for both two dimensio and three, enabling the suppression of picture parts falling outside a specified region th allowing for selective viewing of parts of a large picture.

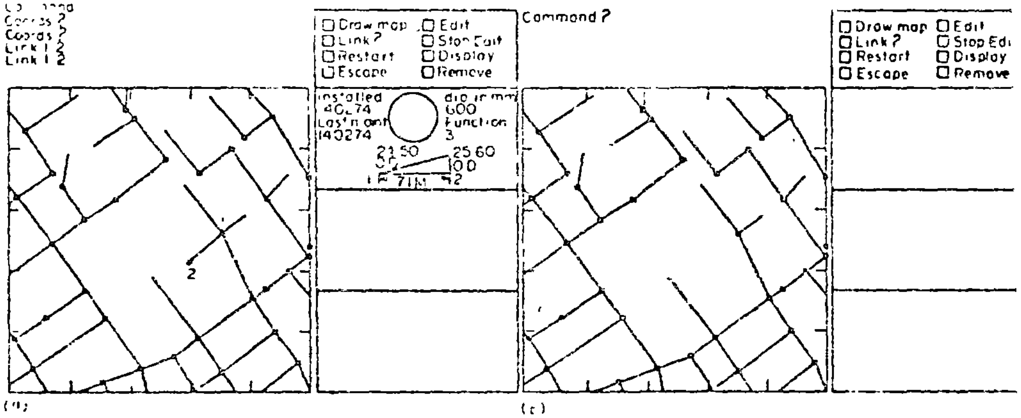


Fig. 10. Removing link
(a) link displayed prior to removal,
(b) network with link removed.

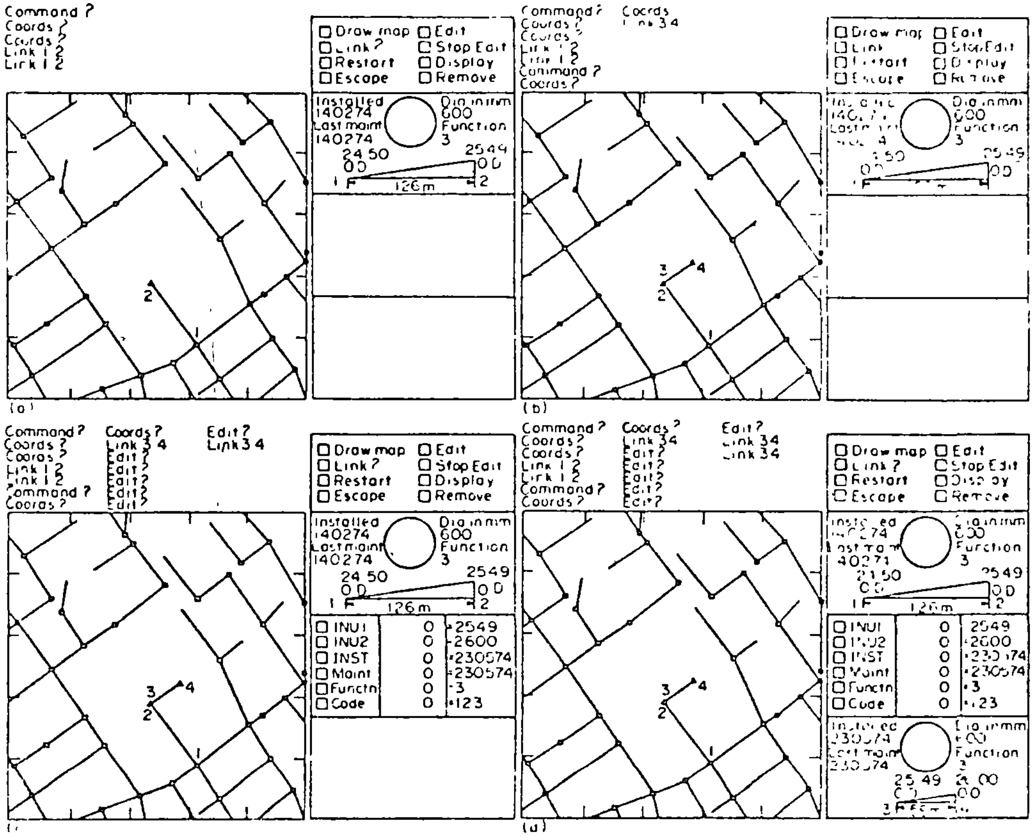


Fig. 11. Adding link.
(a) host link identified,
(b) new link defined,
(c) information added about newly displayed link,
(d) information about new link displayed

Transformations can be applied to all or any subset of picture parts making up a segment. These take the form of translation, rotation, scaling, shearing, reflection, parallel projection or point projection. This means that the user is able to define his picture parts to any scale or orientation that is convenient and then use transformations to position the resulting display as required.

The preceding properties of GINO-F are relevant to all graphics devices but certain extra facilities exist to enable the user to interact with the display tubes. The basis of interaction is the facility to name picture parts. Any picture part, built-in or user-defined, may be given a unique numerical name so that it can be identified by the light pen or cursor. Only named picture parts are sensitive to the light pen or cursor and when a hit occurs on a named picture part, the name is available to the main program. A special program, the interactive handler controls this feedback. Occurrences of hits on named picture parts can be registered and stored by the handler in a list. The list may then be passed to the user's program when required where it is decoded and suitable user-defined action taken. New or replacement picture segments may then be sent to the display to await further interaction. At a more fundamental level the cursor position on the display tube is available for interrogation in the m[illegible]rogram without reference to picture parts. This brief description the display system serves as a mere summary of the definitive specification [2]

DATA FILES

A vital element of the system is the data files on which the program works. Ther three such files: a description of the network, an inventory of the components an digital maps. Each of these files is controlled by a file handler with which the main prog communicates. The file handlers contain algorithms which incorporate a set of rules manipulating the data whilst observing the conventions of the data structure which expr the mutual interrelationships between the stored data. The data structure comprises ac procedures and storage structure. The former extract and insert information from and the latter. The storage structure design needs to anticipate the patterns of access and i tion of data with a view to optimising the access procedure subject to the constraints o local computing environment. Let us see how these principles apply to each of the data

The network file

Service networks take the topological form of branching trees (sewerage) or gr (water). The links and nodes of such networks can have locational identifiers attache them and when spatially referenced in this way they can be mapped and manipul selectively by location. The identification of generic types of spatial data—point, lin area located phenomena—allows us to specify conventions for encoding locational infor tion which permit subsequent cross-referencing. Our example of service networks cle calls for line encoding as a means of recording their locational pattern.

To retrieve both the topology and geometry of the network, the line encoding n record node positions and connectivity. The monitoring function assumes the additio attributes to both the located nodes and their connecting links. There are a numbe alternative conventions which achieve these ends. For continuous unbranching lines (s as contours in a digital terrain model) coordinate strings are attractive. These strings be encoded in three ways. Firstly, each vertex (change in direction) of the string is loc by National Grid coordinates. Secondly, a polar coordinate convention is employed w the string broken down into vectors of given length and direction. Thirdly, the strin described in terms of unit length vectors and only the direction need be given (of wh chain encoding is a variety). For branching lines such as a network it is preferable to enc each node or link as a logical record. In the former case each node is given an alphanum identifier, spatial location and fixed length list of attributes and this is followed by a varia length list of node identifiers to which it is connected, associated with a fixed length list attributes for each link.

Alternatively this file can be inverted to give fixed length logical records represent each link with node identifier, locational reference and attributes of each end node gi together with attributes for the link. Of these possibilities, it is argued that the nature service networks and the monitoring tasks envisaged on them, make the inverted file more satisfactory choice. They have the added advantage of being able to incorpor dynamic information (flows) relating to the network within the defined static data cod design.

The coding design is best explained in terms of our example of the sewage netwo Each logical record in the file represents a link in the network to which is attached ten d items:

(1) National Grid Easting,
(2) National Grid Northing, } of downstream node
(3) Invert level,

(4) National Grid Easting,
(5) National Grid Northing, } of upstream node
(6) Invert level,
(7) Date when link was installed,
(8) Date when link was last maintained,
(9) Function of link (nominal code for foul-sewage or run-off),
(10) Component code of link (for cross-reference with component file).

Initially, this file is created from the digitisation of maps to which the necessary information has been appended. The digitiser output is coerced into the structure described using an independent computer program to translate and rotate the digitiser coordinates into National Grid coordinates and to unscramble the attribute codes entered through the digitiser keyboard. This file is then kept up to date by the SNIP system being described. An example of part of this file is shown below.

The data structure outlined represents the manual form of the file which needs to be held on magnetic disc. The storage structure involves the padding of these short logical records together to give larger physical records (fixed length blocked). The enormity of the

```
500  240  2300  498  225  2305  120652  060173  3  0201
498  225  2305  467  200  2314  200652  200652  3  0202
467  200  2314  426  170  2327  210649  060172  3  0203
426  170  2327  404  155  2335  220652  150371  3  0204
404  155  2335  312  090  2365  290752  290752  3  0205
312  090  2365  271  060  2380  101053  210372  3  0206
271  060  2380  217  040  2400  151153  151153  3  0207
217  040  2400  158  122  2435  201153  010272  3  0208
158  122  2435  125  168  2460  251153  010272  3  0209
125  168  2460  070  246  2484  011253  230472  3  0210
070  246  2484  015  322  2510  151253  010473  3  0211
015  322  2510  066  340  2560  030154  030154  3  0212
217  040  2400  153  016  2420  151153  141156  3  0213
153  016  2420  111  ·000 2430  201153  010174  3  0214
500  355  2400  440  453  2465  200360  010271  3  0316
440  453  2465  394  500  2503  300360  010371  3  0317
498  225  2305  429  319  2350  010460  010460  3  0318
429  319  2350  370  400  2401  100460  190673  3  0319
404  155  2327  352  261  2425  150760  090264  3  0320
352  261  2425  289  342  2456  200760  200760  3  0321
404  155  2335  442  100  2385  090462  090462  3  0322
442  100  2385  481  049  2400  210463  010372  3  0323
481  049  2400  500  000  2440  300462  020967  3  0324
312  090  2365  351  034  2420  120563  130468  3  0325
351  034  2420  375  000  2451  150563  150563  3  0326
271  060  2380  338  000  2420  100453  100453  3  0327
217  040  2400  225  000  2430  230553  060770  3  0328
158  122  2435  038  039  2506  101056  120367  3  0329
038  039  2500  080  012  2530  191056  191656  3  0330
125  168  2460  061  123  2500  211156  211156  3  0331
070  246  2484  080  195  2500  300157  010272  3  0332
070  246  2434  123  265  2485  300357  300357  3  0333
123  265  2485  176  320  2495  130257  130257  3  0334
176  320  2495  260  384  2500  140274  140274  3  0335
015  322  2510  080  310  2540  140274  140274  3  0336
015  322  2510  070  361  2550  140274  140274  3  0337
153  016  2420  161  000  2425  140274  140274  3  0338
440  453  2465  392  418  2475  140274  140274  3  0439
467  200  2314  596  160  2386  140274  140274  3  0440
442  100  2385  360  039  2420  140274  140274  3  0441
481  049  2400  420  000  2440  140274  140274  3  0442
```

file for a bounded area such as a medium sized town is potentially problematic. Since a monitoring exercise is only able to work on a small area at a time it is inefficient to have handle the entire file on each occasion. Since interaction will be on previously defin areal domains it is expedient to divide the data set in this way. For reasons of compatibil with the Ordnance Survey digital maps, each division contains that part of the netwo falling within a 1:1250 map sheet. Since networks are not constrained by the artific boundaries of the map sheet it is necessary to introduce dummy nodes (identified by a mark in the associated data item 3 or 6) which break links crossing the map sheet boundarie The partitioned macro structure is most suited to these storage needs and allows for a members of the file to have the same sequential format and reference to be made to t whole data set or to any specific member.

The file handler controls the access procedures and opens relevant file members on t basis of requests from the main program. It also has to arrange for the updating of t data base held on magnetic disc, a task made more difficult because of the partitioning of t data set and the fact that changes to one member may influence the contents of another. T user is responsible for archiving the data set onto magnetic tape at regular intervals.

The component file

This file contains graphical and engineering descriptions of the links in the form of simple look-up table. The component code attached to each link in the network file is cros referenced with an entry in the table. In terms of components there is a one-to one mappi but the mapping of links on to the component is many-to-one. In the case of our examp of the sewage network the file contains the following information for each code:

1. component code,
2. nominal diameter of the pipe,
3. internal diameter of the pipe,
4. external diameter of the pipe,
5. material from which the pipe has been manufactured,
6. manufacturer,
7. BS number,
8. viscosity coefficient.

The file is ordered in component code sequence and held as fixed length logical recor on magnetic disc. As soon as the program is entered the file is read into core and stored i matrix form. Access to any given component code is by binary search with average searc time tending to $\log_2 N$ for the N component entries. (This is better than the $N/2$ search tim for a linear search.) Hash coding would offer a faster look-up if a suitable hashing functio could be given for the file and which remains invariant during update operations.

In normal cases the file is held on disc and read into core at run time. Consequently the file handler is trivial. If, however, it is not possible to hold the look-up table in cor (because it is too large for the available free core space) than the program must refer to th disc each time a component description is sought. This is only feasible given direct acces entry to the disc file. The advantages of holding this elided information in a separate fil rather than as attributes associated with each link in the network file is that changes to th component file can be made without reference or modification to the network file. Thes changes are achieved using an interactive editor which allows for the removal of logica records, the insertion of new records in their correct sequence and the replacement of data items within any logical record. This phase is outside the main interactive system.

The map file

The developments in the digitisation of large scale maps at the Ordnance Survey are well documented [3,4]. The topographical detail on the map sheet is recorded as spatially located line segments with an associated feature code to indicate precisely what type of detail is referenced. In the SNIP system, the digitised data for each 1:1250 map sheet is held as a member of a partitioned data set on magnetic disc in the way described for the network file. The source data is as supplied by the Ordnance Survey and its manipulation follows ideas reported elsewhere [5].

CONCLUSIONS

This paper has attempted to describe a computer system for the monitoring of network data using visual display tubes interactively. As implied from the detailed explanation of the system, SNIP is operational but requires to incorporate any facilities desired by possible users. A necessary further qualification is that it has yet to be implemented on alternative hardware systems with or without GINO-F being available.

In the current example, the interactive commands are received by the program through the identification of named picture parts on the screen. It is also possible (at times preferable) to use the screen only for graphics and have a linked teletype to enter COMMAND information and to print all alphanumeric responses. In this case the cursor is still used for identifying parts of the display. A whole range of similar modifications are possible given an indication of user demands.

The file handler has three duties. Firstly, it opens the necessary members of the data set covering the area of town specified by the operator in the main program. Secondly, the user may specify selected classes of detail in the output map and the file handler filters the data to satisfy the given classes. In our example only the road outline is drawn because the user specified feature code 14400 (road/pavement border) as the only class of detail required. Thirdly, the file handler has to reconcile the resolution of the output map subject to precision of the original data set, the scale of the output map, and the limitations of the hardware.

Features on the digital map are held as strings of line segments. In transferring these to the screen there is a scale change which, together with the raster size of the storage tube, calls for a spatial filtering of the data set in order to minimise the display file created here as a GINO stream object. A first degree filter is applied which is conditioned by the main program giving information about the scale reduction and the constraints of the hardware. Since there is no updating of this data set, the file handler has no role to play in this respect.

Acknowledgements—Thanks are expressed to the Ordanance Survey, who supplied a magnetic tape copy of the digitised map, and to the Department of Industry's Computer-Aided Design Centre, who provided in-house computing and resources.

REFERENCES

1. Woodsford P. A. The Design and Implementation of the GINO 3D Graphics Software Package, *Software—Practice and Experience*, 1, 335-365 (1971)
2. Computer Aided Design Centre *GINO-F Manual*, Cambridge (1974).
3. Gardiner-Hill R. *Automated Cartography in the Ordnance Survey*, ICA Commission, Paris (1971).
4. Williams E. P. J. *Digitisation of Large Scale Maps*, I.C.A. Commission, Paris (1971).
5. Baxter R. and Shah B. GINO- Graphics Language for Cartography, *Cartograph. J.* **10**, 99–104 (1973).

USO DE PROCEDIMIENTOS DE RAMIFICACION Y DOBLE ACOTACION EN LA SOLUCION DE PROBLEMAS DE LOCALIZACION ESPACIAL*

Simmons ha propuesto y discutido un método orientado a resolver el problema de localización, el cual, a pesar de ser engañosamente simple, puede ser impráctico. Me he interesado en este problema por razones pedagógicas y dedicado esfuerzo considerable a revisar el método de Simmons con un resultado que mejora la solución en tiempo.

También he discutido algunos paralelismos entre los procedimientos heurísticos de la planeación y diseño profesional con el proceso de "branch and bound" —ramificación y doble acotación—. Se tocan muy claramente esos aspectos en esta nota y nos enfocamos por entero a algunos aspectos usuales de la programación del "branch and bound" que se han revelado en este ejercicio.

No se trata de atacar el trabajo de Simmons, pero desde el punto de vista de resolver problemas prácticos, un pequeño incremento en el tamaño del problema requiere un tiempo de computación que excede de las limitaciones, como se muestra claramente en la tabla III.

El problema propuesto por Simmons es el siguiente: Tenemos N cuartos de diferentes tamaños, y una matriz de pagos que relaciona costos de los cuar

---

* Britton Harris. Profesor de Planeación Urbana y Regional de la Universidad de Pennsylvania.

Este problema de asignación fue identificado como un problema de localización por Koopmans y Beckman y la solución fue discutida por Lawler. La solución se facilita por el hecho de que el costo consiste de dos partes que se multiplican y cuyos pares de localizaciones y pares de actividades son independientes.

El problema de Simmons puede ser formulado como un problema de asignación cúbico. En este problema suponemos que el tamaño de los cuartos es específico para cada actividad. Las locaciones j, k y l no son absolutamente fijas pero se refieren al orden en el cual los cuartos son acomodados.

Así:

$$\min_{g,h,i} \sum_{j=1}^{n-2} \sum_{k=j+2}^{n} \sum_{l=j+1}^{k-1} X_{gj} \, X_{hk} \, X_{il} \, C_{gh} \, W_{l}$$

Esta formulación está sujeta a las mismas restrucciones del problema cuadrático precedente.

No hemos experimentado con la solución al problema en esta forma pero si se ha realizado con el procedimiento de "branch and bound".

Una revisión de los medios recomendables para solucionar el problema de Simmons, sugiere algunos puntos de vista de los propiedades del proceso de branch and bound.

En esta revisión hemos identificado cuatro pasos principales característicos en el diseño de un proceso usual y discutido, las maneras de las cuales Simmons y yo las hemos tratado, estos son:

1.- Establecer una útil descripción del árbol de decisiones, las cuales permiten la enumeración de todas las posibles soluciones y las relaciones con otros aspectos del proceso.

2.- Establecer una cota inicial que permita el recorrer muchas ramas del árbol en los primeros pasos. Puesto que es un problema de minimización, esta cota deberá referirse en las discusiones como la cota superior.

3.- Establecer medios para hacer una estimación parcial de la función objetivo en cualesquier punto del árbol. Esta estimación no deberá ser más grande que la función objetivo en el último nodo de la ramificación que se sigue y deberá ser al menos igual a la función objetivo en cualquier nodo terminal del árbol. Los valores de la función objetivo para cualquier nodo no terminal se consideran la acotación inferior, puesto que ellos representan tal acotación para la función objetivo de esa rama. Cualquier valor de la función objetivo en un nodo terminal que sea menor que la acotación superior anterior, lo convierte en una nueva acotación superior. Cuan do todas las cotas inferiores sean más grandes que las cotas superiores corrientes, el problema ha terminado.

4.- Establecer una estrategia de ramificación que seleccione el nodo que deberá ser explorado y que gobierne la siguiente división del universo en subconjuntos.

La representación de Simmons del árbol de decisión está basada en el artificio de localizar cuartos alternos en las terminales opuestas del corredor. Este artificio facilita el establecimiento de las cotas inferiores en los primeros pasos del proceso, puesto que los primeros costos de interacción deberán ser explorados sobre las distancias más largas posibles. He adoptado este procedimiento como inobjetable y totalmente necesario para desarrollar el algoritmo de ramificación y acotación.

He encontrado que Simmons no dedicó mucha atención al segundo requisito de la programación de ramificación y acotación, el temprano establecimiento de una buena cota superior. Este establecimiento afecta la importante estrategia de ramificación como se discute en seguida. Ello acelera el corte del árbol de decisión y reduce el número de pasos de computación necesarios. En un sentido más amplio, cuando el procedimiento de ramificación y doble acotación no puede ser usado de una manera completa, el establecimiento de una cota superior por métodos heurísticos reemplaza la solución del problema. De esto se sigue, que en cualquier caso, la estimación de una solución cerca de la óptima es deseable.

Mis primeros esfuerzos para establecer una buena cota superior estuvieron basados en una extensión relativamente simple de las sugestiones originales de Simmons, usando cambios de técnica y tratando de localizar uno o más cuartos a la vez en lugares superiores. Los resultados de esos trabajos se muestran para ocho problemas típicos, de los cuales cuatro fueron presentados por Simmons, en la Tabla I, columna 1, y pueden ser comparados con la solución óptima en la columna 3. Cambiando técnicas de varios grados de complicación se obtienen bases útiles e interesantes para procedimientos heurísticos, los cuales, una u otra, reemplazan o, como aquí, inician procedimientos de optimización. Una base sistemática para esos métodos fue propuesta en artículos por Reiter y Sherman. Ideas similares han sido aplicadas en el problema del agente viajero por Lin, y una red de optimización por Scott. Una discusión amplia de su aplicación fue presentada por Nugent, Vollman y Ruml con el número de referencias usuales; esos métodos pueden ser usados para los problemas de dos y tres dimensiones mencionadas arriba.

Yo no estuve completamente satisfecho con los cambios iniciales en los resultados y desarrollé algunos procedimientos heurísticos presentados en la segunda columna de la misma tabla, basados en un método sugerido durante una comunicación privada con William T. Mc Cormick Jr. del Instituto de Análisis de Defensa. Este método usa un número de diferentes puntos iniciales, intentando construir una buena solución mediante la localización óptima de cada uno de los cuartos subsecuentes.

Este método heurístico también falla en encontrar el óptimo en un número de casos. Cuando los dos métodos son usados a la vez, primero el heurístico 2 y después un estilo mejorado del heurístico 1, el valor óptimo siempre - fue localizado en todos los ejemplos que se muestran.

No tengo conocimiento de qué desempeño deberá ser esperado de esas combinaciones heurísticas en problemas grandes, puesto que el costo de verificar el óptimo para esos problemas deberá ser excesivo. Debo también puntualizar que en su forma corriente esos métodos requieran de cerca de 10 segundos para resolver un problema de 15 cuartos y el tiempo de cómputo se incrementa de acuerdo a 5^n (esta tasa de incremento puede ser demostrada tanto en el campo teórico como empírico). Un problema de 60 cuartos requiere de aproximadamente tres horas de cómputo. Esto hace que aún estos métodos heurísticos sean imprácticos para problemas grandes, aunque espero que se pueden mejorar, sustancialmente no emprendería este trabajo por ahora.

Un tercer requerimiento en el proceso de ramificar y acotar es el método para generar cotas inferiores, para estimar el desempeño de un sistema parcialmente especificado. Conforme avanzamos en varios niveles del árbol de permutaciones y localizamos más o menos facilidades, ciertos costos pueden ser fijados. Estos son: primero, el costo de interacción entre cuartos que ya han sido localizados y segundo, el costo de interacción entre esos cuartos y aquellos que no han sido aún localizados, desagregando en todo lo posible el orden de los cuartos no localizados.

Simmons y yo usamos métodos equivalentes para estimar esos costos. Diferimos, sin embargo, en como la base de la estimación podrá ser hecha con respecto a los efectos de ordenar los cuartos que aún no han sido localizados. En este aspecto, he hecho dos cambios al procedimiento de Simmons.

Simmons presenta una proposición que es de gran importancia en este proceso. La discusión de esta proposición será tratada en otro trabajo. Este lema prueba que si un número de cuartos está óptimamente localizado con respecto a otro cuarto y bajo las condiciones del problema de Simmons. Esos cuartos pueden ser ordenados en orden decreciente de la relación de interacción del costo al tamaño, como nos movemos lejos de un cuarto fijo. En este proceso, Simmons mantiene el rastro de este orden para los cuartos no localizados, y del costo de interacción, el cual este orden implica con cada uno de los cuartos localizados y considerados separadamente. Sin una prueba formal he extendido este lema para las interacciones de los cuartos no localizados con el conjunto de los localizados. Para cada cuarto no localizado, aquellos cuartos localizados siempre a la izquierda están dando un costo de interacción positivo y negativo, para los cuartos localizados siempre a la derecha, el costo neto total de la interacción se obtiene dividiendo el costo entre el ancho de cada cuarto no localizado, y esos cuartos se localizan en orden decreciente de este índice de izquierda a derecha. Entonces es calculado el costo de interacción con los cuartos siempre localizados.

Este procedimiento involucra substancial adición de trabajo al sacar los cuatos en un nuevo orden en cada vez y en calcular los costos de interacción, pero creo que la reducción en el número de ramas a ser exploradas es bastante substancial, lo que hace que la operación valga la pena. Esta hipótesis no ha sido probada separadamente para todos los otros cambios aquí descritos.

Aquellos cuartos que no han sido localizados en algún punto del proceso, también tienen costo de interacción con cada uno de los otros. Es posible hacer una descripción acerca del nivel mínimo de esos costos de interacción, sin los cuales hemos encontrado el mínimo arreglo, la cual podría, por supuesto, - ser la solución a una pequeña versión de el problema de Simmons. Esta pequeña descripción se expresa de diferente manera: Para los cuartos remanentes - existe un número fijo de pares de interacciones de significancia en el problema. Consideremos que esas tienen el conjunto de mínimo costo a ser encontrado en cualquier lado de la matriz. Al mismo tiempo, esas interacciones toman lugares sobre 1, 2, 3, etc, interviniendo cuartos en cierto número fijo de interacciones. Podemos considerar que cada una de las interacciones de un cuarto está hecha sobre los cuartos más angostos etc., y así se asegura un conjunto de distancias mínimas posibles de la interacción mayor que el número requerido. Si ahora arreglamos las distancias mínimas de interacción en sentido opuesto al mínimo costo de interacción y formamos el producto interno de los dos rectores, tenemos el mínimo estimado para la interacción total. De estas estimaciones yo preparé un avance para tres o más cuartos aún no asignados, pero no he intentado terminarlo. Este dispositivo particular es probablemente más importante en los casos que Simmons llama "problemas difíciles", en los cuales los costos de interacción son uniformemente altos y es más difícil descubrir el óptimo.

El efecto conjunto de mejorar los costos superiores e interiores, es un más rápido corte del árbol de decisión. Esto se ilustra en la tabla II, la cual -- muestra el número de nudos actualmente explorados por el método de Simmons y el mio para un problema particular de 10 nudos. El número de nudos por explorar se reduce, en general, a un tercio.

El cuarto requerimiento, en la construcción del algoritmo de ramificación y acotación, es la elección de una estrategia para ramificar. No me concierne aquí discutir cualquier número de tales estrategias, pero antes de contrastar los méritos de los procedimientos convencionales anteriores usados por Simmons con la programación de "seguir el rastro". Esta, a pesar de su nombre diferente, puede ser, sin embargo, un procedimiento de ramificar y acotar. La programación mediante "seguir el rastro" no incluye ninguna decisión tocante a qué nudos del árbol de decisión serán explorados, pero examina las cotas en todos los nudos de una manera sistemática. Este procedimiento tiene, ciertamente, una gran economía en operación. Cualquier procedimiento alterno que haga uso de una regla de selección para escoger el nudo siguiente a ser explorado, sacrifica algunas de las ventajas del procedimiento de "seguir la huella", e incurre, al final en varias desventajas serias, tales como la necesidad de almacenar los resultados intermedios observados en un gran número de nudos que han sido establecidos y evaluados pero aún no explorados.

La ventaja aparente de seguir una regla de selección, es que esta regla alcanzará más rápidamente una buena cota superior, encontrando una solución que está probablemente muy cerca del óptimo y mejorarla. En un gran número de problemas que parecen heurísticos, los cuales traen consigo una parte del óptimo. Como hemos discutido antes e ilustrado en la tabla I, ésto es particularmente cierto para el problema de Simmons. Puesto que el objetivo de ramificar y acotar, por mucho que se organice, es explorar todas las ramas implícitamente o explícitamente, una estrategia para ramificar que reemplace el procedimiento de seguir la huella, no podrá reducir sustancialmente en número de ramas por explorar si una buena solución es, en principio, recomendable.

En el caso de un buen diseño del procedimiento de seguir el rastro, la mejor ventaja es que los resultados intermedios siempre consiguen ser usados directamente en cada paso del programa al calcular el siguiente. Frecuentemente es completamente posible almacenar esos resultados intermedios para cada nodo a lo largo de la rama que se explora. Para problemas grandes de naturaleza compleja este almacenamiento es extenso pero aún sustancialmente menor que el requerido para un record incompleto en una estrategia de ramificación. Yo creo que del uso consistente de los resultados intermedios, resulta grandemente facilitado nuestro incremento de acelerar las corridas en el problema de Simmons.

Figuras comparativas, en tiempo, de las corridas se dan en la tabla III, aquí, los primeros seis problemas son exactamente aquellos propuestos por Simmons, pero el tiempo mostrado para él está dividido por dos en base a las diferencias entre una 360/65 y una 360/75.

Unas breves palabras pueden ser hechos para observar las aplicaciones en la planeación de estas ideas. Observé problemas de planeación y diseño, y frecuentemente encontré problemas muy grandes de optimización en espacios discretos y con múltiples soluciones óptimas. En tales problemas, un procedimiento completo de ramificación y acotación es prohibitivamente difícil, pero algunos de sus conceptos básicos son aplicables. Primero, especificando la forma de la influencia del problema en todos los pasos finales. Segundo, en cualquier esfuerzo sostenido de planeación existe una cota superior en el mejor diseño, inicialmente ésto podría ser la situación existente. Tercero, puesto que es muy caro explorar un plan al detalle, una buena manera racional de establecer cotas inferiores dentro del costo de los planes es muy deseable; los métodos de juicio usuales para hacer esto son probablemente tardados. Cuarto y finalmente, los procedimientos de exploración que son usados en la búsqueda de un buen diseño en problemas grandes son, en sí mismos, materia de estudio. - Para esta clase de problemas, la programación de seguir la huella, porque, por su naturaleza extensiva es casi inaplicable seguir otro tipo de programación. - Sin embargo, es completamente posible que una combinación de procedimientos heurísticos de exploración y procedimientos de exploración desarrollados para ramificaciones difíciles y problemas de acotación sean aplicados. Existe un fuerte paralelismo entre estrategias de ramificación, estrategias de programación heurística y protocolos de diseñadores, las cuales deberán ser exploradas.

TABLA I

ANALISIS DE METODOS HEURISTICOS / CRITERIOS DE VALOR

No. de Cuartos	Heurístico # 1	Heurístico # 2	Optimo
8	433	440	433
9	1,370	1,349	1,349
10	1,569	1,585	1,569
11	4,389	4,497	4,389
12	6,217	6,003	6,001
13	7,985	8,210	7,985
14	10,735	10,762	10,735
15	16,263	15,857	15,810

TABLA II

NODOS EN UN ARBOL DE PERMUTACION DE DIEZ NIVELES

Número de nodos a ser Explorados

		Ejemplo	
Nivel	Sin cortar	Simmons	Harris
2	45	45	45
3	360	358	320
4	2,520	2,111	777
5	15,120	4,390	750
6	75,600	529	260
7	302,400	201	32
8	907,200	13	6
9	1,814,400	2	2

TABLA III

TIEMPOS DE COMPUTACION EMPLEADOS

No. de Cuartos	Tiempo de solución (Segundos)	
	Simmons ÷ 2	Harris
Problemas difíciles		
8	3.0	.23
9	30	1.81
Problemas medios		
8	1.7	.28
9	11.8	.78
10	32	1.55
11	240	10.20
Problemas grandes		
12		47
13		169
14		679
15		6,400 (est.)

UN SISTEMA DE MODELOS PARA LA PLANIFICACION DE AREAS METROPOLITANAS

Informe elaborado por:

Arqtº Luis Carlos Palacios

ESTE TRABAJO FUE ELABORADO EN COLABORACION CON EL EQUIPO DE INVESTIGACION DE LA UNIDAD DE MODELOS URBANOS AMBIENTALES

INDICE

I. INTRODUCCION

II. MODELO REGIONAL SIMPLE

- 1. Descripción General
  - 1.1 Sistema Geografico y Variables Basicas
  - 1.2 Estructura del Modelo Regional Simple
- 2. El Modelo Económico
  - 2.1 Determinacion de la Inversion
  - 2.2 Determinacion de los Flujos de Inversion y el Empleo.
- 3. El Modelo Geografico
  - 3.1 Determinación de la Poblacion Bruta
  - 3.2 Determinación de las Migraciones y Poblacion Regional.
  - 3.3 Problemas y posibles soluciones

III. MODELO INTRA-URBANO SIMPLE

- 1. Descripcion General
  - 1.1 Sistema de Areas de Análisis

1.2 Estructura del Modelo Intra-Urbano

2. Modelo de Empleo Basico

    2.1 Clasificacion del Empleo

    2.2 Estructura General de Modelo

    2.3 Sub-Modelos de Empleo Básico

    2.4 Restricción de Densidad.

3. Modelo de Edificaciones y Actividades

    3.1 Estructura General del Modelo

    3.2 Sub-Modelo de Edificaciones

    3.3 Sub-Modelo Residencial

    3.4 Sub-Modelo de Servicio

    3.5 Posibles Restricciones.

IV. MODELO DE TRANSPORTE

1. Descripción General

    1.1 Sub-Modelo de Generacion

    1.2 Sub-Modelo de Distribución y División Modal

    1.3 Sub-Modelo de Asignación

    1.4 Sub-Modelo de Costo Generalizado.

INTRODUCCION

El conjunto de Modelos que se describen a continuación es un esfuerzo para utilizar la modelación en la Planificación Integral de los Sistemas Urbanos dentro de un contexto regional y prestando especial importancia a las variables que los organismos gubernamentales podrian controlar (Infraestructura, Servicios, Inversion etc.).

Los Modelos son de dos tipos:

a) Modelos descriptivos, para simular politicas

b) El Modelo Evaluativo, que con un conjunto de criterios y normas evaluan politicas de desarrollo.

Los Modelos descriptivos estan estructurados a tres niveles ó constituyen tres grandes modelos. Ellos son:

1. Modelo Regional
2. Modelo Intra-Urbano
3. Modelo de Transporte.

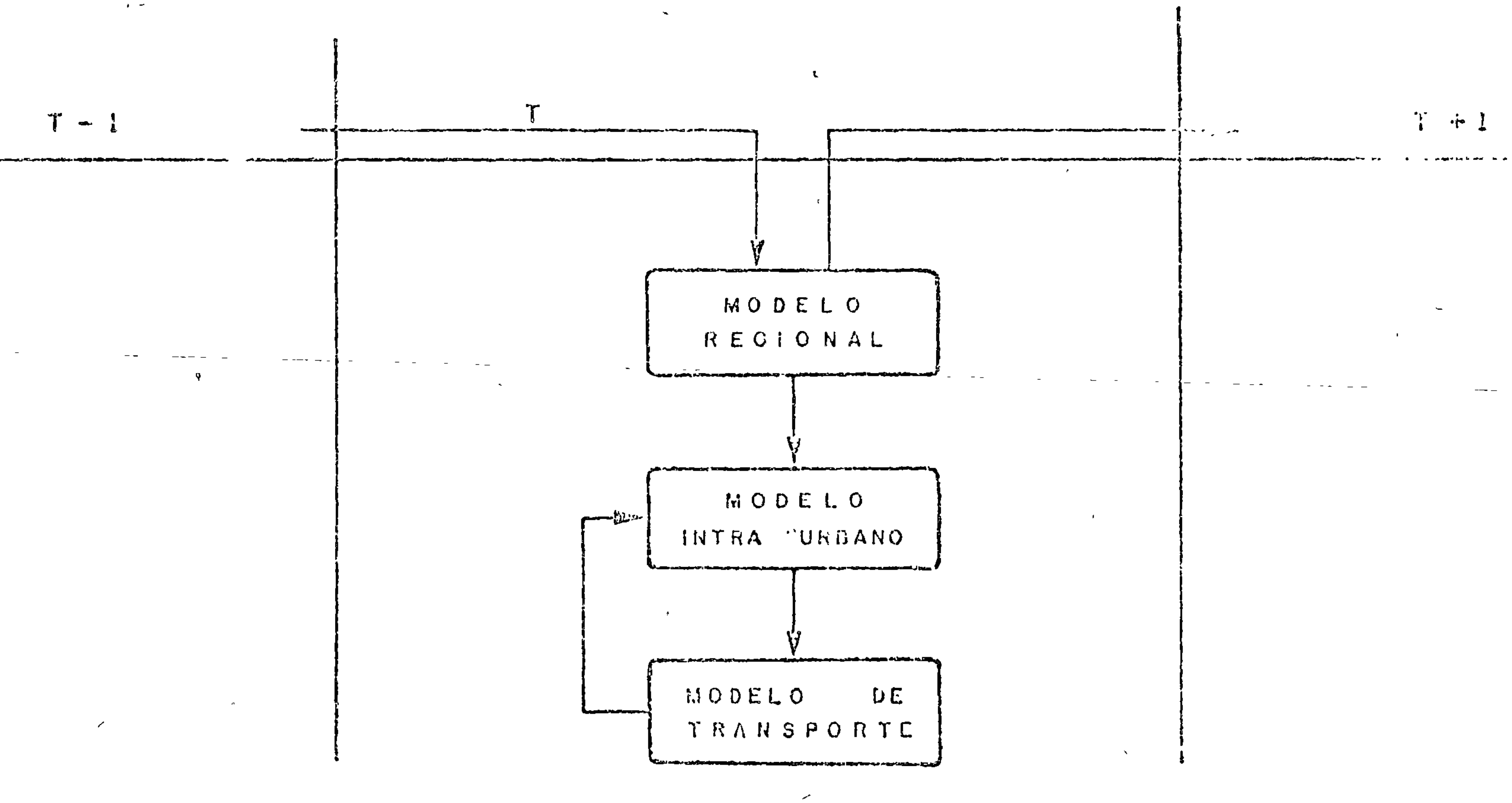
En términos generales el dinámismo de todo el Sistema se encuentra en el Modelo Regional, como puede apreciarse en la Fig. 1.

A continuación pasaremos a realizar una descripción de los modelos mencionados.

En la figura 10, la cual aparece al final del informe puede observarse la estructura global del conjunto de modelos

DINAMICA DEL CONJUNTO DE MODELOS

FIG. 1



II. MODELO REGIONAL SIMPLE

1. Descripción General.

El Modelo Regional Simple es un Modelo sencillo de carácter dinámico. Fundamentalmente se consideran como en el Modelo (1) Complejo, dos aspectos ó submodelos; Poblacion y Empleo.
La población, al igual que en el modelo complejo es una función de series históricas de nacimientos y muertes regionales y de la migración interregional. La migración es considerada como una función de la demanda de empleo en cada región.
El empleo es considerado como una función del nivel de la inversión regional en cada sector económico. Esta inversión en cada región, por sector económico, fluye interregionalmente de acuerdo a las interrelaciones económico-espaciales, provocado un destino final de las inversiones que prestan marcadas diferencias con las características de la inversión inicial.

En resumen dos factores son considerados: Las Características Económicas del Sistema Regional, dada por la distribución Regional de la inversión y la generación de empleo. Y la Característica Demográfica, determinada por condiciones históricas y las migraciones, producidas por la demanda de empleo en cada región.

---

(1) Modelo Complejo de la estructura Económica-Espacial, Regional, Manzanilla Hugo, Palacios, Luis Carlos 1975.

1.1 SISTEMA GEOGRAFICO Y VARIABLES BASICAS

El país se ha dividido en un conjunto de Z regiones (para las primeras pruebas se está considerando el país dividido en tres ó siete regiones). Cada región está representada por su centroide geográfico y está unida a las otras regiones por la red de transporte interregional. Tenemos entonces los siguientes elementos:

a) El número de regiones z

b) La matriz de accesibilidad D, cuyo elemento típico d(i,j) representa el costo de viajes entre la región i y la j.

c) El número de empleos en cada sector n, por cada región i (E (i,n))

d) La población de cada región i (P(i)).

En el desarrollo del Modelo se han considerado regionalizaciones de tres y siete regiones como se indican en los Mapas 1 y 2.

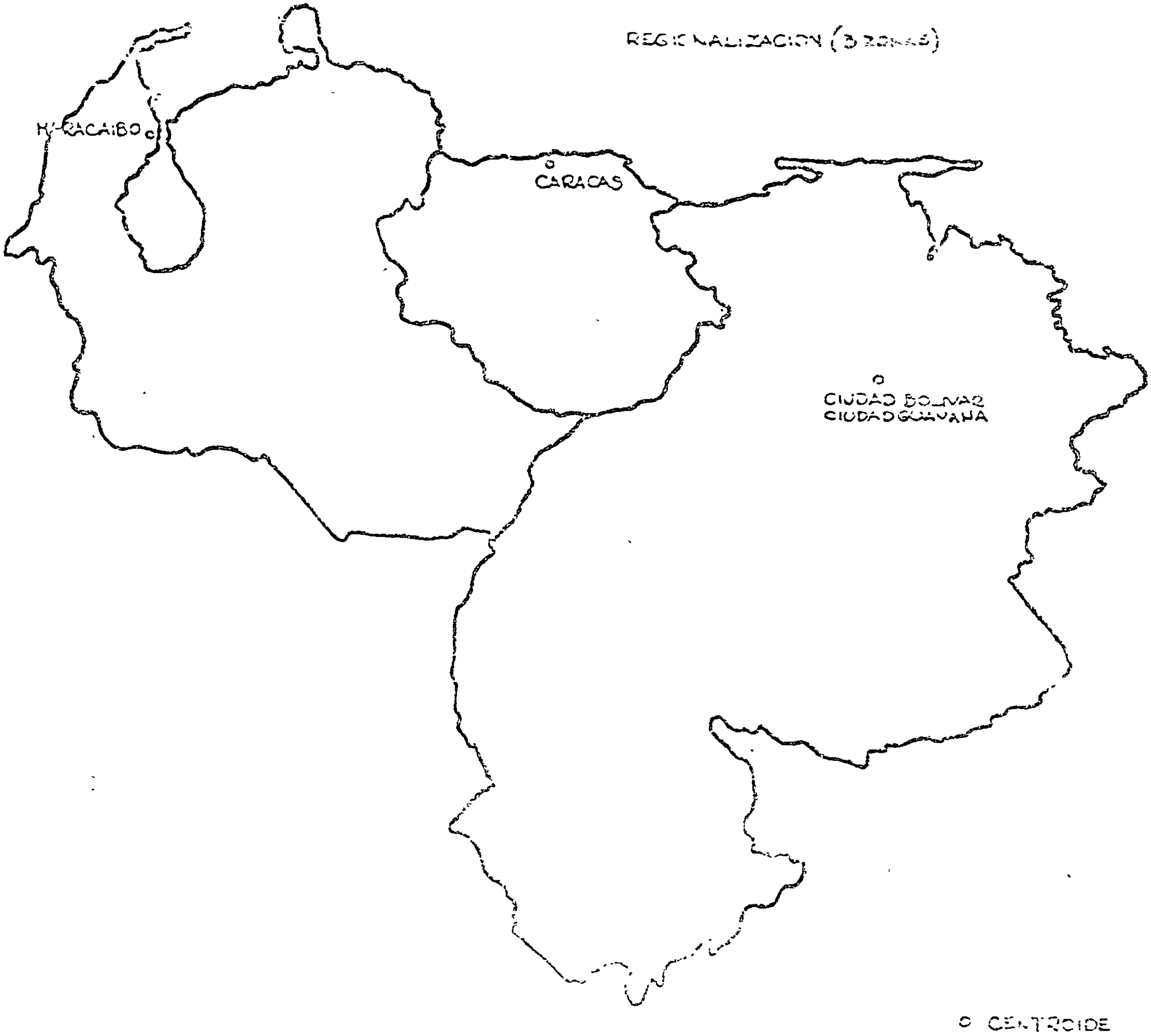
1.2 ESTRUCTURA DEL MODELO REGIONAL SIMPLE

La característica conceptual del Modelo está indicada en la Fig. Nº 2, donde se indica el carácter dinámico del Modelo, lo cual está basada en un Modelo desarrollado por Tomás de la Barra (1974). (2)

Como se puede apreciar en la Fig. 2 el Submodelo de Empleo, basado en los flujos de inversión inter-regional determina el empleo secto-

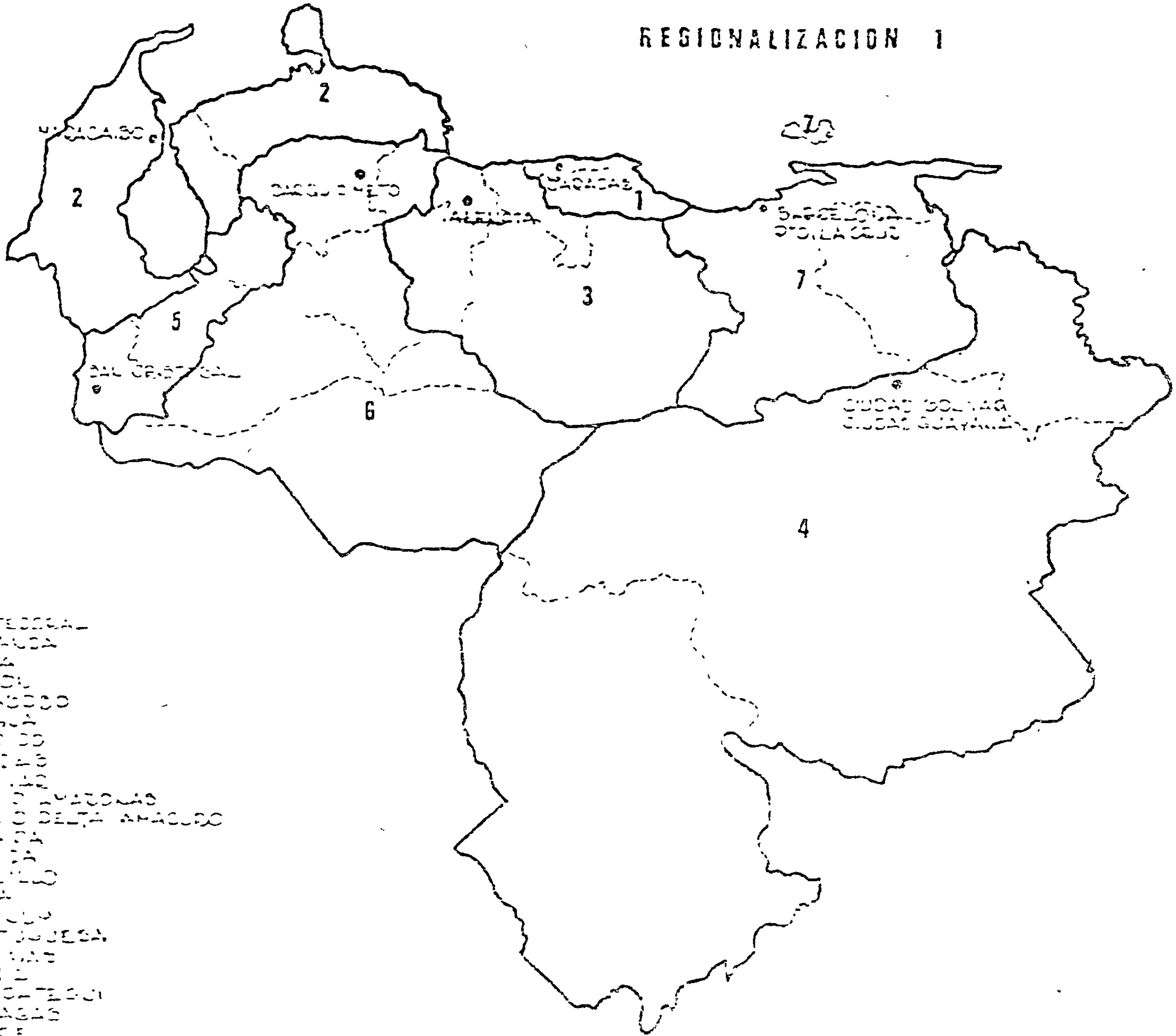
---

(2) Modelo urbano Regional, Tomás de la Barra, 1974



REGIONALIZACION (3 ZONAS)

REGIONALIZACION 1

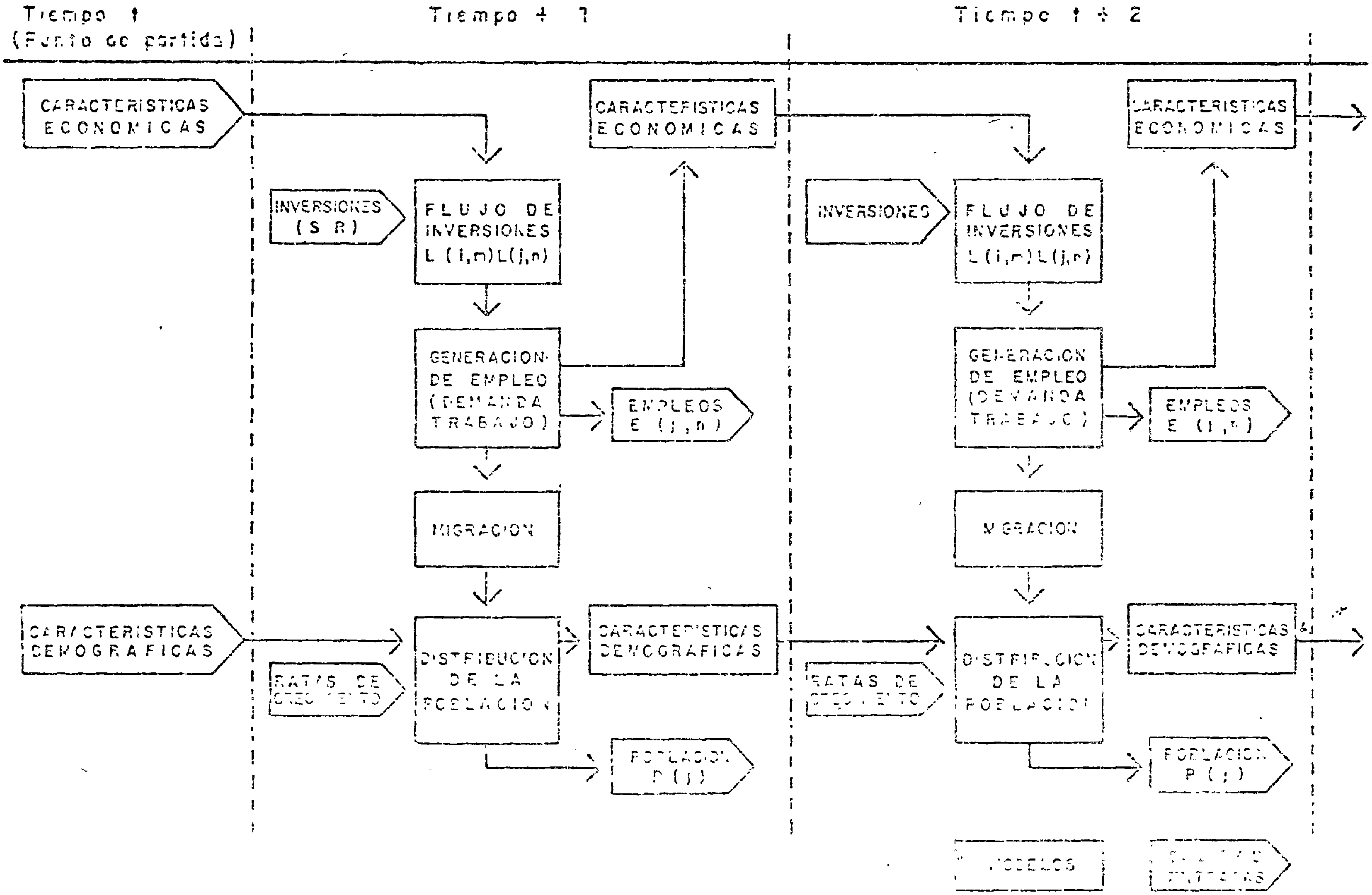


REGION 1 DISTRITO FEDERAL
EDO MIRANDA
REGION 2 EDO ZULIA
EDO FALCON
REGION 3 EDO CARABOBO
EDO ARAGUA
EDO GUARICO
EDO COJEDES
REGION 4 EDO BOLIVAR
TERRITORIO AMAZONAS
TERRITORIO DELTA AMACURO
REGION 5 EDO TACHIRA
EDO MERIDA
EDO TRUJILLO
REGION 6 EDO LARA
EDO YARACUY
EDO PORTUGUESA
EDO BARINAS
EDO APURE
REGION 7 EDO ANZOATEGUI
EDO MONAGAS
EDO SUCRE
NUEVA ESPARTA

• CENTROIDE

FIG 2

ESTRUCTURA CONCEPTUAL DEL MODELO REGIONAL



rial regional y actua sobre el Sub-Modelo demográfico como factor clave en la atracción a las migraciones. A continuación haremos una descripción más detallada del Modelo Regional Simple

2. EL MODELO ECONOMICO

El objetivo de este modelo es simular al empleo por sector económico en cada región. El modelo está basado en el supuesto de que existe una función lineal que relaciona la inversión y el empleo, y un conjunto de funciones simples que relacionan el Producto Territorial bruto con la Inversión.

Formalmente,

$E = I . \delta$

donde,

I = Inversión

δ = Constante

E = Empleo Nacional

2.1 DETERMINACION DE LA INVERSION

La inversión nacional se calcula como una función del producto territorial bruto en el mismo período de tiempo.

$I\ (t+1) = PTB\ (t+1)\ \varepsilon(t+1)$

donde,

$I\ (t+1)$: Inversión en el período $\underline{t+1}$

PTB(t+1): Producto Territorial Bruto en el período t+1

ε(t+1): Proporción de producto territorial bruto destinado a la inversión en período t+1

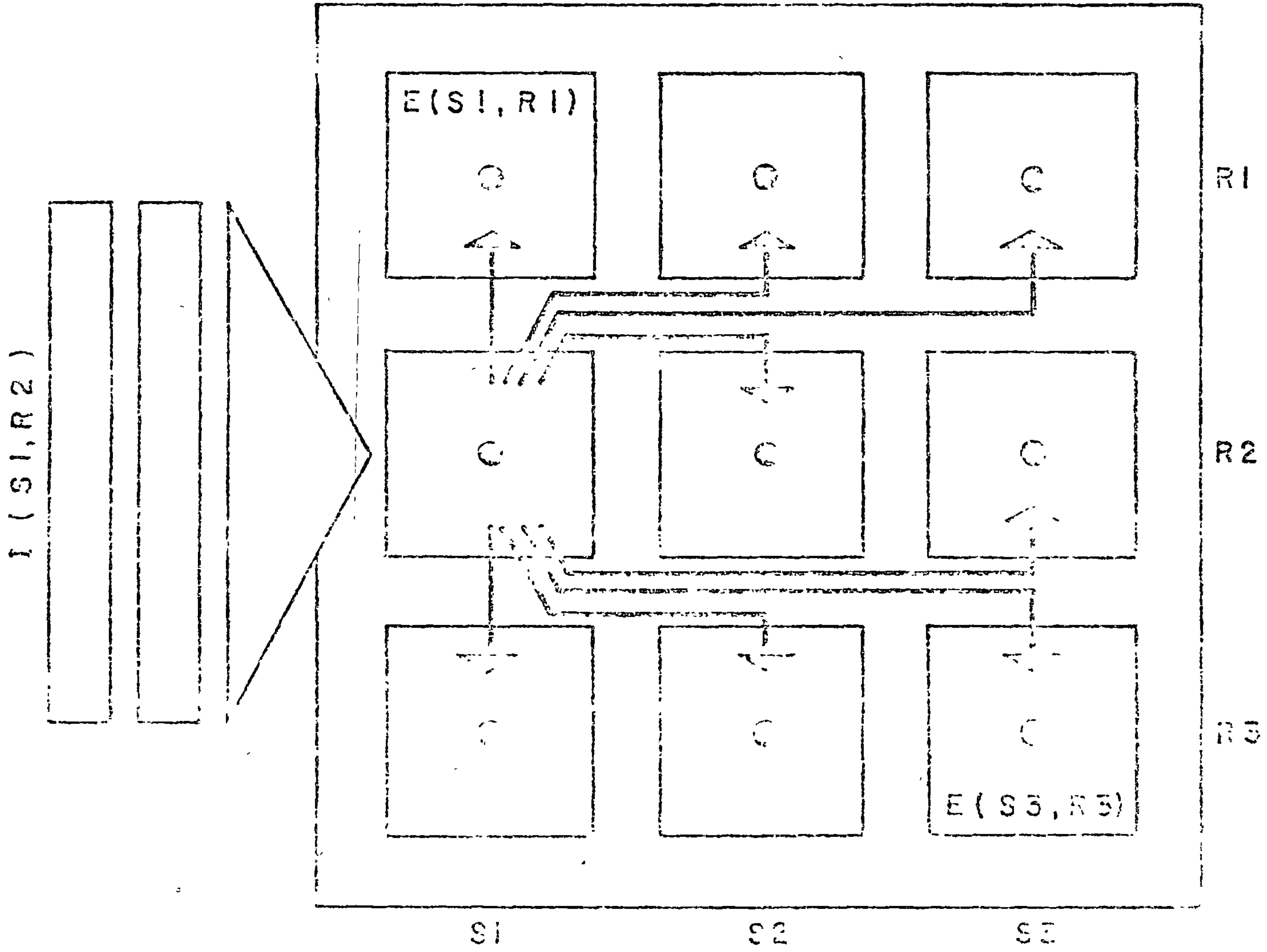
2.2 DETERMINACION DE LOS FLUJOS DE INVERSION Y EMPLEO REGIONAL.

En el Modelo Económico un aspecto de importancia es la determinación de los flujos interregionales e intersectoriales.

Se asume que una inversión realizada, por ejemplo en el sector uno (S1) de la economía y la región dos (R2), no permanece en ese sector y esa región, sino que pór la estructura de las relaciones interregionales, ella fluye a otras regiones y sectores siendo este destino final de la inversión, el elemento generador de empleo. En la Fig. 3 puede observarse en forma diagrámatica como la inversión en sector uno de la región dos (I (S1,R2)), fluye a otras regiones y sectores. El modelo económico ha sido diseñado para que sea validado tomando en cuenta el carácter de las inversiones en el pasado y para probar, en el futuro, el efecto de distintas de políticas de inversión.

La matriz de flujo de inversiones se calcula en el Modelo según dos opciones, aplicando en

FIG-3



ambas el método de maximización de la entropía.

A) En la primera opción el dato exógeno que maneja el Modelo es la Inversión en el sector ó capacidad total de inversiones regional sectorial.

Formalmente,

$$I\ (i,m) = \sum_{n}\sum_{j} F\ (i,j,m,n)$$

donde,

$I\ (i,m)$ = La Inversión en la región $\underline{i}$, Sector $\underline{m}$

$F\ (i,j,m,n)$ = Flujo de inversión de la región $\underline{i}$ a la región j, del sector $\underline{m}$ al sector $\underline{n}$

La matriz de flujo de inversión interregionales e intersectoriales se calcula en esta opción a través de la siguiente expresión:

$$F\ (i,j,m,n;\ t+1) = I\ (i,m;t+1)\quad C\ (m,n)\ .\ \frac{A(j,n;t)\exp(-\beta(n)d(i,j))}{\sum_{j} A(j,n;t)\exp(-\beta(n)d(i,j))}$$

donde,

$F(i,j,m;t+1)$ = Flujo de inversión entre la región $\underline{i}$ y región J, Sector $\underline{m}$ a sector $\underline{n}$, en periodo $\underline{t+1}$

$C(m,n)$ = Matriz de coeficientes de flujo de inversiones sectoriales.

$$C(m,n) = \frac{\sum_{i}\sum_{j} F(i,j,m,n)}{\sum_{n}\sum_{i}\sum_{j} F(i,j,m,n)}$$

$A(j,n;t)$ = La atracción de la región j a inversiones del sector n, en el n periodo t

$\exp(-B(n)\, d(i,j))$ = "Impedancia" o costo de viaje de la inversión que llega al sector n para viajar de la región i a la región j

Es decir el flujo de inversión de una región a otra y de un sector a otro se hace en dos pasos. La matriz $C(m,n)$ distribuye los flujos intersectorialmente, de los sectores m a los sectores n. Luego se regionaliza la inversión por la interacción espacial. En ella la atracción regional $(A(j,n;t))$ depende de la especialización regional en cada sector, y la impedancia es una función de los costos de viaje de la inversión.

La atracción regional sectorial depende de la especialización regional en cada sector, formalmente,

$$A(j,n;t) = \left[\frac{\sum_n E(j,n;t)}{\sum_j \sum_n E(j,n;t)} + \frac{E(j,n;t)}{\sum_j E(j,n;t)} \right]^{G(n)}$$

De la matriz de flujos se calcula el flujo total que recibe la región en cada sector. Formalmente,

$$I(j,n;t+1) = \sum_i \sum_m F(i,j,m;t+1)$$

y con este elemento se calcula el empleo regional sectorial.

$$E(j,n;t+1) = I(j,n;t+1) \; / \; CE(n)$$

donde,

$E(j,n;t+1)$ = Empleo en la región <u>j</u>, sector <u>n</u>, período <u>t + 1</u>

$CE(n)$ = Coeficiente de la inversión-Empleo Sectorial

$\Theta(n)$ = Parámetro de economía de escala sectorial.

B) En la segunda opción el modelo maneja como dato exógeno la inversión bruta fija regional sectorial.

Esta opción se considera debido a que el dato de más facil obtención y de más facil manipulación como política de inversiones es la inversión bruta fija en cada sector, por región. La inversión bruta fija no es la inversión en el sector por región sino el monto de la inversión del sector en la región. Esto es el flujo de inversión que recibe cada sector en cada región. Este dato es más proclive a la manipulación de políticas debido a que es más plausible asumir cual es el monto de la inversión global de un sector en una región y suponer que cada sector en cada región adquiera bienes de capital de los otros sectores en forma poco controlada. Se parte pues de conocer el dato de la inversión que recibió el sector en cada región, la inversión bruta fija, la cual es producto de determinadas políticas.

Formalmente,

$$I(j,n) = \sum_{i}\sum_{m} F(i,j,m,n)$$

Al mismo tiempo se tiene también como dato exógeno la capacidad de inversión en el sector por región, esto es I(i,n), el cual se calcula igual que en la primera opción. (Capacidad de inversión del sector en la región ó capacidad de producir bienes de capital).

Conocida la inversión bruta fija sectorial regional el modelo pasa a determinar de que sector proviene esa inversión (sectorialización) y luego a determinar la procedencia regional (regionalización).

Esto determina el impacto de la inversión bruta fija en cada región y en cada sector.

El impacto de la inversión bruta fija en cada región y cada sector representa la demanda de bienes de capital que determina la inversión bruta fija en cada region y para cada sector. Esta demanda debe ser comparada con la Inversión en el sector por región ó la capacidad de inversión regional anteriormente calculada. Si la demanda excede la capacidad, dentro de ciertos rangos, ello indica que no existe suficiente capacidad productiva de bienes de capital para sostener la demanda que produce la política de inversión bruta fija que se está probando.

Formalmente,

$$F(i,j,m,n;t+1) = I(j,n,t+1).\ C1(m,n)\ \frac{A(i,m;t)\exp(-r(m)d(i,j))}{\sum_{i} A(i,m;t)\exp(-r(m)d(i,j))}$$

donde,

F(i,j,m,n;t+1) = Flujo de inversión de región i a región j sector n, en período t+1

C1(m,n) = Matriz de coeficiente de flujos intersectoriales. Esta matriz es distinta a la C(m,n). En este caso los coeficiente se establecen en relación al total de inversiones o bienes de capital que existe en el sector n

$$C1(m,n) = \frac{\sum_{i}\sum_{j} F(i,j,m,n)}{\sum_{m}\sum_{i}\sum_{j} F(i,j,m,n)}$$

A(i,m;t) = "Atracción" de la región i en el sector m, para que la región j en el sector n adquiera esas inversiones ó bienes de capital.

Esta "Atracción" es similar a la atracción A(j,n,t) es decir una función de la especialización regional y sectorial.

exp(-B(m) d(i,j)): Función de impedancia o costo de viaje de la inversión del sector m para viajar de la región I a la región J

Con esta matriz flujo es factible calcular la demanda de bienes de capital sectorial por región.

$$D(i,m;t+1) = \sum_{j}\sum_{n} F(i,j,m,n,t+1)$$

donde,

D(i,m;t+1) = La demanda de bienes de capital ó de inversión que hace sobre la región i, sector m, período t+1 las inversiones brutas regionales.

Si $D(i,m;t+1) > (I(i,m;t+1) + EPSI(I,m,t+1)$

No es factible la inversión bruta fija establecida en las políticas regionales.

EPSI(I,m,t+1) = Rango de variación de capacidad de inversión del sector m en región I, período t+1

El empleo Regional Sectorial se calcula después de comparar la demanda de inversiones con la capacidad según el coeficiente CE(n).

Formalmente,

$$E(j,n,t+1) = I(j,n;t+1) / CE(n)$$

3. MODELO DEMOGRAFICO

El objetivo de este modelo es determinar los niveles de población en cada región. El cambio de población de cada región depende de la tasa de natalidad y mortalidad regionales y los movimientos migratorios.

3.1 Determinación de la Población Bruta

Las tasas de natalidad y mortalidad calculada por un procedimiento similar al presentado en la versión compleja. Esto es ajustado un polinomio de alto grado a la función de la *tasa acumulada* de natalidad y mortalidad.

Formalmente,

$$B(i,t) = Y_b(t) - Y_b(t-1)$$

$$D(i,t) = Y_d(t) - Y_d(t-1)$$

donde,

$B(i,t)$ = tasa de natalidad anual de año t a t+1

$D(i,t)$ = tasa de mortalidad anual de año t a t+1

$Y_b(t)$ = tasa acumulada de natalidad hasta año t

$Y_b(t-1)$ = tasa acumulada de natalidad hasta año t-1

$Y_d(t)$ = tasa acumulada de mortalidad hasta año t

$Y_d(t-1)$ = tasa acumulada de mortalidad hasta año t-1

3.2 Determinación de las Migraciones y Población Regional

Calculada la población bruta por las tasas de natalidad y mortalidad, es necesario calcular las migraciones interregionales de población para determinar la población regional.
La probabilidad de migrar es considerada como proporcional a un factor de atracción y dependiente de una "Impedancia" a migrar que es una función de la distancia interregional.

Formalmente,

$$Pr: M(i,j;t+1) = \frac{A(j,t)^{0}.\exp(-Bc(i,j))}{\sum_j A(j,t)^{0}.\exp(-Bc(i,j))}$$

donde,

$Pr: M(i,j,t+1)$	Probabilidad de migrar de la región i a la región j en periodo t = 1
$A(j;t)$:	Atracción de la región j durante periodo t Esta pueder ser expresada como

una relación entre empleo y población.

$$A(j;t) = \frac{E(j;t)}{P(j;t)}$$

$\exp(-Bc(i,j))$ Función de costo de migrar de región $\underline{i}$ a $\underline{j}$

0 = Parámetro

La migración de la región $\underline{i}$ a la $\underline{j}$ será:

$$M(i,j;t+1) = P(i,t)(1+B(i,t)-D(i;t)).\ \frac{A(j,t)^{0}\exp(-Bc\ (i,j))}{\sum_{j} A(j,t)^{0}\exp(-Bc(i;j))}$$

y la población será en la región $\underline{j}$ será:

$$P(j,t+1) = \sum_{i} M(i;t+1)$$

3.3 Problemas y posibles soluciones

Hasta ahora de los resultados obtenidos al aplicar el Modelo Demográfico, se cree conveniente profundizar en dos aspectos:

a) Factor de atracción de las migraciones

b) Cálculo de las tasas de mortalidad y natalidad como ajuste de series históricas a un polinomio.

3.3.1 Factor de Atracción de las Migraciones:

Se considera que la atracción como una simple proporción entre el empleo y la población no tomaba directamente en cuenta factores que afectan las migraciones tales como el nivel de desempleo. Por ello se probarán nuevas formulaciones de la atracción como la siguiente:

$$A(j,t)^{0} = ((PA(j,t) - L(j,t))^{-1})^{0}$$

Esta ecuación puede considearse como una aproximación al desempleo de la zona (j),el cual actuaría inversamente proporcional a las migraciones. También se podría utilizar el ingreso per/capita de la región como un factor adicional de la misma.

La atracción se expresaría matemáticamente:

$$A(j,t)^{0} = \frac{E(j,t)\,Y(j,t,)^{0}}{P(j,t)}$$

donde,

$E(J,t)$ - Empleo en j en t

$P(j,t)$ = Población en j en t

$Y(J,t)$ = Ingreso regional per capita.

Estas posibles nueva formulaciones traería como consecuencia algunas adiciones en el Modelo, inicialmente planteado.

3.3.2 Cálculo de las tasas de natalidad y

mortalidad como ajuste de series históricas a un polinomio.
El ajuste al polinomio para el cálculo de las tasas se hizo de la siguiente manera:

$$Y_i(t) = \sum_{j=1}^{m} \text{coef } j \left(\frac{2(X_i - 1)}{N-1} - 1 \right)^{j-1}$$

donde,

N = número de puntos ajustados al polinomio.

Coef j = coeficientes del polinomio

m = grado del polinomio

X_i = intervalo anual

$Y_i(t)$ = valor de la función polinómica en el período t.

El problema se presenta al intentar proyectar tasas fuera del intervalo al cual fue ajustado el polinomio, (61-71), debido a que la funcion del polinomio se comporta de manera adecuada entre los puntos considerados, pero no existe completa garantía de su comportamiento fuera del intervalo ajustado
Una solución adecuada para períodos más allá de 1971 sería, realizar hipótesis acerca de las tasas con un método demográfico convenien

te, y ajustar luego el polinomio a ese nuevo intervalo. Esta solución está en periodo de prueba.

III. MODELO INTRA - URBANO SIMPLE

1. Descripción General

El Modelo Intra - Urbano Simple tiene como objeto simular el sistema urbano dentro del contexto de su región. El Modelo Intra - Urbano considera la ciudad dentro de una de las regiones del país, y las actividades y edificaciones generadas por el modelo son localizadas dentro de la ciudad según sus relaciones en su contexto regional.
Los insumos regionales fundamentales los provee el Modelo Regional.

1.1 Sistema de Areas de Análisis

El primer paso para probar el modelo para la Región Capital fue la definición de división del área en entidades geográficas. Las variables se desagregaron de acuerdo a esta división.
Se definió un área de global de análisis dentro de la Región Capital, y en ella se distingue el Area Metropolitana de Caracas, y unas zonas externas.
El funcionamiento de modelos espaciales es muy sensitivo a los cambios de tamaño de las zonas del área en estudio. La Sub-división del área Metropolitana debe ser en zonas pequeñas donde

la interacción entre zonas sea alta y el radio de ellas sea menor que la distancia promedio de viajes al trabajo. Por el contrario las zonas externas deben ser grandes en relación al viaje a trabajo y deben aproximarse a sistemas homogéneos. Por lo menos el 90% de los viajes deben tener su origen y destino dentro de la misma zona.

En síntesis:

a) El Area Metropolitana tiene una zonificación fina con inter-acción fuerte entre zonas.

b) Las zonas externas son grandes con gran interacción interna y poca entre zonas.

En el Area Metropolitana se adoptó la zonificación existente del Banco de Datos desarrollado por la UMUA para el Modelo Desagregado Urbano. Esta divide Caracas en 30 zonas internas. El resto de la región se dividió en 5 zonas externas aprovechando las divisiones políticas de Censo y tratando de hacer sub-sistemas homogéneos. Las zonas externas son:

31 Departamento Vargas (El Litoral)

32 Distrito Guaicaipuro
   parte Dpto. Libertador (Los Teques, Panamericana).

33 Distrito Independencia (El Tuy)
   Distrito Lander
   Distrito Paz Castillo
   parte Distrito Sucre.

34 Distrito Plaza

Distrito Zamora (Guarenas/Guatire)
35 Distrito Acevedo (Barlovento)
Distrito Brión
Distrito Páez
Ver Mapa 3

Se adoptó como año base 1966, porque existían datos para las 30 zonas internas de Caracas. El problema apareció para los datos de tierra disponible para las 5 zonas externas. La tierra disponible se cálculo en base de mediciones en planos y también de informes de Planeamiento Urbano del MOP y de OMPU. La matriz de distancia se midió en un plano vial de la Región tomando la distancia promedio de viajes al trabajo como la distancia interna de las zonas externas. (Estos cálculos de datos básicos aparecen en la Apéndice 1).

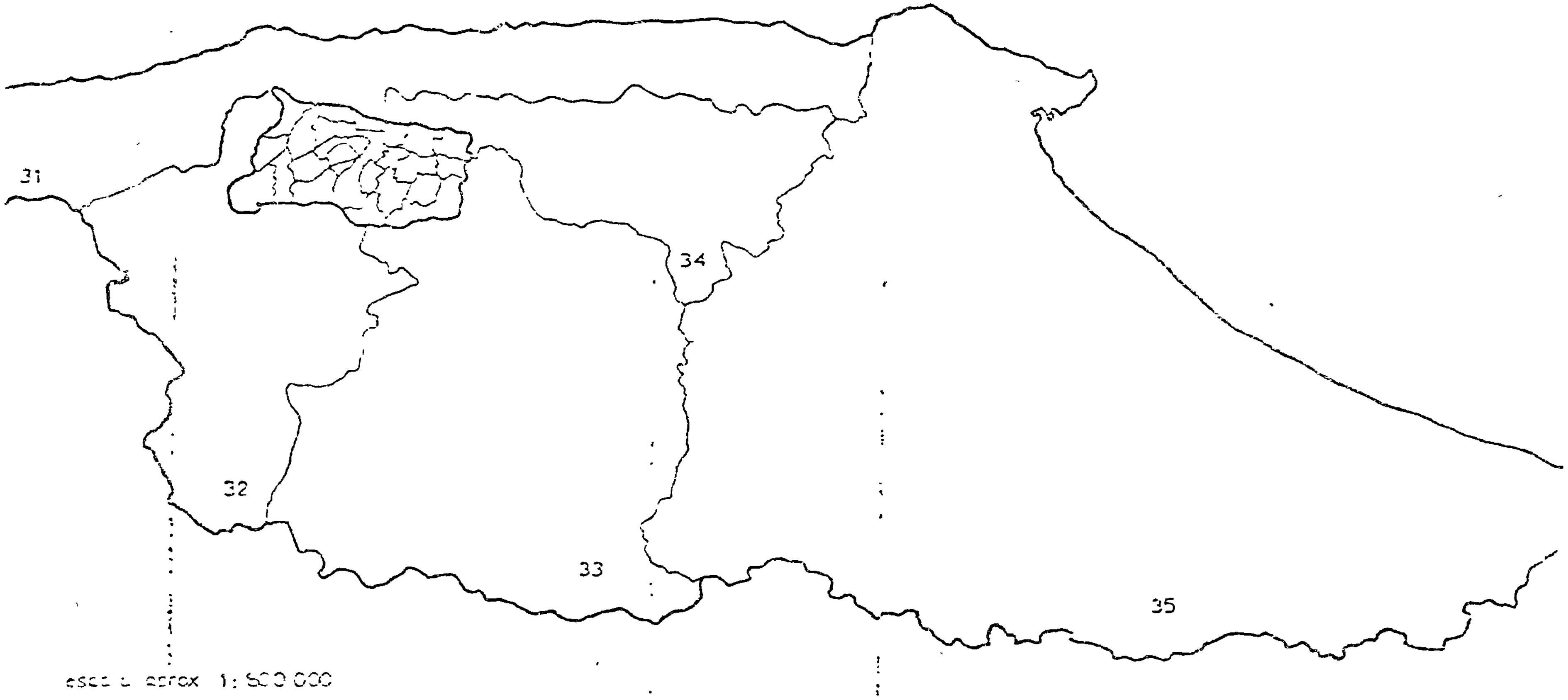
1.2 Estructura del Modelo Intra - Urbano

El modelo consiste en dos sub- modelos:

a) Modelo de empleo básico; que localiza el empleo básico en zonas internas de la ciudad y zonas externas de la región específica.

b) Modelo de edificaciones y actividades; que localiza espacio construído, residentes y empleo en servicio en las zonas antes mencionadas.

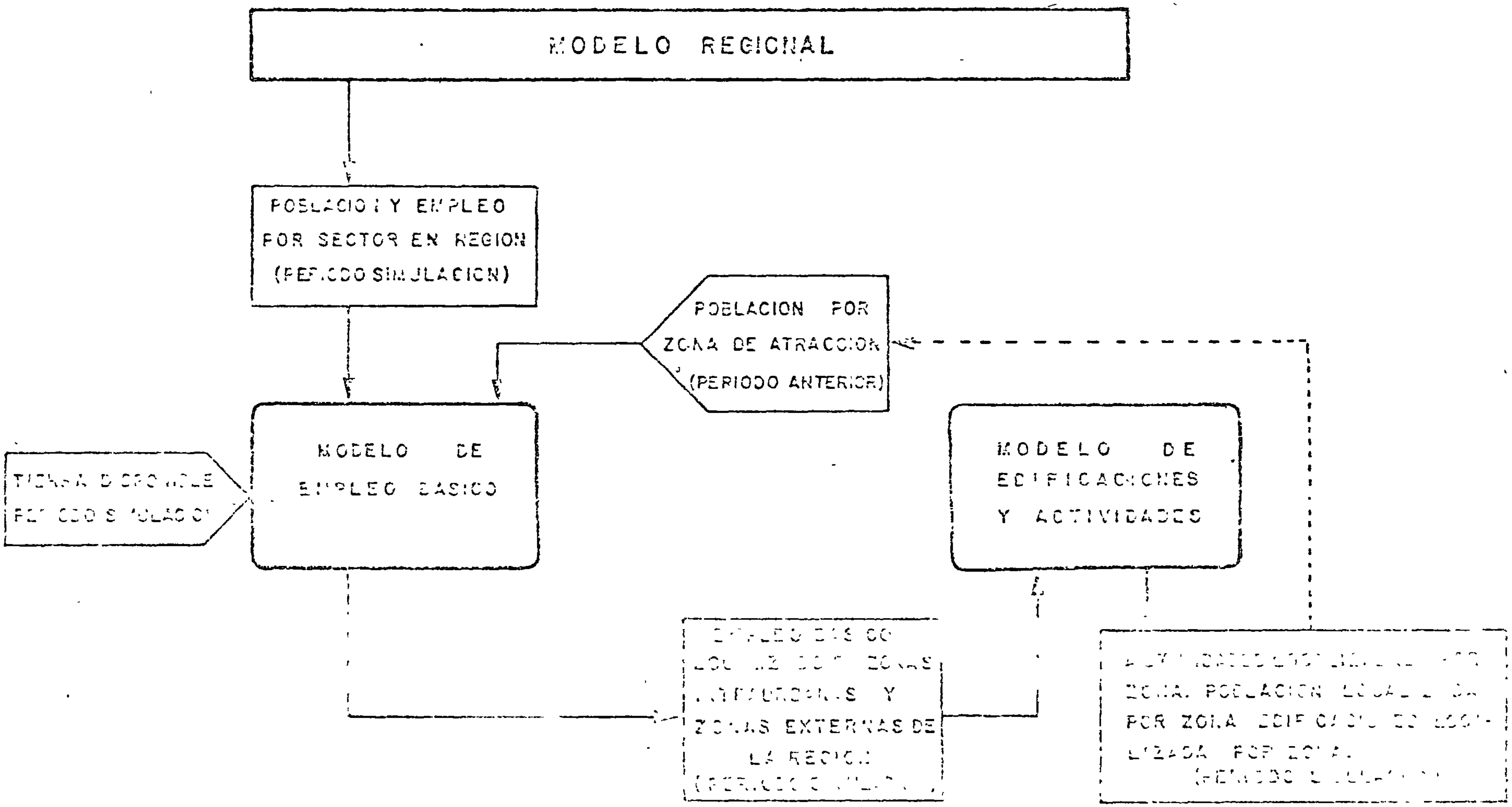
Como se puede apreciar en la Fig. 4, la pobla-



Region Capital - Zonificacion

ESTRUCTURA CONCEPTUAL DEL MODELO INTRAURBANO

FIG. 4



ción y en el empleo en toda la región (en este caso la región capital) es un insumo del Modelo Regional. Luego el Modelo de Empleo básico, con la entrada de la tierra disponible por actividad en la región y la población por zonas del periodo anterior, localiza el empleo básico para cada una de las zonas de la región, incluyendo las zonas internas de la ciudad y las zonas externas. Esta salidad pasa al modelo de edificaciones y actividades el cual produce la generación y localización de las actividades restantes, las edificaciones y la población.

2. MODELO DE EMPLEO BASICO

2.1 Clasificación del Empleo

Al abarcar el modelo a 35 zonas de la Región Capital se presenta la necesidad de revisar el concepto de Empleo Básico tradicional.Para la Región Capital es básico la Manufactura, Transporte (Puerto de la Guaira, Aeropuesto), Agricultura y Gobierno.
El empleo de Transporte localizado en el Puerto y el Aeropuesto puede ser considerado por el Modelo como un dato exógeno.

2.2 Estructura General del Modelo de Empleo Básico

El objetivo del Modelo de Empleo Básico es localizar para cada una de las 30 zonas urbanas y 5 zonas exteriores, el empleo básico. Este empleo

Básico localizado es el insumo al Modelo de Edificaciones y Actividades, el cual es un Modelo de Estructura Urbana.

La Estructura General del Modelo puede apreciarse en la Fig. 5.

En términos generales el Modelo de Empleo Básico está formado por un conjunto de Sub-Modelos, cada uno de los cuales localiza en las diferentes zonas los diferentes tipos de Empleo Básico. Estos Modelos de localización son Modelos de interacción espacial con una sola restricción donde se utiliza el método de maximización de la entropía.

La fórmula general planteada para este proceso es:

$$E(j,n;t+1) = \sum_{j} E(j,n;t+1).L(j,n;t+1)V(j,n;t+1)\sum_{j} A(j,n;t+1$$

Es decir, el Empleo en cada zona j, sector n, se localiza proporcionalmente a la tierra disponible en esa zona para ese tipo de empleo y en función de un factor de atracción específico para ese empleo donde se toma en cuenta la accesibilidad.

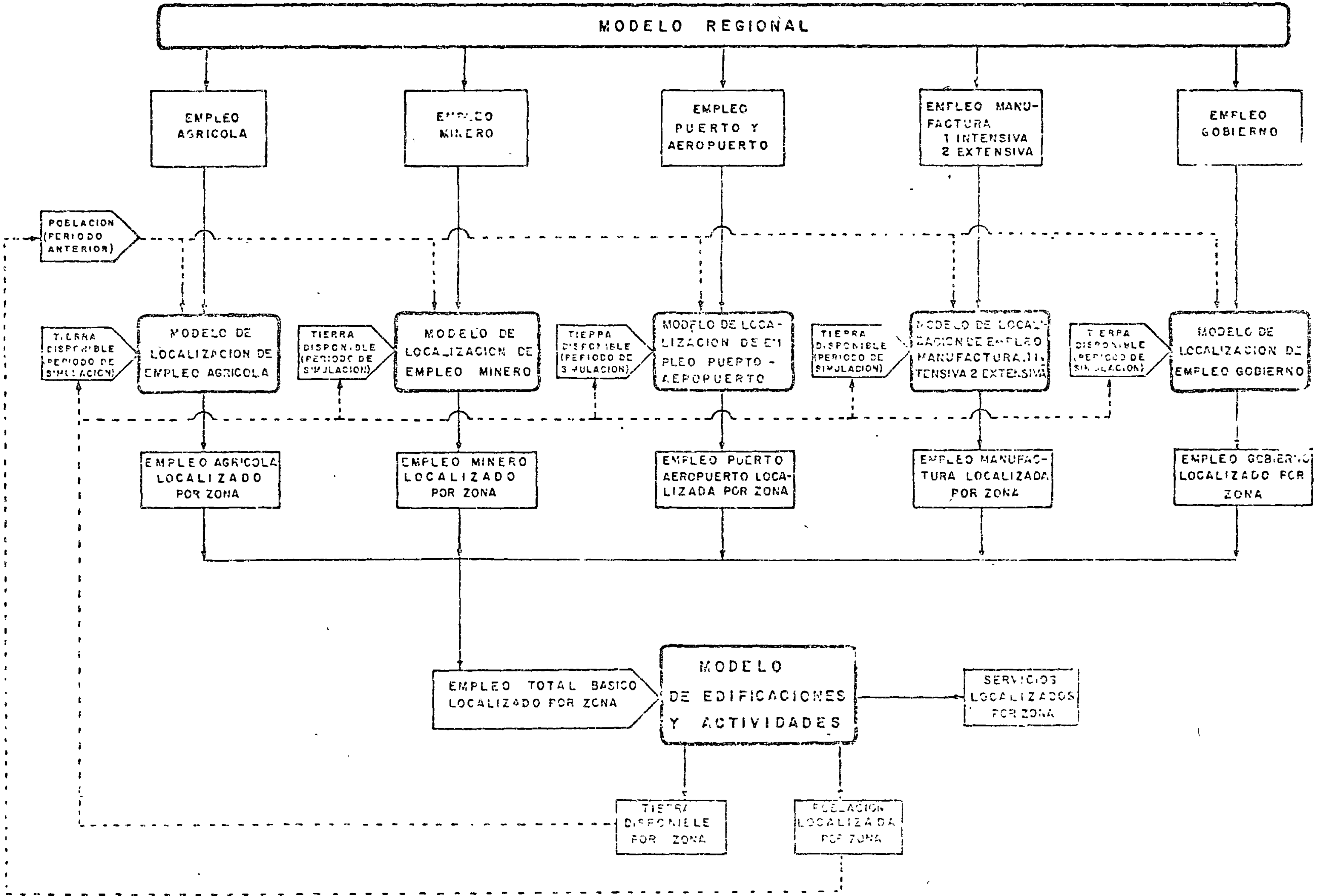
Específicamente,

$E(j,n;t+1)$ = Empleo Básico n en zona j, en período t+1

$\sum_{j} E(j,n;t+1)$ = Cantidad total de empleo de tipo n en la región, en período t+1

ESTRUCTURA GENERAL DEL MODELO DE EMPLEO BASICO

FIG. 5



$L(j,n;t+1)$ = Cantidad de tierra existente en zona j para desarrollo de tipo de Empleo (ó actividad) n (tierra total para ese tipo de desarrollo menos la tierra utilizada en t).

$W(j,n;t+1)$ = Factor de atracción de la zona j, para empleo de tipo n en periodo $t+1$.

Este factor está por lo general determinado de la siguiente manera:

$$W(j,n;t+1) = V(j,n;t) \sum_i X(i;t) \exp(-B(n)d(i,j))$$

donde,

$V(j,n;t)$ = Representa la variable de atracción de la zona j para el sector n, periodo t

$\sum_i X(i;t).\exp(-B(n)d(i,j))$ = La interacción de la zona j con el resto de las zonas i

donde,

$X(i;t)$ = Variable directamente proporcional en la interacción entre zonas i y zona j

$\exp(-B(n)d(i,j))$ = Impedancia entre zonas i y la zona j

$\sum_j W(j,n;t+1)$ = Factor de normalización.

donde,

$$\sum_j A(j,n;t+1) = \left(\sum_j L(j,n;t+1)\, W(j,n;t+1)\right)^{-1}$$

2.3 Sub-Modelos de Empleo Básico

Como anteriormente manifestamos el empleo básico está dividido en Agricultura, Minería, Transporte, Manufactura y Gobierno. Para cada uno de estos tipos de empleos existe un sub-modelo de localización. Asumiremos que el índice <u>n</u> que indica el sector puede tomar 6 valores. Correspondientes a:

n = 1 Empleo Agrícola
n = 2 Empleo Minero
n = 3 Empleo Puerto y Aeropuesto (Transporte Básico)
n = 4 Empleo Manufacturado Intensivo
n = 5 Empleo Manufacturado Extensivo
n = 6 Empleo Gobierno.

Los Sub-Modelos para cada uno de estos sectores son:

Sub-Modelo Agrícola

$$E(j,n;t+1) = \sum_j E_j(j,1;t+1)\, L(j,1;t+1)\, W(j;1;t+1) \sum_j A(j,1;t+1)$$

Siendo el factor de atracción la accesibilidad al mercado consumidor representado por la población en todas las zonas.

$$W(j;1;t) = \sum_i P(i;t)\exp(-B(n)d(i,j))$$

donde,

$P(i;t)$ = Población en zona $\underline{i}$, en periodo t

Sub-Modelo Minero

$$E(j,2,t+1)=\sum_j E(j,2;t+1)L(j,2;t+1)W(j,2;t+1)\sum_j A(j,2;t+1)$$

Siendo el factor de atracción el empleo que Minero existía anteriormente en la zona y la impedancia de viaje.

$$W(j,2;t+1) = E(j,2;t).\sum_i \exp(-B(n)d(i,j))$$

Sub-Modelo de Puertos y Aeropuertos

Con respecto a este tipo de empleo existen dos opciones. Es factible leer exogenadamente esta variable ó derivarla por el Sub-Modelo.

$$E(j,3;t+1) = \sum_j E_j(j,3;t+1)L(j,3;t+1)W(j,3;t+1)\sum_j A(j,3;t+1)$$

Siendo el factor de atracción es el empleo de Puertos y Aeropuertos que existía anteriormente en la zona y la impedancia de viaje.

$$W(j,3;t+1) = E(j,3;t)\sum_i \exp(-B(n)d(i,j))$$

Sub-Modelos de Manufactura

Manufactura Intensiva $k = 1$

Manufactura Extensiva k = 2

$$E(j,4;t+1) = \sum_j E(j,4;t+1)L^1(j,4;t+1).W(j,4;t+1)/\sum_j W(j,4;t+1)$$

Al valor k = 1, la tierra para manufactura intensiva se pondera por un procedimiento distinto al de manufactura extensiva.
El factor de atracción es el empleo industrial intensivo que existía en la zona y la accesibilidad al mercado consumidor.

$$W(j,4;t+1)= E(j,4;t\quad)\sum_i P(i;t)\exp(-B(n)d(i,j))$$

$$E(j,5;t+1)= \sum_j L(j,5;t+1)\ L^2(j,5;t+1).W(j,5;t+1)/\sum_j W(j,5;t+1)$$

El factor de atracción es similar del empleo industrial.

$$W(j,5;t+1) = E(j,5;t)\ \sum_i P(i;t)\exp(-B(n)d(i,j))$$

<u>Sub-Modelo de Gobierno</u>

$$E(j,6;t+1) = \sum_j E(j,6;t+1).L(j;6;t+1)W(i,6;t+1)/\sum_j W(j,6;t+1)$$

El factor de atracción el Empleo anteriormente existe en la zona y la accesibilidad a la población residente, en cierta manera un elemento que mide la "Centralidad".

$$W(i,6;t+1) = E(j,6;t)\sum_i P(i,t)\ .\exp(-B(n)d(i,j))$$

Restricción de Densidad

El Modelo de localización de Empleo Básico tiene opcionalmente un control de densidad que limita la cantidad de empleo básico que puede localizarse en cada zona, y el excedente es localizado en las zonas cuya capacidad no ha sido excedida.
Esta opción tiene objetivos de planificación. En forma esquemática funciona como se indica en la Fig. 6

3. MODELO DE EDIFICACIONES Y ACTIVIDADES

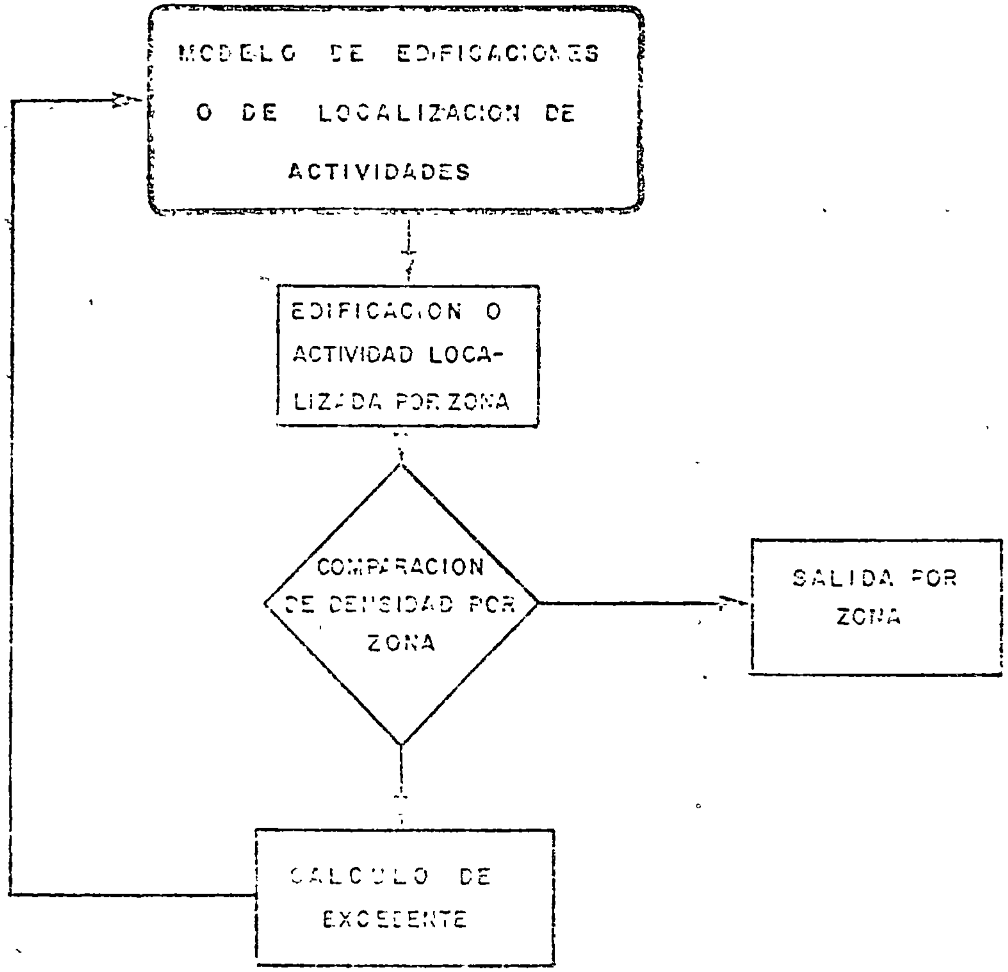
Como se indica al describir la estructura general del Modelo Intraurbano, el Modelo de Edificaciones y Actividades, con la entrada del empleo básico localizado por las 35 zonas definidas, calcula y localiza para cada una de las zonas el resto de las actividades de empleo (empleos y servicio), las edificaciones y la población residente. Este modelo es una versión Venezolana del Modelo Lowry-Garin-Echenique.
El Modelo necesita como entradas especiales elementos similares al Modelo de Empleo Básico, esto es la cantidad de tierra para el desarrollo y al red de Transporte. Estos elementos fueron definidos ya en el sistema de Area de Análisis.

3.1 Estructura General del Modelo

El Modelo comienza localizando en las diferen

CONTROLES DE RESTRICCION

FIG. C

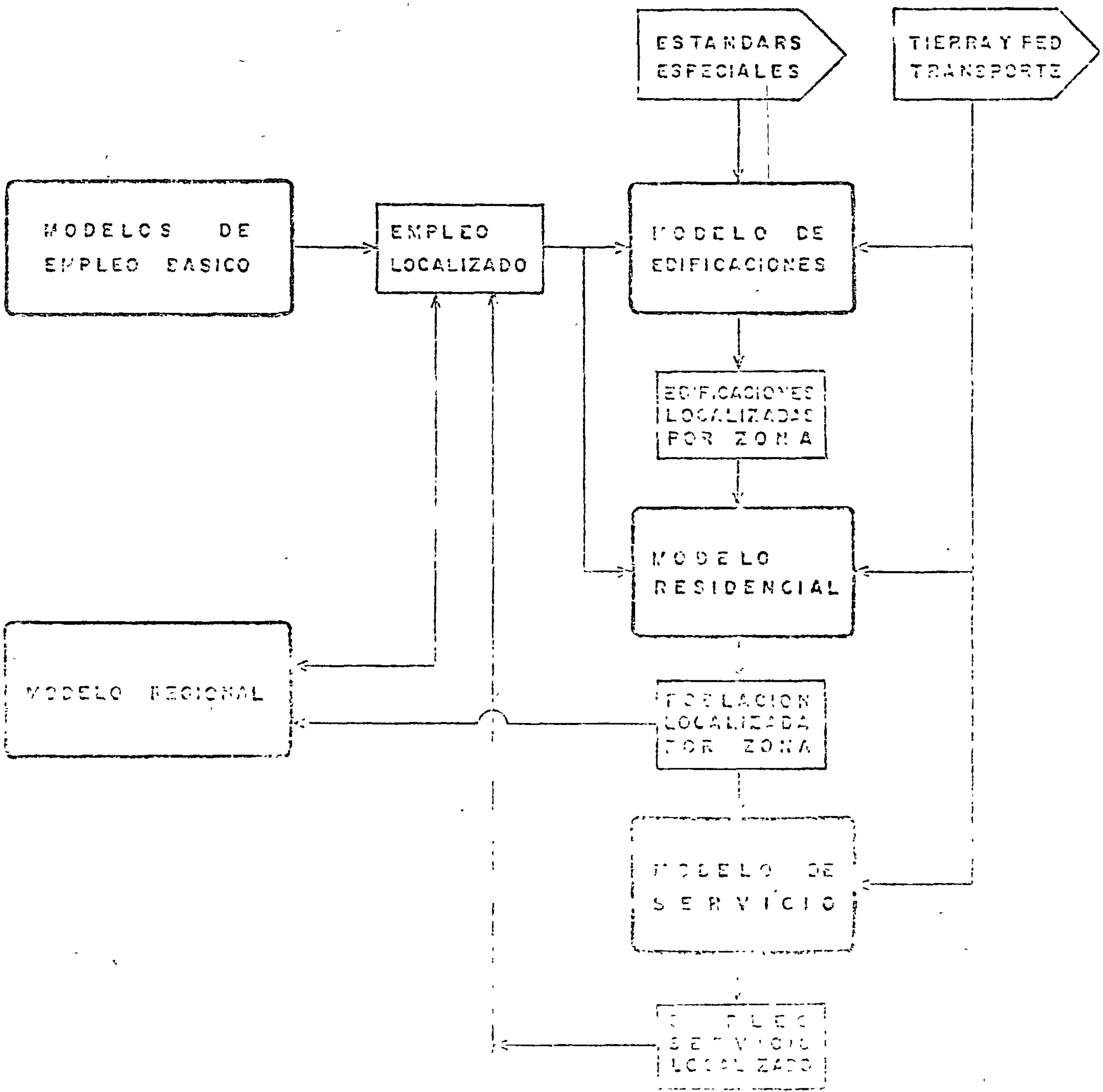


tes zonas la cantidad de edificaciones ó espacio construido generado por el empleo básico como una función de la tierra existente para el desarrollo y las diferentes zonas y su accesibilidad. Este "Stock" ó espacio construido localizado atrae las residentes, los cuales se distribuyen tomando en cuenta también el empleo básico y la accesibilidad. Finalmente el empleo de servicio es generado por la actividad residencial y su localización es función de la localización de los residentes y accesibilidad tomando en cuenta un parámetro de escala que simula la agrupación de este tipo de empleo.
El Modelo es iterativo; el empleo de servicio calculado en la primera iteracción se añade al empleo básico y el Modelo recorre el mismo proceso hasta llegar a un estado de equilibrio.
La población y el empleo total por zona deben sufrir una comprobación de consistencia con el Modelo Regional. Como se indicó en el Modelo de Empleo Básico, el Modelo de Edificaciones y Actividades alimenta al Modelo de Empleo Básico con la Tierra y la Población que necesitará ese Modelo para la próxima iteración.
En la Fig. 7. se aprecia la estructura general del Modelo.
Al igual que en el Modelo de Empleo Básico se puede plantear una fórmula general para los diferentes Modelos. Esta es:

$$Q(i,j,x;t+1)=O(i,x;t+1)W(j,x;t+1)\exp(-L(x)d(i,j))A(i,x;t+1)$$

ESTRUCTURA GENERAL DEL MODELO DE EDIFICACIONES Y ACTIVIDADES

Fig.7



donde,

$Q(i,j,x;t+1)$ = Espacio ó actividad de tipo $\underline{x}$ localizada en zona $\underline{j}$, originada en zona $\underline{i}$ durante el período $\underline{t+1}$

$O(i,j,x;t+1)$ = Generador de espacio construido ó actividad de tipo $\underline{x}$ en la zona origen $\underline{i}$, delante el período $\underline{t+1}$

$W(j,x;t+1)$ = Atracción de la zona $\underline{j}$ a Espacio construido ó actividad tipo $\underline{x}$, durante período $\underline{t+1}$

$\exp(-\beta(x)d(i,j))$ = "Impedancia" para espacio construído ó actividad de tipo $\underline{x}$ entre zona $\underline{i}$ y zona $\underline{j}$

$A(i,x;t+1)$ = Factor de normalización, calculado como:

$$A(i,x;t+1)\ \left(\sum_j W(j,x;t+1)\exp(-\beta(x)d(i,j))\right)^{-1}$$

Al sumar la interacción de localización respecto al origen e introducir un factor de escala, la cantidad total de actividad ó de espacio construído es obtenida en la zona destino $\underline{J}$

$$\sum_i Q(i,j,x;t+1) = \sum_i Q(i,j,x;t+1).V(x)$$

donde $\underline{V}(x)$ es el factor de escala, el cual se obtiene.

$$\underline{V}(x) = \frac{\sum_i \sum_j Q(i,j,x;t+1)}{\sum_i O(i,x;t+1)}$$

$\sum_i \sum_j Q(i,j,x;t+1)$ = cantidad total de espacio construído ó actividad $\underline{x}$

$\sum_i O(i,x;t+1)$ = total de generador de actividad ó espacio construído tipo $\underline{x}$

3.2 SUBMODELO DE EDIFICACIONES

La cantidad de espació construído en cada zona se calcula:

$$F(i,j;t+1) = \sum_j E(i,j;t+1)W(j,t+1)\exp(-B(7)d(i,j))A(i;t+1)$$

donde,

$F(i,j;t+1)$ = Espacio construído localizado en zona $\underline{j}$ con origen en zona $\underline{i}$, durante el período $\underline{t+1}$.

$\sum_j E(i,j;t+1)$ = Empleo total en zona $\underline{i}$ en período $\underline{t+i}$. En la primera iteración este empleo es solo el empleo básico. Luego se añade el empleo de servicio creado por los residentes.

$W(j;t+i)$ = Atracción de la zona Destino, calculado como:

$V(j;t+1) = L(j,x;t+1)$

Es decir la tierra disponible en zona j.

$\exp(-B(7)d(i,j))$ = Impedancia entre la zona i y zona j para las edificaciones.

$A(i;t+1)$ = Factor de normalización

$$A.(j;t+1) = \left(\sum_{j} V(j;t+1)\exp(-B(7)d(i,j))\right)^{-1}$$

Luego el total de espacio construído en la zona j se calcula:

$$F(j,t+1) = \sum_{i} F(i,j;t+1)\ V(1)$$

donde,

$$V(1) = \frac{\text{Cantidad total de espacio construído (dato)}}{\text{Empleo total urbano (Modelo Regional).}}$$

Como puede observarse en la fórmula el espacio construído se localiza como generador por el empleo total en zona origen i cual es atraído a las zonas destino j según la tierra disponible y aplicando una tasa de espacio construído/empleo.

3.3 Sub-Modelo Residencial

La población residencial en cada zona j se calcula:

$$R(i,\ j,\ t+1) = \sum_j E(i,j;t+1)W(j,t+1)\exp(-B(3)d(i,j))A(i,t+1)$$

donde,

$R(i,j;t+1)$ = Empleo residente que trabaja en zona <u>i</u> y vive en zona <u>j</u>, durante periodo <u>t+1</u>

$\sum_j E(i,j;t+1)$ = Empleo <u>t+1</u> urbano en zona de origen <u>i</u> en período <u>t+1</u>

$W(j,t+1)$ = Atracción en zona destino <u>j</u> para población residencial en período <u>t+1</u>. Este se calcula como la cantidad de espacio construído para uso residencial.

$$W(j;t+1) = F(j;t+1) - \left(\sum_j E(i,j;t+1)V(2)\right)$$

donde,

$V(2)$ = la cantidad de espacio construído utilizado por el empleo.

$$V(2) = \frac{\text{Total de Espacio construído usado por empl}}{\text{Total de Empleo}}$$

$\exp(-B(8)d(i,j))$ = "Impedancia" para localización residencial entre zona <u>i</u> y zona <u>j</u>.

$A(i;t+1)$ = Factor de normalización, calculado como:

$$A(i;t+1) = \left(\sum_j W(j;t+1)\,\exp(-B(8)d(i,j))\right)^{-1}$$

luego la población en zona $\underline{j}$ se calcula:

$$R(j;t+1) = \sum_i R(i,j;t+1).V(3)$$

donde,

$V(3)$ = La cantidad de población por empleo residente.

$$V(3) = \frac{\text{Población total (Modelo Regional)}}{\text{Empleo total (Modelo Regional)}}.$$

Como puede observarse en la fórmula en Empleo que trabaja en zonas $\underline{i}$ es atraído a las zonas $\underline{j}$ por el espacio construido residencial y luego un factor de escala transforma el empleo residente en población.

3.4 Sub-Modelos de Servicios

La cantidad de empleo de servicio generado por la población urbana en cada zona se calcula:

$$E(i,j;t+1) = P(i;t+1)W(j,t+1)\exp(-B(9)d(i,j))A(i;t+1)$$

donde,

$E(i,j;7;t)$ = Empleo de servicio que trabaja en zona $\underline{j}$ generado por la población residente en zona $\underline{i}$, en período $\underline{t+1}$

$P(i;t+1)$ = Población residente en la zona i durante período $t+1$

$W(j,t+1)$ = Atracción de la zona j para el empleo de servicio, durante período $t+1$

$$W(j,t+1) = \sum_i E(i,j;t+1)$$

La atracción al empleo de servicio es el empleo total por zona.

P = Parámetro que representa las ecuaciones de escala del empleo de servicio.

$\exp(-B(9)d(i,j))$ = "Impedancia para el empleo de servicio entre la zona i y la zona j.

En la variable $E(i,j,7;t)$ hasta ahora solo hemos calculado la distribución de la población que generará el empleo de servicio.
Para calcular el empleo de servicio se usa la tasa de empleo de servicio/población.

Formalmente,

$$E(j,7;t+1) = \sum_i E(i,j;t+1).V(4)$$

donde,

$V(4)$ = Tasa de empleo servicio-población

$$V(4) = \frac{\text{Empleo total de servicio (Modelo Regional)}}{\text{Población total (Modelo Regional)}}."$$

Calculado el empleo de servicio se añade al empleo básico y el Modelo de Edificaciones y Actividades itera hasta llegar a un estado de equilibrio.

3.5 Posibles Restricciones

Al igual que en el Modelo de Empleo Básico se consideró, siguiendo la tradición de Lowry(1964) conveniente incluir en el Modelo opciones para probar diferentes tipos de restricciones de planificación. Ello puede ser instrumentado por vectores de control para diferentes zonas y actividades ó con restricciones de densidad. Las restricciones de densidad funcionan en forma similar a las del Modelo de Empleo Básico, el cual se indica en forma diagramática en la Fig. 6.

III. MODELOS DE TRANSPORTE

El Modelo de Transporte está ligado al Modelo Intra - Urbano debido a que las variables fundamentales que utiliza son derivadas de la distribución de actividades que realiza El Modelo Intra - Urbano. Esta relación es muy importante debido que los Modelos "Estandar" de Transporte, donde no existe interacción entre la localización de actividades y el comportamiento del sistema de transporte, restan importancia al transporte como instrumento para controlar el desarrollo urbano.

Siendo que el transporte está por lo general bajo el control efectivo de las autoridades ya bien locales o centrales, es evidente que para intentar controlar el desarrollo urbano este instrumento, la planificación del transporte, es importante debido a la factibilidad de implementar políticas en esta área y el impacto del Sistema de Transporte en el Desarrollo Urbano.

La relación entre el Modelo Intra - Urbano y el Modelo de Transporte es una relación de mutua interdependencia. La Distribución de las actividades que realiza el Modelo Intra-Urbano está determinada en parte por la Matriz de accesibilidad entre zonas, la cual se desprende las características de la Red de Transporte. La Generación de Viajes en el Modelo de Transporte depende de la localización de las actividades y un cambio en la localización de las actividades afecta la generación de viajes. Al mismo tiempo en cambio en la generación de viajes ó en el congestionamiento de la Red afecta la Matriz de accesibilidad lo cual afecta la localización de las actividades.

Actualmente el funcionamiento del Modelo de Transporte no está totalmente acoplado al Modelo Intra - Urbano. Debido a esta circunstancia y a la complejidad misma del Modelo de transporte nos limitaremos en el presente informe a realizar una descripción breve de los elementos centrales de su estructura. En un informe subsiguiente se describirán en detalle este Modelo.

1. Descripción General

La relación entre el Modelo de Transporte y el Modelo Intra - Urbano es una relación más estrecha que la que existe El Modelo Regional y el Modelo Intra - Urbano.

La relación entre los dos Modelos está concebida como un proceso iterativo, el cuál puede apreciarse en la Fig. 8.

Si un cambio se produce en la localización de las actividades, la relación entre ellas, específicamente los flujos de transporte cambian entre las diferentes zonas. Esto afecta la congestión de la red de transporte y el tiempo de viaje entre zonas, lo cual conduce a cambiar la Matriz de Accesibilidad que inicialmente tenia como entrada El Modelo Intra - Urbano. Si se introduce la nueva Matriz de accesibilidad en el Modelo Intra - Urbano ello conduce a cambiar la localización de las actividades. Este proceso iterativo se repite hasta llegar a un estado de equilibrio.

El Modelo de Transporte está dividido en cuatro Sub-Modelos fundamentales:

1.1 Sub-Modelo de Generación

En este Sub-Modelo el número de viajes generados (Origen ó Destino) es determinado en cada zona dependiendo de la distribución de las actividades, ó las edificaciones. Los viajes origines y destino se calculan para la hora pico y para 24 horas.

1.2 Sub-Modelo de Distribución y División Modal

El Sub-Modelo de Distribución y División Modal distribuye los viajes generados en el Modelo de Generación entre las diferentes zonas y en los diferentes modos.

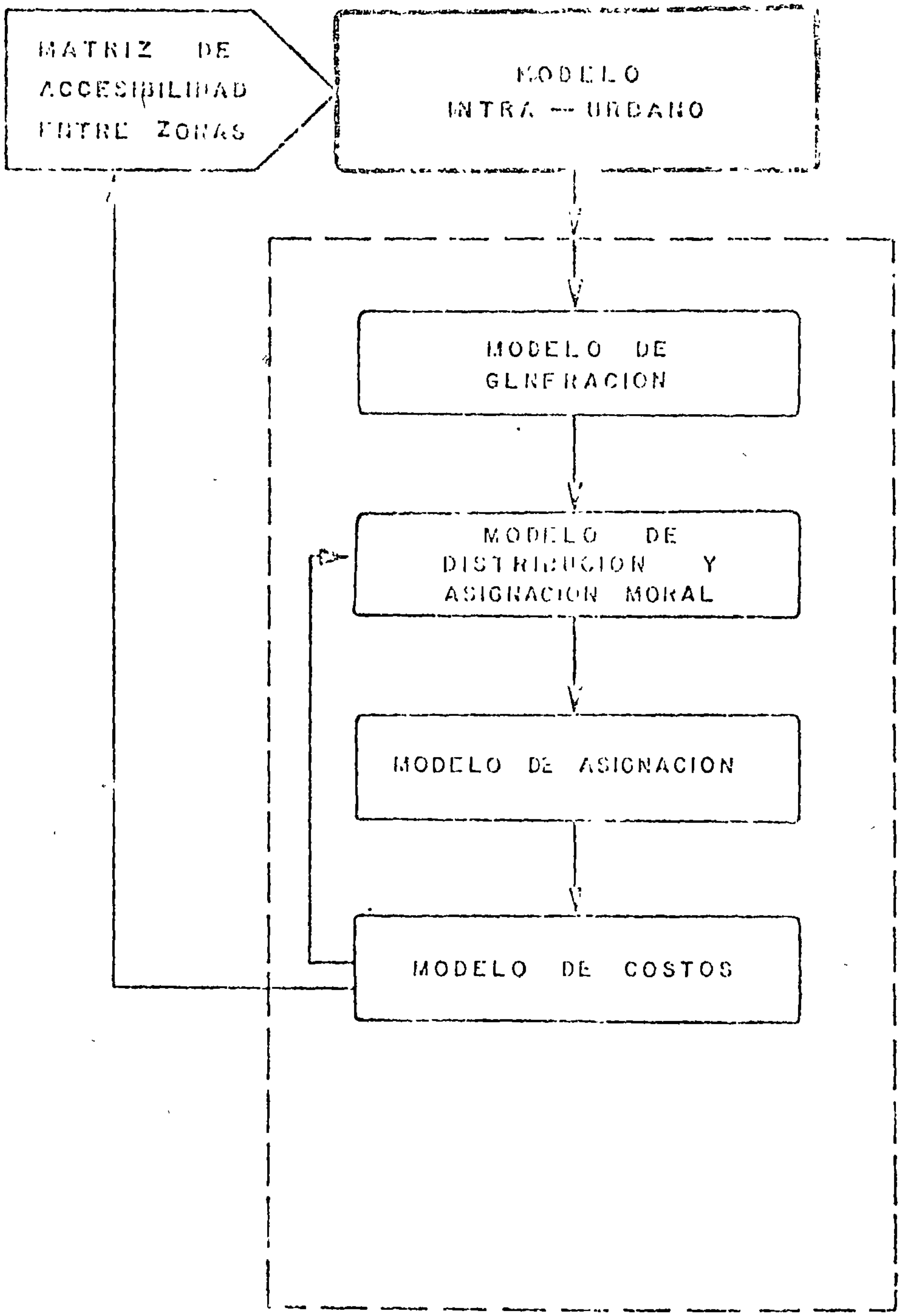
Calcula la Matriz de viajes, cuyo elemento típico T(i,j;k) representa el número de viajes que van de la zona origen i a la zona destino j por el modo k (privado, buses, etc). Esta distribución utiliza un modelo de interacción espacial doblemente restringido el cual se deriva utilizando el método de maximización de la entropía. Para realizar la distribución de viajes es necesario conocer el costo de viaje entre zonas por modos C(i,j,k). En la primera iteración solo se tiene una idea aproximada de este valor pero su valor se ajusta en el Submodelo de Costo generalizado.

1.3 Sub-Modelo de Asignación

En el Modelo de Asignación en él los viajes T(i,j,k) calculados por el Modelo de Distribución y División Modal son asignados a la Red de Transporte, es decir a su conjunto de arcos y nodos. Los diferentes modos de transporte tienen un comportamiento diferente por lo que existe diferente asignaciones para Autos Privados, Buses, Por puesto, Metro, y Peatones. El Modelo de Asignación es de tipo probabilístico y en el caso del Transporte Público, cuyas rutas son conocidas, asigna pasajeros a la ruta. En el caso por Modo Peatonal, no se toman en cuenta restricciones de capacidad y la asignación es por caminos minimales. Para los otros modos la congestión afecta el tiempo del viaje y este a la asignación.

ESTRUCTURA DEL MODELO DE TRANSPORTE

FIG. 8



1.4 Sub-Modelo de Costo Generalizado

Este modelo calcula la "Impedancia" de viajar entre las zonas origen i y zona j, por los diferentes modos k como un costo de viaje generalizado. C(i,j;k). La función de costo es compleja y diferente para los modos. El objeto de ello es poder manipular con propósitos de planificación esta función de costo para obtener divisiones modales que sean, por ejemplo más favorable al transporte público. Por ello se incluye en esta función el tiempo de viaje,tarifas,etc. Después que se ha calculado el costo por modo,se calcula en el cos generalizado entre zonas que es entrada para calcular la nueva matriz de accesibilidad que afectará al Modelo Intra-Urbano. El proceso como se dijo anteriormente itera hasta llegar al equilibrio.

MODELO DE EVALUACION

1. Planificación, Evaluación y Modelos
2. El Proceso de Evaluación
3. Estructura del Modelo (MODEVA)
4. Comentarios Finales.

1. PLANIFICACION EVALUACION Y MODELOS

Si bien es cierto, que dentro del campo del diseño existen tantas teorías de planificación como planificadores, también lo es el hecho de que todas ellas, comparten el esquema ó proceso siguiente:

1. Generación de Alternativas

2. La escogencia de una Alternativa.

La escogencia de una de las alternativas, significa una evaluación entre las distintas alternativas generadas en la etapa anterior.
Existe una tercera etapa, que no siempre es incluída en el esquema anterior que es el "control del diseño a través del tiempo". Esto implica otro tipo de evaluación, el del comportamiento del diseño en forma dinámica.
En la práctica el planificador siempre tiene que escoger entre alternativas competitivas.
Esta escogencia sehace hoy en día en forma subjetiva, utilizando la experiencia acumulada, de éxitos y fracasos. Hay también que tomar en cuenta que muchas de las decisiones críticas urbanas, son de tipo político en el que el planificador tiene poca participación real; muchas veces el trabajo del planificador se limita al estudio del efecto positivo o negativo de una serie de decisiones tomadas a priori.
Con el advenimiento o surgimiento de los modelos matemáticos, el planificador está en mejor situación, y puede prever los efectos de ciertas políticas antes de que sean implementadas. Con los modelos matemáticos se pueden generar muchas alternativas diferentes del sistema urbano, introduciendo cambios, por ejemplo en la red vial, en la localización de actividades, etc, y estudiar los efectos de

estos cambios en el sistema global.

Ahora bien, a la par de las descripciones precisas y cuantificables que del sistema urbano dan los modelos urbanos, es necesario disponer de un sistema que evalúe descripciones en forma racional. Este sistema sería, un Modelo el cual proveerá al planificador de un "sistema de evaluación o instrumento de evaluación". Este sistema de evaluación mediría la eficiencia del funcionamiento de la ciudad, la salud de la ciudad", en términos del nivel de satisfacción de las necesidades de la mayoría de la ciudadanía.

Antes de entrar a discutir este sistema, es necesario aclarar los supuestos con los que se van a trabajar en la formulación del Modelo Evaluativo:

1. Todos los habitantes tienen igual importancia, o sea a mayor cantidad de personas beneficiadas, mejor es la alternativa que se evalúa.

2. Los criterios de evaluacion no son propios del Modelo, sino fijados por el ciudadano común, los planificadores y políticos.

   Con esto se intenta que la ciudad no sea evaluada por los intereses de un mismo sector.

3. En la elaboración del Modelo Evaluativo se prefiere la simplicidad versus complejidad.

4. La evaluación, es un proceso que consiste en un diagnóstico de la situación. Para eso hacen falta modelos que describan la realidad (Modelos de estructura urbana, etc), y un modelo normativo que describa la situación deseable. El aspecto crítico de este proceso es la cuantificación de esta situación considerada como normativa que comparada con la situación real de la medida de deficiencia o anormalidad del Sistema Urbano.

5. Las variables de evaluación se escogerán en función de los factores más determinantes de la estructura urbana y también de aquellos factores sobre los cuales el planificador tiene mayor control.

6. Cada variable de evaluación llevará un costo asociado. O sea, para cada alternativa analizada, se dispondrá de una medida de beneficio a la comunidad y de una medida de costo a los agentes de implementación.

2. EL PROCESO DE EVALUACION

En la primera parte del informe definimos el proceso de evaluación como la comparación del estado real de una variable con el estado normativo, o satisfactorio de esa variable.

Cuando un médico hace un diagnóstico de un paciente lo que hace es comparar el estado real del organismo del paciente con un patrón de funcionamiento aceptable del organismo humano.

Cuando este diagnóstico se hace al sistema urbano, la dificultad aparece al tratar de establecer el patrón de funcionamiento satisfactorio o aceptable del sistema. El nivel de lo que se considera "aceptable" cambia temporalmente y para muchos es el nivel mínimo que hay que proveer a la ciudadanía, para evitar posibles conflictos sociales.

En realidad se cree que un estandard es más que una situación, un objetivo que hay que alcanzar. Por ejemplo, en el caso de ciudades medianas, se puede establecer que el viaje promedio al trabajo debería aproximarse a 2 kilómetros y el viaje promedio a los servicios a 1 kilómetro.

Definir todos los niveles de normalidad para cada variable, (densidades residenciales, empleo, etc), implica un estudio, que debe ser realizado por el organismo o agentes de decisiones que realicen la planificación.

Cuando se entra en el problema de la comparación del estado real y normativo de una variable surge el problema de que no todas las variables tienen el mismo tipo de comportamiento, o una misma estrutura normativa. Por ejemplo, al aumentar el valor de una variable no necesariamente nos acercamos a un óptimo; puede suceder todo lo contrario.

Del análisis realizado se clasificacion tres tipos de variables:

Tipo 1 Variable Puntual

Tipo 2 Variable Decremental

Tipo 3 Variable Incremental.

Estas variables hay qye reducirlas a un sistema normativo común, es decir un sistema en el cual un aumento de valor de la variable signifique un beneficio. Para ello se han elaborado las siguientes fórmulas:

TIPO 1 $$EVA = \frac{S}{|V-S| + S}$$

TIPO 2 $$EVA = \frac{S}{V}$$

TIPO 3 $$EVA = \frac{V}{S}$$

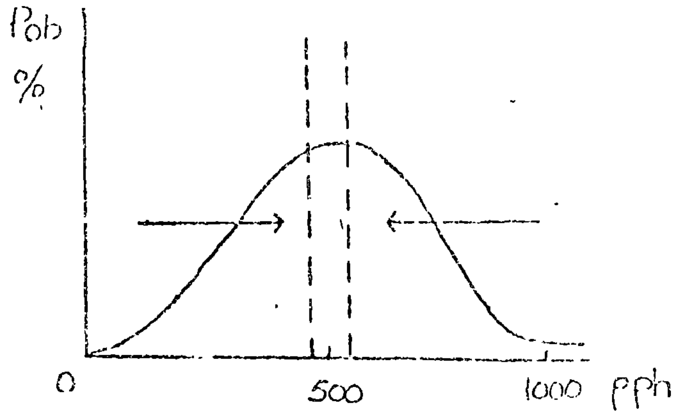
donde,

V = Variable original

S = Estandard asociado a la variable original

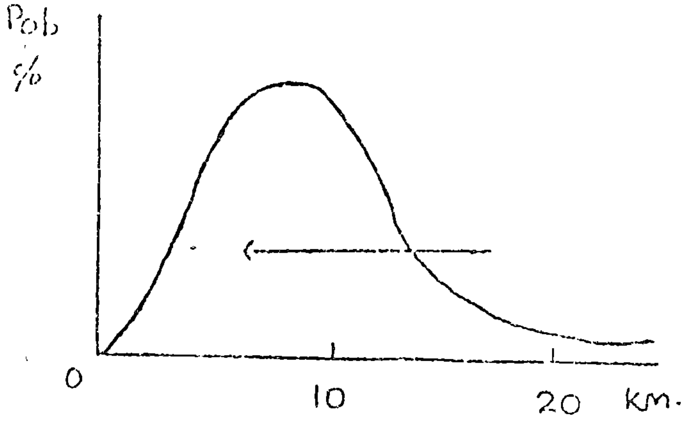
EVA = Variable en un sistema normativo común.

TIPOS DE VARIABLES Y DIRECCION DE MEJORA



TIPO 1 PUNTUAL

Ej. Densidad

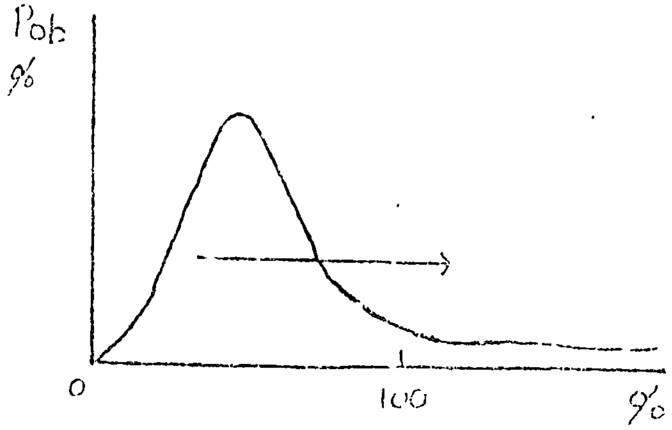


TIPO 2 DECREMENTAL

ej. Accesibilidad

Ruido

Contaminación



TIPO 3 INCREMENTAL

ej. oportunidad

provisión de servicios

la flecha indica la dirección de mejora.

Este proceso de evaluación es muy flexible y se puede aplicar a topo tipo de variable cuantificable. Inicialmente se generó una lista de una doscientas variables, que se estructuraron en función de su causalidad. Debido a las limitaciones de los modelos descriptivos (en este caso los Modelos mencionados en los apartes anteriores), y lo limitado del campo dentro del cual actúan los planificadores se redujeron las variables a una dieciocho (ó veinte) que se consideraron claves para la evaluación de una proposicón de desarrollo urbana. Las variables forman cuatro grupos o familiar de variables que son:

DENSIDADES

1. Densidad residencial neta
2. Densidad residencial bruta
3. Densidad de empleo
4. Intensidad de construcción.

ACCESIBILIDAD:

5. Distancia promedio entre zonas
6. Distancia promedio al trabajo
7. Distancia promedio a servicios.

OPORTUNIDAD:

8. Oportunidad de trabajo
9. Oportunidad de servicios
10. Oportunidad de educación primaria
11. Oportunidad de educación media
12. Oportunidad de asistencia médica.

PROVISION:

13. Provisión de vivienda
14. Provisión de agua
15. Provisión de cloacas.
16. Provisión de electricidad
17. Provisión de parques infantiles
18. Provisión de areas verdes.

Debe hacerse énfasis que esta lista no pretende ser exhaustiva y que el planteamiento teórico del modelo es capaz de incluir otras variables cuantificables.

3. ESTRUCTURA DEL MODELO

La estructura del Modelo evaluativo, llamado MODEVA, puede verse en la Fig. 9. Para cada alternativa de desarrollo el modelo sigue las siguientes etapas:

3.1 Transferencia de las variables de los Modelos Descriptivos.

En Modeva se utilizan todas las variables calculadas en los Modelos descriptivos, especialmente de los Modelos de estructura urbana, ó sea población, empleo, espacio construído, tierra disponible, y distancia entre zonas.

3.2 Leer entradas de Modeva

Los insumos principales de Modeva son:

Variable: Las variables de evaluación, su tipo estandard, y peso.

Capacidades: Las capacidades de los sistemas urbanos

Coeficientes: Los coeficientes de demanda de servicios urbanos.

Costos: Los índices de costos de provisión de servicios.

3.3 Cálculo de las Variables de Evaluación.

El cáluclo de las variables de evaluación se hace mediante cuantro subrutinas especiales, una para cada grupo de variables que son: densidad, accesibilidad, oportunidad y provisión de servicios. Las fórmulas generales de cada subrutina son las siguientes:

Variables; variables de evaluación, su tipo, estandard y peso.

SUBRUTINA	FORMULA
DENSID	$Vi = \frac{Pi}{Ai}$
ACESO	$VI = \sum_{j=1}^{n} \frac{Dij \, . \, HVTij}{\sum_{j=1}^{n} HTVij}$
OPORT.	$VI = \sum_{i=1}^{n} \frac{CAP_I \, . \, \exp(-B.D\, ij)}{\sum_{j=1}^{n} P_I \, . \, C_j(-B.Dij)}$
PROVN	$VI = \frac{CAP\ I}{P_I . C_j}$

donde,

V = Variable

P = Población ó personas

A = Area

D = Distancia

MV = Matriz de Viajes

CAP = Capacidad del sistema urbano

C = Coeficiente de demanda.

3.4 Evaluación

Una vez calculados los valores de las variables por las subrutinas mencionadas, se introducen estos valores en un arreglo Vij. Según el tipo de la variable se transfiere el control a una de las tres instrucciones, que coloquen las variables en un sistema normativo común, y se comparan con un standard.

TIPO J = 1 $\quad \text{EVA ij} = \dfrac{Sj}{|Vij - Sj| + Sj}$

TIPO J = 2 $\quad \text{EVA ij} = \dfrac{Sj}{Vij}$

TIPO J = 3 $\quad \text{EVA ij} = \dfrac{Vij}{Sj}$

donde,

V = variable

S = estandard

i = suscrito de la zona urbana

j = suscrito de la variable.

Este procedimiento, de comparar el valor de la variable con un estandard, forma el aspecto central de proceso de evaluación propuesto. En ello se convierte la matriz Vij de variables de diversas escalas de medicción en una matriz EVAij que posee una estructura probabilística común donde todas las variables se miden en forma similar. Los elementos de esta matriz se multiplican luego por los pesos de la variable y también de la zona.

$$EVA = \frac{EVA\ ij\ .\ PESO\ j\ .\ Pi}{POB\ .\ TOTAL}$$

donde,

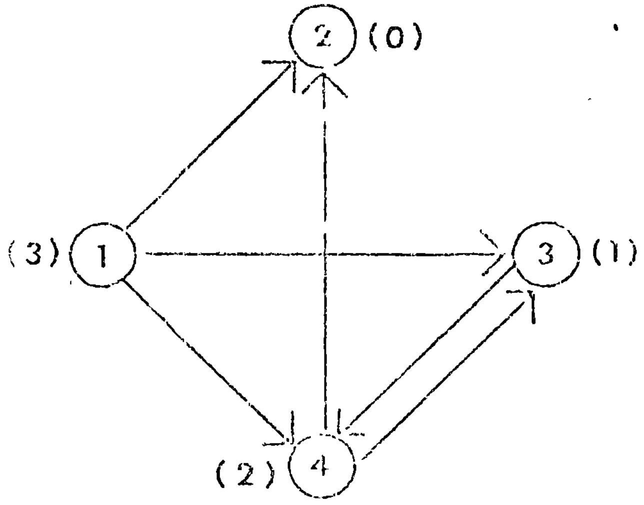
$PESO_j$ = peso de la variable j

P_i = población residencial en zona i

La asignación de los pesos a cada variable es una etapa crítica en el proceso de evaluación.

Si se quiere que el valor normativo asignado represente la opinión de los varios sectores de la comunidad, (Residentes, Políticos, Planificadores) debe hacerse un censo de opiniones de estos grupos y derivar normas y pesos compuestos.

Un camino posible sería el de utilizar la misma estructura de variables que se mencionó al inicio de este informe. La variable que incida o influya sobre un número grande de otras variables puede decir que tendrá un peso alto.



En este ejemplo la variable (1) influye sobre tres variables la (3) sobre una y la (2) sobre ninguna, por lo tanto la variable (1) tiene el peso mayor.

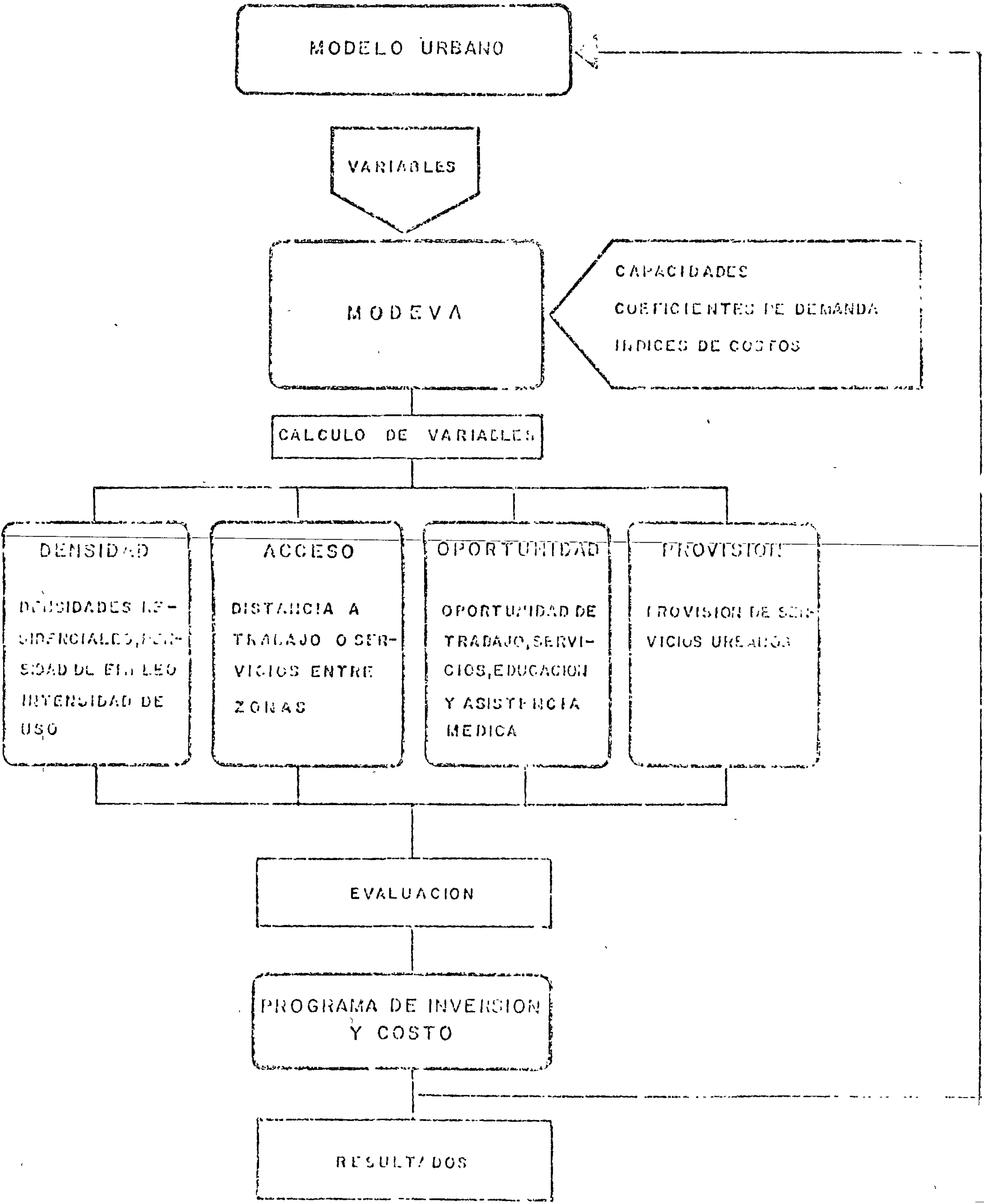
3.5 Programa de Inversiones y Costos

En base a la demanda y oferta de las variables urbanas como vivienda, escuelas,etc, el modelo hace un cálculo de la cantidad de inversión pública necesaria de servicio en cada zona de la ciudad y por cada año. También hace una estimación del costo de la obra en bolívares, para cada uno de los distintos servicios públicos, por ejemplo:

escuelas primarias
escuelas secundarias
vivienda
acueducto
cloacas
electricidad
parques

ESTRUCTURA DEL MODELO EVALUATIVO

FIG S

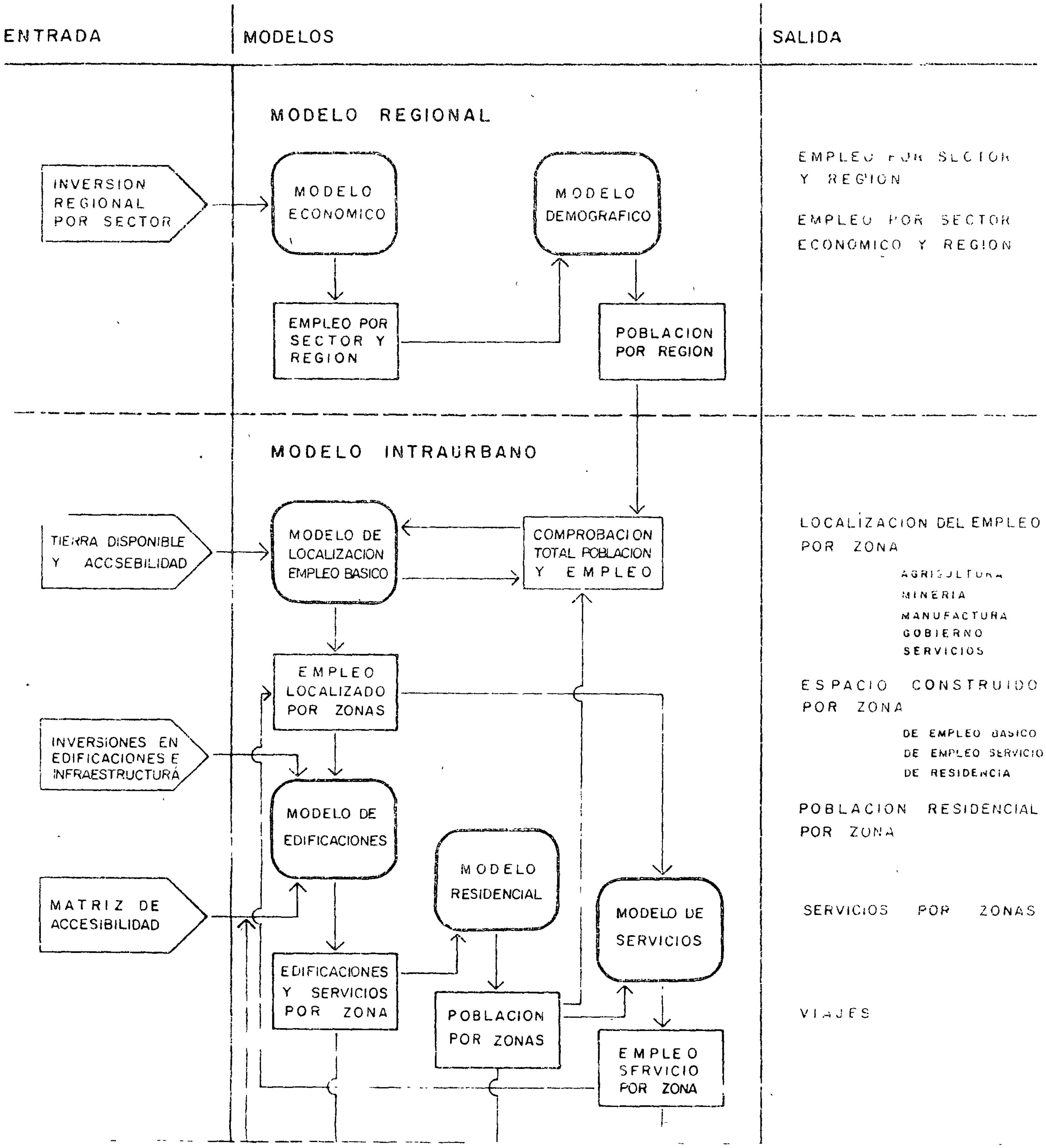


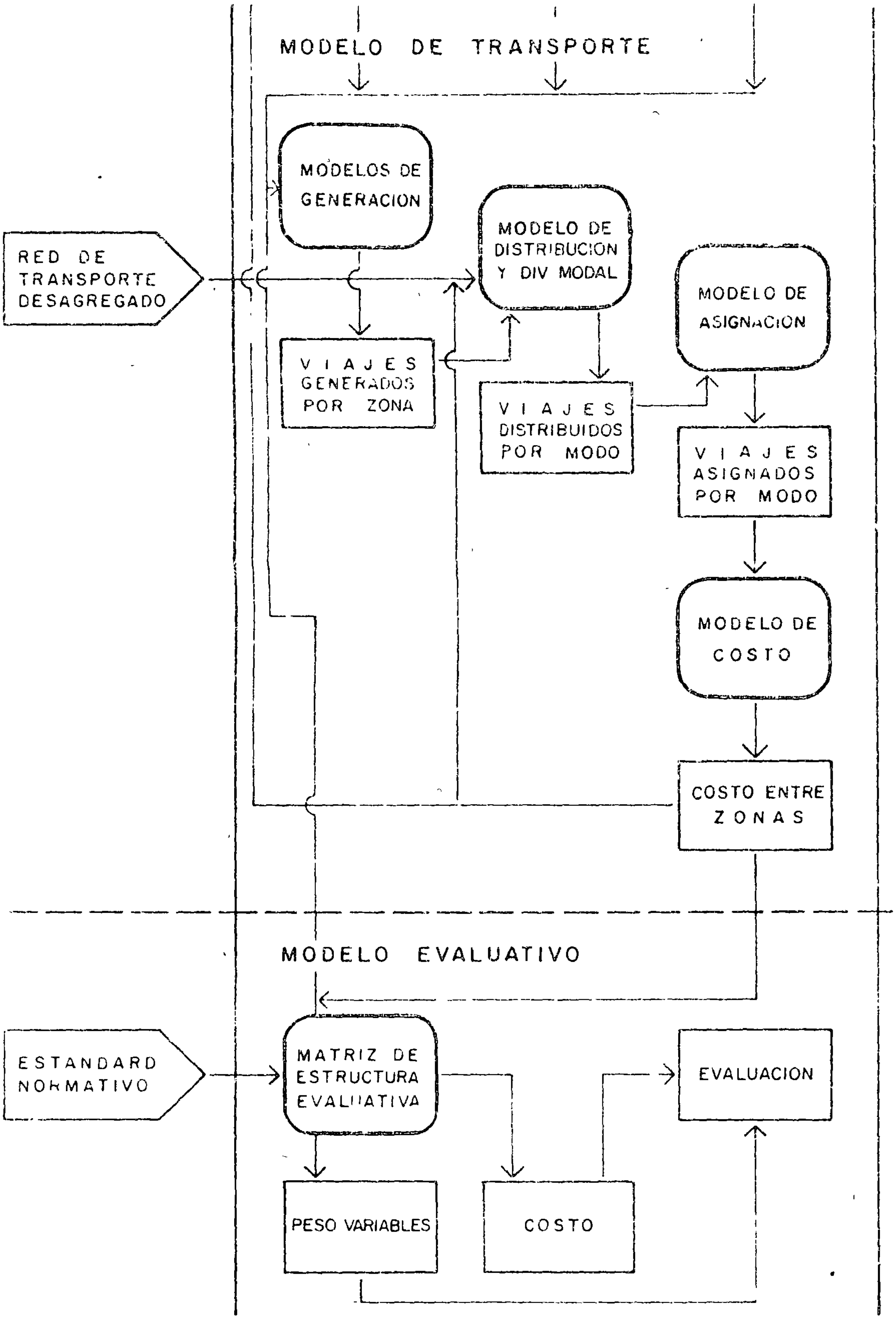
4. COMENTARIOS FINALES AL MODELO DE EVALUACION

Existen una serie de variables o indicadores físicos y socio económicos, que no se estan considerando en este modelo como son: micro-clima, sequedad del sub-suelo, distribución de la vegetación, mobilidad rural, estructura familiar, migración, relaciones sociales, delincuencia, etc , pero el tratamiento de estas variables significa el desarrollo de una serie de conceptos socio-económicos, que llevarían en complicar el modelo, y nuestro interés para esta etapa de trabajo es mantener la estructura lo más simple posible, sin que se pierda por ello la capacidad de evaluar una realidad descrita.

ESTRUCTURA GENERAL DEL CONJUNTO DE MODELOS

fig 10





VIAJES ORIGINADOS POR ZONA Y POR MODO

VIAJES ENTRE ZONA POR MODO

CARRO PRIVADO
BUSES
POR PUESTO
PEATONAL
METRO

TIEMPO VIAJES POR MODO

COSTO DE VIAJES POR MODO

EVALUACION ESPECIFICA POR VARIABLE EN RELACION A COSTO

RESIDENCIAL
SERVICIO
TRANSPORTE, ETC

EVALUACION GLOBAL

Research Note

Edited by Michael B. Teitz

A MATRIX FORMULATION OF THE LOWRY MODEL FOR INTRAMETROPOLITAN ACTIVITY ALLOCATION

Robert A. Garin

Introduction

In the past fifteen years many models have been developed in the field of urban planning. The earlier attempts focused mostly on the transportation system—traffic distribution or assignment—mainly because they were proposed for application in metropolitan transportation studies. More recently, however, increasing attention has been directed toward models for the allocation of activities in urban areas, concerning the distribution of residential, commercial, industrial, and other land uses over space.

A most important one among these, the Lowry Model, uses an iterative procedure to generate a distribution of employment and population in a region from a given distribution of employment in the basic sector. (The basic sector here is defined as consisting of those activities which are locationally exogenous, those which are considered to locate independently of the location of other activities.)

At each iteration, an additional population and the non-basic employment necessary to serve it are generated and distributed over the area. The procedure is stopped as soon as total population in the region reaches some predetermined value. For instance, in the Bay Area Simulation Study test model, only five iterations were needed to satisfy this condition.

However, such a constraint added merely to prevent an infinite number of iterations of the system is arbitrary, it is not necessary for the conceptual framework of the model. The initial purpose of the approach proposed here is to find a more logical way to deal with the infinite number of iterations needed. It soon appeared that other benefits could be obtained from this simple matrix formulation.

Formulation of the Model

Let us consider a region with n zones (numbered from 1 to n). Basic employment is assumed to be given for the moment, and can be represented as an n-vector.

$$E^b = (E_1^b, E_2^b, \quad \ldots, E_n^b)$$

Residential location of basic employees. The residential location of basic employees is found by means of a *journey to home function* a'_{ij}, which distributes the residences of people working in any zone i over all zones j in accordance with observed behavior or some model.

Further multiplication by the labor participation rate in zone j, a_j will give the basic population living in j.

$$P^b = E^b A = E^b A' [a] \qquad (1)$$

where $P^b = (P_1^b, P_2^b, \ldots P_j^b, \ldots P_n^b)$ is the basic population vector,

$A = [a_{ij}] = [a'_{ij} a_j]$,

$A' = [a'_{ij}]$ is the *journey to home* distribution matrix,

$$\text{and } [a] = \begin{bmatrix} a_1 & & & & 0 \\ & a_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ 0 & & & & a_n \end{bmatrix}$$

is the labor participation rate diagonal matrix.

Location of population-serving activities. Non-basic employment necessary to serve population P_i^b is obtained as a fraction of P^b and then distributed over all zones j in terms of *journey to shop function* b'_{ij}.

This operation can be written

$$E^{(1)} = P^b B = P^b [\beta] B' \qquad (2)$$

where the index 1 in $E^{(1)}$ is the iteration number, and where $B = [b_{ij}] = [\beta_i b'_{ij}]$.

$B' = [b'_{ij}]$ is the *journey to shop* distribution matrix,

$$\text{and } [\beta] = \begin{bmatrix} \beta_1 & & & & 0 \\ & \beta_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ 0 & & & & \beta_n \end{bmatrix}$$

is the population-serving employment ratio, diagonal matrix.

Combining (1) and (2) we obtain

$$P^b = E^b A$$

$$E^{(1)} = P^b B = E^b (AB)$$

We can now perform the same sequence of operations—residential location now of all employees, location of population-serving activities—with non-basic employment $E^{(1)}$. This will generate a new non-basic employment $E^{(2)}$.

$$P^{(1)} = E^{(1)} A = E^b (AB) A$$

$$E^{(2)} = P^{(1)} B = E^{(1)} (AB) = E^b (AB)(AB) = E^b (AB)^2$$

It becomes apparent that the successive iterations will provide additional employments and populations

$$E^{(2)} = E^b (AB)^2 \rightarrow P^{(2)} = E^b (AB)^2 A$$

$$E^{(3)} = E^b (AB)^3 \rightarrow P^{(3)} = E^b (AB)^3 A$$

$$\cdot$$

$$E^{(n)} = E^b (AB)^n \rightarrow P^{(n)} = E^b (AB)^n A$$

[illegible] 1966

Robert Garin is presently studying at the Institute of Transportation and Traffic Engineering at the University of California, Berkeley.

stances the vector of basic employment, and hence the population vector as well, stabilizes around a distribution that grows at a constant rate of growth and remains proportionately constant, that is, each sub-region's share of the total population and employment remains the same over time, even though the total itself increases at a constant rate (Moreover, it can be shown that this constant rate of growth is the dominant characteristic root of the matrix G, and that the stationary distribution is the associated characteristic vector)

3. *A Numerical Example of Stationarity*

As an illustration, consider Garin's example. His hypothetical data describe the population and employment distributions at a particular point in time, 1960, for a theoretical town, Lartsville. From his data source [1], we may borrow the hypothetical 1970 distribution of basic employment and assume that it was predicted by the following growth process:[1]

$$E^b_{1970} = E^b_{1960}\,G$$

$$= (4500,0,0,0,0,9000,0,1000,0,0)\begin{bmatrix} 2\ 00 & 0 & 0 & 0 & 0 & .08 & 0 & 02 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .20 & 0 & 0 & 0 & 0 & 1\ 10 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 10 & 0 & 0 & 0 & 0 & 04 & 0 & 1.41 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= (10900,0,0,0,0,10300,0,1500,0,0).$$

Hence we can easily solve for P_{1970} by applying Equation 1 with Garin's values for A and $(I-AB)^{-1}$.

$$P_{1970} = (12053,10912,13520,9412,13761,12158,$$
$$6992,11980,11698,10741).$$

Now consider the stationary rate of growth and spatial distribution It can be demonstrated that raising G to higher and higher powers ultimately yields the relationship:

$$G^{s+1} = v\,G^s, \tag{5}$$

where v, a constant, is the stationary rate of growth [3]. This means that at and beyond $n=s$, the system is stable and

$$E^b_{s+1} = E_s^b\,G = E_1^b\,G^s = v(E_1^b\,G^{s-1}) = vE_s^b, \tag{6}$$

and, therefore,

$$P_{s+1} = vP_s. \tag{7}$$

That is, at stationarity, each sub-region's share of the total population (or employment) remains constant and grows at a rate of v every period Denote this allocation among the sub regions by the vector W where

$$W_{si} = \frac{P_{si}}{\sum_i P_{si}} \tag{8}$$

Then, for our hypothetical city, we have the following proportional allocations over time.

Zone	Basic Employment 1960	Basic Employment 1970	Basic Employment Stat	Total Population 1960	Total Population 1970	Total Population Stat
1	.310	.480	.892	098	107	127
2	.0	.0	.0	092	096	108
3	0	.0	.0	.116	119	130
4	.0	0	0	083	083	085
5	.0	.0	0	121	122	123
6	621	.454	.079	.115	107	093
7	.0	0	0	.058	062	.058
8	069	066	029	106	106	105
9	0	.0	.0	.109	103	091
10	.0	.0	.0	.102	095	080
Total	1.000	1 000	1 000	1.000	1.000	1 000

The stationary ten-year growth rate $v=2.0$ (Notice that during the ten-year period 1960–1970 the growth rate was 1.6.) Thus we may conclude that, as of today (today being 1960 or 1970, depending on whether the G was observed or predicted), the City of Lartsville is heading toward an ultimate rate of growth of 2 0 per decade and an ultimate spatial allocation of people to sub-regions as described above. Tomorrow's "speedometer" reading may indicate a sudden shift.

4. *A Relation Between Population Growth and Employment Growth*

Consider a growth matrix S which when applied to the population vector P_t projects it forward one time period [4], that is,

$$P_{t+1} = P_t S. \tag{9}$$

Each diagonal element of such a matrix denotes the sum of *one*, the proportion of people who remain in a particular sub-region during a unit time interval, and *two*, the natural increase (births over deaths) rate in that sub-region. Off-diagonal elements denote the proportion of people who migrate from one sub-region to another during the unit time interval.

Now consider the relationship between the basic employment growth matrix G and the population growth matrix S. Combining Equations 3 and 9, we find

$$P_{t+1} = E_t^b\,G(I-AB)^{-1}A = P_t\,S \tag{10}$$

But by Equation 1,

$$P_t = E_t^b(I-AB)^{-1}A$$

Therefore,

$$P_{t+1} = E_t^b\,G(I-AB)^{-1}A = E_t^b(I-AB)^{-1}A\,S$$

or

$$E_t^b\,G\,M = E_t^b\,M\,S$$

where

$$M = (I-AB)^{-1}A$$

Hence,

$$G\,M = M\,S \tag{11}$$

and

$$G = M\,S\,M^{-1} \text{ or } S = M^{-1}G\,M \tag{12}$$

Thus, given a population growth matrix S we may find the basic employment growth matrix G, or vice versa This provides a consistency check on both growth processes. It is of interest to note that the stationary growth rates implied by the two growth matrices must be equal.

$$E = E^b + E^{(1)} + E^{(2)} + \cdots + E^{(n)} + \cdots$$

$$= E^b(I + AB + (AB)^2 + \quad + (AB)^n + \quad) \qquad (3)$$

$$P = P^0 + P^{(1)} + \quad + P^{(n)} +$$

$$= E^b(I + AB + (AB)^2 + \quad + (AB)^n + \quad)A \qquad (3')$$

Under certain conditions on the product matrix AB, which are satisfied in our case (see Appendix for more complete discussion) the matrix series (between parentheses) in (3) and (3′) converges to the inverse of the matrix (I − AB). We can now write:

As an illustration,[1] consider the following 10 zone example in which we take simple *journey to home* and *journey to shop* functions:

$$a'_{ij} = b'_{ij} = \frac{e^{-0.1 t_{ij}}}{\sum_j e^{-0.1 t_{ij}}}$$

(where t_{ij} is the travel time from zone i to zone j in minutes) and $\alpha_i = 2.5$ $(i = 1,2, \ldots\ldots, 10)$

$\beta_i = 0.2$ $(i = 1,2, \ldots\ldots, 10)$

Interzonal travel times:

	1	2	3	4	5	6	7	8	9	10
1	2.00	5.25	3.22	10.50	4.94	9.41	15.86	7.94	10.94	12.66
2		1.50	3.75	5.25	5.47	9.97	10.61	8.47	11.47	13.19
3			1.50	7.72	1.72	6.22	13.08	4.72	7.72	9.44
4				2.00	6.00	10.50	5.36	6.00	9.00	10.72
5					1.50	4.50	11.36	3.00	6.00	7.72
6						2.00	15.22	6.85	3.86	3.86
7							2.00	8.36	11.36	13.08
8								2.00	3.00	4.72
9									1.50	1.72
10										2.00

$$A' = \begin{bmatrix} .172 & .124 & .152 & .073 & .128 & .081 & .043 & .095 & .070 & .059 \\ .117 & .170 & .136 & .117 & .114 & .073 & .068 & .085 & .063 & .053 \\ .123 & .117 & .145 & .078 & .143 & .091 & .046 & .106 & .078 & .066 \\ .069 & .118 & .092 & .163 & .109 & .069 & .116 & .109 & .081 & .068 \\ .099 & .094 & .136 & .089 & .139 & .103 & .052 & .120 & .089 & .075 \\ .075 & .071 & .103 & .067 & .123 & .157 & .042 & .097 & .131 & .131 \\ .054 & .091 & .071 & .154 & .084 & .057 & .216 & .114 & .084 & .071 \\ .076 & .072 & .105 & .092 & .125 & .085 & .073 & .138 & .125 & .105 \\ .060 & .057 & .083 & .073 & .099 & .123 & .058 & .134 & .156 & .152 \\ .056 & .053 & .078 & .068 & .092 & .136 & .054 & .125 & .169 & .164 \end{bmatrix} = B'$$

$$(I - AB)^{-1} = \begin{bmatrix} 1.097 & .100 & .117 & .094 & .120 & .097 & .068 & .111 & .101 & .096 \\ .095 & 1.102 & .115 & .098 & .118 & .095 & .073 & .111 & .099 & .089 \\ .095 & .099 & 1.115 & .094 & .119 & .098 & .069 & .112 & .102 & .092 \\ .089 & .099 & .111 & 1.101 & .116 & .095 & .078 & .112 & .102 & .091 \\ .092 & .097 & .114 & .095 & 1.119 & .099 & .070 & .113 & .104 & .093 \\ .089 & .093 & .111 & .092 & .117 & 1.103 & .068 & .113 & .109 & .100 \\ .086 & .097 & .107 & .104 & .114 & .093 & 1.086 & .113 & .103 & .092 \\ .089 & .094 & .111 & .095 & .117 & .099 & .072 & 1.114 & .107 & .097 \\ .087 & .091 & .109 & .093 & .116 & .103 & .070 & .115 & 1.111 & .101 \\ .086 & .090 & .108 & .092 & .116 & .104 & .070 & .115 & .112 & 1.103 \end{bmatrix}$$

A basic employment vector

$$E^b = (4500,0,0,0,0,9000,0,1000,0,0)$$

would then lead to the following distribution of total employment and population.

Zone	Basic Employment	Total Employment	Population
1	4500	5835	7126
2	0	1387	6673
3	0	1640	8383
4	0	1351	6033
5	0	1718	8801
6	9000	10472	8303
7	0	1000	4217
8	1000	2639	7709
9	0	1549	7898
10	0	1405	7352

where $E = E^b(I - AB)^{-1}$;

and $P = EA$.

Journey from Work to Shop

For simplicity we have assumed so far that all shopping trips were originated at the home. That is, a population P_i residing in zone i calls for $P_i \times \beta_i$ population-serving employees, and these were allocated to all zones j, according to a *journey from home to shop function* b_{ij}.

There is no need to maintain this assumption, and we can now consider how shopping trips originated at the place of work can be incorporated into the model For example, we could assume as Lowry did that all trips from work to shop are walking trips and terminate in the same zone where they originate

A number of employees $E_i^{(1)}$, working in zone i will demand $E_i^{(1)} \times \gamma_i$ non-basic employees, which in turn are allocated to their places of work j in terms of a *journey from work to shop function* $C'_{ij}[C]$

In matrix notation.

$$E^2 = E^{(1)}[\gamma]C' = E^{(1)}C \quad \text{where } \gamma = \begin{bmatrix} \gamma_1 & & & & 0 \\ & \gamma_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & \gamma_n \end{bmatrix}$$

is the "employment-serving" employment ratio diagonal matrix and,

$$C = [C_{ij}] = [\gamma_i C'_{ij}]\text{, and}$$

$$C' = [C'_{ij}]$$

is the *journey from work to shop* distribution matrix

Going back to the formulation previously presented, we should now, at each iteration, add non-basic employment generated (1) from home-based and (2) from place of work-based shopping trips. This yields·

$$\begin{aligned} &E^b && P^b = E^bA \\ &E^{(1)} = P^bB + E^bC = E^bAB + E^bC = E^b(AB+C) && P^{(1)} = E^{(1)}A = E^b(AB+C)A \\ &E^{(2)} = P^{(1)}B + E^{(1)}C = E^{(1)}AB + E^{(1)}C && \\ &\qquad = E^b(AB+C)^2 && \\ &E^{(n)} = P^{(n-1)}B + E^{(n-1)}C = && P^{(n)} = E^{(n)}A \\ &\qquad = E^b(AB+C)^n && \qquad = E^{(n)}(AB+C)^nA \end{aligned}$$

And totalling

$$\begin{aligned} E &= E^b(I + AB + C + (AB+C)^2 + \\ &\qquad\qquad + (AB+C)^n + \quad) \\ &= E^b[I - (AB+C)]^{-1} \\ P &= E^bA + E^b(AB+C)A + \quad + E^b(AB+C)^nA \\ &= E^b[I - (AB+C)]^{-1}A. \end{aligned}$$

Evaluation and Extensions

Starting with the same basic assumptions as the Lowry Model, this formulation leads to a very simple and concise expression of the overall relationship between the distributions of population, basic, and non-basic employment. (Equations (4) and (4′).) It makes directly apparent the net effect of various factors, such as labor participation rates, accessibilities, or other factors, that influence the locational behavior of the firm and the individual.

Consequently, it is possible to test different forms for the *journey to home* and *journey to shop* function, by applying equation (4) and (4′) to past or present data. (For example, "least squares," methods can be used to estimate some parameters from these 2 x n equations.) It is worth emphasizing that the only information needed to "calibrate" the model are population and employment—basic and non basic—per zone. Although very useful, origin-destination studies could be spared if time or budget constraints make it necessary.

As for future predictions and planning, whether the *journey to shop* and *journey to work* functions are assumed to change or not, population and employment in all zones of the region can be derived from an expected future distribution of basic industries Or taking the opposite approach to the problem, what policies in terms of public investment or basic industrial location could achieve a proposed pattern of development? More precisely, given a "desirable" population-vector P, one can solve the linear system of equations (4) and (4′) and find whether the corresponding basic employment distribution is feasible; or else determine what public investment in transportation, housing, etc.; that is, what changes in the matrices A and B, would lead to such a pattern.

Finally, extensions of the model can be envisioned First, one could think of introducing more than one class of population (by income, age groups, etc), with different "tastes" for their residential locations and shopping needs, and dividing population-serving activities into several sectors While such an expanded model would be more cumbersome to operate—with n population classes and p activity sectors, each of the matrices A and B would be composed of $n \times p$ submatrices $(n \times n)$—its formulation is not changed, and it could provide more insight into the determinants of development patterns

The other dimension to be added is time. So far the model is static; that is, it relates population and employ ment distribution at time *t* with basic industries locations, accessibilities and other factors at time *t*. Obviously, different activities would have to be included in the basic sector for long-term and short-term analyses. This leads to the belief that much could be gained by introducing the time element into this formulation to build a dynamic, recursive model.

Appendix
Convergence of Matrix Series

Before discussing mathematical conditions for the matrix series (3) to converge, a preliminary remark imposes itself: it is a fact of common experience that urban populations are finite; therefore, if the description of the process by which population-serving employees are generated is to be right, it should converge. We shall give only sufficient conditions.

$$\text{Let } \overline{\alpha\beta} = \underset{k=1,n}{\text{Max}}\ \alpha_K \beta_K$$

$$\text{Then } AB = A'[\alpha][\beta]B' \leq A' \times (\overline{\alpha\beta} I) \times B' = (\overline{\alpha\beta}) \cdot A'B'.$$

But since A' and B' are distribution matrices, the sum of the elements in any row are equal to 1:

$$\sum_j a'_{ij} = 1$$

$$\sum_j b'_{ij} = 1$$

and the same property holds for the product matrix A′ B′. Therefore, we have an upper bound for the row sums of matrix AB.

$$\sum_j (ab)_{ij} = \sum_j \Big(\sum_k a'_{ik}\alpha_k\beta_k b'_{kj}\Big) = \sum_k a'_{ik}\alpha_k\beta_k \leq \overline{\alpha\beta} \qquad (i = 1, \ldots, n)$$

When $\overline{\alpha\beta} < 1$, the spectral radius of the matrix AB, is less than one (Ref 3—Lemma 2.5). Consequently, $I - AB$ is non singular and $I + AB + (AB)^2 + \cdots = (I - AB)^{-1}$ the series on the left converging (Ref 3—Theorem 3.7).

Although an improved upper bound for the spectral radius of AB might be necessary in some cases, the above argument will hold normally in practical applications where $\overline{\alpha\beta} < 1$.

NOTE

[1] Data on travel times and basic employment are those of the theoretical town of LARTSVILLE developed at the Los Angeles Regional Transportation Study.

REFERENCES

1 Lowry, Ira S., *A Model of Metropolis* (Santa Monica, Calif The Rand Corporation, August 1964)

2 Goldner, W and W S Graybeal, *Bay Area Simulation Study Pilot Model of Santa Clara County and Some Applications* (Center for Real Estate and Urban Economics, Univ of Calif, Berkeley, 1965)

3 Varga, Richard S, *Matrix Iterative Analysis* (New York Prentice Hall, 1962)

Author's Note
I wish to acknowledge the help and advice of Professors Andrei Rogers and Michael B. Teitz of the Department of City and Regional Planning, University of California. Any shortcomings of this paper are of course my own responsibility.

A Note on the Garin-Lowry Model

Andrei Rogers

Garin's matrix formulation of the Lowry model [2] succinctly exhibits that model's fundamental character: it is a cross-sectional, static model that is *iterated* to an *equilibrium* solution. As in input-output models, the matrix formulation conveniently allows one to find this solution in one step by inverting the appropriate matrix. Recent research efforts in matrix models of population growth suggest a way of introducing, via recursion, a dynamic dimension to the Garin-Lowry model.[1] In the process, one may identify another kind of equilibrium: that of *stationarity*. This is the distribution that arises after repeated application of an unchanging *recursive* growth process.

1. *The Dynamic Garin-Lowry Model*

Let us denote basic employment at time *t* by the n-vector:

$$E_t^b = (E^b_{t1}, E^b_{t2}, \ldots, E^b_{tn}),$$

and population at time *t* by the *n*-vector:

$$P_t = (P_{t1}, P_{t2}, \ldots, P_{tn}).$$

Then, using Garin's Equation 4′, we have:

$$P_t = E_t^b (I - AB)^{-1} A \tag{1}$$

Now consider a growth matrix *G* which when applied to the basic employment vector E_t^b projects it forward one time period, i.e.,

$$E^b_{t+1} = E_t^b G \tag{2}$$

Then we may combine Equations 1 and 2 to form the following dynamic relationship:

$$P_{t+1} = E^b_{t+1}(I - AB)^{-1}A = E_t^b G (I - AB)^{-1}A, \tag{3}$$

and, if we assume for the moment that *A*, *B*, and *G* remain constant over time,

$$P_{t+1} = E_t^b G^n (I - AB)^{-1}A. \tag{4}$$

It is, of course, clear to everyone that the matrices *A*, *B*, and *G* will not remain constant through time and that therefore in a forecasting effort they would need to be predicted at each recursion of the model. Nevertheless, it is of some interest to derive what may be referred to as a "speedometer" reading of where we are headed if things do not change. In this way we could compare today's "speedometer" reading with yesterday's and see if the "horizon" projection to which we are headed has shifted and, if so, in what direction. This may be done by recursing the model until it reaches a stationary state, that is, an unchanging growth rate and spatial distribution.

2. *The Stationary Growth Rate and Spatial Distribution*

Recall Equation 4 and consider what happens as *n*, the number of time periods, gets very large, that is, as *n* approaches infinity. It can be shown [3] that in such an

Andrei Rogers is Assistant Professor, Department of City and Regional Planning, University of California, Berkeley.

the contrary, the builder and sponsor will agree that the model "works"

In the face of such ambiguities, it is not hard to imagine a reasonable man's refusal to participate in such a probable boon-doggle. But for the reasons indicated at the beginning of this article, I do not anticipate any shortage of sponsorship for model-building projects. It is better to try something—anything—than to merely wring one's hands over the futility of it all. Sponsors and model-builders too can take comfort in the thought that they are building for the distant if not the near future

Above all, the process of model-building is educational. The participants invariably find their perceptions sharpened, their horizons expanded, their professional skills augmented. The mere necessity of framing questions carefully does much to dispel the fog of sloppy thinking that surrounds our efforts at civic betterment. My parting advice to the planning profession is: If you do sponsor a model, be sure your staff is deeply involved in its design and calibration. The most valuable function of the model will be lost if it is treated by the planners as a magic box which yields answers at the touch of a button.

NOTES

[1] Wingo (Ref 15), p 144 Model-building has also been greatly encouraged by the electronic revolution in data-processing and computation, mathematical models have an insatiable appetite for numbers.

[2] An immensely important topic in the field of model-design which is *not* covered by this article is the joint effort of model-builder and client to define the "problem" to which a model offers a possible "solution" On this point, I know of no better reference than a RAND book on systems analysis (Ref 21), particularly the essays of R. D Specht ("The Why and How of Model-Building"), Roland McKean ("Criteria"), and E S Quade ("Pitfalls in Systems Analysis")

[3] Some model-builders would freely substitute "simulate" for my "replicate" All models are intended in some sense to simulate reality, but this usage is the source of some confusion in the literature since "simulation" has acquired another more technical meaning, descriptive of a class of algorithms In this essay, I use the term *only* in the latter sense See below, "Solution Methods"

[4] For example, traffic analysts use zonal interchange models to generate estimates of zone-to-zone traffic flows from inventories of the land uses in each zone One prominent member of the profession is so convinced of the descriptive reliability of these models that he sees no further need for direct surveys of traffic movements (O&D studies)

[5] Philosophers of science view the concepts of "cause" and "effect" with jaundiced eyes For lesser mortals these concepts are most helpful and not at all dangerous so long as they are applied within the framework of a system of interdependence Cf Simon (Ref 23), Ch 1–3

[6] No variable is intrinsically endogenous or exogenous These terms, like the statistician's "dependent" and "independent," merely define the position of a variable within a particular model A further useful distinction can be made between exogenous variables subject to policy control and those which are not, and between endogenous variables of direct interest to the planner and those which are included only because they are necessary to complete the logical structure of the model Cf Sonenblum and Stern (Ref 24), pp 112–114

[7] The fundamentals and applications of linear programming are summarized in very readable form by Baumol (Ref 2), pp 837–853 I cannot find any simple exposition of dynamic programming, but see Bellman (Ref 4), pp vii-xi, for a brief account of the class of problems to which the technique is applicable

[8] "The value of Y is a function of (depends on) the values of U, V, X, and Z, and so forth" For a gentle introduction to the notation and methods of mathematical modeling, Beach (Ref 3) is an excellent source

[9] Some examples, in prose rather than symbols·

Technological The maximum vehicular capacity of a roadway is a function of the number of lanes, the average distance between signals, and the weather

Institutional Disposable family income is a function of gross family earnings and the tax rate

Behavioral The level of housing density chosen by a family depends on disposable family income, the average age of family members, and the location of the work-place of the principal wage-earner

Accounting Total land in use is the sum of land in residential use, in retail use, in manufacturing use, and so forth

[10] The essays in Part II of Zipf (Ref 26) should give the reader a "feel" for the macro-analytic perspective in urban models See also Carrothers (Ref 8), and Berry (Ref 5)

[11] Dyckman (Ref. 9) provides an excellent review of the theory of rational choice in a planning context Any introductory text in economics will describe the micro-analytic underpinnings of demand and supply schedules and will also review a family of market models The most ambitious micro-analytic model ever undertaken in the social sciences is described in Orcutt (Ref 20) For models that embrace more than "economic" man, see Simon (Ref 23) or Lazarsfeld (Ref 16)

[12] Cf Lichfield (Ref 17)

[13] It is my personal conviction—not shared by all members of the fraternity of model-builders—that the macro-analytic approach to modeling urban form and processes shows the greater promise of providing reliable answers to concrete problems of prediction and planning For a contrary view, see the forceful statement by Harris (Ref 13), p 16

[14] Descriptive models of urban form are nearly always static or "equilibrium" models, and are sometimes used for quasi-predictions (comparative statics) For convenient examples, see Harris (Ref 12) or Lowry (Ref 18)

[15] Contrast the emphasis on stocks in the San Francisco CRP model designed by A D Little, Inc (Ref 1) with the emphasis on flows in Bolan, *et al* (Ref 7) or with the several "growth allocation" models described in this issue of the *Journal*

[16] Cf Black (Ref 6)

[17] The reader is warned that Equation 3 is not a general solution for any Y_{t+n}, but merely the simplest expression for Y_{t+1}

[18] An example of the iterative technique is given in some detail in Lowry (Ref 18), pp 12–19

[19] A good bibliography in simulation methods is Shubik (Ref 22) Geisler, *et al*, (Ref 10) offer a quick and readable review of the field, with emphasis on man-machine simulation or "gaming" Grundstein (Ref 11) describes a "community game" for the training of planners and municipal administrators

[20] Special data-problems encountered in modeling urban form and process are discussed by Britton Harris, "An Accounts Framework for Metropolitan Models," in Ref 15, pp 107–127 Also see Steger (Ref 25), pp 1–6

[21] Beach (Ref 3), Part II, provides an especially good introduction to statistical and econometric methods

[22] The convenience of this method is so great that it is often applied to systems containing known non linearities, on the grounds that a linear approximation is better than nothing Simultaneous estimation of the parameters of non-linear systems is possible, but more difficult, the outstanding example among land use models is Karl Dieter's Program Polimetric for fitting in exponential model with a great many parameters (The model, but not the fitting method, is described in Ref 7)

[23] Cf Niedercorn (Ref 19) His model is partitioned into three subsystems, each of which was fit independently The discussion on pp 14–15 illustrates the variety of estimating methods ordinarily required to fit a model See also Harris (Ref 12) for a discussion of the "gradient search" method of estimating parameters.

[24] Hill (Ref 14) reports with unusual thoroughness on a test of this type for the *EMPIRIC* Model developed for Boston by Traffic Research Corporation

REFERENCES

1 Arthur D Little, Inc "A Simulation Model of the Residential Space Market in San Francisco" Paper prepared for the Seminar on Models of Land Use Development, Institute of Urban Studies, University of Pennsylvania, Philadelphia, October, 1964

2 Baumol, William J "Activity Analysis in One Lesson," *American Economic Review*, XXXXVIII (December, 1958), 837–873.

3 Beach, E F *Economic Models An Exposition* John Wiley and Sons, New York, 1957

4 Bellman, Richard *Dynamic Programming* Princeton University Press, Princeton, 1957

5 Berry, Brian J L "Cities as Systems Within Systems of Cities" Paper presented at the Annual Meeting of the Regional Science Association, Chicago, November, 1963.

6 Black, Russell VanNest "Scientific Versus Empirical Projections," *Journal of the American Institute of Planners*, XXVI (May, 1960), 144–145

7 Bolan, Richard S, Willard B Hansen Neil A Irwin and Karl H Dieter "Planning Application of a Simulation Model"

References

Alonso W., *Location and Land Use*, Harvard University Press, Cambridge (1964)

Barras R., T A Broadbent, M Cordey-Hayes, D B Massey, K Robinson, J Willis, J Wm., "An Operational Urban Development Model of Cheshire," *Environment and Planning* Vol 3 115–233 (1971).

Boyce, D E., Day, N D, and McDonald, C., *Metropolitan Plan Making*, Regional Science Research Institute, Monograph Series No. 4, Philadelphia, (1970)

Boyce, D E., C McDonald and A Farhi, *An Interim Report on Procedures for Continuing Metropolitan Planning*, Regional Science Department, University of Pennsylvania, Philadelphia, (1971).

Donnelly, T G., Chapin, F S, Jr., and Weiss, S F., *A Probabilistic Model for Residential Growth*, Center for Urban and Regional Studies, University of North Carolina, Chapel Hill (1964).

Lowry, I S., *A Model of Metropolis*, RAND Corporation, Santa Monica (1964)

Schneider, M., "Access and Land Development," *Urban Development Models*, Special Report 97, Highway Research Board, Washington, D C (1968).

Schneider, M., *Transportation and Land Development*, Creighton, Hamburg Inc., prepared for the U S Department of Transportation, Washington, D.C. (1969)

Seidman, D., *The Construction of an Urban Growth Model*, Delaware Valley Regional Planning Commission, Philadelphia (1969)

3

A Model for the Distribution of Residential Activity in Urban Areas[a]

John D Herbert and Benjamin H. Stevens

Introduction

The model presented here is designed to distribute households to residential land in an optimal configuration. The model was constructed for the Penn-Jersey Transportation Study as part of a larger model designed to locate all types of land-using activity.

Since the model had to be suitable for practical application, a certain amount of conceptual elegance has been sacrificed in favor of operational simplicity. The larger model operates in the following way: the total relevant time period is subdivided into a number of short iterative periods. For each iterative period different types of land-using activity are handled separately. A particular type of activity is distributed in a configuration that is optimal only with respect to all previously located activities.[1] Interactions that are expected to occur among simultaneously locating activities are ignored. We are assuming that they can be ignored if iterative periods are kept short enough to ensure that the number of users located in a single run of the model is small. Operating in this way we are able to achieve computational simplicity and, at the same time, recognize most of the basic interactions among land users.[2]

For the residential model, in a particular iterative period, the number of households to be located and the amount of land that is expected to be available for residential use is forecast exogenously.[3] A linear program is used to produce, for the end of that period, an optimal configuration of the new households on the available land. This configuration is optimal with respect to the configuration of all previously located activities, and constitutes a prediction of the way in which the forecast households will locate.[4]

---

[a] This paper was adapted from a report prepared for the Penn-Jersey Transportation Study and is reprinted with permission from *The Journal of Regional Science*, Vol. II, No. 2 (1960). The authors wish to acknowledge the contributions of Britton Harris, who directed the section of the study under which the model was prepared and who made invaluable criticisms and suggestions throughout its development, Nan Letter and the other members of Penn-Jersey staff who participated in discussions of the problem. The paper also draws on the excellent theoretical development of similar materials in Alonso (1960). The authors accept full responsibility for the model's shortcomings.

Definitions

Household. A household is a person or group of persons with a common budget purchasing a single residential bundle.

Household Group. A household group is a collection of households which have similar residential budgets and similar tastes with respect to residential bundles.[5]

House. A house is the physical structure to be occupied by a single household.

House Cost. House cost is the dollar cost of constructing, operating, and maintaining a house over a specified time period, computed on an annual basis.[6]

Amenity Level. The amenity level associated with a site is the level of psychic satisfaction which a household has an opportunity to enjoy because of certain characteristics of that site.[7]

Trip Set. A trip set consists of the numbers of each type of trip generated by a household.[9]

Trip Pattern. The trips in a trip set when their origin and destinations have been identified.[10]

Travel Cost. Travel cost is the annual dollar cost to the household of carrying out a trip pattern.[11]

Site. A site is the parcel of land assignable to a particular household.[12]

Total Site Rent: Total site rent is the dollar value of the amount received annually by a site owner for the use of the *land* in the site. It is exclusive of the value of the house, amenity level, and travel pattern associated with the site.

Residential Bundle. A residential bundle is a unique combination of a house, an amenity level, a trip set, and a site of a particular size.

Market Basket: A market basket is a unique combination of a residential bundle and a bundle of all other commodities (which we shall call the "other commodities" bundle) consumed annually by a household.

Total Household Budget: A household's total budget is the dollar amount that a household allocates annually to the purchase of a market basket.

Residential Budget. A residential budget is that part of a household's total budget that is allocated annually to the purchase of a residential bundle.

Region. A region is the geographical space within which our model is required to allocate a given number of households to a given supply of land.

Area. An area is a regional subdivision whose characteristics are homogeneous with respect to the costs of construction, amenity costs, and the costs of transportation to other areas.

Conceptual Framework

We assume that the factors which a household considers in choosing an area in which to locate are its total budget, the items which constitute a market basket, and the costs of obtaining those items. For each household group we posit a set of market baskets among which each household in that group is "indifferent."[13] We posit the set which includes, but is not necessarily limited to, the market baskets currently consumed by households of that type.[14] We permit the household to optimize not by selecting a market basket from *all* the conceivable sets from which it could obtain satisfaction but by selecting from the posited set the market basket which maximizes that household's "savings."

These "savings" arise in the following way. A household has a fixed total budget. For a particular market basket the prices of the items in the "other commodities" bundle are given. The residential budget is therefore a residual determined by the size of the total budget and the cost of the "other commodities" bundle. Clearly, it may vary from market basket to market basket. Notice that the character of each of the four items that constitute a residential bundle may vary from market basket to market basket also. Each market basket in the indifference set will have in it a unique residential bundle which has a unique residential budget associated with it. Disregarding site for a moment, the costs of each of the other three items in a residential bundle may vary from area to area. For a particular area, the difference between the residential budget assigned to a particular residential bundle and the cost of the bundle exclusive of the site in it is the maximum amount the household can pay in that area for that site. And it will be the maximum amount that the landowner could extract from the household as site rent. If land were free, it would be a measure of the savings enjoyed by the household because of the locational advantages of the area. We define this difference as the household's rent-paying ability for that site in that area.

Although we have said that a household is "indifferent" among the market baskets in its indifference set, it seems reasonable to suppose that such savings would have a positive marginal utility for the household. Therefore, in the model a rational household will attempt to obtain from its indifference set the market basket in which those "savings" are a maximum. In reality, the functioning of the land market may make it possible for the landowner to draw off the "savings" as rent. Nevertheless, the *attempt* of each household to maximize its savings will result in households being allocated to land in configuration that is optimal from the point of view of all the households that are to be located. This allocation will be optimal in a Pareto sense: no household can move to increase its savings without reducing the savings of some other household and simultaneously reducing aggregate savings. Since we have made savings synonymous with rent-paying ability, an optimal allocation is achieved by the maximization of aggregate rent-paying ability.[15]

The Primal Problem

Notation

U areas which form an exhaustive subdivision of the region. Areas are indexed by the superscripts $K = 1, 2, \ldots, U$

n household groups indicated by subscripts $i = 1, 2, \ldots, n$
m residential bundles indicated by subscripts $h = 1, 2, \ldots, m$
b_{ih} is the residential budget allocated by a household of group i to the purchase of residential bundle h
c^K_{ih} is the annual cost to a household of group i of the residential bundle h in area K, exclusive of site cost
s_{ih} is the number of acres in the site used by a household of group i if it uses residential bundle h
L^K is the number of acres of land available for residential use in area K in a particular iteration of the model
N_i is the number of households of group i that are to be located in the region in a particular iteration
X^K_{ih} is the number of households of group i using residential bundle h located, by the model, in area K.

The allocation model

The primal linear programming model for allocating households to land has the rather simple form.[16]

Maximize

$$Z = \sum_{K=1}^{U} \sum_{i=1}^{n} \sum_{h=1}^{m} X^K_{ih}(b_{ih} - c^K_{ih}) \qquad (1.0)$$

subject to:

$$\sum_{i=1}^{n} \sum_{h=1}^{m} s_{ih} X^K_{ih} \leq L^K \qquad (K = 1, 2, \ldots, U) \qquad (1.1)$$

$$\sum_{K=1}^{U} \sum_{h=1}^{m} X^K_{ih} = N_i \qquad (i = 1, 2, \ldots, n) \qquad (1.2)$$

$$\text{and all } X^K_{ih} \geq 0 \qquad \begin{matrix} (K = 1, 2, \ldots, U) \\ (i = 1, 2, \ldots, n) \\ (h = 1, 2, \ldots, m) \end{matrix}$$

Constraints (1.1) prevent the consumption of land in each area from exceeding the land available. Constraints (1.2) require the model to locate the projected numbers of households of each group. These constraints are equalities because inequalities (of either sense) would not fit the overall requirements of the model. Suppose these constraints were written in such a way that the model were precluded from locating *more than* the projected numbers of households. This would

be logical since we are interested in the situation where a particular number of households are located, not where the model can continue locating households in unlimited numbers until all the available land is used up. On the other hand, it is just as logical to write the constraints in such a way that the model is required to locate *at least* the projected numbers of households. This is particularly important where there are household groups which have negative or zero rent-paying ability in all areas. Without constraint, the model would choose not to locate these households at all, since at best they would not add to, and at worst they would not subtract from, aggregate rent-paying ability. For these reasons, it is difficult to establish a general rule for the sense of the inequalities. Therefore it is preferable, and perfectly reasonable, to make the constraints equalities.[17] The objective function (1.0) to the maximized is, of course, aggregate rent-paying ability.

Households may be allocated to land in the following ways. (1) One type of household may use all the land available in an area. This will occur where that type of household can yield the highest unit rent for the land in the area and there is a sufficient number of such households to fill the area. (2) The land available in an area may not be used up entirely. Partial utilization will occur where the area has strong locational advantages for only one of the household groups and there are not enough households of this type to fill the area, or where the area has strong locational advantages for two or more household groups and these groups, *in toto,* cannot fill the area. (3) The available land in an area may be left vacant if all households have higher unit rent-paying abilities in other areas and can find sites in other areas. (4) The land available in an area may be used by more than one type of household. This will occur where there are not enough households in the group with the highest unit rent-paying ability in the area to fill that area and they are joined by other households with unit rent-paying abilities the same as, or lower than, the highest group but higher in this area than in other areas. Joint utilization can occur also in the unusual circumstance where two or more household groups have identical unit rent-paying abilities in the area and in all other areas where they could outbid other groups for sites.[18]

The Dual of the Allocation Model

The notation of the dual problem is identical to that of the primal except that the solution variables, X^K_{ih} are replaced by.

r^K the annual rent per unit of land in area K ($K = 1, 2, \ldots, U$)
v_i the annual subsidy per household for all households of group i ($i = 1, 2, \ldots, n$). The use and meaning of the subsidy variables will be made clear below.

The dual problem is to minimize

$$Z^1 = \sum_{K=1}^{U} r^K L^K + \sum_{i=1}^{n} \nu_i(-N_i) \qquad (2.0)$$

subject to

$$s_{ih} r^K - \nu_i \geq b_{ih} - c_{ih}^K \qquad (K = 1, 2, \ldots, U), (i = 1, 2, \ldots, n) \quad (h = 1, 2, \ldots, m) \qquad (2.1)$$

$$\text{all } r^K \geq 0 \qquad (K = 1, 2, \ldots, U).$$
$$\nu_i \gtrless 0 \qquad (i = 1, 2, \ldots, n).^{19}$$

In most linear programming models, the dual presents a problem in interpretation. The existence of the dual is a mathematical fact. But often it also contains information and provides insights which are as important as those provided by the primal itself. This is particularly true in the case of the present model. If we look at the objective function (2.0) and neglect for a moment the second summation, we can interpret the first summation as the total land rent.[20]

It may seem peculiar to minimize total land rent in the dual when we are maximizing aggregate rent-paying ability in the primal problem. It can be shown that the optimal solution of the primal problem must be exactly equal to the optimal solution of the dual. But there is also an important economic interpretation of the dual objective. Suppose all land in all areas were owned by a monopolist. Then the minimization of site rent will minimize the returns to this monopolist. Alternatively, land could be widely held by individual holders. We would then be minimizing returns to the rentier class as a whole. To put it another way, we are obtaining sites for households as cheaply as possible within the constraints of the model.

This is not necessarily a desirable goal if it causes inequities to land owners. But notice that the constraints (2.1) prevent the unit rent on each site from falling below the unit rent-paying ability of *any* household that might locate on that site.[21] This means that the individual landowner can receive at least as much per unit as the highest bidder for his land is willing to pay. This will create certain problems when the household group which can bid highest does not actually purchase the land because it has an even higher unit rent-paying ability elsewhere. It is this latter case, and certain other cases, in which the "subsidy" variables become important.

Bear in mind that a household which can bid the highest unit rent in a particular area is not necessarily of the "wealthiest" household group. Unit rent-paying ability depends upon both total rent-paying ability and size of site

purchased. "Poorer" households using small sites may be the highest bidders, per unit of land, in a particular area. Thus "subsidies," in the model, may be assigned in some cases to "wealthy" households.

Interpretation of "Subsidies" and Rents

The foregoing serves to introduce the idea that the "subsidy" variables, ν_i, are basically a mathematical device.[22] Now let us consider their economic meaning. Suppose, for example, that all household groups save two have been located. Suppose further that all areas have been filled save one. Assume that one household group can bid more per unit for the land than the other. This high-bidding group will then establish the rent level in the area. Because areas are not divisible within the model and because of the nature of linear programming, the same unit rent must then be charged for *all* units of land in the area.[23] The households with the lower rent-paying ability must be located somewhere and this is the only area left. Since their rent-paying ability is insufficient to meet the high rent level established in the area, households of this second group must be "subsidized." And, as indicated above, the same situation would prevail even if the group with the higher rent-paying ability in this area actually locates in another area or areas.

This simplest interpretation of the subsidy variables is not hard to grasp. However, the existence of these variables creates problems in interpretation which are more difficult to handle. Notice that the ν_i are specific to household *groups*. Therefore, once any household in a group receives a subsidy, all of that group must receive the same subsidy. It might be possible to argue that this would be a realistic consequence of an egalitarian public policy, but the argument would be a weak one. For suppose that all "high income" households but one could be accommodated in a particular area. This leftover household would have to locate in another area. If the locational advantages of the first area were extremely large but other areas were highly unsatisfactory to "high-income" households, we might find that this leftover household had a *lower* rent-paying ability in all other areas than competing households of other groups. Since this household must be located somewhere, it will then have to be subsidized. But in the process, all high-income households would be subsidized, raising their rent-paying abilities in the first area to unnecessarily high levels.

We could avoid this problem by disaggregating households and having a constraint (1.2) for each individual household. But this would make our problem extremely unwieldy. A better approach would be to disaggregate areas after a run of the model. This would involve splitting any area in which two or more different household types locate into two or more subareas each occupied by a single household type.[24] Rent levels could be adjusted, ostensibly by taxing away from the landowners the "extra" rents made possible by the subsidies. Rents would then be equal to rent-paying abilities for all household types and subsidies, of the type we have been discussing, would be eliminated.[25]

Areal disaggregation would enable us to approximate a continuous model in which each small parcel of land is held separately and can carry a different rent from neighboring small parcels. The linear programming approximation to the continuous model is especially important in any practical application since the latter is very difficult to solve.[26] Furthermore, the linear programming model may actually be the more realistic of the two. A truly continuous model assumes a level of information and sophistication on the part of both landowners and households which is not likely to exist in practice. From the landowners' point of view it is hard to imagine a real situation where there are marked variations in residential rent among contiguous sites; from the households' point of view, it seems unlikely that contiguous sites will be regarded as having distinctly different locational cost and advantages.

We still have not dealt with one important problem. Suppose that a certain ousehold group has the highest rent-paying ability in an area but cannot fill it. Assume that all other households have higher rent-paying abilities in other reas and can be completely accommodated in these areas. This would mean that he land constraint (1.1) for the area would remain an inequality at an optimal olution. In accordance with the general structure of linear programming models, his means that the rent variable corresponding to this area should be equal to ero. In economic theory, land, as a resource for this household type, is then not eally scarce and should carry a zero rent. Yet the dual constraints (2.1) for he area require that the land rent be no less than the unit rent-paying ability of ny household which might locate there. We appear to have a contradiction nless the households which make the highest bid have zero or negative unit nt-paying ability in the area. If their rent-paying ability is strictly positive we ppear to find all units of land in the area earning a positive r^k even though ome of the units are not consumed at all. This would seem to make the value of e dual objective function larger than the value of the primal objective function.

We may resolve this conceptual dilemma if we recall that the "subsidy" riables, r_i, need not be positive. Therefore, we can fulfill both the condition of ro rent on unfilled areas and the condition of unit rent greater than or equal unit rent-paying ability in all areas. The negative "subsidy" can be regarded as "tax" on the households (equal to the rent-paying ability of the households the area which they occupy but do not fill) to be used as a "subsidy" to the ndowner in the area to "pay" him for the land actually consumed.

By now it may have occurred to the reader that we have obtained the ormation necessary to construct a classical rent surface for the region. Given a ap of the region we could place a block on each area with a height proportional the unit rent in the area. The surface thus constructed would have discontinuities. But perhaps a discontinuous surface is actually more realistic than

the smooth surface of classical theory. A multimodal metropolitan region with an irregular topography and transportation system and a mixed pattern of land uses, should not be expected to exhibit smooth rent gradients.

The sharpness of the rent discontinuities can be reduced if we can assume that households need not bid the full amount of their rent-paying abilities. If they are able to win land by just outbidding competing households, they can save the difference between rent-paying ability and actual amount bid.[27] The landowner may be able to extract the full rent if, as we suspect, he is the more powerful bargainer. But if there are actually a large number of small competing landowners, their bargaining power may not be much greater than that of the households. Then true net rents may be as high as the rent-paying abilities of the households which occupy the land, as low as slightly above the highest bid of competing households, or somewhere in between, and will depend upon the pattern of land ownership and the vagaries of the land market.[28]

Some Limitations of the Model

Many of the limitations of the model are implicit in its construction. They include the following. (1) Data problems: it will often be difficult to obtain consistent data on household budgets and tastes, amenity levels and costs, etc. For the present, in the light of preliminary surveys, we are assuming that it will be possible to obtain data which, though crude, will be adequate to make the model operational. (2) Forecasting problems: successful use of the model depends upon obtaining accurate forecasts of the numbers and characteristics of location-seeking households and of the amounts of available land. (3) Restriction of the choices of a household to a single indifference set: this restriction may prevent the model from achieving a true optimum, since households cannot shift to indifference sets yielding higher levels of satisfaction. Locational savings may or may not constitute an adequate proxy for the satisfactions that might be gained through continuous substitution. (4) Optimization over a number of iterative periods: the model optimizes for individual iterative periods. It may not be rational to assume that an aggregation of such optima will constitute an optimum for the aggregate time period. (5) The simulation problem: the failure of the model to take into account interactions (other than competition for land) among households locating simultaneously is a serious enough limitation to warrant further discussion.

The problem can be illustrated by considering the cumulation of households around concentrations of job opportunities. Recall that we use a probability interchange model to determine travel cost. This requires the development of a set of expected trip costs for each area on the basis of the number and location of opportunities for trip ends. These expected costs are heavily influenced by the existence of large employment centers which provide top-end opportunities for large numbers of workers

The difficulty is that too many households may locate in the areas near large employment centers and not enough in the areas surrounding smaller employment centers At an optimal solution we may find that there are not enough jobs in a large center to accommodate all household wage-earners so that some of them will have to travel relatively long distances to find jobs This is more a problem with the model than with the real world although there is some evidence to indicate that, given two centers of employment, households tend to orient themselves toward the larger one even if they are not sure to find employment there It may be that near the larger center they expect to have a wider range of choice of employment and be in a better position to find another job should they be laid off.

Nevertheless, the model may give us erroneous results because of the cumulation phenomenon There are several ways of handling the problem but none of them is completely satisfactory. One approach would be to assign households to jobs in advance of the model run For example, suppose we were dealing with a metropolitan area with one major and one minor center of employment for workers of a particular household group Assuming one worker per household, we could then break this group into two parts in proportion to the number of jobs available at the two centers One part would have its transport costs computed as if only the major center existed The smaller part would have transport costs relative to the minor center Such an approach would give strong locational advantages to approximately the right proportions of households near each of the two centers although a successful solution would still not be guaranteed.

An alternative approach would be to make travel costs a function of the number of households locating in particular employment areas Unfortunately, this would require a nonlinear program which is much more difficult to solve. It might provide a theoretical solution to the problem, but not a practical one.

Finally, an approach might be developed out of repeated use of the model. If we find that cumulation occurs, we might be able to determine experimentally how probabilities should be altered or weighted so as to produce a more "realistic" solution.

Alternative Approaches

It is appropriate at this point to discuss the question of alternative models. We have used a linear program because it appears to afford the best compromise between theoretical elegance and practical applicability A nonlinear programming model would be able to handle the cumulation problem as well as problems such as the effect of traffic congestion on travel costs, the effect of growing residential densities on household decisions, and the effects of changing land use on amenity levels and costs But it will be some time before the new computational techniques for nonlinear programming become economically feasible for large-scale problems

We specifically reject multiple regression models because they tend to project the past into the future in a way which fails to recognize the basic structural interrelationships of activities and land uses [29] For somewhat the same reason we feel that simulation models do not provide the most promising direction for further research However, current developments in simulation models for residential decisions indicate that they will give much better results than multiple regression techniques And it is difficult to ignore the possibility that residential patterns may be more random than rational.[30]

Finally, it is important to recognize Alonso's unique contribution in the development of a rigorous theoretical model of residential location His method is difficult to apply directly, but we feel that our linear programming model provides an analogous approach that is both acceptable and workable.[31]

Policy Implications

The model is applicable to public policy decisions concerning zoning, transportation, redevelopment, public housing, segregation, and other areas of public interest For example, the model could be run without zoning restrictions Then restrictions could be applied to determine their effects, if any, on rents and residential patterns. Similarly, the effects of altering transportation costs could be tested by multiple runs of the model.[32]

Where the model gives high rents in areas with little available land and close to existing blighted areas, it may indicate that these areas are ripe for redevelopment. Moreover, the existence of high rent levels would indicate a low cost of writing down the land.[33] Some areas may exhibit extremely low rent levels in comparison with surrounding areas. These may be ripe for redevelopment to raise amenity levels (and reduce amenity costs). Because the model can identify the "savings" that households might enjoy, not only on vacant land but also in partially improved and built-up areas, it should be of interest in problems of conservation and rehabilitation as well as redevelopment. In all of these applications it could be used as one type of cost and benefit analysis.

We may wish to use the model to consider real subsidies to households which have negative rent-paying ability in all or most areas Tests could be made to determine whether it is more effective to raise these households' location budgets through direct subsidies (thereby shifting them into another household group) or to reduce their location cost through indirect subsidies (e g through public housing) The dual variables, together with knowledge of the total and locational budgets of such households, can provide a measure of how large such subsidies should be

Notice that the model can recognize racial segregation and similar policies if they have measurable dollar consequences for rent-paying abilities. It can do this in much the same way as it recognizes the amenity value of particular areas. For example, we can ascribe to a predominantly white area high "amenity" costs for particular types of Negro households seeking particular locational bundles.

Whenever we use the model as a guide to policy-making, we must be careful that we do not ascribe spurious meanings to its output. There may be social welfare considerations which are noneconomic and which are not reflected in the maximization of aggregate rent-paying ability. And these noneconomic considerations should be taken into account by the policy-maker. He may decide that a true social optimum can be achieved only by deviating from the pattern produced by the model. If he does, the model can tell him how much the deviation will reduce aggregate rent-paying ability. In other words, an "optimal" residential pattern can be found against which to test the effects of alternative programs of public expenditure and control. Although it was designed for use by a transportation study, the model has implications and applications which make it useful in a much broader range of decision problems.

References

Alonso, W., A Model of the Urban Land Market: Location and Densities of Dwellings and Businesses, unpublished doctoral dissertation, University of Pennsylvania, microfilmed (1960). A brief description of this model is to be found in Alonso, W., "A General Theory of the Urban Land Market," *Papers and Proceedings of the Regional Science Association,* Vol. VI (1960).

Carroll, J. D., and Bevis, H. W., "Predicting Local Travel in Urban Regions," *Papers and Proceedings of the Regional Science Association,* Vol. III, pp. 183-197 (1957).

Chicago Area Transportation Study, *Allocating Population to Small Areas,* Chicago (1957).

Dorfman, R., Samuelson, P. A., and Solow, R., *Linear Programming and Economic Analysis,* McGraw-Hill, New York (1958).

Gale, D., *The Theory of Linear Economic Models,* McGraw-Hill, New York (1960).

Garrison, W. L., "Toward a Simulation Model of Urban Growth and Development," paper prepared for the Symposium on Urban Studies, Lund (1960).

Marble, D. F., "Transport Inputs at Urban Residential Sites," *Papers and Proceedings of the Regional Science Association,* Vol. V, pp. 253-267 (1959).

Stevens, B. H., "A Review of the Literature on Linear Methods and Models for Spatial Analysis," *Journal of the American Institute of Planners,* Vol. XXVI, No. 3, pp. 253-259 (August, 1960).

Stevens, B. H., and Coughlin, R. E., "A Note on Inter-Areal Linear Programming for a Metropolitan Region," *Journal of Regional Science,* Vol. 1, No. 2, pp. 75-83 (Spring, 1959).

Notes

1. For a particular type of activity in a particular iterative period, previously located activities comprise all activities located in previous periods plus activities of other types previously located (by other elements in the larger model) in the same period.
2. For example, if we take an extreme case with an iterative period of one week, the number of land users that will be located in that period is likely to be small. It seems reasonable, both conceptually and realistically, to assume that they will make their locational decisions largely independent of one another. However, interactions between users located in a particular week and those located in previous weeks will be recognized, with the result that a vast majority of the important interactions are taken into account. For Penn-Jersey we envisage an iterative period of at least a year, which will certainly introduce inaccuracies, but we can achieve any level of accuracy that we desire, at the cost of increasing computational complexity, by decreasing the length of the periods.
3. The residential model can handle land that is vacant, partially improved, or completely built-up. For expositional simplicity, the discussion will be limited to vacant land unless otherwise noted. Forecasts are exogenous in the sense that they are made outside the overall model (by techniques that are beyond the scope of the present discussion) but can be modified for a particular iterative period to recognize the configurations produced by the model in previous periods.
4. Linear programming is not ordinarily regarded as a predictive tool. However, if we have a prediction of the number of households that is to be located and the model locates them in a realistic configuration then we can use the model to predict configurations. Since the configurations it produces are optimal in a specific economic sense the model may be both predictive and prescriptive.
5. We are assuming that households that have similar tastes give rise to similar phenomena on the sites they occupy. E.g., households with similar tastes with respect to travel patterns generate the same numbers of each kind of trip, regardless of the number of persons in the household.
6. We hope to obtain house costs from contractors' estimates. Clearly, the construction, operating, and maintenance costs for a particular type of house may vary from area to area, and will be dependent in part upon the topography, soil, and microclimate in each area.
7. The site characteristics contributing to the amenity of a site have yet to be selected. It is expected that they will include characteristics such as the

existing households in the area in which the site is located the levels of public services in the area, the age of the area other land uses in the area, planting on the site, and views from the site

8 Tentatively it is proposed that estimates of amenity costs be obtained in the following way Realtors and developers will be asked to identify for each area in the region the "premiums" they could add to or subtract from the selling price of a particular type of house offered to a particular type of household because of the amenity of that area Where these premiums reflect the accessibility of an area as well as its amenity level, we anticipate that the effects of accessibility can be eliminated For a particular type of household and house the area in which that premium is a maximum will be regarded as having zero amenity cost for that type of household purchasing the residential bundle in which that type of house occurs The corresponding amenity cost in any other area will be the difference between the premium associated with it and the maximum premium Since premiums may be negative, amenity costs in some areas may be so high that they prevent the model from locating particular types of households in those areas It is worth noting here that the model can recognize, in a similar way, segregation and other policies which affect the prices households are willing to offer for houses

9 By kind of trip we mean, for example, a trip to a particular type of employment, recreation, shopping, etc The origins and destinations of the trips in a set are not specified since we assume that the alternative trip sets considered by a household will be independent of that household's location. There is some evidence in support of this assumption (Marble, 1959). It will be seen below that the assumption may not be unreasonably restrictive since we provide households with a number of alternative trip sets from which to choose

10 The destinations of the trips in a trip set will vary with the location of the household generating the set These destinations will be forecast exogenously by a probability interaction model (Carroll and Bevis, 1957).

11 Travel costs, including out-of-pocket and imputed time costs, can be estimated from data on the transportation system.

12 Land not covered by a house but under the same tenancy or ownership as the house is included in the site associated with that house Where a parcel of land has more than one house superimposed upon it, as in the case of an apartment building, the site assignable to each household is found by dividing the total area of the parcel by the number of households occupying it

13 The household is "indifferent" among the baskets only in a limited sense which will be made clear in the subsequent discussion

14 This is based on the assumption that households have, in the past, come close to achieving optimal levels of satisfaction However, where empirical evidence suggests that market imperfections have precluded optimization we can add to the indifference set market baskets that could be chosen by the households in a market free of imperfections Obviously there are conceptual weaknesses involved in the use of empirical evidence for the identification of indifference sets We assume that it is possible to construct operationally acceptable sets that are based on such evidence without being tied rigidly t it The model will not permit the indifference set that is relevant for a particular household to change during an iteration But this does not preclude the possibility of allowing taste changes to occur from one iteration t another

15 In a Henry Georgian single-tax economy, the maximization of aggregate rent-paying ability would provide a maximization of public revenue In a socialist system, if land were free, maximization of aggregate rent-paying ability would provide a maximization of consumer's surplus

16 Readers with limited knowledge of linear programming are referred to Dorfman, *et al* (1958) and the bibliography in Stevens (1960) Although there are Umn variables to be determined, it is possible to eliminate many of them in advance of the computation of the program. We can do this for each household in each area by disregarding all residential bundles that yield less than the maximum unit rent-paying ability for that household in that area These would have to be eliminated in any case in the process o maximization, a prior removal of them can reduce computational time and effort considerably.

17 Actually, it is more likely to be necessary to force the model to locate households with zero or negative rent-paying ability than to restrict the number of households which may be located This is reflected in the minus signs which appear on both sides of constraints (1 2) If these constraints were written as inequalities they would read:

$$\sum_{h=1}^{m}\;\sum_{k=1}^{U}\; -X_{ih}^{K} \leqslant -N_i$$

But without the minus signs the inequalities would be of the opposite sense In a maximization problem the inequalities *must* be of the "⩽" form Therefore the minus signs are necessary. They could be removed when constraints (1 2) are changed to equalities But the interpretation of the dual variables is somewhat easier if the minus signs are left in the primal

18 This is the degenerate case in which there is no unique optimal allocation of the households in the groups which fulfill this condition. A further degenerate case can occur where a particular household group has the same unit rent-paying ability in a number of areas, none of which it can fill complete

19 An inequality (1 1) in the primal corresponds to a nonnegative variable r^A in the dual But an equation (1 2) in the primal corresponds to a variable r_i, whose sign is unrestricted in the dual Thus the r_i, can be positive or negative See Gale (1960)

20 The second summation is the total of "subsidies" paid to households We will see later how these subsidies add to the rent-paying ability of households and thereby to the rental income of landowners But notice that the

total value of the subsidies is *subtracted* from the total land rent (and could, in fact, be taxed away from landowners without altering the optimal configuration). Therefore, the value Z^1 to be minimized is actually *net* land rent.

21 Neglecting the r_i we could divide both sides of (1.1) by s_{ih}. Then unit rent (rent per acre) must be no less than unit rent-paying ability (on the right). Since this must hold for every household-bundle combination in an area it then must hold for the combination which would yield the highest unit rent.

22 For each constraint in the primal problem, there must be a variable in the dual and for each variable in the primal there must be a constraint in the dual. The U land constraints (1.1) are associated with the U variables $r^\wedge$. The n constraints (1.2) have a similar correspondence with the n variables v_i. The requirement that all households should be located therefore makes the "subsidy" variables a necessary part of the model.

23 Areas must be fixed in size and number for any run of the model. They may be subdivided or otherwise adjusted after the run.

24. Generally two or more household types will locate in the same area no more times than there are numbers of household groups since the solution to a linear programming model need contain no more nonzero variables than constraints. Our primal problem contains U land constraints and n household group constraints. Therefore no more than $U + n$ variables need be other than zero at an optimal solution. It should be clear that if at least one household locates in every area, no more than n areas can have a second household type.

25 Subsidies to households with negative rent-paying ability in all areas available to them might still appear, much as subsidies to low-income households may be necessary in the real world.

26 Cf. Alonso (1960).

27 The household could use these savings to purchase additions to any component of its market basket other than site or travel without altering the optimal spatial configuration.

28 The discontinuities can be reduced in another way. If we make the areas smaller and smaller (and hence more and more numerous), the rent surface will be smoother. As indicated above, however, there is both a conceptual and a practical limit to this process.

29 E.g. See Chicago Area Transportation Study (1957).

30 Cf. Garrison (1960).

31 Alonso (1960).

32 Cf. Stevens and Coughlin (1959).

33 And the rent variables may be of use in establishing fair compensation in other types of land condemnation proceedings.

4

A Land-Use Plan Design Model[1]

Kenneth J. Schlager

Postwar advances in applied mathematics and electronic computation have stimulated great interest in the application of mathematical models and data processing systems to urban and regional planning. Significant progress has been made in the application of these techniques to urban transportation planning and more recently a number of research projects aimed at the development of land-use models have been initiated. There seems little question that the long-range potential impact of these methods will be revolutionary, but some critics have questioned the relevance of current planning models to the real problems of planners. The obvious question is: What problems are current models able to solve?

Even a brief review of current land-use planning models will reveal a strong emphasis on explaining and predicting human behavior. Quite correctly, many of these models include the word forecasting somewhere in their title description.[2] Such an approach conceives of the urban complex as a phenomenon to be explained scientifically and as a changing configuration that can be predicted in the same way that the solar system can be predicted from the theories of physics. Indeed, such an approach is well designated as applied social physics. The philosophy underlying this approach is the natural result of the direct transfer of the methodology of the physical sciences.

Plan Design: The Central Problem

A contrasting viewpoint conceives of the urban complex as a subject for design. In this approach, the plan is a conscious synthesis of urban form to meet human needs. Rather than serving as a negative restraint on undesirable aspects of human behavior, the plan serves as a positive force for the directed development of the community.

This design viewpoint is not new. It has provided the basis for architectural and engineering achievement for centuries. What is new, or at least overlooked

---

[1] Reprinted with permission from the *Journal of the American Institute of Planners*, Special Issue, Vol. XXXI, No. 2, Urban Development Models: New Tools for Planning, Britton Harris, guest ed. (May, 1965).

[illegible] the possibility of using [illegible] applied mathematics and electronic computation for plan design. Design and not explanation and prediction becomes the primary problem for solution.

The subject of this paper is land-use plan design and a land-use plan design model now under development. Design, however, is but one of a sequence of functions in the planning process. For this reason, the introduction of the design model will be preceded by a discussion of the role of mathematical models in a specific land-use transportation planning sequence.

A system diagram illustrating the functional relationships in the planning process is shown in Figure 4-1. Although this diagram specifically represents the planning sequence related to the formulation of a regional land-use transportation plan, it is typical of other planning sequences.

The first function in the planning sequence is that of forecasting population and employment, as a basis for determining future land-use requirements. In the current Land Use Transportation Study of the Southeastern Wisconsin Regional Planning Commission, new methods of socioeconomic forecasting are being investigated in an attempt to provide more accurate and comprehensive employment and population forecasts. These new techniques, which center around the Regional Economic Simulation Model, are the subject of another paper[2] and will not be discussed in detail here. Whatever the method used, population and employment forecasts must be provided as the output of the first step of the planning sequence.

In the second function, aggregate land-use demand requirements are determined by applying a conversion coefficient usually designated as a design standard to each employment and population category. Such a multiplication and summation will result in a detailed classified set of aggregate demands for residential, industrial, commercial, and other land uses. These aggregate demands provide one of the primary inputs to the third function, plan design.

Plan design lies at the heart of the planning process. The land-use plan design function consists essentially of the allocation of a scarce resource, land, between competing and often conflicting land-use activities. This allocation must be accomplished so as to satisfy the aggregate needs for each land use and comply with all the design standards (derived from the plan objectives) at a reasonable cost.

The plan selected in the design stage of the planning process must be implemented in the real world. Private decisions of land developers, builders, and households may run contrary to the land pattern prescribed in the plan. This problem of plan implementation is the function of the third stage of the planning process illustrated in Figure 4-1, land-use plan implementation test.

Land-use plan implementation is simulated in the Land Use Simulation Model by detailed representation of the decision processes of households and business firms influential in land development. Public land-use control policies and public works programs are exogenous inputs to the model. In practice, a number of experimental simulation runs must be performed with different land-use control policies and public works programs until a set of policies and programs are determined that result in the implementation of the target land-use plan. The feedback on the diagram between land-use development and land-use plan design accounts for the changes that will probably be needed in the plan design to make it realizable. The output of the third stage of the process is a land-use plan capable of practical implementation.

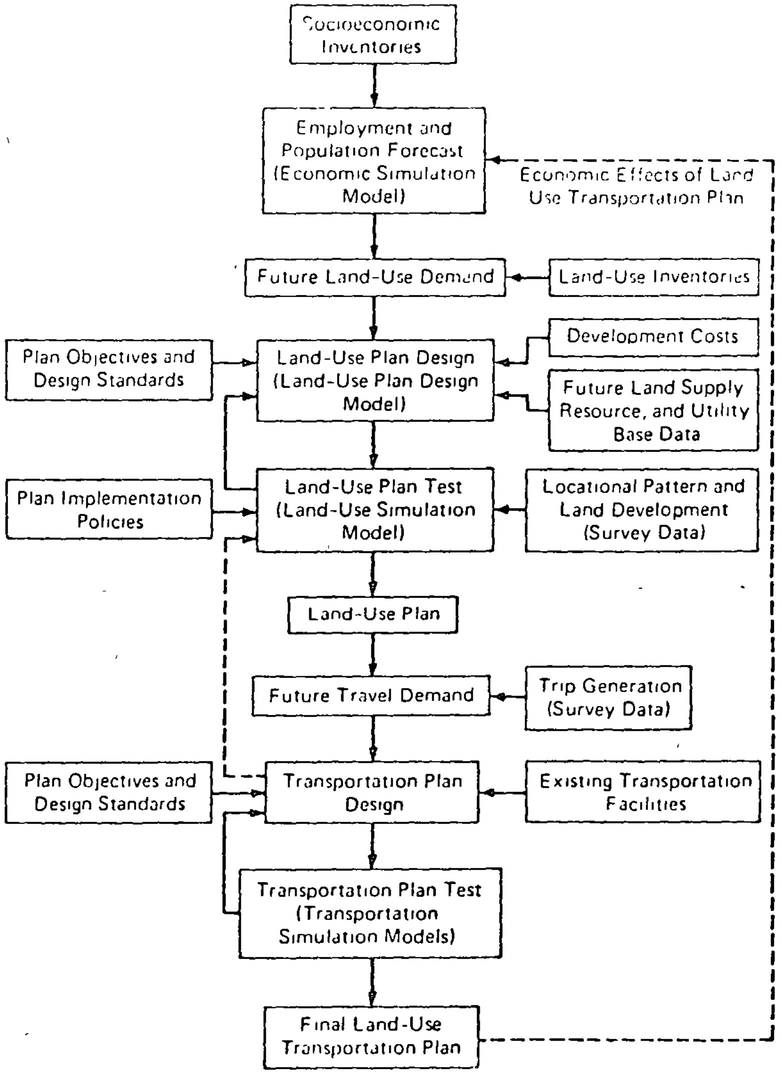


Figure 4-1. Land-Use, Transportation Study Planning System Diagram.

The remaining stages of the planning process depicted in Figure 1-1 relate the development of a transportation plan. The primary inputs to a transportation system are the trips generated by a given land use. For this reason, the land-use plan shown in the diagram is input to the transportation plan design. No models are indicated in the transportation plan design function since none exist to my knowledge. Trip distribution and traffic assignment models ay be used to test the plan intuitively designed by the transportation planner. s a result of model simulation, the transportation plan network is revised ntil a satisfactory system is developed.

Although each iteration in the planning process is important to the final alization of a creative and practical plan, the vital role is played by plan design ce it is the focal point of all preceding and succeeding planning activity.

The Land-Use Plan Design Process

To appreciate the need and requirements for a land-use plan design model, 's necessary to examine closely the design process in general and the land-use n design process in particular. Analytical discussion of the design process is e. Most of the literature on design is based on intuitive and artistic concepts or les that have predominated in certain periods of history.

An exception to this general scarcity of literature is a recent work by xander which defines the design problem in terms of a "fit" between the blem statement and its solution:

> s based on the idea that every design problem begins with an effort to achieve ess between two entities: the form in question and its context. The form is solution to the problem; the context defines the problem. In other words, n we speak of design, the real object of discussion is not the form alone, but ensemble comprising the form and its context.[3]

Achieving this fit between the form and its context is not a simple task, for many requirements that make up the context of the design problem often act in a complex manner. Attempts to satisfy one design requirement often to a violation of another. Faced with such complexity, the designer may mpted to ignore the real design problem and substitute a traditional design. ough such an approach may be acceptable in a political sense, the original blem remains unsolved.

Difficulties in the design process derive primarily from the inability of the n designer to manipulate simultaneously a large number of interacting design bles. Mathematics, particularly in its newer forms such as modern bra, provides a powerful tool for the manipulation of these relationships for ective solution of design problems.

To be useful in design synthesis, mathematical formulatio with two conditions related to Alexander's definition of a "sele one. "It must be possible to generate a wide enough range of p lutions symbolically." Two, "It must be possible to express a solution in terms of the same symbolism."[4] While Alexander d direct solution of selection problems by means of mathematica definition provides useful criteria for the systematic formulatio problems.

Land-use plan design, despite its admitted complexity, po inherent characteristics that meet Alexander's requirements of problem. The first requirement, involving the generation of a w native solutions symbolically, is naturally achieved in land-use reason of the common measure of all land-use plans, the land plans for areas ranging from the smallest subdivision to multi-s be expressed symbolically by three sets of variables:

(1) The type of land use (quality variables)
(2) The density of land use (quantity variables)
(3) The geographic location (location variables).

Typically, the land area concerned will be subdivided into of equal area. The location variable is determined by the geogra of the zone in question. For each zone, the types and densities may be expressed as a measure of the activities in that zonal ar of detail provided will depend on the coarseness of the grid. Fo a zone may be as small as individual residential lot parcels. In l may be counties or even states. The key point to be observed i plans may be expressed by these three classes of variables.

The grid nature of the coordinate system does not limit th rectangular plans. On the contrary, the most complex and irreg expressed with the designated variables if an appropriate grid s

The second condition, relating to the symbolic relationsh alternative forms and design requirements, is also complied wi plan design problem. All design requirements or "standards" r the set of acceptable land-use plans. For a design model, these may be divided into two primary classes:

(1) Requirements that restrict the minimum or maximum nu land use or a relationship between land uses *within a grid* standard). Examples of these requirements are the exclus plan areas from development in a given grid zone (maxim or prevention of the simultaneous development of both n residential land in the same grid zone (relationship standa

(2) Requirements that restrict a relationship between land uses *between grid zones* (interzonal standard), for example, the need to provide an elementary school within a specified distance (or time) of all residential units

In either class the design requirement can be expressed symbolically as an algebraic equality, or more often an inequality, using the three classes of variables noted above. Again compliance with this condition, like the first, is possible because land-use planning is concerned with a single measurable resource, land. That these claims of symbolic design alternative generation and requirements-alternatives comparison are authentic will become more apparent as the design model methodology is explained further

It is useful at this point to provide a specific, succinct statement of the land-use plan design problem indicating the nature of both the design requirements and the design alternatives. To an experienced urban planner the problem will certainly not be new, since it is the same basic problem that he has been concerned with intuitively during his past design experiences. The problem, as stated below, may seem excessively quantitative, and the emphasis on minimal costs may appear unnecessary, but fundamentally it is the same problem of urban form design that has challenged man since cities were found useful. In brief, the problem of the designer of urban form is:

(1) Given design requirements expressed as:
    (a) A set of design standards in terms of restrictions on land-use relationships that may exist in the plan
    (b) A set of needs or demands for each type of land use based on forecasts of future urban activity
(2) Synthesize a land-use plan design that satisfies both the land-use demands and design standards considering the current state of both natural and manmade land characteristics, at a minimal combination of public and private costs

The conceptual basis for minimal costs, it must be emphasized is not to provide a cheap plan but to avoid unnecessary expenditures of precious resources as long as the design standards and land demands are complied with in the plan design.

Intrazonal design standards may take the form of limitations on density or restrictions of the types of land use that may coexist within a zone. An example of an interzonal design standard would be the provision of a regional shopping center within a certain travel time of every residential area. Land-use demand requirements would restrict the set of acceptable plans to those that provided the aggregate total of each land-use need over the entire design area. The current state of the land whether developed or in a natural state, is a primary considera-

tion in plan synthesis because of the relationship of the land to both the design standards and the costs associated with new or renewed development

The Design Model

Two related mathematical techniques will now be discussed as possible frameworks for a land-use plan design model. The first technique, linear programming, has a record of successful accomplishment in other fields and has efficient, highly developed computational procedures. Dynamic programming, the second and newer technique, while not as productive in previous applications or standardized computational procedures, is less restricted in its assumptions and, potentially at least, is a more flexible framework for a land-use plan design model.

Both linear and dynamic programming are sometimes classified as subsidiary fields under the general title of mathematical programming. Such a general classification is desirable, inasmuch as both fields have as their objective the solution of problems involving the optimization (maximization or minimization) of some objective, such as cost, within the restrictions of certain constraints such as design standards. The techniques involved differ considerably, however, with linear programming imposing rather severe restrictions on the nature of both the objective and constraints while dynamic programming is almost unrestricted in its formulations of both the objective and constraint functions. Linear programming models, on the other hand, can usually be solved by the use of standardized computational procedures, while dynamic programming usually provides at least a serious challenge and often insurmountable obstacles to an efficient computational solution. With either technique, the sheer size of many land-use plan design problems bring with them what has been called the "curse of dimensionality," which militates against any simple "brute force" approach to solution.

The linear programming formulation of the land-use plan design model problem is straightforward. The objective function relates to the cost of developing land for a given land use

$$c_t = c_1 x_1 + c_2 x_2 + \dots + c_n x_n$$

where the variables (x) may represent residential, industrial, or other land uses in given areas and the constants (c) the costs of developing this land. Land-use categories may be subdivided into subsidiary classes such as single family residential, multifamily residential, and the costs may be related to the topography and soil characteristics of the area. With each subdivision, of course, the number of variables grows larger, and the computation time for a m[illegible] solution is are used. In practice, a compromise must be made between the desire for detail and reasonable solution times. With the [illegible] development in computer technology, however, this problem will [illegible] decreased significance in the coming years.

The equality and inequality constraints in the linear programming formulation of land-use plan design include the following:

(1) The total demand requirement for each land-use category (equality constraint)

$$d_1x_1 + d_2x_2 + \dots + d_nx_n = E_k$$

where

E = regional land-use demand requirement for each land use
d = service ratio coefficients which provide for supporting service land requirements, such as streets, which are necessary for primary land-use development

2) Maximal (minimal) limits on land uses within a zone:

$$x_1 + x_2 + \dots + x_n \leqslant F_{m}$$

where

F_{nm} = upper limit on land use n in zone m

3) Interzonal or intrazonal land use relationship constraints

$$x_n \leqslant Gx_m$$

where

G = ratio of land-use n allowed relative to land-use m with land-uses m and n in the same or different zones.

The land-use demand equality constraint (1) follows a standardized format th one equation for each primary land-use category. Since some land uses such single family residential are usually subdivided further according to lot es, the number of demand equations in a typical design model may exceed 20 ationships. It is important to emphasize that only primary land uses such as idential, industrial, agricultural, and recreational land are directly determined ice ratios incorporated in the d parameters account for secondary land uses h as local streets and parks.

The second and third categories of constraints reflect the design standards and may take a wide variety of forms. The maximal constraint will usually reflect a density standard, but it also may provide for the exclusion of an unsuited soil type for a given type of land use. Land-use relation [illegible] from design standard restrictions on coexistent land uses within a zone or in adjacent zones. Accessibility standards for employment and shopping areas will also be reflected in this type of constraint.

The above constraint relationships reflect the types encountered so far in experimental plan design model runs in test areas. Other constraint forms may b needed when a complete regional plan design is attempted, but they may be easily included as long as they are linear, continuous constraints. Nonlinear discontinuous constraints are not possible with linear programming and account fo the primary disadvantages of the method.

For a region subdivided into about 30 zones, the size of a typical linear program for a land-use plan design is about 60 equality and inequality constrain and 400 variables. Computer time on an IBM 1620 computer is about three hours. On larger systems, such as the IBM 7090, it would take less than 30 minutes.

Model Application

Some initial experience with applications of the model will now be detailed in order to provide the reader with an idea of the input data requirements and computational characteristics of the model.

Four primary sets of input data are required for model operation:

(1) The costs of unimproved land and land development for each primary land-use activity for each type of soil
(2) The aggregate demand for each primary land-use activity
(3) Design standards which reflect the plan objectives and restrict the set of acceptable plans by limiting interzonal land-use relationships
(4) The current land inventory, which will include both land-use activities by area and soil characteristics.

Land-development cost data may be obtained either by engineering estimates or by statistical analysis of recent land development in the area. The former approach has been used in the initial tests of the model in the Waukesha city pilot area. Collection of land-development cost data is always expensive and in many cases difficult or even impossible to obtain. Land developers are usually extremely reluctant to reveal their costs, and the cost data obtained is of uneven quality since many developers do not maintain complete records. For all these [illegible] engineering cost estimates are usually preferable if competent professional experience is available

In the Waukesha area separate land-development cost estimates were made for five sizes of residential lots with their associated service land uses such as streets, neighborhood shopping, schools, and parks. Additional cost estimates were made for industrial, regional shopping, and regional park land uses. These were not gross estimates but detailed analyses of the costs of each improvement related to both the land use and the type of soil involved. All estimates were subdivided into their component parts, each with its individual cost.

Separate cost estimates were prepared for each of three classes of soil. Soil data were obtained from a comprehensive soil survey made in southeastern Wisconsin as part of the Land Use Transportation Study. Unimproved land costs presented a special problem since they could not be obtained from engineering estimates. Assessed and equalized land-value data were obtained from each of the communities and were adjusted on the basis of prices realized in recent land transactions in the area.

Initial tests of the model used historical aggregate land-use demands for 1950-1962 to provide comparisons between actual and "optimal" land development in the area. Typically, however, a design application of the model will require forecasts of future land-use demands, which may be obtained by applying design standards to forecasts of population and employment in the region of interest.

The various forms of design standards usually provided were described in the previous section. In current tests of the model, design standards were limited o the exclusion of development from areas such as flood plains, along with the rovision of service ratios for the amounts of secondary land (streets, parks) equired to support the primary land uses. Design standards for the regional and-use plan are still in preparation and will be used in model tests as soon as hey become available.

An inventory of both current land-use activities and soil characteristics is ritical for model application. In current tests, developed areas were eliminated rom consideration for future land development. It is possible, however, to onsider redevelopment in the form of urban renewal as a set of alternatives in ne design. For this approach, redevelopment costs would be required. Through ne use of the soil inventory, it was possible to assign a development cost to each ubarea in the test area.

Proper presentation of the Land Use Plan Design Model output is an import-nt consideration in achieving acceptance of its design by planners and vernmental officials. Initial model outputs were in tabular form and were eaningful only to someone familiar with the operation of the model. Improved esentation was later achieved by tabular designation of the intensity of each

land-use activity in each zone. Printed output was supplemented by colored land-use maps manually prepared from the tabular print-out.

Available Mathematical Models

Although linear programming provides a reasonably satisfactory framework for a land-use plan design model, it possesses certain inherent disadvantages that restrict its usefulness in design. The primary limitation is the need for continuous rather than discrete values for the land-use variables. Land-use design choices are by nature usually discrete rather than continuous. The basic element of residential land use is the subdivision rather than the lot. Industrial land-use units tend to be industrial parks rather than individual factory sites, much less land acreage. While it is possible to round off the linear programming solution to satisfy these natural discrete levels, such a solution does not usually correspond with the associated discrete optimal combination. A second limitation of linear programming is the need for both a linear objective function and linear constraints. The linear objective function is not a severe limitation, because the inaccuracies introduced by a linear approximation of costs are usually less than the errors of cost estimation. In the few instances where known nonlinear cost functions occur, such as in the plant capacities of areawide facilities for water supply or sewage treatment, the cost break may usually be satisfactorily approximated by a multivariable series of linear cost variables.

Nonlinear constraint relationships present a more serious problem. Certain design standards are inherently nonlinear, and a linear approximation sometimes provides an unsatisfactory substitute. When a design model is not able to provide satisfactorily for a design standard, it loses most of its usefulness.

Dynamic programming, another member of the mathematical programming family, has the potential for removing the two primary restrictions inherent in linear programming. Although dynamic programming may be used to solve the same land-use plan design problem, it is based on a different class of mathematical procedures, which are capable of handling discrete and nonlinear objective functions and constraint relationships.

Richard Bellman of the Rand Corporation was the originator of dynamic programming and has developed the theory and application of this multistage approach to decision making to a high degree in the last decade. A large number of classes of dynamic programming processes have been formulated for problems in production scheduling, rocket trajectories, and feedback control systems, but the class of process of primary interest in design is the allocation process.[5] In a dynamic programming model, the basic cost and design relationships are similar to those defined for the linear programming model, but the method of

[illegible] data and permits the use of more complex and [illegible] relation-
ps

The Design Concept in Urban Planning

The ultimate contribution of this paper will depend on its success or the
of it in accomplishing at least a partial reorientation of land-use model
elopment toward design. Although the importance of forecasting land-use
elopment was indicated by the role of the Land Use Simulation Model
lly described earlier in this paper and detailed in another recent paper,[6] the
gers and limitations of nondesign-oriented models that are only remotely
ed to the synthesis of better urban and regional plans should be apparent.
The need for design models in urban planning is fortunately accompanied
reater possibilities for their success. Industrial applications of mathe-
cal models in normative functions such as production scheduling and
mal product design have been conspicuously more successful than attempts
nulate human behavioral patterns in a market. Quite simply, it is much
r to use a model to tell people what they *should do* than to explain *what*
are doing. Given the fantastic complexity of the modern metropolis, would
t be well to emphasize model development in areas that promise both a
ficant contribution and a high probability of success?
The image of design in urban planning as a remnant of a bygone age of the
beautiful" must be replaced by a new design concept based on the creative
esis of complex plans using all the tools provided by modern technology.

ostscript: Commentary
n Five Years of Development Experience

The original concept of a Land Use Plan Design Model was originated in the
er of 1964 at the Southeastern Wisconsin Regional Planning Commission
RPC) by Kenneth J. Schlager and was further developed in detail for the
ation of the foregoing original paper in the JAIP in May of 1965. This
d paper stirred enough interest to serve as the basis for a research proposal
Department of Housing and Urban Development. The proposal resulted
search grant to SEWRPC under the Urban Planning Research and
stration program with the project designation of Wisconsin PD-1. This
h has proceeded through two phases and is currently at the beginning of a
nd concluding phase. This commentary will relate some of the experience
during this research program and reflect some thoughts on the future
al and limitations of urban design models.
mmary, the purpose of Phase I of Wisconsin PD-1, which began in

October of 1966 and ended in January of 1968, was to develop a Land Use Plan Design Model and apply it to a small community as a test vehicle. Two important lessons were learned in Phase I. The first was learned early in Phase I and influenced the course of action. The second became apparent too late to have any effect on Phase I but became a primary determinant of the Phase II Program. These two were:

(1) Since the nature of urban design is such that it deals with discrete objects (schools, hospitals, shopping centers, etc.) rather than areas of land, the use of linear programming as the model algorithm became undesirable if not infeasible.
(2) The computer programs for reducing raw data to meaningful model inputs is as challenging a task as the computer program for the model itself.

The impact of the first lesson was felt early so that little time was spent with linear programming, and the model algorithm was formulated to deal directly with discrete elements designated as "modules" which were placed in geographic areas known as "cells" subject to certain constraints and with the objective of minimizing development and operating costs. The model algorithm, which used a steepest descent search approach, operated fairly satisfactorily in Phase I but its limitations became quite apparent with the larger scale applications of Phase II.

Data-reduction problems, the second lesson, were painfully obvious later in Phase I in which model input data was provided through a series of unrelated data processing programs. This piecemeal approach, with its clumsy inconvenience, served as the incentive for an integrated data-reduction program package in Phase II.

Phase II, which began in July of 1968 and ended in September of 1969, was concerned mainly with the application of the model on a larger scale to the development of a regional land-use plan. Since the SEWRPC is a regional planning agency and had previously developed a regional land-use plan by traditional methods, the plan developed by the model could be subject to some searching comparisons.

Since the Southeastern Wisconsin region is large in size (2,689 square miles), Phase II was dominated by data-reduction activities. With the experience of Phase I in mind, a great deal of effort was directed toward the development of an integrated data-reduction system. Integration was enlarged such that the data-reduction program and the model operating programs blended into one system with raw module, constraint, cost, and topographic data as inputs and a land-use plan as an output.

In Phase II with the data-reduction problem under control, the first opportunity arose to see the results of model operation with realistic input data. Although the results of these initial model runs were encouraging in that they resembled a plausible plan considering the goals of constraint observance and [illegible] need for future improvement of the model were indicated in wo areas

1) The model algorithm seemed slow in operation and did not provide a true optimal (or a near optimal) solution
2) The design constraints, which are intended to reflect the objectives of the design other than minimal cost, require a great deal of thought and study

The original model algorithm assigns modules to areal cells in a series of inary partitions. The whole design area is first partitioned into two subareas, nd each module is assigned to one of these two subareas based on constraint bservance and cost minimization. Each subarea is in turn divided in half, and odules are assigned to each half again. This process continues until the smallest rea is assigned a single cell. This approach leads to "holistic" errors in that odules are assigned early in the process to a set of cells that later prove to be onoptimal. Since such errors cannot be corrected, they result in permanent mis- location of modules contrary to an optimal plan. After a great deal of investiga- on of possible improvements in the present algorithm, a new approach to gorithm formulation seemed desirable.

The other problem concerning constraint formulation was less a mathe- atical problem than a value problem. Quite simply, the professed goals, jectives, and design criteria do not reflect the real goals, objectives, and design iteria. In the language of social psychology there is a "hidden agenda" of in requirements. The problem is to uncover this hidden agenda of real design quirements and convert them into constraints for input to the model.

Results to date and the nature of existing problems indicated the need for a o-pronged attack to make the model a useful tool for planners, architects, d engineers:

- A program to document the existing model and associated data-reduction programs in the form of a User's Manual understandable by the practitioner such as the planner, architect, or engineer
- A research program to explore fundamental questions of algorithm opera- tion and constraint formulation.

Fortunately, financial support has been obtained for both programs. A lopment program to document the present model with some minor improve- its will be funded by the United States Department of Housing and Urban elopment under the Wisconsin PD-1 Program to the Southeastern Wisconsin onal Planning Commission. A second research-oriented program is being [illegible] by the National Science Foundation through Marquette University ether these programs have the potential of reducing the design model research [illegible] to a system [illegible] in the practical world of design

Notes

1 *Review of Existing Land Use Forecasting Techniques*, Traffic Research Corporation, presented to the Boston Regional Planning Project, [illegible] (1963)
2 Schlager, Kenneth J., "Simulation Models in Urban and Regional Planning," *Technical Record*, Southeastern Wisconsin Regional Planning Commission, Waukesha, Wisconsin, II, No. 1 (1964)
3 Alexander, Christopher, *Notes on the Synthesis of Form*, Harvard University Press, Cambridge, pp. 15–16 (1964)
4. *Ibid.*, pp. 74-75.
5. Bellman, Richard E., and Dreyfus, Stuart E., *Applied Dynamic Programming* Princeton University Press, Princeton (1962).
6 Schlager, *op. cit.*

a consideration of the simple correlations between all pairs of variables and arrives at a generalized coefficient of determination R^2.

Especially in the urban context, multiple regression is plagued by the fact that many variables are spatially distributed in a correlated fashion. When a large number of such correlated independent variables are considered as a basis for explaining other phenomena, the situation gives rise to a mathematical difficulty called collinearity or multi-collinearity. The existence of collinearity casts many doubts on conventional statistical analysis, and creates severe operating problems. One approach to reducing the number of variables and the interaction between them while preserving, if desired, the influence of a large number of variables, is based on the use of component analysis or factor analysis. The actual nature of these techniques is in detail highly abstract and mathematical. In concept, however, they provide a way in which the appearance of a number of interrelated variables in the metropolitan area can be reduced to a more limited number of independent factors. In demographic analysis, for example, it has been found that a large proportion of the differences between neighborhoods in a city may be explained on the basis of only three independent or nearly independent factors which are related to status, density and family size, and segregation. Similar component or factor analyses provide a powerful means for organizing large masses of data and identifying underlying patterns of interrelations among variables. Their use is largely descriptive and does not frequently enter directly into further computations.

For purposes of further organizing knowledge gained from factor analysis and other statistical analysis, some effort has been directed toward cluster analysis. This technique, as yet imperfectly developed, is directed at providing systematic means for grouping areas or lines of business or population classes for further systematic treatment.

An important technique for solving certain problems in optimizing or in market behavior is called linear programming. This technique will find by iterative methods an allocation which maximizes or minimizes a linear objective function. The allocation may be an allocation of construction to subareas, an allocation of families to dwelling units, or of budgets to projects. The objective function would be some measure of cost or achievement determined by multiplying each allocation by an associated coefficient and summing over all cases. The solution to a linear programming problem has a dual solution which provides information about costs and benefits assignable to different aspects of the allocation arrived at. In particular, the imputed values derived from the use of resources are sometimes called shadow prices.

No attempt has been made to provide an exhaustive glossary or discussion of these problems. The interested reader is referred to articles in this *Journal* and to standard textbooks.

NOTE

[1] For a more complete discussion of some of these points, see my "Plan or Projection: An Examination of the Use of Models in Planning," *Journal of the American Institute of Planners*, XXVI (November, 1960), 265–72.

AN OPPORTUNITY-ACCESSIBILITY MODEL FOR ALLOCATING REGIONAL GROWTH

George T. Lathrop and John R. Hamburg

The authors present in this article the results of the first stages in the development of a model of urban growth which depends heavily on transportation concepts and which makes use of a modification of Morton Schneider's intervening opportunity model (ordinarily applied to trip distributions). In requiring the operator of the model to estimate activity densities, this paper recalls Hamburg's contribution to the Journal *of six years ago, but the methods used for distributing growth appear to be much more realistic. The authors express the hope that levels of activity densities may in the future be established within the model. The paper is especially noteworthy for the attention it gives to the problems of designing a flexible and easily used model and of establishing good communication between the computer and the analyst in the* *form of legible and usable outputs. Many of these results are reproduced here and the design criteria for the model are summarized at the conclusion of the article.*

B. H.

John R. Hamburg is Director of the Bureau of Planning, Subdivision of Transportation Planning and Programming, New York State Department of Public Works. The Bureau is responsible for preparing comprehensive transportation plans for metropolitan areas in upstate New York. Formerly a staff member of the Chicago Area Transportation Study and the Detroit Metropolitan Area Traffic Study, Mr. Hamburg has also been a consultant to transportation studies in Pittsburgh, Wichita, and Oahu, and a visiting lecturer at the University of North Carolina, Cornell University, and the Salzburg Seminar in American Studies. He is an affiliate of AIP.

As a staff member of the Bureau of Planning, Subdivision of Transportation Planning and Programming, George T. Lathrop has worked on transportation plans for metropolitan areas in upstate New York since 1962. He holds the MCP degree and the Certificate, Traffic Engineering, Bureau of Highway Traffic, from Yale University. He is a provisional member of AIP.

Systematic and comprehensive transportation planning has come to depend more and more upon objective techniques involving the use of high-speed computers to deal with problems of data processing, data analysis, traffic simulation, and network evaluation. The selection and design of multi-million dollar transportation plans depend on the simultaneous evaluation of many diverse factors, including traffic volumes, operating costs, construction costs, land costs, accident characteristics, and travel costs for thousands of transportation links and millions of people. Subjective evaluations and intuitive speculations have come to play a much smaller role in planning transportation systems than ever before.

This trend is also apparent in city and regional planning. While it is clear that design and aesthetic characteristics have a very important role to play in planning cities of the future, it is equally clear that the analysis and manipulation of the massive details that make up a functioning city of a million inhabitants require computer technology. Nowhere is this more apparent than in the projection of land and transportation requirements for a metropolitan region to a time 20 to 25 years in the future.

The opportunity-accessibility model attempts to bring into an objective framework methods and concepts which have been commonplace in planning for several years.[1] Their application, however, has often been subjective, and their replication by other professionals has been difficult. Obviously, the manual application of holding capacity, access, and density in estimating growth allocation is not new, but their use has been limited largely to this sort of subjective process and to small geographic areas.

There are empirical as well as theoretical shortcomings in the model as it now exists. It does, however, incorporate some of the more significant factors thought to be associated with the growth and functioning of urban regions into a flexible program which produces spatial arrangements that correspond quite closely to observed patterns. To the extent that the model simulates urban growth, it is extremely useful in providing the measured statements necessary for planning transportation facilities.

The model has been developed and is being tested by the Transportation Planning and Programming Subdivision of the New York State Department of Public Works. Its purpose is to allocate future estimates of activities (expressed in the form of trip destinations, in this instance) to small geographic areas in the Niagara Frontier Area centered on Buffalo, New York. This geographic allocation, in turn, is the input to the traffic assignment model which is used in the testing, evaluation, and design of alternative systems of transportation facilities.

THE CONCEPT OF THE MODEL

The model is an opportunity model. In essence, the spatial distribution of an activity is viewed as the successive evaluation of alternative opportunities for sites which are rank-ordered in time from an urban center. Opportunities are defined as the product of available land and density of activity (units of activity per unit area of land).

$$A_j = A\,[e^{-\ell O} - e^{-\ell(O+O_j)}]$$

where A_j = the amount of activity to be allocated to zone j,

A = the aggregate amount of activity to be allocated.

ℓ = probability of a unit of activity being sited at a given opportunity,

O = the opportunities for siting a unit of activity rank-ordered by access value and preceding zone j, and

O_j = the opportunities in zone j.

The use of the negative exponential formulation following an access search across an opportunity surface presumes that the settlement rate per unit of opportunity is highest at the point of maximum access or, most usually, the center of a region. This elementary presumption is consistent with both empirical observations and the bulk of the theory dealing with the economics of land use.

THE NOTION OF OPPORTUNITIES

The concept of an opportunity for siting a unit of activity involves both land and a measure of the intensity of use of that land. Land use intensity or density has been treated as an equilibrium of the price of land and transport costs.

A historical analysis of density must consider changing transportation costs, changing building costs with particular emphasis on the costs of first floor area versus multi-storied floor area, and changing requirements or preferences for location among competing activities. In addition to the difficulties that these considerations impose, there is the problem of structural rigidity of the physical region in terms of buildings and transportation facilities. These represent substantial investments which change only slowly.

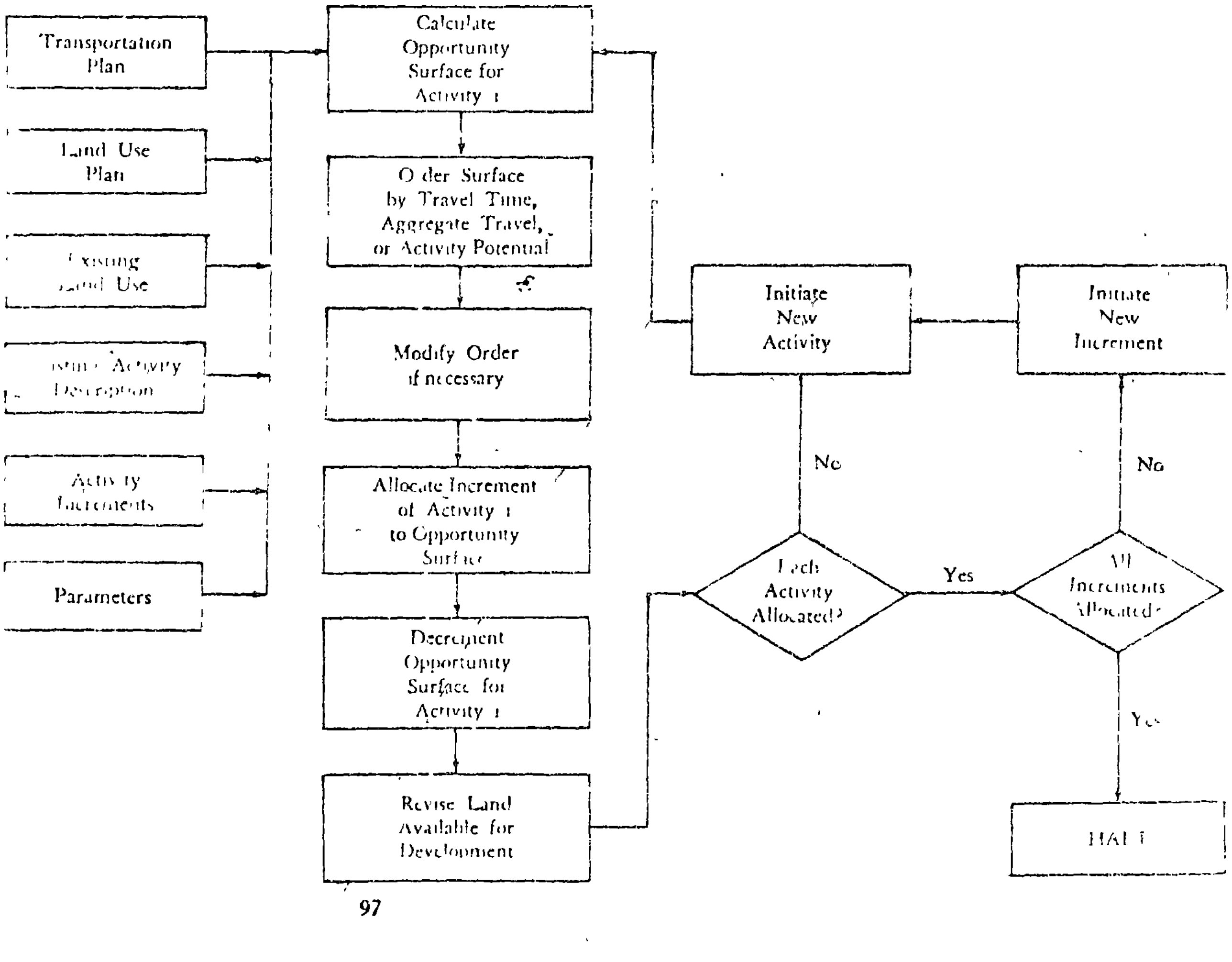
Largely because of the difficulties involved in simulating the intensity of land use, the present density has been selected as an appropriate measure which is introduced independently into the model. This independence allows the use of alternative densities, whether they are analytically derived, guessed, or planned. Preferably, these values should be an integral part of the model, but that will have to be done another time.

THE PROBABILITY OF SITING—"ℓ"

This parameter is the probability that a unit of activity will settle or be sited at a unit opportunity. For a given surface of opportunities, the larger this value, the more tightly packed the region will be. The smaller the value of "ℓ" the more scattered or sprawling the settlement pattern will be. Thus it is a measure which describes, within the constraint of the density-land opportunity surface, the relative importance of central positioning within the region.

The model distributes growth increments across an opportunity surface which has been rank-ordered by time path value to the center. After each increment of growth is allocated, the available land is reduced by the amount of land required to site the increment of activity, the opportunity surface is decreased, and the activity inventory by zone is updated. (Figure 1 shows the sequence of operatio

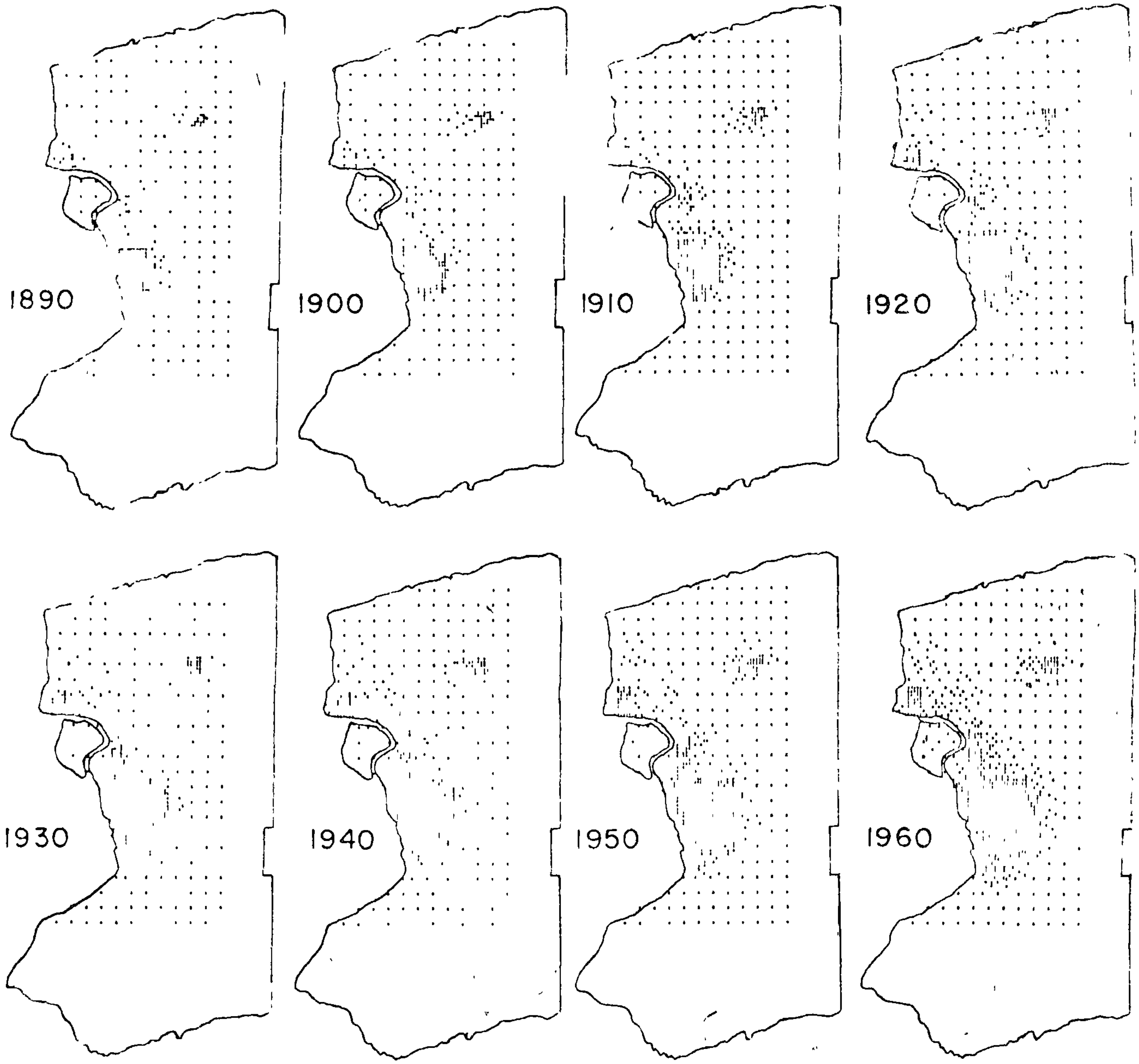
FIGURE 1 *Sequence of Operations within the Model*



within the model.) If competing activities are ignored for the moment, the use of an "ℓ" with large values would tend to settle each unit of activity at the first opportunity encountered. Thus growth would simply be a process of using land completely in ever-increasing bands of access from the center. Very small values of "ℓ," on the other hand, would tend to scatter activity across the region. While the center would still dominate, the pattern of settlement would be very sparse. As "ℓ" approaches zero, the notion of a region simply disappears.

There appears to be some crude historical correspondence to a decreasing "ℓ," presumably as a result of changes in the transportation technology (especially the widespread use of the auto). Thus the transition from rural to urban was more abrupt at the edges of earlier cities. Land in a given time-ring tended to be substantially used up before successive time-rings were settled. Currently, the demarcation is typically in a broad band which may be several miles in width.

FIGURE 2 *Virginia Frontier Population Growth Simulation with Prototype of Land Use Model*



Lathrop
Hamburg

An analysis of the City of Chicago population settlement pattern reveals a lessening of the slope of land saturation at increasing distances from the Loop through time

Figure 2 shows a simulation of the growth of the Niagara Frontier region produced with a prototype of the present model The decreasing concentration of activity is apparent in these allocations, which are quite similar to the growth that actually occurred in the region

THE NOTION OF ACCESS

The obvious impact of transport costs on the development of a region has long been recognized. The inclusion of some measure of accessibility into any model proposed to simulate present growth or estimate future growth is imperative. However, the form and weight that access should have in the model are not so obvious; in fact, this is the major area which requires clarification.

The model starts from an overly simplified notion that growth begins at a center and proceeds outward. The supply or surface of opportunities for growth will be examined in the order of travel time required to reach any location on that surface. This concept is neither new nor especially unique.

Experiments with the model prototype revealed that the settlement pattern was very sensitive to the transportation facilities Figure 3a shows a settlement pattern which might hypothetically have resulted from a transportation system composed of five high-speed facilities radiating out from the center. This pattern has been noted in real cities and is especially conspicuous in the stellate pattern of Chicago which is superimposed on the radial commuter lines.

Other hypothetical networks also produced plausible patterns. For example, a simple grid of transportation facilities with equal speeds gives a square settlement pattern rotated 45° with respect to the grid (Figure 3b). If the central X- and Y-dimension facilities have a speed advantage over the other facilities, the sides of the square are pulled in, and the settlement pattern approaches the shape of a four-pronged starfish (Figure 3c). Some midwestern plains cities do in fact correspond to this pattern or very similar ones.

If only a single transportation line offers high-speed service, the linear form of the city emerges. This is common to cities which fall in a valley with the main street running parallel to the ridges. Here, of course, the topography itself (in the form of the opportunity surface) tends to reinforce the linear form of this settlement pattern (Figure 3d).

The notion of access, as developed here, should not be confused with the accessibility notion that a given location is related to all other locations by the sum of the quotients formed by each location's activity, divided by its time or distance to the given zone raised to some power. This gravity concept, or propensity for interaction of access, is a distinctly different measure An option for calculating this measure has been included within the model, although to date the results of using this alternative measure of access have not been compared.

The major point to be emphasized here is that by using minimum time paths via the transportation network, the model has a much more refined measure of the opportunity surface than airline distance to the central business district.

Thus, within the limitations of the allocation procedure, it is possible to incorporate the effects that specific transportation improvements would have on the settlement pattern The more complex question of feedback between land use and the transportation system can then be handled from a much better position.

THE MODEL AS AN AID IN EVALUATING LAND USE POLICY

The preceding description of the model illustrates its utility as a basis for forecasting the future distribution of people and trip-making. It is not necessary, however, to so limit its use. The model can be used to examine the regional growth that might result from certain policies with respect to land development

DENSITY CONTROLS

Density or the intensity of the use of land is fundamental to the structure of urban regions* The contrast between New York City and Los Angeles is the usual example of density extremes that have arisen in different regions. A considerable range of densities can be found even within a single urban region The density of population in the center of Buffalo is 93,250 persons per square mile of net residential land, in contrast to the 10,000 persons per square mile of net residential land found in the suburbs.

Although the model can shed no light on the issue of whether high densities are good or bad (and city planners can be found on both sides of this question), or on the question of the extent to which planners can effectively control densities, it can be used to examine the development which might occur with a prescribed density surface as opposed to the extrapolation of present densities.

OPEN SPACE CONTROLS

Another policy often considered in shaping the urban settlement pattern is the use of controlled open space. An open space plan can be entered into the model as a separate input, zone by zone, and the resultant hypothetical settlement pattern will emerge in visual form. This approach may be used in conjunction with density controls or completely independent of them. Again, however, a rapid,

FIGURE 3 *Settlement Patterns Corresponding to Different Transportation Assumptions*

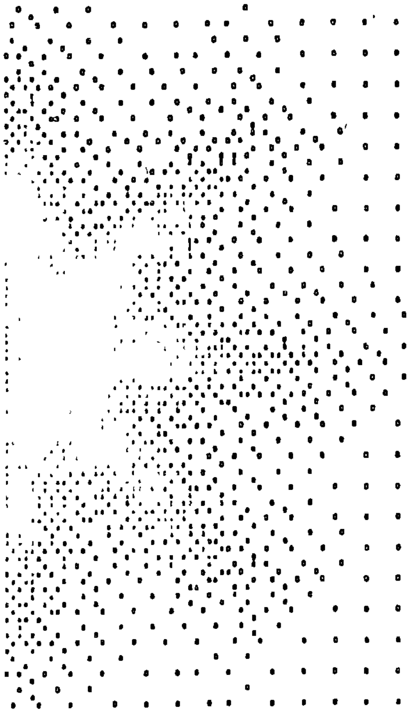


FIGURE 3a

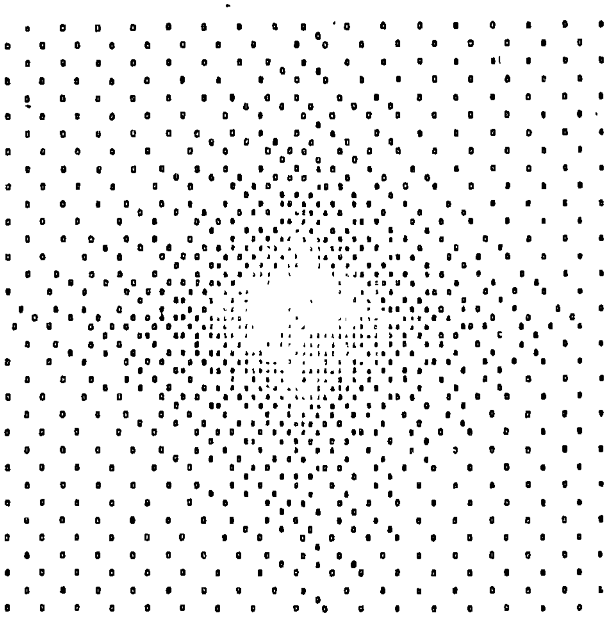


FIGURE 3b

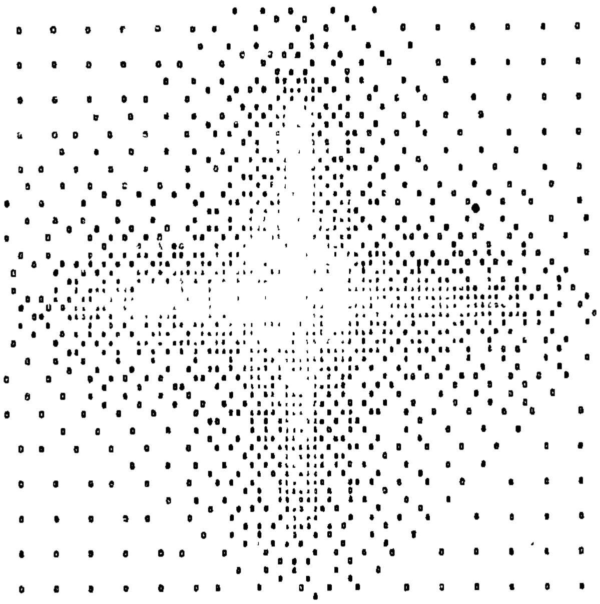


FIGURE 3c

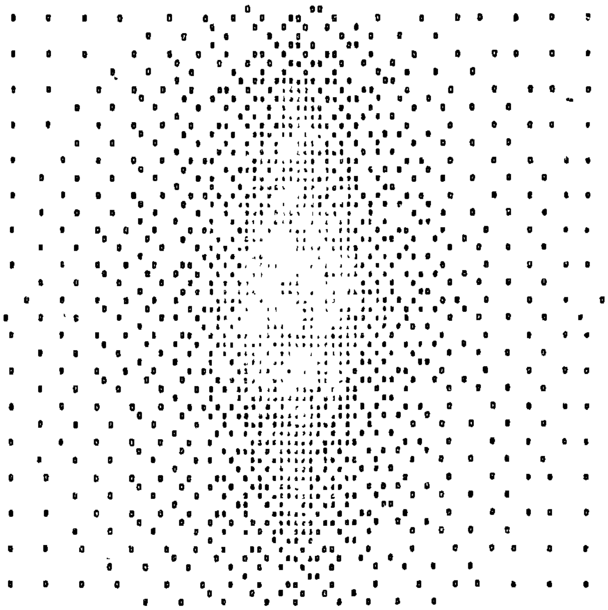


FIGURE 3d

FIGURE 4 *Influence of Planning Assumptions on Distributions by the Model*

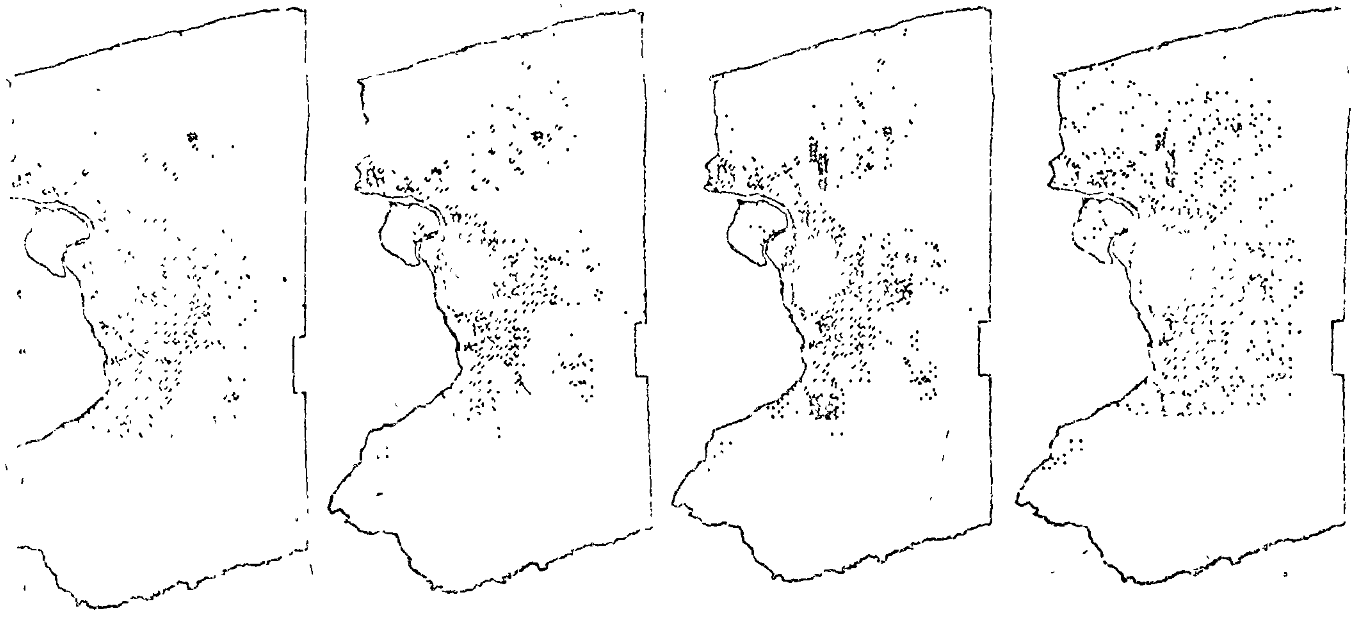


FIGURE 4a *Graphic Summary of Planning Assumptions*

visual picture of the region is obtained, just as though one had done a broad-stroke sketch plan. The difference, however, is that a measured statement is provided of how many people and trips will be located in each zone of the region. Thus quantified, the settlement can be converted into loads on the proposed transportation network; and the resulting transportation costs, including costs of construction as well as travel, can be used to evaluate the transportation plan. Eventually, of course, one would wish to identify costs of land development as well as transportation in order to evaluate alternative plans.

It should be clear that the use of the model is not restricted to forecasts alone. Specific plans for redevelopment, shopping centers, the central business district, open space, density controls, and highways can be entered into the model. Such plans are considered as givens, and the model then estimates the population and travel distributions that would accompany these plans.

An example of the inclusion of planning decisions or assumptions is illustrated in Figure 4. Figure 4a illustrates two assumptions:

1. That a strong open space program would reserve nearly all currently vacant land in five areas between present radiating corridors of development.

2. That three high-density sub-centers would be encouraged through density controls north, east, and south of the major regional center, Buffalo, and that these centers would be connected to Buffalo and Niagara Falls by a high-speed, non-stop transit facility. (Two of these centers are located in presently developed areas where remaining vacant land would be developed at higher density. The third, north of Buffalo, is in an area nearly completely vacant.)

Figure 4c shows the allocation pattern resulting from inclusion of the first assumption. In the allocation shown in Figure 4d the second assumption was included alone, while the pattern in Figure 4e resulted from the reservation of open space and the introduction of high density sub-centers with transit service. Figure 4b provides contrast as an allocation made with neither of the assumptions.

Graphic displays such as these are extremely useful for quick comprehension of the spatial allocation of a growth increment. But concrete measures of the differences in the allocations are necessary for land use and transportation planning and for evaluation of the impact of various alternative plans.

Again, to cite a brief example, the amount of land brought into use illustrates the results obtained from the different assumptions.

Amount of Land Brought into Use

	Square Miles
No controls assumed	217
Open space only	250
New centers only	177
Both assumptions	201

The figures for trials including the open space assumption do not include the 40 square miles of reserved vacant land.

It is interesting to note that the inclusion of the open space assumption resulted in a larger quantity of land being brought into use in each case. This result is understandable, for land held out of development by the open space reservation would have developed at densities higher than land which is farther from the center and is now brought into use. A more sophisticated approach than mere extension of present densities for newly developing sub-areas would have altered this result.

SUMMARY

Qualitative examination of these hypothetical results provides persuasive evidence that the settlement pattern has historically responded to access as we have measured it. Quantitative verification and calibration is difficult because of the lack of historical data on land use and transportation networks. An attempt is being made, however, to measure the extent to which population growth since 1940 can be explained by use of the model. This requires the assumption that net activity densities have remained fairly constant during this period.

The use of external, independent estimates of activity density is not satisfactory. Eventually, the determination of these densities must be incorporated within the model.

The operating and output flexibility of the model are a great advantage in research. Evaluating the extent of variation in growth attributable to additional variables or to different statements of access is an immediate goal. In this sense, the model can be considered an experimental technique which will allow the urban growth process to be further measured and, hopefully, better understood.

CRITERIA FOR MODEL DESIGN

1 The model should be based on a theoretical statement of the mechanisms of land development Although it need not simulate individual decisions within the land market, it should give results which correspond to the real world

2. The model should be incremental and recursive Ideally, data on past land use and transportation systems should be used to simulate the present development pattern If this ability is lacking increments of growth should be layered on the present structure

3 The model should be relatively simple A finite number of land uses and a minimum number of sub-sets of households should be required to minimize the difficulties of data acquisition and handling

4 Ideally, the calculation of activity density should be contained within the model Alternately, the model should readily accept exogenous densities.

5 The model should accept alternative measures or indices of access. This condition provides the flexibility required for situations in which one measure is particularly appropriate to a given activity type while a different measure of access is best suited to other activity types

6 The model should be able to accept data from redevelopment, urban renewal, or new-town plans This operation can be handled as a preliminary updating (internal to the model) of the land use and activity base or it can be done within the main frame of the model

7 The model should be capable of being calibrated For example, it should be possible to simulate past growth, or at least to calibrate the model parameters using the present structure

8 Provision for sensitivity analysis should be considered in the design of the model. It is vital to be able to evaluate the effect of unit changes in a given parameter on all facets of the allocation produced by the model

9 The output of the model should permit easy and rapid comprehension of allocation results, with particular emphasis on a simple graphic description of settlement patterns This graphic output is particularly important for the comprehension and evaluation of alternative model inputs Tabular outputs which can be used in calibration, sensitivity analysis, and allocation evaluation are also an obvious requirement

10 Output from the model should be directly usable in existing traffic assignment procedures to minimize the difficulty and time involved in applying the results of the operation of the model

The ultimate contribution of this paper will depend on its success, or the lack of it, in accomplishing [illegible] conversion of land use model development toward design. Although the importance of forecasting land use development was well stated by the [illegible] of the Land Use Simulation Model [illegible], criticized earlier in this paper and detailed in another recent paper,[6] the dangers and limitations of non-design-oriented models that are only remotely related to the synthesis of better urban and regional plans should be apparent.

The need for design models in urban planning is fortunately accompanied by greater possibilities for their success. Industrial applications of mathematical models in normative functions such as production scheduling and optimal product design have been consistently more successful than attempts to simulate human behavioral patterns in a market. Quite simply, it is much easier to use a model to tell people what they *should do* than to explain *what they are doing*. Given the fantastic complexity of the modern metropolis, would it not be well to emphasize model development in areas that promise both a significant contribution and a high probability of success?

The image of design in urban planning is a remnant of a bygone age of the "city beautiful," must be replaced by a new design concept based on the creative synthesis of complex plans using all the tools provided by modern technology.

NOTES

[1] Traffic Research Corporation, *Review of Existing Land Use Forecasting Techniques*, presented to the Boston Regional Planning Project (Toronto, 1963).

[2] Kenneth J. Schlager, "Simulation Models in Urban and Regional Planning," *Technical Record*, Southeastern Wisconsin Regional Planning Commission, Waukesha, Wisconsin, II, No. 1 (1964).

[3] Christopher Alexander, *Notes on the Synthesis of Form* (Cambridge: Harvard University Press, 1964), pp. 15-16.

[4] *Ibid.*, pp. 74-75.

[5] Richard E. Bellman and Stuart E. Dreyfus, *Applied Dynamic Programming* (Princeton: Princeton University Press, 1962).

[6] Schlager, *op. cit.*

A GROWTH ALLOCATION MODEL FOR THE BOSTON REGION

Donald M. Hill

The model described in this article is based on extremely simple concepts and represents the first case of a complete model covering all aspects of urban location. Basically, the future growth of each of a number of activities in a number of subareas of the Boston Metropolitan Region is projected in short steps, with each step depending upon the results of previous steps. Thus, in concept the development of each activity in each subarea influences future development by way of competition for land and changed accessibilities. The initial analysis used in establishing the parameters for the model is based on multiple correlation, and shows extremely close correspondence between the estimates generated by the model and actual historical events. When the model is used as a projection device, it is particularly interesting to note that current trends do not continue, but are changed by later events. The decline of the central city is checked, as is the rate of growth of the inner suburban ring. This model is particularly flexible in that it can accommodate the influences of a very large number of variables. It may be contrasted with the Pittsburgh model as discussed in Steger's article in that it contains very little explicit analysis of locational behavior and depends primarily on a blanket interpretation of past events.

[illegible]

Urban and regional planners are developing sensitive and systematic methods for evaluating the functional effectiveness of alternate development and transportation plans. As a result of the systematic approach to planning, mathematical models are rapidly gaining in prominence as highly useful tools of analysis and prediction. This is particularly true in transportation planning where the application of mathematical models has gained widespread acceptance as a technique for analyzing and predicting person movements or traffic flows in a large number of urban areas.[1]

In recent years, increasing attention has been given to developing systematic methods in the form of mathematical models to predict patterns of urban and regional population, employment and related variables.[2] This work has been undertaken in part to avoid basing sophisticated traffic models on inputs of land use data which are produced largely by trend-projection methods modified, if possible, in accordance with expectations that urban development is influenced by transportation and other public policies. Trend projections, modified according to expected impact of transportation and other public policies, are not easily reproducible except by the same planning team. Further, the actual preparation of these projections must span many calendar months. The objective of a mathematical model is to forecast patterns of urban or regional development systematically and efficiently. The model described in this article can be operated recursively in series with a traffic model to predict the development of a region and its traffic in either short 5-to-10 year or long 15 to 20 year time intervals.

While the potential usefulness of a land use model is evident, the difficulties of developing it are formidable. Some planners argue that regional development cannot be predicted meaningfully because the process of development includes certain major decisions that cannot be determined in advance: for example, decisions about the location of regional shopping centers and government installations. Others maintain that the many small decisions which influence regional development—for example, family decisions on household locations—are not based entirely on rational considerations. Nevertheless, one may argue that development studied at a regional scale appears to demonstrate stable relationships, and that the number of individual decisions made during a time interval of 5 or 10 years is large enough to provide a solid basis for statistical appraisal of these relationships and for their inclusion in a mathematical model to describe development. Only through extensive appraisal and testing of mathematical models can we answer many of the questions raised here.

Encouraging evidence which demonstrates that a land use model can be developed is described in the remainder of this article. In May, 1963, a project of land use model development was initiated on behalf of the Boston Regional Planning Project. The model is to be used to develop a comprehensive development plan for the Greater Boston Region, an area of approximately [illegible] square miles, housing 3,400,000 people in 152 cities and towns.[3]

The essence of the model's development is the scientific method comprising theoretical reasoning and empirical testing. The rationale of the model is premised on the following observations:

(1) that the locational preferences of population, employment, and other activities are highly interrelated;

(2) that the development of land for various uses is influenced by numerous exogenous causal factors.

A model in this context is a mathematical technique capable of predicting detailed subregional development patterns of urban activities based on existing patterns of urban development, on externally forecast regional growth, and on exogenous policy considerations concerning transportation, open space, zoning policies, public utilities, and regional growth.

The model described below is comprehensive in that it locates all types of activities in a simultaneous fashion. Further, the model can be applied in a time-recursive manner to predict development in a step-wise fashion over several consecutive 5 or 10-year intervals. This is in contrast with models or techniques

Donald M. Hill is a staff member of Traffic Research Corporation in Toronto serving currently as Project Director of TRC for consulting services to Metropolitan Toronto and Region Transportation Study. He has contributed to TRC studies in Boston, Washington, and Edmonton. He holds a bachelor's degree in mathematics and physics and a master's degree in statistics from the University of Toronto.

which produce a direct forecast for a target date 20 or 30 years in the future with no intervening forecasts. A forecast model which disregards the immediately previous order of development may predict sudden and perhaps oscillating changes in land use which are not reflected in real life, or may tend to reproduce straight line trends of historical periods, or both.

NATURE OF GROWTH ALLOCATION MODEL

The function of the model is to distribute or allocate externally-supplied growth forecasts of the various activities among the subregions comprising the study region.

The underlying concept of the model is that the development patterns of urban activities are interrelated in a systematic manner which provides a reasonable basis for their prediction. The model provides the formal mathematical mechanism for evaluating the extent of interrelations between activities. The only restriction imposed by the model is that the interrelationships must be expressed so that the influences of variables (activities) are additive. Accordingly, the model incorporates a linear form. Any desired combination or transformation of variables may be introduced to describe the urban activities whose locational pattern we wish to measure and predict.

To describe the model, it is convenient to define a number of quantities as follows:

The region is divided into a number of small areas called *subregions* (see Figure 1).

The purpose of the model is to predict the amounts of several urban activities in each subregion at the end of a given forecast period. These activities are called *located variables*, signifying that the task of the model is to allocate given regional totals of these variables ~~at the end of the~~ forecast period to the subregions comprising the region.

It has been found that the locations and intensities of several urban activities are related to development patterns of one or more variables in a causal manner. That is, the presence or absence of these activities in a subregion, or the ease of access to them from a subregion, may be said to influence the amounts of one or more located variables in each subregion. These influencing activities are called *locator variables*.

The concept of the model may be stated as follows:

The change in the subregional share of a located variable in each subregion is proportional to: *one*, the change in the subregional share of all other located variables in the subregion; *two*, the change in the subregional share of a number of locator variables in the subregion; and *three*, the value of the subregional shares of other locator variables. This concept is expressed by the following equation system

$$\Delta R_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} a_{ij}\Delta R_j + \sum_{k=1}^{M} b_{ik}(Z_k \text{ or } \Delta Z_k) \qquad (1)$$

where: i or $j = 1, 2, \ldots i, \ldots N$ number of the located variable (a total of N equations).

$k = 1, 2, \ldots, k, \ldots M$ number of the locator variables

$\Delta R_{i \text{ or } j}$ = change in the level of the ith or jth located variable over the calibration or forecast time interval

Z_k = level of locator variable k at the beginning of the calibration or forecast time interval

ΔZ_k = change in the level of the kth locator variable over the calibration or forecast time interval

a_{ij}, b_{ik} = coefficients expressing the interrelationships among variables

There is one such equation for each located variable. The a and b coefficients are determined by simultaneous regression analysis of the data from two past points in time (that is, the model is calibrated on the basis of a past period of time).[3]

After the coefficients are determined the equations are used to estimate future subregional values of each located variable by substituting in each equation the pertinent values of the locator variables for that subregion and solving the equations simultaneously for the subregional located variables. To obtain the forecast in absolute values rather than relative values, the subregional shares at the end of the forecast interval are multiplied by the externally forecast control figure for the total of each located variable in the study region.

DEVELOPMENT OF MODEL Development of the model required detailed analysis of "cause-and-effect" relationships between development patterns of all land use categories, as well as detailed analyses of the independence and interdependence of locational groupings of urban activities at the subregional level. Four methods of study were followed in order to achieve the necessary understanding of urban development:

A graphic analysis of a large number of relationships between subregional growth rates of population and employment, and a large number of causal (locator) variables which were thought to be important.

An investigation of the changes and trends in population and employment that took place during the 1950–60 decade.

here

An analysis of population and employment data (using factor analysis techniques) to determine to what extent variables can be grouped according to their tendency to locate in proximity to one another, or to exhibit similarities in their influence on the locational tendencies of other variables.

within the overall [illegible]

The development of multiple regression equations which demonstrate the relationships between growth rates of population and employment and a large number of variables that have significant effects upon locations of population and employment.

The findings of these studies provided the insight necessary to formulate the model for use as a prediction tool.

The equation system (1) comprises one mathematical relationship for each activity (located variable) to be located. The particular set of activities selected for inclusion in the model were:

1. White-collar population (PW). This is the resident population which participates in the white-collar labor force (workers and dependents).[5]

2. Blue collar population (PB). This is the resident population which participates in the blue-collar labor force.

3. Retail plus wholesale employment (RW), covered by the Massachusetts Division of Employment Security.

4. Manufacturing employment (M) covered by the Massachusetts Division of Employment Security.

5. All other employment (OE), including employment not covered by Massachusetts Division of Employment Security.

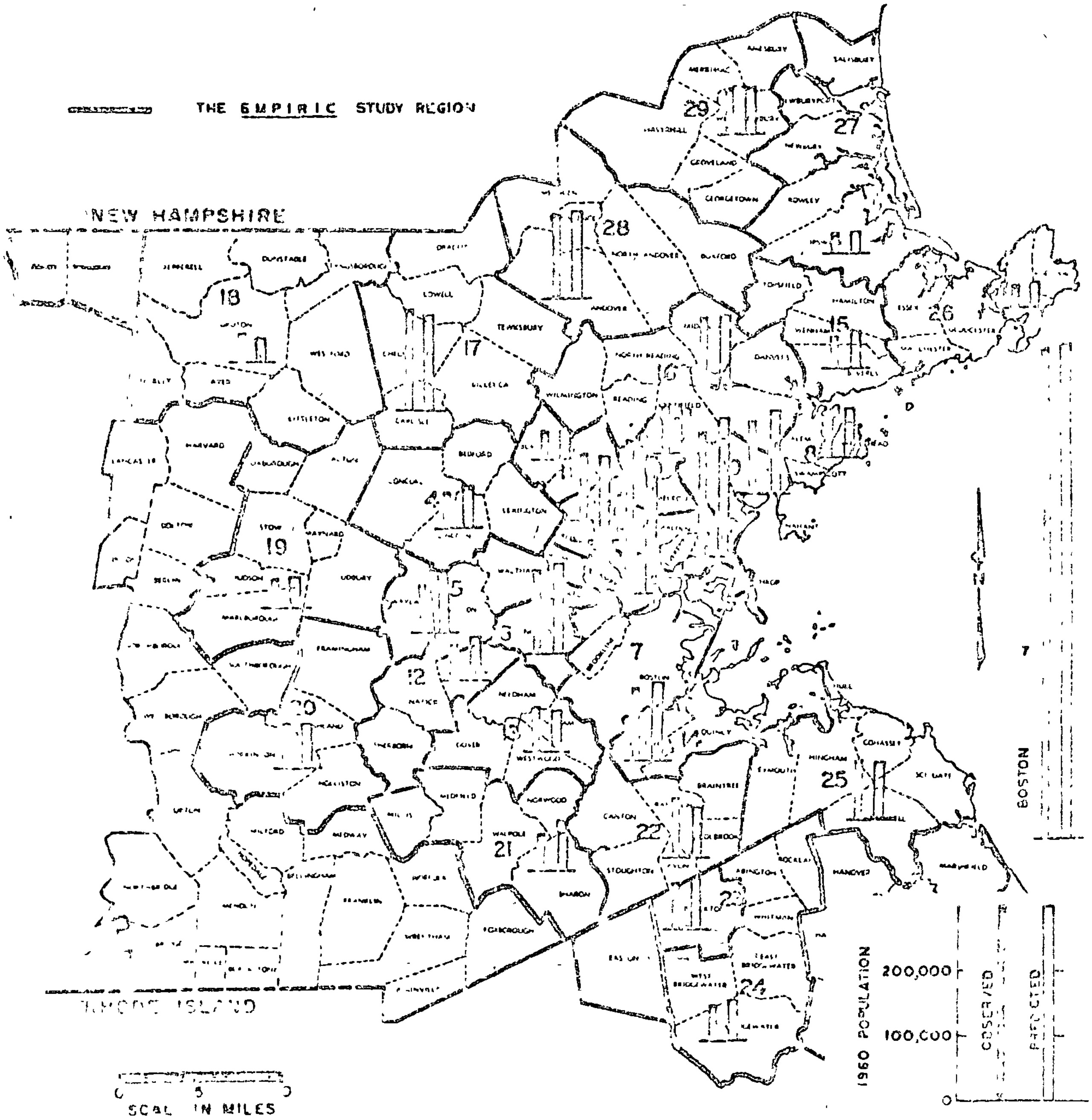
The breakdown of population and employment into 2 and 3 categories, respectively, was based on special investigations using component analysis techniques.[6] The findings demonstrated that two socio-economic classes of population showed distinct residential preferences for different subregions. The division of total employment isolates special locational preferences of firms included in each of the three categories, while at the same time allowing the three categories to sum to total employment.

The equation system (1) is a flexible formulation which permits the inclusion of many different types of exogenous or locator variables. This flexibility made it possible to calibrate the model with three different sets of locator variables:

SET I Intensities of land use (densities)[7] Zoning practices[8] Automobile accessibilities[7] (within the region)
SET II Same as Set I plus Transit accessibilities[9]
SET III Same as Set II plus Quality of water service[10] Quality of sewage disposal service

CALIBRATION OF MODEL Calibration tests were carried out by applying the model to forecast 1960 subregional development from the base information of 1950. The coefficients of the equation system (1) were established by means of regression analysis which deal with simultaneous equations. The objective of the analysis was to determine a set of coefficients to enable the model to reproduce the 1960 development patterns with minimum error.

IGURE 1 *Comparison of Observed 1960 Population Levels with Those Predicted by the* MPIRIC *Land Use Model (Application II)*



The model was calibrated for each of the three sets of locator variables described earlier. The conditions of the calibration are described below for the second set of locator variables (referred to as Application II):

(1) The region was divided into 29 subregions as shown in Figure 1. A large share of the population and employment were contained in the central Boston area in 1950 and 1960 (subregions 1, 2, and 7)

(2) The base date for projecting 1960 land uses was 1950. Subregional data for population and employment for 1950 were used with the following input variables to predict 1960 subregional values of population and employment

(2.1) 1950 densities of two classes of population and three classes of employment.

(2.2) Zoning practices during 1950 to 1960.

(2.3) Changes in the highway transportation system between 1950 and 1960.

(2.4) Changes in the mass transportation system between 1950 and 1960.

(3) The located variables were the two classes of population (white-collar and blue-collar population) and three classes of employment: manufacturing, retail plus wholesale, and all other employment. These classes of population and employment were then aggregated to form the additional output variables: total population and total employment.

Figure 1 presents the observed and projected 1960 subregional values of population. The accuracy with which the model simulated subregional development in the Boston Region in the decade 1950 to 1960 is shown by the close correspondence of the subregional population presented for each subregion in Figure 1. Similar results were obtained with employment. The model was able to forecast both growth and decline throughout the region, including the loss of approximately 100,000 people and 30,000 jobs in the City of Boston. Accordingly, the model appears to deal satisfactorily with both situations of decline in the older cities and of growth in the new suburban cities and towns.[11]

Three numerical indices were calculated to measure the aggregate correspondence attained between observed and predicted 1960 subregional values. These indices summarize the accuracy obtained over all subregions, and result in a single accuracy measure for each output variable. The indices used are the Root Mean Square Error (RMS), the RMS Error Ratio, and the Coefficient of Determination (R^2).[14] The summary measures calculated using the subregional values presented in Figure 1 and others are contained in Table 1.

The results of the calibration demonstrate that the test model explains a high percentage of the change throughout the 1950–60 decade in the Greater Boston Region. The low RMS errors illustrate the small chance of making erroneous planning decisions on the basis of a prediction using the model. An example is that an error of 3,000 resident population is approximately equivalent to less than 600 person-trips during a peak rush hour. These 600 person trips could be accommodated easily by one lane of a minor collector city street.

The results of calibrations with the two additional sets of locator variables (Applications I and III) illustrate the effect upon the model's forecast accuracy of decreasing or increasing the number and types of input variables. The conditions under which the two additional applications were made are the same as those given for the previous Application II, except for numbers and types of input variables.

As described previously, Model Application I deletes transit accessibilities as input variables from those used with Model Application II, while Model Ap-

TABLE 1 *Summary Reliability Measures*

Model Application	*Output Variable*	*RMS Error (people or jobs)*	*RMS Error Ratio*	*R^2*
II	*Total Population*	3439	0.03	0.99
	White collar	3031	0.06	0.99
	Blue collar	1686	0.03	0.99
	Total Employment	1512	0.03	0.99
	Retail & wholesale	564	0.06	0.99
	Manufacturing	1128	0.08	0.99
	Other	955	0.05	0.99
I	*Population*	7825	0.07	0.99
	Employment	3631	0.08	0.99
III	*Population*	1259	0.01	0.99
	Employment	1306	0.03	0.99

AIP

plication III adds water supply and sewage disposal services as input variables to those used in Model Application II. The effect of this decrease and increase in the number of input variables is reflected in the three summary measures of the accuracy of 1960 subregional values predicted with Model Applications I and III (See Table 1). The results show that when increased information is available on transportation systems and other planned variables, predictions are more accurate. It would appear that the more comprehensively we are able to plan in advance the various modes of transportation and other policy measures, the more accurately we will be able to predict the future locations of urban activities.

VALIDATION OF MODEL

Several tests were conducted to validate the model as a useful forecast tool. One test involved investigating the effect on forecast accuracy of varying the length of the historical forecast period. Other tests involved measuring the sensitivity of model forecast values to the lengths of future forecast periods and the size and number of subregions.

EFFECT OF VARYING THE HISTORICAL FORECAST PERIOD

Each application of the model was tested with data for the five-year period, 1950-1955. The values of the locator variables were calculated for 1950 and all highway, transit, water, sewage, and zoning changes for the period 1950-55 were accounted for by the appropriate locator variables. The 1955 population and employment, by category, were forecast from the 1950 base, using the model applications calibrated for the 1950-60 period, as described above.

The resulting RMS Errors, RMS Error Ratios, and Coefficients of Determination are shown in Table 2. A comparison of these summary measures with the measures of Table 1 indicates that the model forecasts for the 1950–55 period with similar accuracy as for the 1950–60 period. The results of Application I for the 1950–55 period are more accurate than for 1950–60; the results of Applications II and III are not as accurate for 1950-1955 as they are for 1950–1960.

EFFECT OF VARYING LENGTH OF FORECAST PERIODS

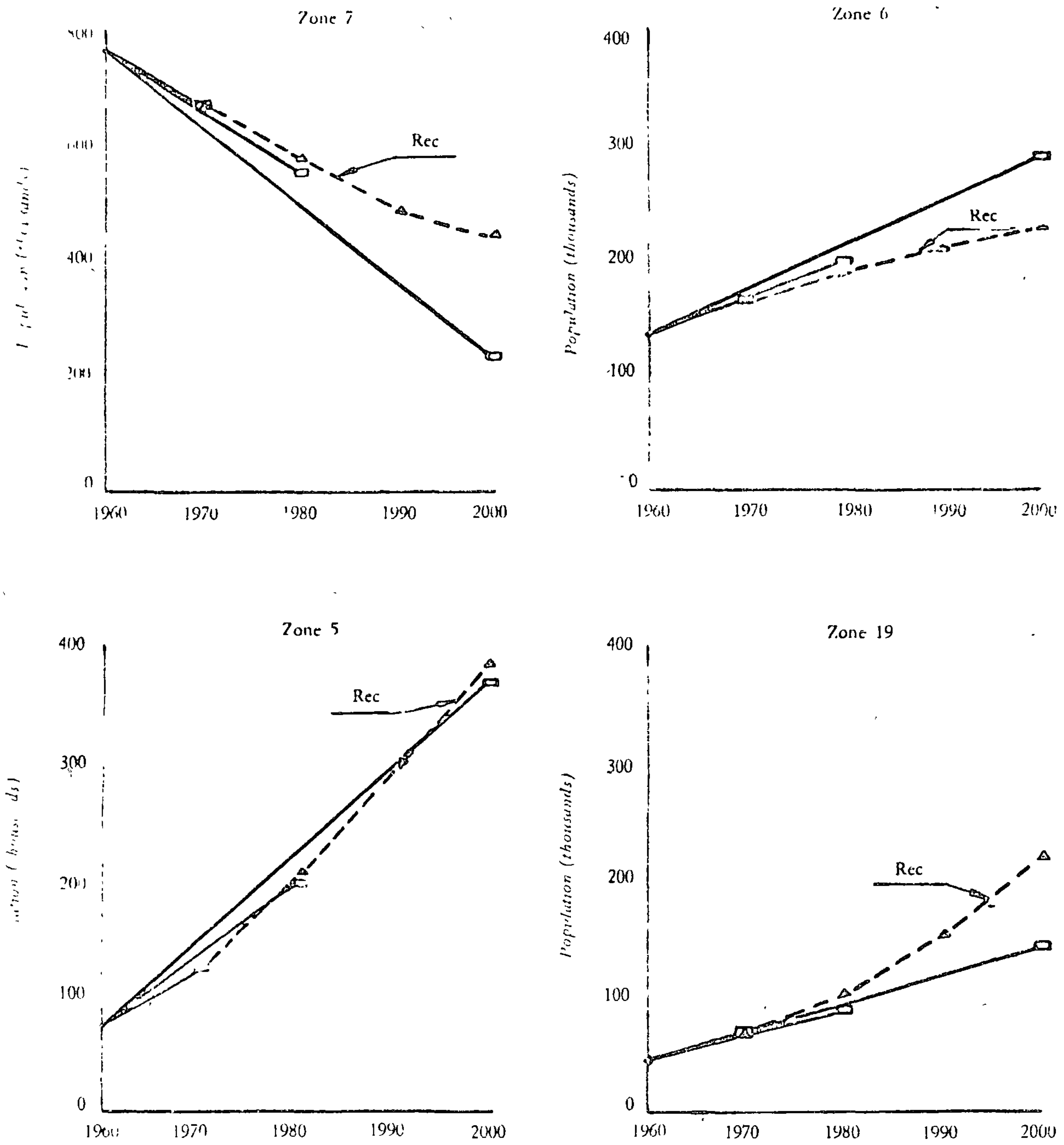
With any forecast model, errors generally increase as the length of the prediction period increases. The further into the future one tries to forecast, the greater are the predicted subregional growths (or declines) as a percentage of the original subregional activity levels. Hence, even if the percentage errors in the projected subregional growth remained constant, the percentage errors in the projected subregional values would increase.

Three single projections were made with model Application I: from 1960 to 1970, 1960 to 1980, and 1960 to 2000. The population results for selected subregions are illustrated in Figure 2. In addition to the single (non-recursive) tests, recursive ten-year projections were conducted with Application I. In recursive forecasts, the model is applied sequentially for relatively short forecasts, with the results of each forecast serving as a base for the succeeding forecast. At the end of each ten-year projection period, new accessibilities and densities were calculated, based on the same 1960 transportation network, but using the

TABLE 2 *Summary Measures for All Model Applications of 1950–55 Forecast*

Model Application	*Output Variable*	*RMS Error*	*RMS Error Ratio*	*R²*
II	*Total Population*	2243	0.02	0.99
	White collar	Not available—not census year		
	Blue collar			
	Total Employment	3171	0.08	0.99
	Retail & wholesale	649	0.07	0.99
	Manufacturing	2478	0.18	0.98
	Other	1128	0.06	0.99
I	*Total Population*	3425	0.03	0.99
	Total Employment	3134	0.08	0.99
III	*Total Population*	2297	0.02	0.99
	Total Employment	3120	0.07	0.99

Figure 2. Population Projected to 1970, 1980, and 2000 in Single Projections and Population Projected Recursively in Ten Year Intervals to 1970, 1980, 1990, and 2000



newly predicted subregional activity levels. (Information on future transportation networks was not available for this test.) The findings are presented in Figure 2, and they demonstrate a curvilinear trend.

Recursive projections are expected to predict subregional growth and declines more accurately than a single forecast. Recursive forecasts are more sensitive to changes in policy measures, land utilization, and accessibility characteristics, which serve as projection inputs more frequently. Accordingly, the divergence of recursive projections from a linear trend probably reflects future development rates more closely, and may be expected to exhibit a generally lower forecast error than single non-recursive projections.

TABLE 3 *Error Measures Obtained with 29, 123, and 134 Subregion Projections 1950 to 1960*

Number of Subregions	*Activity*	*RMS Error*	*RMS Error Ratio*	*R²*
29 Subregions	*Population*	1259	.01	0.99
	Employment	1306	.03	0.99
123 Subregions	*Population*	5673	.22	0.99
	Employment	3715	.36	0.99
134 Subregions	*Population*	3479	.14	0.99
	Employment	2755	.29	0.99

EFFECT OF VARYING SIZE AND NUMBER OF SUBREGIONS

In an attempt to measure the model's accuracy with an increase in the number of subregions and a consequent decrease in subregion size, model Application III was calibrated with data for the 123 cities and towns in the study region. In addition, the model was calibrated with data for the 123 cities and towns, but with the City of Boston divided into 12 subregions to give a total of 134 subregions.

Table 3 presents values of RMS Errors and RMS Error Ratios for total population and total employment. The RMS Errors appear to increase both absolutely and relatively as subregions become smaller. Also, with the more uniform distribution of subregional activities in the case of the 134-subregion system, the forecast accuracy is greater than the 123-town test. However, it remains less accurate than the 29-subregion tests.

The decreasing accuracy of the model with the increased number of subregions results from a number of factors. It should be noted that the set of locator variables was not altered in accordance with the smaller subregional size, although the values were calculated for the smaller subregions. It is to be expected that new sets of variables which reflect locational decisions on a finer areal basis would reduce forecast errors with sets of smaller subregions. Clearly, more tests will be required with new variable sets to explain adequately the effects on model forecast accuracy of increases in the number of subregions

CONCLUSION

The research, development, calibration, and test work has been successful and the model is practical and operational. The test model has achieved a level of accuracy that compares favorably with the accuracy obtainable from the best of present-day traffic forecasting models.

Results so far suggest that the model may enable the planner to simulate a chain of events in urban development starting with the variables under public control (such as the transportation system and zoning regulations) and ending with an efficient and desirable pattern of residential and industrial development.

Work is continuing on the development and application of land use prediction models by Traffic Research Corporation on behalf of the Boston Regional Planning Project. A fully developed model is to be used by the Planning Project in conjunction with travel forecasting techniques to evaluate several alternative design policies of transportation systems in future years for the Boston Region. This model will permit forecasts to be made for larger numbers of activities (10 and more) and for 600 subregions.

Author's Note During the course of this study and preparation of this paper cooperation and assistance were received from many quarters In particular I wish to acknowledge the generous assistance given by Richard S. Bolan, Walter B. Hansen, Hans G. von Cube, Neal A. Irwin, Daniel B. Brand, and Karl H. Dieter

NOTES

[1] W L Mertz, *Review and Evaluation of Electronic Computer Traffic Assignment Programs*, Highway Research Board Bulletin No 297 (1961), B V Martin, F W Memmott, A J Bone, *Principles and Techniques for Predicting Future Demand for Urban Area Transportation* (Cambridge Massachusetts Institute of Technology, Department of Civil Engineering, Research Report No 38, 1961)

[2] *Review of Existing Land Use Forecasting Techniques*, prepared by Traffic Research Corporation for the Boston Regional Planning Project (July, 1963)

[3] R S Bolan, W B Hansen, N A Irwin, K H Dieter, "Planning Applications of a Simulation Model," Paper presented to the New England Section of the Regional Science Association, October, 1963

[4] A description of simultaneous regression analysis for systems of linear equations is presented in John Johnston, *Econometric Methods* (New York - McGraw Hill, 1963) For the purpose of the pilot studies reported here, the system of simultaneous equations comprising the model was assumed to be identified exactly Accordingly, the regression technique of indirect least squares was applied In production models to be developed subsequently, the assumption of exact identification will be relaxed Consequently, other estimation methods will be applied two-stage least squares, limited information single equation method, or simultaneous least squares

[5] White-collar labor force included professional, technical, managerial, clerical, and sales workers Blue-collar labor force included craftsmen, foremen, operatives, service workers, laborers, and occupation not reported.

[6] Component analysis is described in Harry H Harman, *Modern Factor Analysis* (Chicago University of Chicago Press, 1960)

[7] The density of activities in this instance is calculated by dividing the appropriate subregional activity levels by the subregional effective area, that is, the area suitable for development by any of the located variables The areas selected comprised the land in use in 1960 for residence, commerce, manufacturing, and wholesaling, plus the agricultural or vacant land which was either suitable for development (having less than 15 percent slope and not swampy) or which was under development for commercial or industrial uses in 1960 Data availability was a prime factor in the selection of this definition of effective area

[8] Zoning practices were measured in terms of the net densities at which development occurred during the calibration period (1950-1960)

[9] Measurement of a subregion's accessibility to a given urban activity is formulated as the sum of the activities in all subregions multiplied by an inverse function of the travel times separating the subregions (the gravity model formulation) Measurement of accessibility was similar to the procedure outlined by Walter G. Hansen in, "How Accessibility Shapes Land Use," *Journal of the American Institute of Planners*, XXV (May, 1959), 73-76

[10] The 1950 and 1960 ratings of quality of water supply and of sewage disposal service were based on the following scale

1 Metropolitan District Commission (MDC) system
1 5 Partial MDC System
2 Supplied by community
3 Not publicly supplied

When a combination of the above four possibilities occurred for a subregion, intermediate values to the 1, 1/5, 2, or 3 rating were calculated, weighted by zonal areas

[11] This finding contrasts with the assumption applied in developing a mathematical model for the Baltimore Region "The contexts of the model process in the center of the City of Baltimore and the rest of the region are assumed to be different and are described by separate models" See T R Lakshmanan, 'A Model for the Distribution of Urban Activities Formulation, Evaluation and Reformulation, Paper presented at the Seminar on Models of Land Use Development Institute for Urban Studies, University of Pennsylvania, October, 1964

[12] Root Mean Square Frror is a summary measure which expresses the deviation from corresponding observed values of the predicted values produced for each located variable Statistical theory indicates that for about 67 percent of the subregions, the observed value will not differ from the predicted value by more than plus or minus the RMS Error. The RMS Error Ratio is the ratio of the RMS Error, defined above, to the arithmetic average of the observed activity values

Coefficient of Determination (R^2) is a third summary measure for each located variable which represents the proportion of the sum of squared deviations of the observed subregional located variable levels from the mean subregional level, that is explained by the model

---

A MODEL FOR SIMULATING RESIDENTIAL DEVELOPMENT

F. Stuart Chapin, Jr

This article is the outgrowth of a long period of experimentation with models of residential and other development which focus on the individual decision-maker—the householder in search of a home, and the builder attempting to cater to his desires. The first part of the paper describes the means by which this process is simulated in a computer. Later parts of the paper indicate how this view of consumer behavior has led naturally into a consideration of households' activity patterns and into research on these patterns. In conclusion the author indicates by implication that the integration of these two approaches may lead in the direction of measurements of accessibility and of taking account of the competition for desirable sites, of consequent increases in the intensity of utilization, and of the rents of these sites Thus this article, in conjunction with others, illustrates a remarkable convergence between research starting at the level of individual decision units and coming to consider the larger framework, and research starting from gross changes in the metropolitan configuration and coming to give increasing consideration to the actual decision process.

B. H.

To develop an operational model of the residential development process poses some complex problems. This is due in part to the variety of agents involved

D.B.R. REPRINT No. 483

Commonwealth of Australia
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANIZATION
Reprinted from "Management Science Theory"

MANAGEMENT SCIENCE
Vol. 16, No 3, November, 1969
Printed in U S A

A General Planning Model*

A general econometric model for allocation of a set of activities to a set of spaces is proposed. The activities may vary in type and in magnitude A_i, and the zones may vary in shape and location and in capacity Z_j and may be filled with a single activity, a mix of activities or only partially filled. The objective function may include the benefits and costs of establishing and operating an activity i in a space j, and the benefits and costs of interactions between activities along paths between spaces.

If any portion a_{ij} of any activity i may be allocated to any zone j, the planning problem may be posed: Select the set $\{a_{ij}\}$ to maximize a measure of merit $U(a_{ij})$ where

$$U(a_{ij}) = \sum_{i}^{N}\sum_{j}^{M} a_{ij}\, b_{ij}\, A_i + \sum_{i}^{N}\sum_{j}^{M}\sum_{k}^{N}\sum_{l}^{M} a_{ij}\, a_{kl}\, b_{ijkl}\, S_{ik}\, R_{jl}$$

subject to

$$\sum_{i}^{N} a_{ij}\, A_i \leqq Z_j, \qquad \sum_{j}^{M} a_{ij} = 1, \quad a_{ij} \geqq 0$$

in which

S_{ik}	is the interaction level between activities i and k,
R_{jl}	is the length of the path between zones j and l, by any preallocated route
b_{ij}	is the sum of the benefits less the costs of establishing and operating one unit of activity i in zone j
b_{ijkl}	is the sum of the benefits less the costs of establishing and operating one unit of interaction S_{ik} along one unit of path R_{jl}
N	is the number of activities
M	is the number of zones.

The problem is a quadratic programming problem with NM independent variables a_{ij} and $N + M$ linear constraints. Various solution techniques are available, and further linearized techniques have been developed. The model optimally allocates activities to regions on the basis of the comprehensive cost benefit function above. The model has been extended to include an additional dimension, time, allowing optimal times of commencing or changing of activity levels to be also determined.

Potential applications include engineering and economic planning at various levels from national, regional and urban development to industrial and building layouts and chemical or electronic systems design. Scheduling problems may also be handled. Current applications are to urban planning, hospital layouts, and building design.

In a city, for example, the activities include residential living of various densities, industrial, recreational, service and commercial activities; the zones include the various land areas suitable for establishment of these activities; the interactions include flows of people, materials, energy, and information; and the paths include highways, residential streets and pedestrian paths, rail, pipe and wire networks. The benefits and costs may be as comprehensive as assessment capabilities allow. Non quantifiable benefits are handled also using sensitivity analyses.

---

* Received August 1969

Papers describing these extensions, solution techniques and planning applications are being prepared, and further applications are being developed.

J. F. BROTCHIE
Division of Building Research
CSIRO, Melbourne, Australia

14 A Land-Use Plan Design Model

14.1 Introduction

This model has been prepared by Kenneth J Schlager[1] for the Land-Use Transportation Study of the Southeastern Wisconsin Regional Planning Commission in 1965, and has been tested using data from the Waukesha area, Schlager has called his model a 'Plan Design' model, which could be identified as a decision model, according to the classification used here.

The process of evaluation of the different alternative programmes is done endogenously by the model itself using mathematical programming techniques The object is to minimize total public and private investment costs subject to certain constraints Only major land uses such as agriculture, industry, residential, open space, etc are handled by this model

14.2 Underlying concepts

The underlying concepts of this model have their basis in the analytical studies of Christopher Alexander[2] who indicates that difficulties in the design process derive primarily from the inability of the human designer to manipulate 'simultaneously a large number of interacting design relationships Schlager[1] remarks:

> While Alexander does not pursue the direct solution of selection problems by means of mathematical techniques, his definition provides useful criteria for the systematic formulation of such problems.

In relation to the following conditions connected to a 'selection problem' pointed out by Alexander,[2]

(1) 'It must be possible to generate a wide enough range of possible alternative solutions symbolically.'
(2) 'It must be possible to express all criteria for solution in terms of the same symbolism.'

Schlager indicates that the first requirement is naturally achieved because all land-use plans may be expressed symbolically by three sets of variables,

(1) The type of land use (quality variables)
(2) The density of land use (quantity variables)
(3) The geographic location (location variables)

The land area under study is subdivided into a grid of 'zones' of equal area The geographic co-ordinates of the zone determine the location variable For each zone, the types and densities of land uses may be expressed as a measure of the activities in that zonal area

The requirements of the second condition, which deals with the symbolic relationship between alternative forms and design requirements, are divided into two primary classes

(1) Requirements that restrict the numerical value of a land use or a relationship between land uses *within* a grid zone Examples of these requirements are the exclusion of flood plain areas from development in a given zone grid, or prevention of the simultaneous development of both industrial and residential land in the same grid zone
(2) Requirements that restrict a relationship between land uses *between* grid zones, for example, the need to provide an elementary school within a specified distance (or time) of all residential units

Then, given these design requirements in terms of restrictions on possible land-use relationships, and a set of demands or total needs (based on previously prepared forecasts for the area under study), the problem is to (Schlager[1]).

> Synthesize a land use plan design that satisfies both the land use demands and design standards considering the current state of both natural and man-made land characteristics, at a minimal combination of public and private costs

He tries to emphasize that minimal cost does not necessarily mean a cheap plan, but only to avoid unnecessary expenditure of precious resources

14.3 The linear programming formulation

Schlager only gives an outline of the mathematical formulation of his model, but in the following analysis a more complete description of the linear programming will be attempted based on the information provided. The original notation will be modified and expanded in order to fulfil this intention

The following notation will be used

m = number of zones of equal area which form an exhaustive subdivision of the land area under study; zones are indicated by the subscripts $i = 1, 2, \ldots m$

n = number of land uses categories such as single family residential, multi-family residential, industrial, agricultural, etc considered by the designer, Land uses categories are indicated by the subscripts $j = 1, 2, \ldots n$

x_{ij} = number of acres of zone i to be allocated to land category j

c_{ij} = cost of developing an acre of zone i for allocation of land use category j; cost may be related to the topography and soil characteristics of the area

E_j = total demand for land use category j

d_j = service ratio coefficient which provides for supporting service land requirements, such as streets, which are necessary for development of land use category j

F_i = limit in the amount of area from zone i that can be allocated to land uses

G_{jk} = ratio of land use category j allowed relative to land use category k, with land uses categories j and k in the *same* zone

G_k^j = ratio of land uses categories j allowed relative to land use category k with land uses categories j and k in *different* zones.

The mathematical structure of this model can be identified as a 'transportation problem' which is solved using linear programming techniques The linear programming formulation has the following form

Minimize
$$Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}x_{ij} \tag{14.1}$$

subject to
$$\sum_{i=1}^{m} d_j x_{ij} = E_j \qquad (j = 1, 2, \ldots n) \tag{14.2}$$

$$\sum_{j=1}^{n} x_{ij} \leqslant F_i \qquad (i = 1, 2, \ldots m) \tag{14.3}$$

$$x_{ij} \leqslant G_{jk} x_{ik} \qquad (i = 1, 2, \ldots m) \quad (j = 1, 2, \ldots k, \ldots n) \tag{14.4}$$

$$x_{ij} \leqslant G_k^j x_{hk} \qquad (i = 1, 2, \ldots h, \ldots m) \quad (j = 1, 2, \ldots k, \ldots n) \tag{14.5}$$

The objective function (14.1) relates to the cost of developing land for a given land use category Constraint equations 14 2 ensure that the total demand requirement for each land-use category is satisfied by the model, that is, the allocations add up to the forecast total Because only primary land uses such as residential, industrial, agricultural and recreational land are directly determined, service ratios incorporated in the *d* parameters account for secondary land uses as streets and parks

Constraint equations 14 3 limit the use of land within a zone, it will usually reflect a density standard Constraint equations 14 4 and 14 5 indicate restrictions on coexistent land uses within and between zones This type of constraint will also reflect accessibility standards for employment and shopping areas The introduction of more constraints will depend on the characteristics of the study, and of course, will have the effect of restricting the set of possible solutions Constraints can be added as long as they are linear and continuous

14.4 Inputs

Four main classes of input data are required

(1) Costs of unimproved land and land development by land use and type of soil
(2) Forecast total demand for each land use
(3) Design standards (e g densities) and inter- and intra-zonal limitations on land-use relationships
(4) The current land inventory including land uses and soil types

Land development cost data obtained by statistical analysis of recent land development generally are very expensive and difficult to obtain Engineering cost estimates are usually preferable if competent professional experience is available In the Waukesha area, separate cost estimates were prepared for each of three classes of soil Assessed and equalized land value data on the basis of prices realized in recent land transactions were obtained for each of the zones It is possible to consider redevelopment in the form of urban renewal, and for this approach, redevelopment cost would be required.

The aggregate demand for each primary land use activity is input to the model and may be obtained by applying design standards to forecasts of population and employment in the area under study.

The design standards reflect the plan objectives and restrict the set of acceptable plans, also they limit the land use relationships that may exist in the plan Basically they are the densities assigned to each zone, the service ratios for the amount of secondary land (streets, parks) required to support the primary land uses, and the ratios of coexistent land uses within and between zones

An inventory of both current land use activities and soil characteristics is one of the most important inputs The soil inventory makes it possible to assign a development cost to each zone in the study area

14.5 Limitations of the linear programming

Schlager introduces his model using linear programming techniques and points out two major limitations

(1) Because land-use choices are by nature usually discrete, land use variables should take discrete values rather than continuous ones Linear programming solution could be rounded off to satisfy these natural discrete levels, but such a solution does not usually correspond with the associated discrete optimal combination
(2) The need for both a linear objective function and linear constraints is the second limitation of linear programming

The use of a linear objective function as an approximation to non-linear costs is not a severe one, because the inaccuracies introduced by this appr are usually less than the errors of cost estimation. But linear constraint relationships provide an unsatisfactory substitute to design standards which are generally non-linear If these design standards are not represented properly, the model loses most of its usefulness. In order to remove these limitations, the author of this model suggests the possible application of dynamic programming as a computational procedure to be used in the model

14 6 Comments

Although many comments have been presented along the analysis of this model, two major points will be emphasized here

The first one is related to the conceptual basis of minimizing costs Cost can be an interesting factor to be employed in the objective function, if it is representative of all the elements that shape the urban spatial structure; but in the particular way that this model handles these elements, it is doubtful whether the representation required is manifested There is little doubt that development cost is important, but to consider it as the basic factor which (under certain restrictions) defines an optimal allocation of different land uses, shows an opposite viewpoint from the comprehensive and systematic approach to planning advocated here

The second point refers to an apparent absence of the dual problem Many interesting features about relevant ways of introducing changes into the land-use plan could be obtained by the use and interpretation of the dual variables, which apparently are not considered in the framework of the model. Changes in the land-use plan can probably appear during the implementation stage, where dual variables could give a deeper insight in the nature of the changes required

New models have been developed which attempt to overcome these deficiences. One is the model[3] for the Irvine New Town in Britain, which was carried out by the Israel Institute of Urban Studies and P A. Management Consultants Ltd. in 1968. The purpose of the model was to maximize the value of the welfare goal function (objective function) within the given constraints, by the efficient allocation of population and dwelling units to the different zones and development stages of the plan.

The Irvine New Town model recognizes that there are criteria which are not measurable in financial terms, and thus the optimal plan cannot be easily obtained. Therefore, several alternative directives are specified for different levels of these criteria and for each alternative, and the best possible solution is found which is regarded as efficient. From the list of alternative efficient plans, that plan which best suits the planner's criteria is chosen 'The selection of this plan is not simple, because the plans differ from one another, not only in the value of their goal function but also in other criteria which are unmeasurable This process of selection requires the deep consideration not only of planners, but also most importantly of public authorities' [3]

The construction of the welfare goal function included the following criteria for the residential areas

The market value of housing
Construction costs of housing
Site development costs of residential areas
The alternative agricultural value of the land
Land value premia due to environmental conditions
The value of the time that the population will spend in travelling for all purposes, etc

The Irvine New Town model makes use of the dual variables, and thus it is possible to indicate by how much the total value of the goal or objective function increases as a result of marginal relaxation of the constraints Since some of the constraints are subject to the decision of the public authorities, the information given by the dual variables allows the authorities to re examine their former decisions and modify them.

Another very interesting experience in England which used linear programming techniques, was started towards the end of 1969 by the Greater London Council in order to assist the overall planning of Greater London [4] In this model, the objective function is more safely regarded as an expression of one of the important factors which influence development It is never considered the sole important factor. Apart from those factors included in the constraints, other important factors, such as the utilization of vacant land, the density of housing, the total cost, the availability of land for redevelopment, etc., were also included in the model by means of objective functions It should be stressed that, although many objective functions are contained in the model, only one can be operative at any time.

Young[4] points out

> One use the model does *not* have is that of calculating what *will happen* in the future Its uses are, in fact, all concerned with calculating what *needs* to happen if objectives are to be met.

The reader will find very easy to follow the description of the GLC model on pages 5–15 of Reference 4.

References

1 SCHLAGER, K J. 'A land use plan design model' *Journal of the American Institute of Planners*, 1965, May, pp. 103–111.
2 ALEXANDER, C. *Notes on the Synthesis of Form*. Cambridge Harvard University Press, 1964
3 P.A MANAGEMENT CONSULTANTS LTD. and ISRAEL INSTITUTE OF URBAN STUDIES LTD 'Irvine New Town planning study' 1969, September
4 YOUNG, W. 'Planning a linear programming model' *GLC Intelligence Unit Quarterly Bulletin*, 1972, June, no 19

Models of Urban Land-Use Development

Leslie J. King

This paper reviews the general problem of constructing models for the forecasting of urban land-use development. In particular, the different analytical approaches adopted in constructing such models are discussed and the extent to which computer facilities were utilized is noted. This general review brings into focus several of the basic issues involved in modeling complex urban systems and prompts some questions as to the overall utility of social planning and engineering.

Particular emphasis is placed on those land-use studies and models that are essentially positive in their approach – that is, those that deal mainly with explanations of what has happened in urban land-use patterns and with the prediction of future developments consistent with these explanations or understandings. But some more challenging and provocative normative analyses – that is, studies of what land-use patterns should or could be consistent with certain objectives, constraints, and standards – have been proposed, for example, in the work on land-use design being pursued by the Southeastern Wisconsin Regional Planning Commission (Southeastern Wisconsin Regional Planning Commission, 1968). Much of this work is still only in the development phase, but again, it prompts an interesting set of questions concerning the philosophy and strategy of social planning. Some of these questions, for example, concerning the nature of the objective functions to be optimized and the value system implied, will be considered at a later point in this paper.

Land-Use Forecasting

The past efforts at forecasting urban land-use development appear to have been prompted by at least two major considerations:

(1) It has been established analytically in some contexts (for example, trip-generation studies and financial analyses) and suspected for many others (for example, the overall level of social welfare in the city) that the pattern and nature of land uses are significant predictor variables.

(2) The recognition that land-use patterns in any urban area are influenced, in turn, by many complex and interacting forces that are dynamic both over

time and space. Within the city, for example, particular areas are specialized in certain land-use activities, but these functions may change over time as competing land uses wax and wane in their fortunes.

Given these considerations, it is perhaps not surprising to find that most of the major urban land-use studies have been set within the context of transportation planning programs. It is really only in these types of programs that serious consideration has been given in the past to systems analysis of the urban complex. This emphasis upon land-use and transportation systems will undoubtedly continue, but it is also possible that land-use forecasting will assume additional future importance in the context of planning and shaping the form of the city (the relevance to zoning controls is obvious) and in predicting future requirement levels for municipal and other services.

The "typical" approach in land-use modeling is to assume first that certain factors or values are given (these are the so-called exogenous factors in some of the models), the model then generates certain levels of employment, and/or economic activity, and/or population by use of mathematical functions or statistical estimators. These derived levels are then allocated spatially to different areas of the city, again by the use of particular allocation functions, and finally, the future land uses in each area are derived by the use of land-requirement functions for the different allocated variables.

In very crude form then, the problem involves the three components shown in Figure 1-1. There are the data inputs which involve the exogenously determined values and the data to be used in estimating the different parameters of the model. (This latter phase is often referred to as the calibration of the model.) Then there is the model which performs the generation and allocation of the variables of interest. As discussed later, this model may be of many different forms, perhaps a set of rules and procedures for a game simulation on the one hand, or a set of complex mathematical functions on the other. Finally, there are the outputs which are the predicted levels of development for each land-use activity and each subregion of the city.

Obviously, the above comments grossly oversimplify both the modelling process and the nature of the land-use development models. The published literature, in fact, provides ample evidence of different ways of viewing these components. Britton Harris (1968), for example, on the basis of extensive experience in this field of research, emphasizes the following six major dimensions of land-use modelling:

(1) *descriptive* or empirically based statements versus *analytic* or deductive formulations,
(2) *holistic* approaches which attempt to deal with the total urban environment versus *partial* analyses focusing on certain selected aspects of the urban complex,

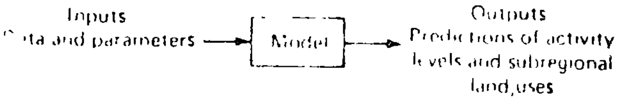


Figure 1-1 Land-Use Modeling Process.

(3) *macro* or aggregative models versus *micro* and behavioral models,
(4) *static* versus *dynamic* formulations,
(5) *deterministic* models which do not allow for chance elements in the land-use process, versus *probabilistic* models which attempt to deal with such elements, and
(6) *simultaneous* versus *sequential* solutions.

Harris presents an excellent review of the different models of commercial, residential, and industrial land uses within the framework of his discussion of the dimensions of modeling.

Kilbridge, et al. (1969) have also recently examined the dimensions of urban planning models. Their review is more general than that of Harris, and they stress (1) the *subject* of the model, (2) its *function*, (3) the *theory* on which the model is based, and (4) the *method* by which the model uses the theory.

Again, Wilson (1968), in discussing models of urban planning, constructs a "hierarchical relevance tree" setting out the tasks of such models and their interrelationships. He recognizes three broad levels of policy, design, and understanding. Within these categories, the following eight tasks are identified:

Policy	(1) Action, (2) Goals, (3) Evaluation
Design	(4) Plan formulation, (5) Design techniques, (6) Problem formulation
Understanding	(7) System models, (8) Techniques

Lowry (1968), in discussing seven of the more highly developed land-use models, deals with a broader theoretical question, namely, the extent to which the different models relate to a generally accepted theory of the market mechanism for urban land. Lowry's contention is that "the market processes of transactions between willing buyers and willing sellers determine the spatial organization of urban activities." He defines an investment function by means of which owners appraise the merits of site improvements, and an evaluation function relating to the price a particular establishment will be willing to pay for a certain site.

Lowry's subsequent review of the seven models considers the extent to which these functions are treated – either implicitly or explicitly – in the models. Lee (1969) outlines an excellent discussion of the conceptual frameworks, the techniques, the model constructions, and the particular applications which have characterized urban land-use analyses to date. In the same vein, Lamb (1967) has reviewed some of different land use models with emphasis upon the analytical frameworks employed in them.

The reviews by the authors mentioned above are illuminating and provocative. But, with the exception of the ones by Lee and Lamb, they deal essentially with broader issues than those concerning this paper. The concern here is mainly with the methods and forms of land-use modeling, with some related features of computer utilization, and with selected problems associated with these approaches.

The discussion in this paper is strongly influenced by the comments of Geisler et al. (1962) on the continuum of systems-analysis techniques. They note that increasing abstraction in analysis involves moving from left to right along the scale shown in Figure 1-2.

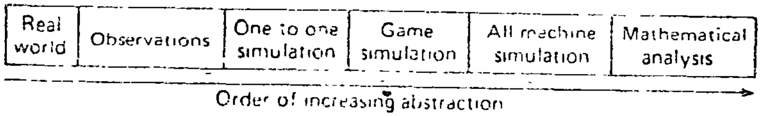


Figure 1-2 Spectrum of systems analysis.

Most of these levels are apparent in the literature of urban land-use studies. At the lower level, there have been numerous qualitative statements based wholly on empirical observations. For example, the idealized statements of city form known as the concentric zone theory (Burgess, 1925), the sector or wedge hypothesis (Hoyt, 1939), and the multiple nuclei theory (Harris, 1945) are all empirically based and together represent fairly low levels of abstraction.

In Figure 1-2 there are three levels involving some form of *simulation*, and some different applications of these analyses to urban land-use studies are reviewed in the following section. So-called "simulation studies" include a variety of different approaches and levels of abstraction. At one extreme are "hand simulations" such as those performed by the Northeastern Illinois Planning Commission (1968). In that study, the patterns of Chicago's urban growth from 1965 to 1990 were forecast using five different models — the "dispersed regional city design," the "finger design," the "multi-cluster design," the "satellite cities-greenbelt design," and the "current trends development design." Each model was considered in turn and its implications were noted and evaluated by the researchers. The models were essentially qualitative statements of urban form, and no computer systems analysis was involved. By contrast, the simulation model developed by Crecine (1964) is far toward the other end of the systems-analysis

continuum. In this study, "simulation" implies a particular approach using the computer to obtain the solution for a set of mathematical equations.

In this paper, hand simulations are ignored in regard to their content, and only those simulations involving the use of computers, in one way or another, are reviewed.

One-to-one simulations (the second level of abstraction in Figure 1-2) do not appear very relevant in urban land-use studies. In such simulations, reality is replicated, but at a different scale. The Link trainer used in flight training, or the wind tunnel testing of models, are essentially one-to-one simulations. Obviously, analogous models of the city would be prohibitive in cost even if they were useful in research and training, which is most doubtful. Urban situations are typically competitive ones, often involving conflict and stress, and reflecting a complex range of decisions made by both the private and public sectors. Hence, game simulations appear more appropriate, and three attempts at developing such simulations are reviewed in the following section.

Game simulations involve inputs from teams of players who act out certain decision-making roles within the constraints of a set of rules. By contrast, all-machine simulations involve samplings from probability functions, and the sampled values are the inputs to the models. Some of these probabilistic approaches to land-use modeling are also reviewed in the next section.

At the highest level of abstraction are the mathematical models for which the solutions are derived analytically. In land-use modeling there have been different forms of mathematical models employed. Some have involved sets of simultaneous linear equations, other linear programming formulations, and still others, nonlinear differential equations and recursive programming. These efforts are also considered in the following section.

The importance of computer facilities varies from level to level. In game simulations, a computer typically serves as the central banker or information bank and monitors the operation of the game. In the City I game (Washington Center for Metropolitan Studies, 1968), for example, an IBM 1130 is involved. The development of software for these game simulations is costly, therefore, there are high computer expenses involved in the running of the game. All-machine simulations depend even more heavily upon computer facilities, and in the case of models such as that developed at the University of North Carolina, there is considerable expense involved in both program adaptation and machine runs.

The mathematical analyses, on the other hand, involve computers in a less intimate manner than do the simulation models. Problems of statistical estimation of parameters and solution of equations are easily programmed for the computer, and the time and costs involved usually are not as great as with the simulation studies. These comparisons are noted in more detail in the following review of specific models.

Models of Urban Land-Use Development

The major structural features of different land-use models are briefly reviewed below within the framework suggested by Figure 1-2

Game Simulations

By definition, a simulation is an attempt to replicate or to create a likeness of reality But, since the effort is directed toward an understanding of the real-world system in question there is always some simplification of reality involved Orcutt (1963, p 22) has noted that "simulation of a social system involves building and operating a model designed to represent those features of the system which are deemed to be significant in view of the objectives behind the simulation "

With regard to game simulations in particular Geisler, et al (1962) note that their purpose is "the study of decision rules in the context of a given organization and environment " From the point of view of urban land-use modeling, this means an emphasis on the decision-making processes which determine the city's land-use pattern The complexity of designing such game simulations for urban problems is obvious, for one thing, urban land-use patterns are in part determined by public sector decisions (for example, in highway planning, education, municipal services and zoning) and also by private sector decisions (for example, in the location and development of businesses and residences and in travel patterns) Further, the intersectoral relationships – for example, between the public and private sectors, between employers and workers, and between productive activities are exceedingly complex and hence difficult to untangle and to allow for in the game simulation. Notwithstanding the conceptual problems, the design problems, and the high development and operating costs of game simulations, some notable progress has been made along these lines by at least three groups These are discussed below Kibel (1970) has provided a much more detailed and informative review of the topic of urban gaming

Cornell Land Use Game (CLUG) The efforts at game simulation by the Cornell planning department have been directed by Dr Alan G Feldt (1965) Essentially, the model allows for the development of an urban community, beginning with an open area in which the land is owned only by the game operator The players begin with small amounts of capital and they may use this to bid for land and to develop the properties they purchase The model allows for residential commercial and industrial development some simple input-output relations concerning payments and purchases among these sectors and for property assessments to pay for the cost of municipal services A number of variations on the basic game are possible (Feldt, 1966)

Feldt (1965) notes several significant points with respect to this game simulation

(1) It is "fundamentally a communications device intended for an educational milieu whether that be in a formal classroom or in actual planning practice " The CLUG model is already widely known in both educational and professional circles in this country

(2) The game generally allows for the development of a city of approximately 250,000 population over 25 or so rounds This takes about 15 to 20 hours, including instruction time But, beyond this the operation of the game becomes too awkward in terms of the amount of information to be processed As reported to date, this game simulation has not made much use of computers the CITY I game reported on below is in many respects a computerized version of CLUG

(3) An attempt to modify the model to fit a real-world situation, specifically Syracuse, New York revealed among other things that "the scale utilized in the elementary version of CLUG was too gross to allow for detailed handlin of patterns of intraurban land-use development " For example, the basic residential unit in the early model consisted of 1000 employed persons and around 4000 total population

No recent literature on this model seems available, presumably, work on the project is continuing

METRO Urban Game Simulation This game simulation is directed by Dr Richard Duke and has been developed with support from Michigan State University, the University of Michigan, and the Tri-County Regional Planning Commission (Duke, 1964) The game simulation is a fairly complex one and is based on some earlier efforts in this field by the same group. Their *METROPOLIS I* was a hand simulation which was subsequently programmed for a computer as *METROPOLIS II* All three models are tailored to the particulars of the urban community of Lansing Mich , and data for this city are used to provide many of the parameter estimates.

Some brief comments are made here on three features of the game simulation namely the player roles the types of decisions which have to be made and the specific simulation models which are involved

The players each have two roles one as a member of a government team (either central city or suburb or urbanizing township) and the other is a membe of a professional association team (politicians planners land developers and educators)

Decisions have to be made by each team concerning budgets issues and policies These decisions are important inputs to the simulation models of response, macroeconomic and demographic growth and population and economic-firm redistribution.

The game simulation thereby involves sequences of activities and interactions both between the players themselves and between the players and the computer.

Duke (1966) stresses that the aims of the METRO project are to

(1) simulate growth patterns that would occur naturally, and enable their comparison with planned growth patterns,
(2) illustrate the kinds of information which are available to decision makers,
(3) similarly inform decision makers about the analytical techniques and models that are available for evaluating and implementing decisions, and
(4) to provide information on the implications for urban development of alternative action programs

Environmetrics The original CITY I model developed by this group under the direction of Dr Peter House (Washington Center for Metropolitan Studies, 1968) was a computerized version which incorporated elements of both the CLUG and METROPOLIS models The game simulation was programmed for an IBM 1130 computer Typically, it involved nine teams of four to five players each, and there were two main groups in the game – the public and the private sectors. The former controlled by a "bureaucracy" comprising an elected mayor and two councillors and the department of zoning highways, education, public works and safety, and finance Eight of the teams were thereby represented on this bureaucracy, while the ninth remaining team functioned as the "mass media" The other players on the eight teams and the "mass media" team functioned as the private sector

The public sector had to work out a budget and plan for different public-sector activities – for example, providing fire and police services and building highways and schools In addition, they acted on requests from the private sector concerning zoning changes etc

The private sector could purchase and develop properties as they chose A number of industrial, commercial, and residential developments were possible, and each of these involved certain development costs service and transportation costs taxes, and income A set of intersectoral relations were specified in the model

As in the METRO project, the computer provided projections of population and employment levels, and subregional allocations of these variables

There were a number of more subtle possibilities in the model, including provisions for borrowing money from the central computer for renovating or demolishing properties, and for variations in pricing policies associated with service establishments

The successor to the City I game developed by House and his colleagues is the City II model which incorporates further refinements A regional model has also been developed

These game simulations are not strictly land-use forecasting techniques By their very nature, they do not involve a set of mathematical solutions and depending to some extent upon the inclinations of the players, they may or may not emphasize questions concerning the land-use pattern of the city They are all nonzero-sum games, which means that the gains by any one team do not have to be balanced by the losses of another, and certainly in the case of CITY I game it is possible for the players to emphasize the role-playing activities without any realistic concern for the spatial patterning of the urban complex But the pedagogical values of these game simulations for policy-makers, planners and educators cannot be stressed enough They give a player a "feel" for the complexity of relationships and decision-making processes which dictate the land-use pattern of the city Also, it would seem quite possible to use these game simulations under fairly well-controlled rules to consider the implications of certain alternative policies and programs Duke (1966) has noted this possibility with regard to the METRO project.

Machine Simulations

Simulations which focus on what are assumed to be random or chance elements in the determination of urban land-use patterns are considered here These models involve variables whose behaviors are best described by probability distributions Random samplings are made by the computer to simulate these probability distributions, and the solutions to the models are inferred from the behavior of these random numbers It is because these models involve the use of randomly chosen numbers that they have been called "Monte Carlo" methods

There are three qualifications which should be made at this point The first is a technical one and need not concern us further It is that Monte Carlo methods are much more general and powerful than their application in simulating probabilistic processes (which is the role described here) might suggest. Hammersley and Handscomb (1964) make this point very clear Second, in urban land-use modeling a number of other machine simulations have been proposed that are not properly classified as Monte Carlo models In fact, B Harris (1961) has noted four conditions, any one or combinations of which might justify the use of simulation techniques:

(1) a mathematical solution is impossible because too many variables are involved
(2) the relationships between variables may not be simple linear ones,
(3) the model is dynamic and the important lags are long ones, and
(4) the processes involved are stochastic (probabilistic)

There are studies by Crecine (1964), Ellis (196[illegible]), Schlager (1965), the Arthur D. Little Company (1966), Forrester (1969), and the Center for Real

Estate and Urban Economics (Berkeley) (1968) that all use machine simulation techniques especially in solving sets of difference equations to obtain land-use and related forecasts. Some of these models are described in the section below on mathematical analyses.

The third qualification is that Monte Carlo models represent an essentially low-level approach to the study of probabilistic systems, and again there are other land-use studies, notably by Curtis Harris (1968) and the Arthur D. Little Company (1966), that develop more formal stochastic models of land-use processes. These also will be mentioned later.

In the context of systems analysis of urban land-use, Monte Carlo models are associated especially with the University of North Carolina planning group (Donnelly, et al., 1964). Their model simulates the conversion of open space into residential land use. A map of the study area is divided into a large number of square grid cells and for each cell available for development an "attractiveness measure" is computed. A given number of new residential units have to be located in each prediction period, and these are assigned to the different cells by the use of random numbers and probability functions which reflect the distribution of attractiveness measures.

This North Carolina model is highly aggregative (emphasizing development), and it is restricted to only residential land use. Carlson (1968) has reviewed the operational aspects of the model and reported on the results of a questionnaire survey of 175 planning agencies in the United States concerning its use. Problems loomed large concerning a lack of staff programmers capable of adapting the model and the inability or unwillingness to collect the data required in computing the attractiveness measures. Carlson, however, is optimistic as to the utility of the model, and he provides some useful guidelines for its adaptation.

A Monte Carlo approach to predicting future land-use changes, which is similar in its design to the North Carolina work, has been used by Morrill in studying the expansion of the residential urban fringe (Morrill, 1965) and the negro ghetto (Morrill, 1965).

Mathematical Analyses

Models employing sets of mathematical equations that are solved to yield forecasts of economic activity and land-use requirements for different subregions of the city are considered below. The types of mathematical equations vary from model to model and the methods of obtaining solutions to these equations may involve mathematical analysis or some form of simulation.

The classical approach to mathematical prediction, at least in the physical sciences, has been to structure a set of differential equations. Such equations express the rate of change in particular variables as functions of the changes in other variables and they can be solved to yield values for the system variables for any points in time in the future (or in the past). In the social sciences, differential equations have been used in studying phenomena such as population growth and changes in different economic variables and it is perhaps not surprising that attempts have been made to structure differential equations for the prediction of urban land-use levels.

The best documented attempt along these lines was the work done on the POLIMETRIC model in Boston. This model which has since been discarded by the Boston group has been described by Irwin and Brand (1965).

"Basically, it is comprised of a series of nonlinear differential equations of the following form:

$$\frac{dR_{i\ell}}{dt} = f\left[R_{i\ell}\left(\sum_{p}^{L} M_{p\ell} \sum_{p=1}^{L} M'_{\ell p}\right)\right].$$

where

i = number of the located variable ($i = 1, 2, \ldots, i, \ldots, N$)
ℓ = number of the subregion ($\ell = 1, 2, \ldots, \ell, p, \ldots, L$).

Stated in words the rate of change over time t of activity i in subregion ℓ is a function of the present level of the activity in subregion ℓ, ($R_{i\ell}$), plus a function of all movements of the activity i from all other subregions p into subregions ℓ,

$$\left(\sum_{p=1}^{L} M'_{\ell p}\right)$$

minus all movements of that activity i from the subregion ℓ out to all other subregions

$$\left(\sum_{p=1}^{L} M'_{p\ell}\right)$$

These in- and out-movements are called in-migrations and out-migrations.

This model appears to have been discarded largely because of problems inherent in the data requirements (Irwin and Brand, 1965). The matrix of in- out-migrations which is suggested by the model presumably would involve extremely difficult estimation problems. The model subsequently was modified by the Delaware Valley Planning Commission (Seidman, 1969) as the basis of their residential and manufacturing location submodels, but in these versions the differential equation form was not retained.

Aside from its use of differential equations to predict changes over time, the POLIMETRIC model was somewhat distinctive in its attempts to handle nonlinear relationships. By contrast, most of the other well-known land-use models deal with linear relationships, often in the form of multiple-regression equations. The EMPIRIC model developed for the Boston area illustrates this point. This model distinguishes between certain output or *located variables* (specifically, white-collar and blue-collar population, retail and wholesale employment, manufacturing employment, all other employment, total resident population, total employment) that are to be predicted for each subregion in the city and the predictor or *locator variables* (namely, intensities of land use, automobile and transit accessibilities, quality of water and sewage-disposal systems). The model is based on the notion that "the change in the subregional share of located variable *i* in subregion ℓ is proportional to (1) the change in the subregional share of all other located variables in subregion ℓ, (2) the change in the subregional share of a number of locator variables in subregion ℓ, and (3) the absolute value of the subregional shares of other locator variables." (Irwin and Brand, 1965)

In equation form, the model is

$$\Delta R_i = \sum_{\substack{j=1 \\ j \neq i}}^{M} a_{ij}\, \Delta R_j + \sum_{k=1}^{M} b_{ik}\, (Z_k \text{ or } \Delta Z_k),$$

where

ΔR = change in located variable ($i, j = 1, 2, \ldots, N$)
Z_k = value of kth locator variable at start of forecast period
ΔZ = change in locator variable

and a_{ij} and b_{ik} are coefficients estimated from data for 1950 and 1960.

There is one such equation for each located variable, and "the equations are used to estimate future subregional shares of each located variable by substituting into each equation the pertinent values of the locator variables for the subregion and solving the equations simultaneously for the subregional located variables" (Hill, Brand, and Hanson, 1965). These shares are then converted into absolute levels through multiplication by the exogenously determined control levels for each of the located variables.

The EMPIRIC model is operational and is currently used for projection and analysis in Boston; it will likely be applied further in other cities.

Another well-known model involving sets of equations is the one

developed by Lowry (1964). This model uses an iterative procedure to forecast the spatial distribution of population and employment in a city given an exogenously determined level of basic employment, that is, employment in "export" industries which "are relatively unconstrained in local site selection by problems of access to local markets."

The Lowry model consists of nine structural equations which generate retail employment and number of households in the city, allocate these totals among the subregions of the city by use of functions in which accessibility indices appear, and compute the amount of land required for retail establishments in each subregion.

The Lowry model was developed originally for Pittsburgh, and subsequently it appears to have been adapted to the needs of other groups, particularly the Bay Area Simulation Study (Center for Real Estate and Urban Economics, 1968 Garin and Rogers (1966) have discussed possible alternative formulations of the model in matrix algebra terms, while Crecine (1964) has developed a time-oriented version of the model. Crecine's work apparently was prompted by dissatisfaction with three of the characteristics of the Lowry model. First, the original model is a static-equilibrium one and assumes that, in any particular forecast period, all retail establishments and households can move. Second, the households in the model are not differentiated by type, and finally, the model relates to a "region" rather than to the particular boundaries of a city. In Crecine's TOMM model, therefore, only a portion of the establishments and households can move in any time period, households are differentiated by income, housing, and social characteristics, and city census tracts are used as the areal units for forecasting. Many of the equations in the model now become difference equations relating the levels at one time period to those of the previous time periods.

There are other models involving sets of equations which might be cited. The Activities Allocation model developed by the Delaware Valley Regional Planning Commission (Seidman, 1964), for example, actually involves a set of seven submodels which are run sequentially for five-year recursion periods for the nine county Philadelphia region.

In all of the above models, whether they are linear or nonlinear in form and whether they are solved iteratively to arrive at an equilibrium situation or recursively to project the amount of change occurring in the future, there is no explicit consideration of an overall objective function and related constraints and the types of normative solutions which these features would suggest. There hav been attempts, however, to structure normative models of urban land use by way of mathematical programming techniques.

Perhaps the best known of these attempts is the Herbert-Stevens (1960) model. This involves a linear program in which households are distributed spatially so as to maximize the aggregate "rent-paying ability." The objective function for the model is

$$\max Z = \sum_k \sum_i \sum_h x_{ih}^k \left(b_{ih} - c_{ih}^k\right),$$

where

Z = aggregate rent-paying ability
k = subscript for regions
i = subscript for household groups or types
h = subscript for residential bundles or packages of characteristics
x = solution variable for number of households
b = the residential budget (including transportation)
c = annual residential cost, exclusive of site cost.

The solution was subject to three constraints:

$$\sum_i \sum_h s_{ih} x_{ih}^k \leqslant L^k \tag{1}$$

$$\sum_k \sum_h x_{ih}^k = N_i \tag{2}$$

$$\text{all } x_{ih}^k \geqslant 0, \tag{3}$$

where

s = site area
L = land area available for residential use
N = exogenous number of households to be located

As in any linear programming solution where the primal problem involves maximization, there is a dual problem of minimization. In this case, it is total aggregate costs which are minimized.

The model, then, provides for the maximization of the sum of the budget

residuals available for land rent. But, in fact, this solution proved elusive, and subsequent modifications of the model sought only to maximize the "bid rent," defined as "a budget residual covering the entire residential package of site and structure (but not the cost of transportation)" (Lowry, 1968). Because of the detailed data requirements of this model, few planning agencies have applied it in practice (Hemmens, 1968).

Linear programming also has been used in the Southeastern Wisconsin Regional Planning Commission (SEWRPC) Land-Use Simulation Model (Southeastern Wisconsin Regional Planning Commission, 1966). In the residential and industrial sectors of this model, linear programs are formulated that provide for the minimization of land-development costs. In addition, the residential land-development process over time is handled by way of recursive linear programming, in which the solution of a program for one time period provides the parameters for the succeeding linear program (Schlager, 1966). Housing demand is thus allowed to build up over time.

In its early work on the land-use design model (as distinct from the land-use and economic simulation studies), the SEWRPC group contemplated the use of linear and dynamic programming. However, these efforts have given way to a model based on an alternative form of mathematical analysis, linear graph theory. (Southeastern Wisconsin Regional Planning Commission, 1968)

For completeness, one approach to the mathematical modeling of urban land-use systems is mentioned briefly. This involves the formulation of stochastic models, that is, models of probabilistic processes operating over time. As part of the San Francisco CRP study, a Markov chain analysis of the deterioration of housing units was undertaken (Wolfe, 1967). The states of the model were different levels of housing quality, and there were transition probabilities for the movements from one state to another. Deterioration was an absorbing state, and the behavior of the system over time with respect to this state could be studied.

Harris (1968) has outlined a more general stochastic model for residential development. A parcel of land may be developed or undeveloped, and given m parcels of land, there are thus 2^m states of development for the whole area. The model is semi-Markovian in the sense that there is a waiting time, in itself a random variate, associated with the move from one state to another. More recently, Bourne (1969) has suggested the use of a transition probability matrix in conjunction with regression equations as a means of allocating land-use development. For each subarea of the city, regression analyses yield estimates of the levels of new construction. Also derived for each subarea is a matrix of transition probabilities describing the change over time from one land-use type to another. The regression estimates are combined with these probabilities to predict the future land-use structure of the subarea.

Characteristics of Land-Use Models

The history of the attempts to forecast urban land development is comparatively short, most of the efforts having been made in the current decade. Most of the models reviewed above, then, are essentially first- or second-generation models and the technical features and shortcomings of the models are not yet documented in much detail. But there are five characteristics which can be noted, starting with computer-system requirements.

Computer-System Requirements. A recent survey by Hemmens (1968) has provided some valuable information on this point. As part of a questionnaire survey of large planning agencies, the extent and level of computer use was probed. Hemmens reports that

> "Most of the agencies which reported their computer usage utilize more than one computer system. Typically, they use a small computer which is operated by the agency itself or by another public agency, and they rent time on a large computer from a service bureau or other vendor."

The reliance upon service bureaus is more pronounced in the case of those agencies heavily committed to land-use forecasting by use of the mathematical models discussed above. The Bay Area Transportation Commission models are to operate on a CDC 3800 computer with 65K memory, the Delaware Valley RPC models and the Boston EMPIRIC model require IBM 7094 or equivalent capability, the SEWRPC land-use simulation was originally programmed only for an IBM 1620, but the later models are designed for the IBM 360 system.

Given the complexity of the simulations which may be involved and/or the large number of multiple-regression equations which may have to be solved, it is clear that the use of sophisticated land-use forecasting models will require access to computers of at least the IBM 7090 series level and, increasingly, of the 360 series level. Even allowing for access to such computer facilities, the operating costs of these models are high. The metropolis model, for example, requires more than 15 minutes of IBM 7090 time per time period, while the activities allocation model of the Delaware RPC requires as much as 50 minutes of IBM 7094 time for one run (Lamb, 1967).

Aggregation Problems. In most of the urban land-use studies there are at least three forms of aggregation problems which have to be resolved.

The first is the level of spatial aggregation, which has to do with the number of subregions comprising the city for which forecasts are to be made. Lowry in his empirical analysis of the Pittsburgh data used 650 one-square-mile tracts; the EMPIRIC model, on the other hand, was tested originally for 29 subregions representing the Boston region and later for 123 and 134 subregions. In the context of estimating parameters and analyzing relationships, the decision as to the level of spatial aggregation is not unimportant. It is clear, for example, that parameter estimates for one level of aggregation will not be applicable at another and consequently, many computations will have to be repeated if forecasts are required for these different levels. Again, relationships which hold at one level may not be as significant at another. This was apparent in some of the sensitivity analyses on the EMPIRIC model (Irwin and Brand, 1965). These kinds of considerations, which in one sense are problems of spatial filtering, are discussed in the spatial-statistics literature (King, 1968), and in the statements on ecological fallacies in theory construction (Goodman, 1959). Fleet and Robertson (1968) have at least drawn attention to the issues in the context of tri-generation studies.

A second problem of aggregation has to do with the number of variables employed in the model. Most of the models are highly aggregative in this respect, the EMPIRIC model, for example, deals with only seven located variables and the Lowry model deals with even fewer. Obviously, the utility of the models would be enhanced by greater detail in the number of forecasted variables, but problems of data availability loom large in this regard. In speaking of the variables used in land-use and urban development studies, there is also another important aspect to the problem of aggregation. This has to do with the macro-level on which the studies focus, and the fact that questions of individual behavior are ignored. The need for disaggregation along these lines will be discussed in greater detail in the subsequent section on "underlying theory."

Finally, there is aggregation over time with respect to the length of the forecast periods. Most of the models employ five- or ten-year periods, which appear satisfactory in view of the typical goal for most planning agencies of developing a master plan for some future date near the turn of this century. Again, disaggregation with respect to this feature would pose serious problems in regard to data for parameter estimation and forecasting.

Data Availability. The urban development models, for the most part, are particularly demanding regarding data requirements. Carlson (1968) in his survey related to the potential use of the North Carolina model, reported that only 9 of the 135 agencies responding had already collected all the necessary data for the model, while 23 percent of the respondents felt that the data requirements were excessive. The main difficulty in all the models is that the data have to be available for the areal subregions of the city, and there must also exist some historical data for the city, to obtain estimates of the model's parameters (that is, to calibrate the model). The latter point emphasizes again one of spatial-aggregation problems, namely, that the model parameters typically are estimated from historical data for the city as a whole and are then used in obtaining forecasts for very detailed subregions of the city.[1]

Alonso (1968) has drawn attention to some important problems concerning data quality He cautions against measurement error and the compounding of such errors in modeling situation, and questions whether, in urban land-use studies the models might not have outrun the capacity of the data Alonso suggests that if this is the case, then the quality of the data might have resulted in a deterioration of the predictions Alonso offers the following "rules of thumb" for model building

(1) avoid intercorrelated variables,
(2) add where possible,
(3) if addition is not possible, multiply or divide,
(4) avoid as far as possible, taking differences or raising variables to powers, and
(5) avoid as far as possible models which proceed by chains

Underlying Theory Some comments on the theoretical bases of urban land-development models are in order, because as Lowry (1968) notes, "in choosing a model for a particular purpose, the planner will do well to understand what is left out as well as what is left in."

For the most part, the models reviewed in this report have dealt with macro-level variables and relationships Indeed, with but few exceptions, the models have been structured along the lines of macroeconomic models, with the emphasis being placed typically upon problems of estimation and forecasting, rather than upon questions related to the underlying theory This is true, for example, of the EMPIRIC, the BASS, and the Lowry models Not surprising, Lowry (1968) can find little explicit consideration of a theory of the urban land market in his review of seven well-known models

It is becoming clear, however, that these aggregative statements are not enough and that increased attention must be given to the theory of individual behavior and the nature of the decision-making processes B Harris, et al (1968) in summarizing the conclusions of the Dartmouth conference notes that, "there is a strong but not unanimous feeling among model-builders that one direction for improving the accuracy with which models reproduce the real world lies in the expansion of studies of the behavior of decision units " Once this possibility is admitted however, the problems of data availability are increased many times By the same token, it is well established that the use of aggregative spatial data does not allow for meaningful statements to be made about individual behavior and that if the new direction in modeling is to be pursued then individual survey data tabulated by subregions and cross classified must be obtained It is important to note that considerable progess is being made in the direction of developing behavioral models of urban spatial structure which should have important implications for the modeling of urban land use The papers by Cox and Kaiser in Part II for example, are illustrative of these efforts

Lowry's (1968) criticism of existing land-use models concerning their [illegible] disregard of any theory of land market mechanisms can be extended also to other theoretical topics of urban spatial structure Reference has already been made in the preceding paragraph to the paper by Cox, and the work that he and others are pursuing on the spatial dimensions of urban political activity as represented by this paper is clearly relevant to the planning of land-use decisions Similarly, the papers by Casetti and Weicher in Part II represent areas of theoretical work which to date have been poorly acknowledged by the modelers of urban land-use patterns Casetti's paper, in fact, should be viewed in the content of a fairly extensive literature on urban rent theory, residential utility theory, urban population densities and spatial equilibrium analysis Much of this work is referenced in Casetti's own paper, but it seldom is considered in the modeling of land use patterns along the lines discussed earlier.

In general the argument could be made that the urban land-use models developed to date have not been strongly *spatial* in character Admittedly, they have sought to allocate land-uses by subregions and they have in different cases employed certain distance-decay relationships in handling accessibility questions and travel patterns, but for the most part, they have ignored the quite extensive and varied literature on urban spatial structure developed especially by the economists, geographers regional scientists, and sociologists

The third point with regard to the theoretical bases of the land-use models has to do with the lack of any broader contexts within which the models are set. That is to say the models are often neatly structured as regards the particular analytical questions they were designed to solve, but there is generally lacking any consideration of the relationships and feedback loops between the analysis and other facets of the social planning task Clearly, a land-use allocation model must take for granted certain goals and objective functions and in turn, the solutions which it yields may be only some of the alternatives confronting society These questions of goal formulations, of defining appropriate utility functions, of evaluating and choosing between alternatives, and of deciding upon means and policies whereby implementation takes place are themselves proper subjects of study and topics for theoretical reasoning It is these considerations which scholars such as Boyce, Day and McDonald (1969) and Harris, B (1970) are pursuing.

One final point which should be made concerning the theoretical bases of the urban-land-development models is that they deal with the city essentially as a *closed system* The models generally take as given certain exogenous forecasts of regional employment and population In the cases of most of the models discussed up to this point this feature imposes no really serious constraint since the models are developed for large metropolitan regions and the possible errors stemming from a lack of closure and probably less serious than those associated with the internal workings of the model But for small urban areas, the exogenous factors are certain to be much more important in a relative sense and the internal urban forecasts may be rendered invalid by only a small variation in the exogenous levels The notions of spatial linkages and spatial hierarchies are relevant in this context, and the applications of sophisticated urban land-development models [illegible] be consistent with these patterns.

Application of Models In a review of the applications of the above models to specific planning problems, two interesting points emerge. First, it is clear that much of the model development has been accomplished by "in-shop" research but that this work often has not been tied too closely to the immediate problems faced by the planning group in question as regards developing a master plan for its region. This is illustrated by the experiences of SEWRPC. This group has published some of the more intriguing and advanced mathematical statements of urban land-use forecasting but, in fact, their own regional land-use plan for 1990 was developed by conventional, nonanalytic methods and the land-use simulation model was used merely in testing the consequences of alternative policies (Southeastern Wisconsin Regional Planning Commission, 1968). The more recent work of this group on the land-use design model is purely an in-shop research project sponsored by HUD.

Second, many planning agencies, in some cases those of large metropolitan areas, are not involved as yet either in model-building or analytical land-use forecasting. The work of the Northeastern Illinois Planning Commission illustrates this point. Their recommended land-use plan for Chicago for 1990 was derived from a hand simulation in which the consequences of five different qualitative models of urban form were evaluated.

The limited extent of the application of the land-use models is borne out in several recent reports. Hemmens (1968) in his survey of some 34 major planning agencies received responses from 26, and of these, only 16 "reported on either current usage or active development of models." Hemmens notes the difficulties these agencies experience in developing the models and making them operational from the point of view of staffing and data facilities. Further, he stresses the lack of communication between agencies and the related absence of any serious cumulative work on the different models. Boyce, Day and McDonald (1970), in a survey of the plan-making process and evaluation methodologies associated with the work of the 13 largest planning programs, notes that as of 1969 only four of them — for Baltimore, Boston, Philadelphia, and the Twin Cities — have actually used computer models of urban growth and development to elaborate a set of alternatives.

The work on urban development models, then, has still a long way to go before operational packaged models are available for wide dissemination and use by small planning agencies. The increasing involvement of several commercial consulting firms in this area of land-use planning will possibly facilitate the development of these standard model procedures, although, typically, these privately developed models suffer from a lack of exposure in the published literature. On a national level HUD could provide a valuable service by promoting the development of a land-use-model package similar to the transportation-network-analysis package provided by the Bureau of Public Roads.

Given the likelihood that these packaged models will be available in the not-too-distant future, and assuming that the numbers of agency personnel better trained in analytical techniques will slowly increase, then it might seem premature for small planning agencies to contemplate seriously the implementation of expensive and "low-reliability" forecasting schemes at this time.

Conclusions and Recommendations

This paper was prepared originally as part of a specific report to the planning agency of a city of under 100,000 population. The recommendations included in that report are reproduced here in summary of some of the points made above. In structuring such recommendations for small metropolitan centers it is important to keep in mind certain background considerations.

(1) It is unlikely that small planning agencies either will have, or can afford, the necessary resources of hardware and manpower for the development, implementation, and monitoring of complex, computerized models of urban land use. At present, these models are far too expensive, both in development and operating costs, for small agencies to experiment with.

(2) The overall state of the art in land-use modeling is not that advanced, and no agency or consulting firm can provide a readily adaptable, packaged land-use model which small cities might be able to use.

(3) The sensitivity of land-use model projections to changes in the exogenous factors has not yet been investigated in detail. Therefore, for urban economies in which the external spatial linkages are strong, the relevance of existing land-use-forecasting techniques is not clearly established.

(4) The previous point notwithstanding, small urban centers with specialized economies may be comparatively easy to model as regards the relevant endogenous variables and the important relationships.

The following general recommendations might be emphasized.

(1) Interested planning agencies undertake a serious evaluation of the overall goals towards which their efforts in land-use analysis are to be directed. If the aim is simply to derive parameters for input to transportation programs then one set of procedures will be adequate. If the goal is to design land-use plans for the city consistent with certain objectives and constraints, then other approaches will be necessary. Specification of goals is critical.

(2) Agencies might begin simply with the development of regression analyses for the prediction of land uses by census tracts or traffic zones within the city. These models could utilize standard stepwise regression programs, could

emphasize the effects on land-use patterns of a few easily obtained variables, and could be calibrated with historical data These efforts would at least provide a framework for organizing existing land-use data and depending upon the results, they may offer a convenient approach to forecasting and sensitivity analysis There are some simple techniques such as shift analysis which also could be incorporated into the regression analyses

3) Agencies seriously consider the pedagogical value to be derived from the participation of their personnel and selected community leaders and citizens in game-simulation sessions conducted by Environmetrics or the METRO project group

References

Note Since this paper was prepared a number of other related statements ave appeared in journals such as *Land Economics, Environment and Planning,* nd *The Journal of the American Institute of Planners*

Alonso, W , "The Quality of Data and the Choice and Design of Predictive odels," Highway Research Board Special Report, No 97, Highway Research ard Washington, D C , pp 178-192 (1968)

Arthur D Little Corp , "Model of San Francisco Housing Market," San ancisco CRP C-65400 (1966)

Bourne, L S , "A Spatial Allocation-Land Use Conversion Model of Urban owth, Journal of Regional Science 9, pp 261-172 (1969)

Boyce D E , Day, N D , and McDonald, C , Metropolitan Plan Making, gional Science Research Institute Monograph Series No 4, Philadelphia 970).

Burgess, E W , "The Growth of the City An Introduction to a Research ject " in *The City*, R E Park et al , University of Chicago Press, Chicago 47-62, (1925).

Carlson E D , "Operational Aspects of a Probabilistic Model for Residential owth " Environmental Policies and Urban Development Thesis Series," versity of North Carolina No 10 (1968), 89 pp

Center for Real Estate and Urban Economics, "Jobs, People and Land Bay a Simulation Study " Special Report No 6 Berkeley (1968)

Greene, J P "IOMM," Pittsburgh CRP Tech Bull No 6 (1964), 18 pp

Donnelly, T G , et al , "A Probabilistic Model for Residential Growth," *Research Monograph*, University of North Carolina (1964)

Duke, R D *Gaming-Simulation in Urban Research*, Institute for Development and Services, MSU , East Lansing (1964)

Duke, R D , The METRO Urban Game Simulation An Experiment in Learning" *Proceeding of the Fourth Annual Conference on Urban Information Systems* Berkeley pp 142-154 (1966)

Ellis, R H , "Modeling of Household Location A Statistical Approach," Highway Research Record No 207, pp 42-52 (1967).

Feldt, A G , "The Cornell Land Use Game," *Miscellaneous Paper No 3*, Center for Housing and Environmental Studies, Cornell (1965).

Feldt, A G , Current Developments in Heuristic Gaming at Cornell University," *Proceedings of the Fourth Annual Conference on Urban Planning and Information Systems*, Berkeley, pp 160-167 (1966).

Fleet, C R , and Robertson, S R , "Trip Generation in the Transportation Planning Process," Highway Research Record, No 240, pp 11-31 (1968)

Forrester, J W , *Urban Dynamics*, The M I T Press, Cambridge (1969)

Garin, R A , (comment by A Rogers), "A Matrix Formulation of the Lowry Model for Intrametropolitan Activity Allocation," J Amer Inst Planners, 32, pp 361-366 (1966)

Geisler, M A Haythorn, W W., and Steger, W A., "Simulation and the Logistics Systems Laboratory," Rand Corp Mem RM-3281-PR (1962)

Goodman, L A , "Some Alternatives to Ecological Correlation," Amer J , Sociology, *64*, pp 610-625 (1959)

Hammersley, J M , and Handscomb, D C , *Monte Carlo Methods*, Methuen and Co , Ltd , London, pp 1-9, (1964).

Harris, B , "Quantitative Models of Urban Development Their Role in Metropolitan Policy-Making," in *Issues in Urban Economics*, H S Perloff and L Wingo (Eds), The Johns Hopkins Press, Baltimore, pp 363-412 (1968)

Harris, B , "Some Problems in the Theory of Intra-Urban Location," Operations Research, *9*, pp 695-721 (1961)

Harris, B , et al , "Construction of Models" (Panel discussion), Highway Research Board Special Report, No 97, Highway Research Board, Washington, D.C , pp 193-216 (1968)

Harris, C C , "A Stochastic Process Model of Residential Development," Journal of Regional Science, *8*, pp 29-39 (1968).

Harris, C.D , and Ullman, E L , "The Nature of Cities," The Annals of Sociology, *242*, pp 7-17 (1945)

Hemmens, G C , (Ed), "Urban Development Models," Highway Research Board Special Report, No 97, Appendix A, pp 253-262 (1968)

Herbert, J D , and Stevens B H , "A Model for the Distribution of Residential Activity in Urban Areas," Journal of Regional Science, 2, pp 21-36, (1960)

Hill, D M , Brand, D , and Hansen, W B , "Prototype Development of a Statistical Land Use Prediction Model for the Greater Boston Region," Highway Research Record *114*, pp 51-70 (1965)

Hoyt, H , *The Structure and Growth of Residential Neighborhoods in American Cities* FHA, Washington, D C (1939)

Irwin, N A and Brand, D , "Planning and Forecasting Metropolitan Development " Traffic Quarterly pp 520-540 (October, 1965)

Kibel B M , *The Evolution of A Dynamic Planning Model* " unpublished Ph D dissertation University of California Berkeley (1970)

Kilbridge, M D O Block, R P , and Teplitz, P V , "A Conceptual Frame-

CHAPTER 10

The spatial distribution of activities

10.1. INTRODUCTION

In Chapters 6 to 9 we have worked with a fairly restricted set of techniques, mainly building on matrix operator-probability, or spatial interaction concepts. In the field of locational analysis (the study of the spatial distribution of urban activities) a much wider range of techniques is available, and it is useful to return to the model-builder's check list of Chapter 4 by way of introduction.

Models of the distribution of activities can be built for a wide variety of purposes. We emphasized earlier that it was important to be clear about purpose, and we shall see here that this issue can indeed affect choice of technique for model development.

The next item on the check list related to choice of variables. Clearly, we are interested in variables which describe activities by location, and the determinants of the associated patterns and processes. We have to be particularly clear as to whether the prime interest is in the activities themselves, the associated physical stock, or both. An example from the population–residential-housing sector will illustrate the issues involved here. We may be primarily interested in residential activity, the spatial distribution of different types of people, the types of houses they live in and so on, or we may be interested in housing primarily in order to set up a house-building programme. Ideally, we should be interested in both simultaneously, but available time and effort may prohibit this. It is extremely easy to confuse the two interests, especially since, at any one time, the population cross-section coincides in some way with the housing stock cross-section. However, the behaviour in time of people and stock is quite different: people move around more frequently than the pattern of the housing stock changes. Typically, one home will be utilized by many families during its lifetime. The distinction which is being made can usefully be considered as relating to aspects of an economic supply and demand system: people demand housing at a location and some houses are supplied there; what is observed at one point in time is the inter- [illegible] of supply and demand curves. This discussion is beginning to anti- [illegible] this under the heading of 'Theory' below, however, and will be pursued [illegible] there. Once again, we can observe that choice of variables in relation to the [illegible] system of interest may affect the model-building techniques to be used.

At this point, it is convenient to discuss another distinction between types of variable which is particularly crucial in a *spatial* distribution model: this relates to whether we focus primarily on the distribution of an activity (or stock) across the zones of our spatial system, or on the set of activities which use up the capacity of a particular zone. In relation to Figure 10.1, it is a matter of whether we focus primarily on rows or on columns. Generally speaking, we will adopt the row or activity location focus as distinct from the column or land-use focus. In any partial study of one activity, we should bear in mind that we should not neglect the fact that it is competing for land (and possibly other resources) with other activities. Each activity location model should ideally be part of a more general model which incorporates a set of land-use accounts. Such issues are taken further in Chapter 11 where we discuss the building of general models.

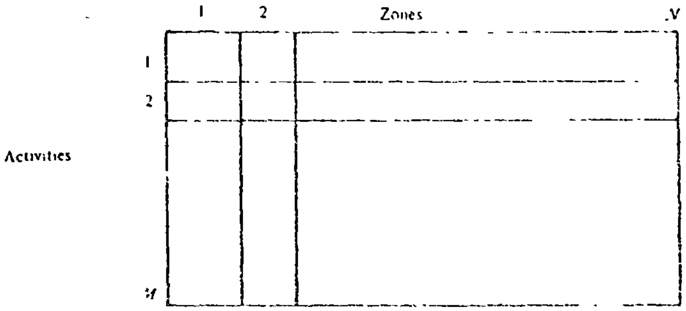


Figure 10.1 Zone-activity cross tabulations (Based with permission, on I. S. Lowry, *Urban Development Models*, Special Report 97, Highway Research Board, Washington, D.C., U.S.A., p. 128, Figure 2. 1968)

The third item on the model-builder's check list was the identification of variables which were under the planner's control. This is partly a matter of planning powers in relation to the system in which the model builder is interested, partly a matter of the model-builder's purpose and interest, and partly a matter of interpretation. These matters cannot be resolved except in relation to particular examples. However, one matter can be usefully recalled at this stage: the question as to whether our models are to be normative or positive. The convention adopted in this book is that it is better to think of all models as positive models, built for analysis, which may be capable of (conditional) predictions which can be utilized within some essentially normative planning process. This argument simply recapitulates that of Chapter 2, but it is particularly appropriate to mention it in relation to location models. In practical terms, the model builder has to decide which [illegible] are *endogenous* to the model and which *exogenous*.

The fourth item in the check list referred to aggregation issues. Since we are interested in the spatial distribution of activities, our urban system must be divided into zones. These may be very small, or larger, depending on the purposes of the exercise, and once again, this decision relates to other decisions which will finally decide the model-building technique to be used. It is the sectoral aggregation/disaggregation question which is a more tricky one for the locational analyst. This is the question as to whether an essentially micro or macro approach should be adopted. Throughout this book we have attempted to develop a *comprehensive* approach to model building in the sense of trying to represent all the activities of the whole study area in their entirety. Within this, we would like to build in as much detail as possible; indeed, in the limit, it is possible to deal with all individuals, organizations and units of physical stock in a comprehensive micro approach. However, this is practically impossible for large systems. It has been argued elsewhere (Wilson, 1972) that an *essentially* micro approach involves modelling the behaviour of an individual (or other micro unit) *in his environment*, taking this environment as given. Such an approach inevitably ceases to be comprehensive and so, in this sense, we are primarily interested in macro or aggregative models. We should note, however, as we shall see in relation to specific examples below, that such models can be given micro-scale interpretations, usually in probabilistic terms, and much can be learned by insisting that the theories which underpin our models are maximally consistent with corresponding micro-behavioural theories. One further comment on this point is in order: micro-analysis at the individual level may often be a powerful tool for the identification of problems. But, when policies are implemented to solve problems, however they were identified, then it is usually important to treat the system-wide impacts of these policies and this can only be accomplished in a comprehensive framework. This is perhaps one of the main reasons for the stance adopted in this book. A deeper discussion would take us beyond the scope of this book, but the interested reader can pursue it further in the paper cited earlier (Wilson, 1972). For the time being, however, we shall assume that we are concerned with macro or aggregative models, and the techniques presented are chosen with this in mind. In some associated fields, such as the study of industrial location, much of the model building has been what we have called essentially micro, and will be largely neglected in this chapter.

The fifth item in the check list concerned the treatment of time. In the building of location models, the important 'time' question relates to whether the model builder is attempting to represent the total cross section of one or more activities at some point(s) in time, or the marginal change between two points in time. This again turns out to have a fundamental bearing on [illegible]. The perfect answer to the question would be to seek a fully [illegible] model in which time was a continuous variable which had been [illegible] for a very long time period, so that both cross-sections and marginal changes were predicted by the model. This is rarely possible in practice, if only because of the lack of suitable time series data. In fact, many cross-section models have been developed for data which refer to one point in time only, and thus are not even tested on a comparative static equilibrium basis, while the rarer marginal models usually relate to only a small number of points in time, often only two. However, the answer to the question depends to some extent on the interest and purpose of the model builder. It is good to be able to model a total cross-section, and sometimes lack of time series data means that that is the only possibility, but a model builder may be interested only in the development of new housing in the next five years, in which case a marginal model would be more appropriate.

A model is a formal representation of a theory and, as ever, the most important task for the model builder is to seek out and develop an adequate and appropriate theoretical basis for his model. We have already discussed a number of preliminary theoretical questions: questions of a prime interest in people or stock, in activity or land use, the micro-macro question and so on. More detail will be added in the introductory parts of each sub-section which follows, where the kinds of hypotheses and theories to be represented in the models are discussed. The discussion on this topic, therefore, will be postponed.

The next item in the check list related to techniques for model building. A number of techniques can be distinguished, though, as we shall see, they overlap and complement each other in various ways. They include

(1) algebraic models, including as special cases spatial interaction models and entropy maximizing models,

(2) matrix operator models,

(3) linear programming models,

(4) differential equation/difference equation models,

(5) simulation models,

(6) econometric models,

(7) models based on economic theory,

(8) ecological models,

(9) models based on the theory of games.

These are not listed in any particular order. It is difficult, indeed, to think of any 'logical' order for such a list. (It is almost impossible to list them in increasing order of 'difficulty' for example, as, in most cases, there are simple and complicated models based on each technique.) The essence of each approach will be explained in broad terms in Section 10.2, and details will only be presented in relation to particular examples in later sub-sections. The preceding discussion has raised issues concerned with population activity/physical stock foci, demand/supply relationships, micro/macro questions and cross-section/marginal foci. If particular techniques contribute to one focus rather than another, this will be noted in the next section.

10.2. SOME AVAILABLE TECHNIQUES

Algebraic models express relationships between variables using the operations of elementary algebra coupled with the notion of a function. The relationships may involve notions of causality, but this is not necessary. A specific example of algebraic models is the family of *spatial interaction models* outlined in Chapter 6. We have already discussed one such model in detail, the retail sales model in Chapter 4, which the reader will now recognize as a production-constrained spatial interaction model. The essence of using the concept of spatial interaction in a location model is that the activity which is being located has a strong interaction with some other activity. In the case of the retail sales model, the activity being located is retailing and this interacts strongly with the resident population and its spending power (which is, in this case, assumed to be given). When a spatial interaction model is used in this way, the given activity is the constrained part of a singly constrained model (it could be either production- or attraction-constrained) and the other activity, relating to the other end of the interaction, is obviously unconstrained, the total volume of that activity located in a zone is obtained by adding all the flows into (or out of) the zone.

The spatial interaction model has mostly been used to represent residential location in relation to the distribution of workplaces, and the utilization of services in relation to the distribution of population. (These two models can also be coupled together, as we shall see in Chapter 11.) Most typically, it has been used to model population activity rather than stock and the total cross-section rather than marginal change, though there is no fundamental limitation associated with the technique. In theoretical terms, it can be seen as a set of equations which represent a model of demand–supply interaction, and the modeller has to be particularly careful in his treatment of supply-side variables within the model. Further, as we noted in Chapter 6, spatial interaction models can be seen as special cases of *entropy-maximizing models* and the statistical averaging interpretation of such models helps to relate micro and macro population of associated modelling tasks (Wilson, 1972).

The class of models just briefly discussed is only a small sub-set of the class of algebraic models. Algebraic model building will always represent one of the most fruitful techniques for the representation of theories. Its main weakness relates to the treatment of time associated with such models, though this can easily be rectified by adding the tools of the differential calculus or by the use of difference equations, as we shall see later.

Matrix operator models are really special cases of algebraic models. We saw in the demographic models of Chapter 7 that the matrix formulation of these models is an optional extra. However, if matrix inversion is involved, as in the input–output model of Chapter 8, or any other special property of matrices, then it would be extremely clumsy not to use matrix concepts. The technique has been of relatively limited application in locational analysis, but could become increasingly important if multi-region versions of the demographic or economic models of Chapters 7 and 8 are applied to systems of interest made up of very small 'regions', or zones.

Linear programming models in locational analysis are often confusing because it is not always clear whether the resulting model is positive or normative, since, in many cases, some measure of welfare or utility is being maximized. Following our earlier comment on this question, we will consider only models in which the objective function characterizes behaviour, and the resulting distribution of activities, may or may not be optimal in some normative sense. That is a matter for separate investigation. (Linear programming techniques may be used additionally as part of a planning process in which a behavioural model has been used separately to assess impact, but the term 'model' will not be used in this context.) The technique can be employed when it is possible to construct an objective function which, when maximized subject to linear constraints, reproduces behaviour. It has been used most commonly in studies of residential location and industrial location, but also in relation to land use.

Differential equations or *difference equations* are concerned with change, with the period of change being measured by some independent variable, the variable being continuous in the first case, discrete in the second. Typically, that variable will be time in our case, though it could in principle be a spatial variable. In effect, the body of mathematics associated with these concepts simply enables us to handle time properly, when we are otherwise capable of doing this. (We may be fundamentally incapable of it, because of data limitation, for example.)

There are circumstances when a model can be specified as a set of rules which enable numbers to be operated on, usually in the computer, though the rules and the consequences of applying them cannot be written down as a set of algebraic equations. Such mathematics was called algorithmic in Chapter 5. When a model is assembled in this way, it is algorithmic, but not necessarily *an* algorithm: it may consist of several algorithms, or algorithms plus equations, and it can usefully be called a *simulation* model, though the term 'simulation' is sometimes used much more widely. Sometimes, the simulation technique lends itself naturally to a problem. This happens, for example, when the underlying theory consists of a set of statements involving conditional probabilities. More generally, we might say that we resort to simulation techniques for situations which are too complicated to be handled by more straightforward algebraic techniques. We shall see examples below, from the residential location field. Of course, the notion of using simulation techniques has recently been given a substantial boost in the urban and regional modelling field with the work of Forrester (1968, 1969).

Econometric modelling techniques are used to study relationships between variables using the methods of statistical analysis The relationships usually have some basis in economic theory, though 'econometric' seems to be used more widely now to describe a certain kind of statistical analysis. The relationship between the variables are often of a rather simple kind. they are usually either linear, or can be transformed into linear form by taking logs, for example It is this simplicity (mainly linearity, and additiveness) which limits their usefulness in locational analysis, but examples will be presented below

Of the remaining three technique areas listed at the end of Section 10 1 each represents a way of looking at a problem in locational analysis: *economic*, *ecological* and *game-theoretical* The resulting models use a combination of the techniques mentioned earlier, but are each constrained by the rules of the viewpoint adopted In some cases, the rules and viewpoints are so highly developed that it is useful to consider them as a model-building technique The economist's method of locational analysis is typically based on the theory of consumer's behaviour, suitably aggregated; the ecologist's approach (in the somewhat old-fashioned Chicago sense in which we are using the notion here) is based on such concepts as invasion and succession, the game theorist hypothesizes a locator playing a game against other locators, or against the market Each approach has been used, and examples will be presented below of all but the last The difficulty of the economic approach is that of aggregation from an essentially micro viewpoint to a comprehensive one without losing the essence of the economic theory by turning it into an oversimplified econometric model The ecological approach has been hampered by the difficulty of formalizing and mathematicizing it. The game theory approach is hamstrung by the virtual impossibility of obtaining an effective aggregation, which is why no example is given here

Finally, we should note that the techniques can be obtained in various ways, indeed we have already had examples of this in the way in which the 'viewpoint' techniques can be combined with the rest.

10.3. METHOD OF PROCEDURE

In Chapter 3 (in Figures 3 1 and 3 2), we identified an interest in the location of population activities and the location of economic activities, and (in Figure 3 3) set these in a demand/supply relationship with each other There are various ways in which we could proceed We could take particular population activities and particular economic activities and discuss the associated models Alternatively, we could define sub-systems which incorporated both demand and supply mechanisms, for example, we could take a residential system which incorporated the population's demand for residences *and* the mechanisms of housing supply There are other ways of cutting up the whole system, of course, and some methods of locational analysis do not fit into either of the frameworks which have been mentioned In the last analysis, we must be concerned with the total system, the general model, and we shall return to this in Chapter 11 Meanwhile, we will adopt the first of the two frameworks and consider in turn the demand side, the location of population activities, in several sections, and then the supply side, the location of economic activities, in one section only. Several sections are devoted to residential location (10.4-10 8), and three to the utilization of services (10 9-10.11). The supply side, about which less is known in modelling terms, is discussed in Section 10 12

This division has the substantial advantage that it clarifies at the outset two of the issues raised in Section 10 1, concerned with population activity/physical stock foci and demand/supply relationships Although there is some overlap, and occasionally supply considerations will be mentioned in a demand section, and vice versa, in the main Sections 10 4-10 11 are concerned with population activity and demand, and stock is one aspect of all supply aspects in Section 10 12 In relation to the third (macro/micro) issue, we have already decided that our goal of comprehensiveness forces us to accept an aggregative position, though macro/micro relationships will be discussed whenever possible This leaves us with the cross-sectional/marginal and choice-of techniques issues and these are discussed in particular sections as appropriate.

10.4. RESIDENTIAL AND WORKSHOP LOCATION: BASIC HYPOTHESES

We can begin our discussion of the hypotheses to be built into our models by focusing on individuals or households. We assume that the appropriate micro-unit for the study of residential location is the household (more or less by definition), and for the study of workplace location, the individual Occasionally, we shall have difficulty in reconciling the two viewpoints in a single model but, for the time being, we shall use 'individual' and 'household' almost interchangeably, and assume that the intended meaning is clear from the context We assume that each individual has a need for shelter, subsistence goods (including services) and an associated income with which to purchase both

Since most income is derived from employment, this at least links 'home' with 'workplace' There is obviously also the physical linkage, which involves an assumed disutility of travel and, as part of this, an expenditure on the journey to work which reduces the disposable income available for shelter and subsistence goods For these reasons, it is difficult to separate the study of residential and workplace location

We can assume that a household makes some broad decision about the allocation of its income between housing, subsistence goods, journey to work, and possibly other [illegible] such as 'savings'. (This follows the notions of a utility tree developed by Strotz (1957).) The amount of money available for housing (and related services since it is assumed that a residential package is being purchased) will be determined partly by available total income, partly by preferences and partly by what is on offer. It is simpler to think this analysis through for one-worker households, but real life is complicated by the existence of no-worker and multi-worker households.

Given these broad decisions, the individual or household can then carry out a search among available opportunities. This may be a joint search for house and job, or for house alone (given a job) or for a job alone (given a house). The pattern of available opportunities is, of course, continually changing through time. Conceptually at least, we can assume that the search is structured in relation to type of house (size, age and condition, tenure and so on), location (and, in particular associated linkages to work, services and so on), environment (physical and social), and price (which will reflect all the preceding items in the list).

We can recall at this point that we are trying to be comprehensive. Individuals are competing in a complicated market process which involves other individuals and the suppliers of housing. The market (or sub-markets in relation to different tenure categories) operates with a complex set of institutional rules.

In summary, then, our models must reflect and reproduce the individual's choice of housing in relation to his income and preferences (which means that the models must distinguish person type), house type, location, environment and price within a complicated 'market' operation. In this analysis, for the present, we are taking the supply side as given.

These hypotheses can now be sharpened (and sometimes simplified) so that we can build associated models. A number of examples are discussed below. The techniques used, in turn, are (i) spatial interaction modelling, (ii) linear programming/economic theory/econometrics, (iii) simulation and (iv) ecological modelling.

10.5. SPATIAL INTERACTION MODELS OF RESIDENTIAL LOCATION

In the following discussion, we shall assume that we have a study area divided into zones of a suitable size and labelled by indices such as i and j. Conventionally (and this is *merely* a convention) we shall associate i with the residential zone and j with the employment zone. We can now begin to discuss spatial interaction models of residential and workplace location. We will work from the simplest, rather unrealistic, models to much more complicated, realistic, ones. In so doing, we are, more or less, giving a historical account of the development of such models. The theoretical basis of the models can be discussed rather briefly, partly because we are building on the theoretical framework laid down in Chapter 6, and partly because a detailed account of the entropy-maximizing basis is given in Chapter 4 of another book by the author (Wilson, 1970-A).

The simplest spatial interaction model of residential location is that developed by Lowry (1963, 1964) as part of a more general model. He assumed a given distribution of jobs, that is, that E_j, the number of jobs in each zone j, is given and that workers are allocated to residences around their workplaces according to a simple spatial interaction hypothesis. If T_{ij} is the number of people who live in zone i and work in zone j, he assumed, in our notation

$$T_{ij} = gE_j f(c_{ij}) \tag{10.1}$$

where c_{ij} is the cost of travel from i to j, f a decreasing function, and g a constant. Diligent readers of Chapter 6 will recognize this as an unconstrained spatial interaction model, in which there is no residential (i-zone) mass term: that is, the i-zone mass term is assumed to be 1.

The fact that the model is unconstrained means that its outputs (the T_{ij}s) are inconsistent with the assumptions in that it is impossible to calculate a value of g which will ensure that (assuming N zones)

$$\sum_{i=1}^{N} T_{ij} = E_j \tag{10.2}$$

is satisfied. We can also argue that there should be a residential mass term as some residential zones are more attractive than others. Thus, however we eventually decide to measure it, we should incorporate a mass term W_i to represent the relative attractiveness of zone i.

Thus, we can transform the model of equation (10.1) into an attraction constrained model, including a mass term, by writing

$$T_{ij} = B_j W_i E_j f(c_{ij}) \tag{10.3}$$

where

$$B_j = 1\Big/\sum_{i=1}^{N} E_j f(c_{ij}) \tag{10.4}$$

B_j replaces g and its value, obtained from equation (10.4), ensures that the constraint (10.2), which is one of our assumptions, is satisfied.

A very simple, implicit, assumption is being made in this model about the supply side: that

$$\sum_{j=1} T_{ij}$$

houses are made available in zone i to meet demand. An alternative procedure is to assume that we are given the number of houses H_i available in each zone i. We now have an additional constraint

$$\sum_{j=1}^{N} T_{ij} = H_i \tag{10.5}$$

to add to equation (10.2) and Chapter 6 tells us that we have a production-attraction constrained situation. The corresponding model is

$$T_{ij} = A_i B_j H_i E_j f(c_{ij}) \tag{10.6}$$

where

$$A_i = 1\Big/\sum_{j=1}^{N} B_j E_j f(c_{ij}) \tag{10.7}$$

and

$$B_j = 1\Big/\sum_{i=1}^{N} A_i H_i f(c_{ij}) \tag{10.8}$$

The fact that A_i and B_j are calculated in this way ensures that equations (10.2) and (10.5) are satisfied. What sort of animal have we now got? If this is a location model, it appears to run counter to our earlier statement (in Section 10.2) that spatial interaction models used as location models are singly constrained and that the unconstrained 'end' provides an estimate of the location variable. From equation (10.3), for example, we would take

$$\sum_{j=1}^{N} T_{ij}$$

as an estimate of the residential location pattern. However, in the doubly constrained model this quantity is given as H_i. In what sense is the equation (10.6)–(10.8) model a location model? It remains a location model for the population involved. It tells us which people live where. This reflects another part of an earlier discussion that it is important to distinguish between people and stock. In the doubly constrained model we are taking the stock as given (for the moment; it may be estimated in another, supply side, model) and we are allocating people jointly to residences and workplaces.

It is also possible to develop hybrid models which are partly singly and partly doubly constrained (cf. Wilson, 1969). This would be appropriate, for example, if it were considered that for most zones, say the set Z_2, the simple competitive mechanism of the singly constrained model operated, while for a few zones, say in the set Z_1, there are population constraints, perhaps due to planning policy. This situation may apply to a situation of rapid suburban growth in a city where population in central areas is strictly controlled by

redevelopment capacity. Let

$$Z = Z_1 \cup Z_2 \tag{10.9}$$

be the total set of zones. Then the constraints are

$$\sum_{i=1}^{N} T_{ij} = E_j, \qquad j \in Z \tag{10.10}$$

and

$$\sum_{j=1}^{N} T_{ij} = P_i, \qquad i \in Z_1 \text{ only} \tag{10.11}$$

The hybrid model can then be written

$$T_{ij} = A_i B_j P_i E_j f(c_{ij}), \qquad i \in Z_1, \qquad j \in Z \tag{10.12}$$

$$T_{ij} = B_j W_i E_j f(c_{ij}), \qquad i \in Z_2, \qquad j \in Z \tag{10.13}$$

where

$$A_i = 1\Big/\sum_{j=1}^{N} B_j E_j f(c_{ij}), \qquad i \in Z_1 \text{ only} \tag{10.14}$$

and

$$B_j = 1\Big/\left\{\sum_{i \in Z_1} A_i O_i f(c_{ij}) + \sum_{i \in Z_2} W_i f(c_{ij})\right\} \tag{10.15}$$

One other kind of hybrid model can be usefully mentioned. We have already hinted at the beginning of this section that there may be different kinds of locational behaviour; this could be explicitly represented as different types of constraint. So far, we have not distinguished person type at all. Suppose now that n denotes type of locational behaviour, as follows:

$n = 1$: locationally unconstrained, seeking both house and job;
$n = 2$: production constrained, seeking job only;
$n = 3$: attraction constrained, seeking house only;
$n = 4$: production-attraction constrained, seeking neither house nor job.

Let T_{ij}^n be the number of persons of type n living in zone i and working in zone j. Then, from an assumption, we would expect to be given T_{**}^1, T_{i*}^2, T_{*j}^3 and T_{ij}^4. Alternatively, instead of T_{ij}^4, we may be given T_{i*}^4 and T_{*j}^4 and assume that T_{ij}^4 can be estimated from a production-attraction constrained model. Define

$$P_*^1 = T_{**}^1 = \sum_{i=1}^{N} \sum_{j=1}^{N} T_{ij}^1 \tag{10.16}$$

$$H_i^2 = T_{i*}^2 = \sum_{j=1}^{N} T_{ij}^2 \tag{10.17}$$

$$E_j^3 = T_{*j}^3 = \sum_{i=1}^{N} T_{ij}^3 \tag{10.18}$$

$$H_i^4 = \sum_{j=1}^{N} T_{ij}^4 \tag{10.19}$$

and

$$E_j^4 = \sum_{i=1}^{N} T_{ij}^4 \tag{10.20}$$

and assume these are all given, so that equations (10.16)-(10.20) are constraints. Suppose we are also given (in the spirit of the model given by equations (10.6)-(10.8)) quantities which can be called H_i^* and E_j^* in other words, the total quantities of housing and jobs by zone, irrespective of the locational behaviour type of the occupants. Then we can define quantities H_i' and E_j' as follows:

$$H_i' = H_i^* - H_i^2 - H_i^4 \tag{10.21}$$

$$E_j' = E_j^* - E_j^3 - E_j^4 \tag{10.22}$$

H_i' is the quantity of housing in zone i being competed for by those seeking houses (in locational behaviour categories 1 and 3) and E_j' is the equivalent quantity for jobs in zone j. This means that we can impose additional constraints on the T_{ij}^ns as follows:

$$\sum_{j=1}^{N} T_{ij}^1 + \sum_{j=1}^{N} T_{ij}^3 = H_i' \tag{10.23}$$

and

$$\sum_{i=1}^{N} T_{ij}^1 + \sum_{i=1}^{N} T_{ij}^2 = E_j' \tag{10.24}$$

These constraints will now lead us to an interesting set of linked spatial interaction models. Let us consider the different categories in turn. Category 1 is constrained by equations (10.16), (10.23) and (10.24). The first of these would normally demand a proportionality factor K, the second A_is and B_js which we could call A_i^1 and B_j^1. Category 2 is constrained by equations (10.17), which needs an A_i^2, and (10.24) which needs a B_j^2. Category 3 is constrained by equations (10.18), which needs a B_j^3, and (10.23) which needs an A_i^3. Category 4 is independent: either T_{ij}^4 is given directly, or we use equations (10.19) and (10.20) as constraints to generate a journey-to-work model. Let us assume that we take the latter assumption. Note that A_i^1 and A_i^3 relate to the same constraint, (10.23), as do B_j^1 and B_j^2, (10.24). In practice, this means that we can take

$$A_i^3 = A_i^1 \tag{10.25}$$

and

$$B_j^2 = B_j^1 \tag{10.26}$$

Then, the models can be written as

$$T_{ij}^1 = K A_i^1 B_j^1 H_i' E_j' f_1(c_{ij}) \tag{10.27}$$

$$T_{ij}^2 = A_i^2 B_j^1 H_i^2 E_j' f_2(c_{ij}) \tag{10.28}$$

$$T_{ij}^3 = A_i^1 B_j^3 H_i' E_j^3 f_3(c_{ij}) \tag{10.29}$$

$$T_{ij}^4 = A_i^4 B_j^4 H_i^4 E_j^4 f_4(c_{ij}) \tag{10.30}$$

where

$$K = P_*^1 \Big/ \sum_{i=1}^{N} \sum_{j=1}^{N} A_i^1 B_j^1 H_i' E_j' f_1(c_{ij}) \tag{10.31}$$

$$A_i^1 = 1 \Big/ \sum_{j=1}^{N} \left[K B_j^1 E_j' f_1(c_{ij}) + \sum_{j=1}^{N} B_j^3 E_j^3 f_3(c_{ij}) \right] \tag{10.32}$$

$$B_j^1 = 1 \Big/ \sum_{i=1}^{N} \left[K A_i^1 H_i' f_1(c_{ij}) + \sum_{i=1}^{N} A_i^2 H_i^2 f_2(c_{ij}) \right] \tag{10.33}$$

$$A_i^2 = 1 \Big/ \sum_{j=1}^{N} B_j^1 E_j' f_2(c_{ij}) \tag{10.34}$$

$$B_j^3 = 1 \Big/ \sum_{i=1}^{N} A_i^1 H_i' f_3(c_{ij}) \tag{10.35}$$

$$A_i^4 = 1 \Big/ \sum_{j=1}^{N} B_j^4 E_j^4 f_4(c_{ij}) \tag{10.36}$$

$$B_j^4 = 1 \Big/ \sum_{i=1}^{N} A_i^4 H_i^4 f_4(c_{ij}) \tag{10.37}$$

These equations, (10.31)-(10.37), ensure that constraints (10.16), (10.23), (10.24), (10.17), (10.18), (10.19) and (10.20), respectively, are satisfied.

A digression on dynamics is appropriate at this point. Up to the model just developed, the elementary models discussed were incontravertably cross-sectional models. All people are involved, and the whole housing stock. It is implicitly assumed that, at any time, an equilibrium position is attained which is described by the model equations. In the latest model, however, a time period is at least implicitly referred to: that period which determines the locational behaviour categorization of the population For example, the population of those seeking houses consists of new households and potential movers *in some time period* If the time period is quite short, then the bulk of the population will be in category 4 (and the category 4 equations are, in effect, simply the journey-to-work equations of Chapter 9) Categories 1–3 are then concerned with marginal change. Provided that we can estimate the quantities which we have so far assumed to be given, P^1_*, H^2_i, E^3_j, H^4_i, E^4_j, H^*_i and E^*_j for a time period, say t to $t+T$, then we have the basis of a dynamic model However, it seems worthwhile to develop further the spatial interaction model before we proceed to build a dynamic one.

We stated in our hypotheses at the beginning of this section that it was important to build person type and house type into the model So far, we failed to do either In the spatial interaction models presented so far, we either implicitly assumed that all people are identical and all houses identical or, more charitably, that the models have performed a lot of averaging around means in making their predictions. It is a relatively straightforward matter to remedy this situation by disaggregation by person type and house type, and we can also build in home price and the fact that it varies with type and location

Suppose at this stage that the main person-type variable we wish to indicate is income, w, with house-type, we need not commit ourselves to a specific meaning, we simply specify type k Our main variable is then T^{kw}_{ij}, the number of w-income people living in a type-k house in zone i, working in zone j. Suppose further that we are given H^k_i, the number of type-k houses in each zone i, E^w_j, the number of income w jobs in j, p^k_i, the price of a type-k house in zone i, and c_{ij}, the cost of travelling from i to j Implicitly, for simplicity, we are assuming one worker per household, and that all income is derived from employment

We can now consider building a spatial interaction model with H^k_i and E^w_j as mass terms, and c_{ij} as travel cost (cf Wilson, 1970-B). This might lead us to suggest something of the form

$$T^{kw}_{ij} = A^k_i B^w_j H^k_i E^w_j f_w(c_{ij}) \tag{10.38}$$

which is a doubly constrained model for each k, w combination We have assumed a cost function which varies with w However, this contains no reference to price p^k_i, and has no mechanisms to ensure that households live in houses (a) they can afford and (b) sufficiently high-priced for their income This can be achieved with the following device Let q^w be the *average* amount which a w-income household spends on housing *after journey to work costs have been deducted* Then consider a term of the form

$$e^{-\mu^w[p^k_i - q^w(w - c'_{ij})]^2} \tag{10.39}$$

where c'_{ij} is the money part of c_{ij} and $q^w(w - c'_{ij})$ is the average amount

available for housing for a w-income household living in i and working in j. If this differs significantly, in either direction, from p^k_i, then the expression in square brackets in (10.39) rapidly becomes large, and the expression (10.39) itself becomes small So we add this multiplicatively to the right-hand side of equation (10.38). This gives, as a revised form of model,

$$T^{kw}_{ij} = A^k_i B^w_j H^k_i E^w_j e^{-\beta^w c_{ij}} \times e^{-\mu^w[p^k_i - q^w(w - c'_{ij})]^2} \tag{10.40}$$

Now, a w-income household will only be assigned to a type-k home in zone i if p^k_i does not markedly differ from $q^w(w - c'_{ij})$. Just how markedly it differs is determined by the magnitude of μ^w. We thus now have a person-type, house-type, disaggregated version of the model presented as equations (10.6)–(10.8). A^k_i and B^w_j must be calculated to ensure

$$\sum_j \sum_w T^{kw}_{ij} = H^k_i \tag{10.41}$$

$$\sum_i \sum_k T^{kw}_{ij} = E^w_j \tag{10.42}$$

so that

$$A^k_i = 1\Big/\sum_j \sum_w B^w_j E^w_j e^{-\beta^w c_{ij}} e^{-\mu^w[p^k_i - q^w(w - c'_{ij})]^2} \tag{10.43}$$

and

$$B^w_j = 1\Big/\sum_i \sum_k A^k_i H^k_i e^{-\beta^w c_{ij}} e^{-\mu^w[p^k_i - q^w(w - c'_{ij})]^2} \tag{10.44}$$

The model given by equations (10.40), (10.43) and (10.44) now has an interesting mechanism, but it is still a cross-sectional model. Further, it has the unrealistic simplifying assumption of one worker per household We must thus proceed in two steps firstly, we must indicate how to remove the simplifying assumption and secondly, how to make the model dynamic The first task is a relatively straightforward one provided that we assume that it is primarily the income of the head of the household which determines expenditure on housing We then add a superscript m to T^{kw}_{ij}, making it T^{kwm}_{ij}, where $m = 1$ indicates a worker who is head of household, $m = 0$ a worker who is not Non-workers would still have to be added in separately Let r_i be the average number of workers per household in zone i, so that there are H^k_i places in type k houses for heads of households and $(r_i - 1)H^k_i$ for non-heads We assume that both heads and non-heads compete for the same jobs A similar method could be used for non-workers, heads or not[1]

We then have to build models which satisfy the constraints

$$\sum_j \sum_w T^{kw1}_{ij} = H^k_i \tag{10.45}$$

$$\sum_j \sum_w T_{ij}^{kw0} = (r_i - 1)H_i^k \qquad (10.46)$$

$$\sum_i \sum_k \sum_m T_{ij}^{kwm} = E_j^w \qquad (10.47)$$

The reader can easily check that the appropriate model is

$$T_{ij}^{kw1} = A_i^{k1} B_j^w H_i^k E_j^w \mathrm{e}^{-\beta^{w1} c_{ij}} \times \mathrm{e}^{-\mu^w [p_i^k - q^w (w - c_{ij})]^2} \qquad (10.48)$$

$$T_{ij}^{kw0} = A_i^{k0} B_j^w (r_i - 1) H_i^k E_j^w \mathrm{e}^{-\beta^{w0} c_{ij}} \qquad (10.49)$$

where

$$A_i^{k1} = 1\Big/\sum_j \sum_w B_j^w E_j^w \mathrm{e}^{-\beta^{w1} c_{ij}} \times \mathrm{e}^{-\mu^w [p_i^k - q^w (w - c_{ij})]^2} \qquad (10.50)$$

$$A_i^{k0} = 1\Big/\sum_j \sum_w B_j^w E_j^w \mathrm{e}^{-\beta^{w0} c_{ij}} \qquad (10.51)$$

and

$$B_j^w = 1\Big/\Big\{\sum_i \sum_k A_i^{k1} H_i^k \mathrm{e}^{-\beta^{w1} c_{ij}} \mathrm{e}^{-\mu^w [p_i^k - q^w (w - c_{ij})]^2}$$

$$+ \sum_i \sum_k A_i^{k0} (r_i - 1) H_i^k \mathrm{e}^{-\beta^{w0} c_{ij}}\Big\} \qquad (10.52)$$

This is the model we now have to put into a dynamic framework by distinguishing the four types of locational behaviour which led to the aggregated model of equations (10.28)–(10.37). We can add a further superscript, n, with the same meanings as before. We can define P_*^{mn}, H_i^{kn}, E_j^{wn}, H_i^{k*}, E_j^{w*}, H_i^k, E_j^w by analogy with P_*^n, H_i^n, E_j^n, H_i^*, E_j^*, H_i and E_j. The constraints to be satisfied by T_{ij}^{kwmn} are then

$$\sum_i \sum_j \sum_k \sum_w T_{ij}^{kwm1} = P_*^{m1} \qquad (10.53)$$

$$\sum_j \sum_w T_{ij}^{kw11} + \sum_j \sum_w T_{ij}^{kw13} = H_i^k \qquad (10.54)$$

$$\sum_j \sum_w T_{ij}^{kw01} + \sum_j \sum_w T_{ij}^{kw03} = (r_i - 1)H_i^k \qquad (10.55)$$

$$\sum_i \sum_k \sum_m T_{ij}^{kwm1} + \sum_i \sum_k \sum_m T_{ij}^{kwm2} = E_j^w \qquad (10.56)$$

$$\sum_j \sum_w T_{ij}^{kw12} = H_i^{k2} \qquad (10.57)$$

$$\sum_j \sum_w T_{ij}^{kw02} = (r_i - 1)H_i^{k2} \qquad (10.58)$$

$$\sum_i \sum_k \sum_m T_{ij}^{kwm3} = E_j^{w3} \qquad (10.59)$$

$$\sum_j \sum_w T_{ij}^{kw14} = H_i^{k4} \qquad (10.60)$$

$$\sum_j \sum_w T_{ij}^{kw04} = (r_i - 1)H_i^{k4} \qquad (10.61)$$

$$\sum_i \sum_k \sum_m T_{ij}^{kwm4} = E_j^{w4} \qquad (10.62)$$

In the above,

$$H_i^k = H_i^{k*} - H_i^{k2} - H_i^{k4} \qquad (10.63)$$

and

$$E_j^w = E_j^{w*} - E_j^{w3} - E_j^{w4} \qquad (10.64)$$

The corresponding model equations are:

$$T_{ij}^{kw11} = K^1 A_i^{k11} B_j^{w1} H_i^k E_j^w \mathrm{e}^{-\beta^{w11} c_{ij}} \times \mathrm{e}^{-\mu^{w1} [p_i^k - q^w (w - c_{ij})]^2} \qquad (10.65)$$

$$T_{ij}^{kw01} = K^0 A_i^{k01} B_j^{w1} (r_i - 1) H_i^k E_j^w \mathrm{e}^{-\beta^{w01} c_{ij}} \qquad (10.66)$$

$$T_{ij}^{kw12} = A_i^{k12} B_j^{w1} H_i^{k2} E_j^w \mathrm{e}^{-\beta^{w12} c_{ij}} \qquad (10.67)$$

$$T_{ij}^{kw02} = A_i^{k02} B_j^{w1} (r_i - 1) H_i^{k2} E_j^w \mathrm{e}^{-\beta^{w02} c_{ij}} \qquad (10.68)$$

$$T_{ij}^{kw13} = A_i^{k11} B_j^{w3} H_i^k E_j^{w3} \mathrm{e}^{-\beta^{w11} c_{ij}} \times \mathrm{e}^{-\mu^{w3} [p_i^k - q^w (w - c_{ij})]^2} \qquad (10.69)$$

$$T_{ij}^{kw03} = A_i^{k01} B_j^{w3} (r_i - 1) H_i^k E_j^{w3} \mathrm{e}^{-\beta^{w01} c_{ij}} \qquad (10.70)$$

$$T_{ij}^{kw14} = A_i^{k14} B_j^{w4} H_i^{k4} E_j^{w4} \mathrm{e}^{-\beta^{w14} c_{ij}} \qquad (10.71)$$

$$T_{ij}^{kw04} = A_i^{k04} B_j^{w4} (r_i - 1) H_i^{k4} E_j^{w4} \mathrm{e}^{-\beta^{w04} c_{ij}} \qquad (10.72)$$

where

$$K^1 = P_*^{11}\Big/\sum_i \sum_j \sum_k \sum_w A_i^{k11} B_j^{w1} H_i^k E_j^w \mathrm{e}^{-\beta^{w11} c_{ij}} \times \mathrm{e}^{-\mu^{w1} [p_i^k - q^{w1} (w - c_{ij})]^2} \qquad (10.73)$$

to ensure that (10.53) is satisfied with $m = 1$,

$$K^0 = P_*^{01}\Big/\sum_i \sum_j \sum_k \sum_w A_i^{k01} B_j^{w1} (r_i - 1) H_i^k E_j^w \mathrm{e}^{\text{[illegible]}} \qquad (10.74)$$

to ensure that (10 53) is satisfied with $m = 0$,

$$A_i^{k11} = 1\Big/\Big\{\sum_j\sum_w K^1 B_j^{w1} E_j^{w'}\, e^{-\beta^{w11}c_{ij}} \times e^{-\mu^{w1}[p_i^k - q^{w1}(w - c_{ij})]^2}$$

$$+ \sum_j\sum_w B_j^{w3} E_j^{w3}\, e^{-\beta^{w13}c_{ij}} \times e^{-\mu^{w3}[p_i^k - q^{w3}(w - c_{ij})]^2}\Big\} \quad (10\ 75)$$

to ensure that (10 54) is satisfied,

$$A_i^{k01} = 1\Big/\Big\{\sum_j\sum_w K^0 B_j^{w1} E_j^{w'}\, e^{-\beta^{w01}c_{ij}} + \sum_j\sum_w B_j^{w3} E_j^3\, e^{-\beta^{w03}c_{ij}}\Big\} \quad (10.76)$$

to ensure that (10 55) is satisfied

$$B_j^{w1} = 1\Big/\Big\{\sum_i\sum_k K^1 A_i^{k11} H_i^{k'}\, e^{-\beta^{w11}c_{ij}} \times e^{-\mu^{w1}[p_i^k - \bar{q}^{w1}(w - c_{ij})]^2}$$

$$+ \sum_i\sum_k K^0 A_i^{k01}(r_i - 1) H_i^{k'}\, e^{-\beta^{w01}c_{ij}}$$

$$+ \sum_i\sum_k A_i^{k12} H_i^{k2}\, e^{-\beta^{w12}c_{ij}}$$

$$+ \sum_i\sum_k A_i^{k02}(r_i - 1) H_i^{k2}\, e^{-\beta^{w02}c_{ij}}\Big\} \quad (10.77)$$

to ensure that (10 56) is satisfied,

$$A_i^{k12} = 1\Big/\sum_j\sum_w B_j^{w1} E_j^{w'}\, e^{-\beta^{w12}c_{ij}} \quad (10\ 78)$$

to ensure that (10 57) is satisfied

$$A_i^{k02} = 1\Big/\sum_j\sum_w B_j^{w1} E_j^{w'}\, e^{-\beta^{w02}c_{ij}} \quad (10\ 79)$$

to ensure that (10 58) is satisfied,

$$B_j^{w3} = 1\Big/\Big\{\sum_i\sum_k A_i^{k11} H_i^{k'}\, e^{-\beta^{w13}c_{ij}} \times e^{-\mu^{w3}[p_i^k - q^{w3}(w - c_{ij})]^2}$$

$$+ \sum_i\sum_k A_i^{k01}(r_i - 1) H_i^{k'}\, e^{-\beta^{w03}c_{ij}}\Big\} \quad (10\ 80)$$

to ensure that (10 59) is satisfied,

$$A_i^{k14} = 1\Big/\sum_j\sum_w B_j^{w1} E_j^{w1}\, e^{-\beta^{w14}c_{ij}} \quad (10\ 81)$$

to ensure that (10 60) is satisfied,

$$A_i^{k04} = 1\Big/\sum_j\sum_w B_j^{w1} E_j^{w4}\, e^{-\beta^{w04}c_{ij}} \quad (10\ 82)$$

to ensure that (10 61) is satisfied, and finally

$$B_j^{w4} = 1\Big/\Big\{\sum_i\sum_k A_i^{k14} H_i^{k4}\, e^{-\beta^{w14}c_{ij}} + \sum_i\sum_k A_i^{k04}(r_i - 1) H_i^{k4}\, e^{-\beta^{w04}c_{ij}}\Big\} \quad (10\ 83)$$

to ensure that (10 62) is satisfied Clearly, building this model even theoretically is a mammoth and complicated exercise. However, it is not conceptually overcomplicated. The reader who has been able to follow the building of the aggregated quasi-dynamic model given by equations (10 27)-(10 37) and of the disaggregated but comparative static model given by equations (10 40), (10.43) and (10.44) should be able to recognize the basis of the composite model which has just been derived.

We can now begin to see how we have the basis for a dynamic residential location model by putting time labels on the variables which are the exogenous variables for the spatial interaction part of the model P_*^{11}, P_*^{01}, H_i^{k*}, E_j^{w*}, H_i^{k2}, H_i^{k4}, E_j^{w3} and E_j^{w4} Suppose we are interested in a time period of t to $t + T$. At time t we will have a distribution of housing stock $H_i^{k*}(t)$ and jobs $E_j^{w*}(t)$. At time $t + T$, the overall distribution becomes $H_i^{k*}(t + T)$ (as the net result of demolitions, changes and new building) and $E_j^{w*}(t + T)$, (as a result of changes in the economy). The total population may change in such a way relative to the housing stock that we may anticipate a change in r_i, say from $r_i(t)$ to $r_i(t + T)$. Suppose the total population changes from $N(t)$ to $N(t + T)$, then our main task is to divide $N(t + T)$ into locational categories, 1–4. We might assume that we have the following kinds of information

(1) initial population and distribution, $T_{ij}^{kwnm}(t)$,

(2) something on in-and-out migration from homes (including effects of births and deaths);

(3) something on in-and-out migration from jobs (including effects of births and deaths);

(4) something on household formation

Clearly, we are now beginning to involve the other components of a general model So that we are not taken too far out of our way, the remainder of this part of the discussion is postponed until the next chapter However, there is one other important aspect of model closure which must be mentioned, relating to the set of house prices p_i^k Once again, we can only raise the issue in a preliminary way and we return to it in Chapter 11

Recall that, in Chapter 6, we defined accessibility to activities X_j, for residents of i We can define

$$Q_i(X) = \sum_j X_j\, e^{-\beta^x c_{ij}} \quad (10\ 84)$$

where β^s depends on the activity X. Following this principle, we could define

$$Q_i^w(J) = \sum_j E_j^w e^{-\beta^w c_{ij}} \tag{10.85}$$

as the access to w-income jobs for residents of i, and

$$Q_i^w(s) = \sum_j W_j^s e^{-\lambda^{sw} c_{ij}} \tag{10.86}$$

as being the accessibility to services of type s for residents of i in income group w. It is also useful to define net residential density as

$$D_i = \sum_k H_i^k / L_i \tag{10.87}$$

where L_i is residential land in zone i. We might then hypothesize a relationship of the form

$$p_i^k = \lambda^k a(1 - e^{-bD_i}) + \sum_w g^w Q_i^w(J) + \sum_s \sum_w h^{sw} Q^w(s) \tag{10.88}$$

In other words, we hypothesize that type-k house prices in zone i are the product of a factor related to type of house, a factor which reflects increasing residential density (up to a limiting point) and the various accessibilities. Since the distribution of services, represented here as W_j^s, will be partly determined by the distribution of population (as discussed further below), then we see that population distribution, housing distribution and employment distribution all help determine p_i^k. This will clearly help us in formulating a general model. This particular formulation has been related explicitly to the comparative static form of the residential model. However, we can assume that the same hypotheses hold for the dynamic version. H_i^k in equation (10.87), for example, would simply be replaced by $H_i^{k*}(t)$.

It may now be useful to summarize the rather complicated argument of this section.

(1) We began with the simplest possible spatial interaction model as first used by Lowry. It was given by equation (10.1).

(2) We saw that this was inconsistent with respect to the assumptions which had been made about the distribution of employment E_j, and also that it did not contain a residential attractiveness term. These two considerations led to the model given in equations (10.3) and (10.4).

(3) It was then argued that a different assumption could be made about the constraints of the supply side, and that a distribution of houses H_i could be used as an alternative to the attractiveness factors W_i. This led to the doubly constrained model given in equations (10.6)–(10.8).

(4) We then saw how to develop a hybrid model, partly singly constrained with the mechanisms described in (2), and partly doubly constrained, using the mechanisms described in (3), the latter so that planners' policy constraints can be reflected. This model was given by equations (10.12)–(10.15).

(5) Different types of locational behaviour were then introduced, and we saw that this led to the quasi-dynamic model presented in equations (10.27)–(10.37).

(6) The next step was to introduce person type (especially income) and house type with price varying by type and location. A disaggregated version of the comparative static model described in (3) was built and presented in equations (10.40), (10.43) and (10.44).

(7) We then showed how another type of hybrid model could be built, one which removed the simplifying assumption of one worker per household. Linked models represented the residential location behaviour of heads of households and other workers. This model was presented in equations (10.48)–(10.52).

(8) Finally, we built a model which combined the features of (5), (6) and (7): a disaggregated quasi-dynamic model which distinguished heads of households and other workers. This was presented in equations (10.65)–(10.83).

(9) We then discussed what would be involved in incorporating the model thus developed in a general model of urban development, and how to close it with respect to the estimation of house prices. This discussion will be continued in Chapter 11.

10.6. MODELS OF RESIDENTIAL LOCATION USING ECONOMIC THEORY, ECONOMETRICS AND LINEAR PROGRAMMING

The economist brings a considerable number of tools of analysis to bear on the problem of building residential location models. First, and in some ways foremost, of these is the notion of land rent, which goes back to Ricardo or even Adam Smith. This is the notion that some plots of land are more productive than others, because of differences in fertility or location etc., and that the value of such differences can be collected by the landowner in the form of rent. In discussing rent, and associated concepts such as price, we shall follow Alonso (1964). He in turn follows Ratcliff (1949), quoting him as follows.

'It will be convenient to use the term "price" in its generic sense and to include under this term the market expression of contract rent, sales price and cost of ownership. These three values move together, though with unevenness. Sales price, the price that a buyer is willing to pay after considering alternatives, represents the present or discounted value of future rental values. The cost of ownership is a function of both contract rent and sales price; the owner must recognize a cost of occupancy that is at least as great as the rental income he might otherwise be receiving if he were to rent out his property, and no smaller than the total of interest on the investment, taxes, maintenance and depreciation, which total, in the long run, is in balance with the rental value' (Ratcliff, 1949, pp 347–348) We will, therefore, usually use 'price' (and not 'rent') in this sense Many economists have been primarily concerned with the price of residential land, and have separated this from the price of housing which goes with it Generally speaking, we shall use 'house price' to refer to the appropriate bundle of housing services which are, or can be, purchased, and which would, of course, include land price Thus, having established the nature of the concept of price, one of the prime interests of the economist is to predict it within a model.

In the typical economy, there are two kinds of animal: consumers (individuals and households) and producers ('firms'). The economist brings to bear on this situation his theory of consumers' behaviour, and the theory of the firm They involve notions of demand and supply and market processes which ensure that supply matches demand at an appropriate price.

In the very broadest terms, this works as follows Suppose there are three goods, 1, 2 and 3 (which might be housing, transport and 'all others', represented by 'money' possibly), and an individual r can buy quantities x_1^r, x_2^r and x_3^r at prices p_1, p_2 and p_3, out of his total income I^r, all of which is spent We assume that the individual's *preferences* can be recorded in the form of a utility function

$$U^r = U^r(x_1^r, x_2^r, x_3^r) \tag{10.89}$$

The individual then behaves so as to maximize this function, subject to his budget constraint

$$\sum_k x_k^r p_k = I^r \tag{10.90}$$

This maximization problem can be solved to give the quantity of each good purchased Formally, we can write

$$x_k^r = x_k^r(p_1, p_2, p_3, I^r) \tag{10.91}$$

since the quantity purchased will depend on each of the prices and the individual's income This is the individual's *demand* for the kth good Under suitable circumstances, we can aggregate over individuals r, to obtain the aggregate demand

$$x_k^* = \sum_r x_k^r \tag{10.92}$$

We now consider a firm s Its productive capability will be defined by its production function It will purchase factor inputs y_1^s, y_2^s and y_3^s, say (which may be labour, land and capital) at prices p_1^f, p_2^f and p_3^f If it produces good k, then its production function can be formally written as

$$x_k^s = x_k^s(y_1^s, y_2^s, y_3^s) \tag{10.93}$$

If it sells at price p_k, then its profit is

$$G = x_k^s p_k - \sum_e y_e^s p_e^f \tag{10.94}$$

It is ordinarily assumed that the firm behaves so as to maximize its profits, G in equation (10.94), subject to (10.93) as a constraint We can solve for the y_e^ss in term of p_k (and the factor prices which we assume given) to give

$$x_k^s = x_k^s(p_k, p_1^f, p_2^f, p_3^f) \tag{10.95}$$

We can, under suitable conditions, aggregate to give

$$x_k^* = x_k^*(p_k, p_1^f, p_2^f, p_3^f) \tag{10.96}$$

and do this for each good k Then, by setting supply equal to demand, (equations (10.92) and (10.96)), we can solve the resulting equation system for the prices p_k, for given factor prices, and then compute the quantities of the goods manufactured and consumed

This presentation is perhaps grossly oversimplified The reader who is dissatisfied with it can consult any standard economics text and find the discussion taken much further, more rigorously However, it does provide a basis for explaining the difficulties confronting the urban economist, and the methods which have been suggested to date for solving these difficulties Suppose, as suggested earlier, that in equation (10.89) for the individual consumer, x_1 represents 'housing', x_2 'transport' and x_3 'other'.

The first difficulty is concerned with making the institutional framework realistic, especially on the supply side If we consider that a number of firms exist, in a perfectly competitive economy, to produce housing, then this is far from being realistic The public sector is involved in a major way, and will be behaving according to mechanisms other than profit maximization All parties will operate in a framework of complicated legal constraints.

A second difficulty concerns the nature of the housing good itself The theories sketched earlier work best if goods can be produced in continuously varying amounts In the case of housing, there are substantial indivisibilities, and the existence of a previously developed stock housing is to some extent a capital good as well as a consumption good

These difficulties, though great enough, are nothing compared to the most important of all we saw in our earlier discussion in Section 10.4 that an important quality of housing as a good is its location, which determines connections to many other activities (and hence, in part, transport expenditure) and social and physical environment It is this difficulty which has most troubled urban economic theorists: how to give the theories of economic behaviour a spatial dimension. For the housing consumer, the essence of the difficulty can be seen in terms of equations (10.89)–(10.91). These must now be made *location specific*. Thus suppose L labels the location, then (10.89) becomes

$$U^{rL} = U^r(x_1^{rL}, x_2^{rL}, x_3^{rL}, I^r) \tag{10.97}$$

and equation (10.91) will become

$$x_k^{rL} = x_k^{rL}(p_1^L, p_2^L, p_3^L, I^r) \tag{10.98}$$

since the prices (at least for housing and transport) must now be considered to vary with location. *Formally the situation is even worse, as demand at a location will also be a function of prices at all other locations.* Alonso describes the ensuing difficulties very well:

'Consequently, the demand curves for the same individual will vary with his location. Since part of the problem of finding the market solution consists of finding individual locations, not knowing the locations of individuals we would not know which of their demand curves to use to build up our market demand curve.' In other words, it is no longer possible to aggregate to obtain equation (10.92). There is a corresponding difficulty on the supply side. It means that we can no longer model the market operation by setting supply equal to demand as we did in the earlier analysis.

Our first task therefore is to explore how economists have built a spatial dimension into their analysis and then we can see what this means for building models.

The basis of the current solution to the problem has its origins in the work of von Thunen (1826) on agricultural land use. This model has been formally developed by Dunn (1954) and Isard (1956). In the whole of this part of the argument, we are restating the work of Alonso (1964). We begin by explaining the structure of the argument in very broad terms, and then connecting back to our earlier discussion. Even the more detailed exposition will be a considerable simplification of Alonso's work and the interested reader should consult his 1964 book directly.

The essence of the argument connects to the notion of 'rent' and runs as follows. We have a land area which has a market at its centre and is otherwise used for agriculture. A farmer's profits result from the sale of his crop, less his factor costs, *less his transport costs*. The last named feature means that, for a given crop, land nearer the centre can command a rent relative to more distant land. The market would ensure that the farmer paid this rent to the landowner. In a competitive situation, any particular farmer would be able to offer a landlord at a location a rent, his *bid-rent*, for that location. The landlord then lets to the highest bidder. The actual land price surface is then the envelope of the whole set of bid-rent surfaces.

Alonso has extended the bid rent concept to be applicable to the urban firm and to the resident. In the case of the firm, land rent payable (henceforth called land price) is one of the cost terms subtracted from revenue in the estimation of profits. It is then possible to estimate a bid-price for a given location for a given level of profits. There is an equivalent procedure for the resident: land price now appears as one of the terms in the budget equation, and we can estimate a bid-price at a location for a given level of utility. Thus, there are now three competing users of land: farmers, firms and residents. Their bid-prices have been constructed in such a way as to be comparable. Landlords at each location now charge land prices which coincide with the highest bid made.

Once this conceptual breakthrough has been achieved, the rest of Alonso's analysis is concerned with showing how the market-clearing process works so that an appropriate equilibrium position is achieved.

Alonso has to make a number of simplifications in order to be able to complete his analysis. The most important is to assume that jobs are concentrated in the centre of the city, so that 'distance from the CBD' serves as an adequate measure of location, and commuting costs as a measure of transport costs. Also, he is primarily concerned with *land* rather than the whole *bundle* of housing services.

We can now indicate how these ideas work with the simplest equations which were introduced earlier. However, we make no distinction between farmer and firms. We also try to avoid the restrictive centrality assumption. Further, we distinguish between the land-rent component of housing price and the rest. Let $x_1^r(L)$ be the quantity of housing which could be purchased by individual r at location L, at price $p_1^H + p_1^R(L)$, where p_1^R is the rent component; let $x_2^r(L)$ be the corresponding expenditure on transport at price $p_2(L)$, and x_3^r at price p_3 (not assumed dependent on location), the rest. Then equations (10.89) to (10.91) (assuming for the moment that we need only consider prices at one location) become

$$U^r(L) = U^r(x_1^r(L), x_2^r(L), x_3^r, I^r) \tag{10.99}$$

$$x_1^r(L)(p_1^H + p_1^R(L)) + x_2^r(L)p_2(L) + x_3^r p_3 = I^r \tag{10.100}$$

and

$$x_1^r(L) = x_1^r(p_1^H, p_1^R(L), p_2(L), p_3, I^r) \tag{10.101}$$

$$x_2^r(L) = x_2^r(p_1^H, p_1^R(L), p_2(L), p_3, I^r) \tag{10.102}$$

$$x_3^r(L) = x_3^r(p_1^H, p_1^R(L), p_2(L), p_3, I^r) \tag{10.103}$$

The last equation shows that x_3^r will, of course, be a function of L. For a given value of U^r, say U_0^r, equation (10.99) determines an indifference surface in the variables x_1, x_2, x_3 and, perhaps one should say, for we are not yet sure what kind of animal it is, L. We can then use equation (10.100) to estimate a bid rent, $p_1^R(L)$, assuming other prices to be given, and x_1, x_2 and x_3 being determined when U_0^r is given. Thus, for each individual r, we have a bid-rent at each location for each level of utility U_0^r

For the firm, land is a factor of production and so its price, $p_2^s(L)$ becomes a function of location, L The price at which it can sell its product may also be location dependent, as may the cost of its labour, and so the equivalent of equations (10 93) to (10 95) is

$$x_k^s(L) = x_k^s(y_1^s(L), y_2^s(L), y_3^s) \tag{10.104}$$

is the production function, and the profit level is

$$G = x_k^s(L)p_k(L) - y_1^s(L)p_1^s(L) - y_2^s(L)(p_2^B + p_2^R(L)) - y_3^s p_3 \tag{10.105}$$

Note that, without loss of generality, we can use the same notation, $p_2^R(L)$ for the land rent part of the buildings/land cost, as for residents p_2^B is the buildings part of the cost Then, a given level of profits G_0, determines a surface in $y_1^s(L)$, $y_2^s(L)$, y_3^s and $p_2^R(L)$. Thus, for a given $y_1^s(L)$, $y_2^s(L)$ and y_3^s, we can determine the bid-rent $p_2^R(L)$ which would give a level of profits G_0.

The market-clearing process now involves (i) for each L, the landlord at L searching bids in order to select the highest, and (ii) for each resident and each firm in turn, a search among their bids to find the available (acceptable) one which gives the highest utility or the highest profit. (This market clearing process in land rent takes place within an environment where other aggregate supply demand relationships, for example on x_3^*, have been determined and satisfied) One of Alonso's substantial achievements was to ensure that, under his restricted conditions, such a market-clearing aggregation could be defined and would operate in a unique way The advantage of the more general formulation presented here is twofold. no assumption of centrality is needed and, in principle, we can consider L to label plots of given, finite size though there is then a corresponding difficulty in giving $x_1^r(L)$ or $y_2^s(L)$ a meaning (One solution is to make L a double label, L_1, L_2, where L_1 represents the total holding of a given landowner, and L_2 the whole set of sub-divisions within it These sub-divisions will overlap, of course, and the final accepted sub-division will be determined by the bidding process)

Perhaps the single main point of Alonso's work, no matter how it is formulated, is that the mechanism of the market-clearing process has to be made explicit before aggregation can be satisfactorily carried out Other authors, who have started with more or less the same tools as Alonso, do not seem to have been as successful in this respect Wingo (1961 p 85) develops a demand function for individual consumption of residential land consumption, but then assumes that it can be aggregated in some simple way, more or less as equation (10 96) was derived from (10 93) Muth (1961, 1969) also begins with the same tools and develops a set of equations to describe the equilibrium position of the individual in the housing market However, he avoids the aggregation issue by simply discussing different kinds of behaviour for different kinds of people (particularly by income and race) He also devotes a considerable amount of effort to a study of the equilibrium of housing producers Yet another approach can be found in the work of Mills (1969). He avoids the aggregation problem by concentrating on an aggregate production function approach to the analysis problem

The reader can pursue these various approaches in whatever detail he likes, beginning with the references cited above. There he can also find a guide to earlier work by urban economists, an account which will not be repeated here Alonso's work seems to be the most fully developed at the present time, and it provides a micro-analytical framework against which many model-building attempts can be compared (including the spatial interaction models which we will pursue in Section 10 8) There is, however, perhaps one major disappointment with all this work based on economic theory: with one possible exception (the Herbert- Stevens model to be discussed further below), no very exciting operational models have developed from the work. Many insights and qualitative analyses have been obtained, but few effective models. One possible reason for this is that most economists in the field, having established their theoretical framework, then resort to essentially linear or log linear econometric models for their empirical work Alonso (1964, p 126), for example, presents only one empirical equation, estimated for Philadelphia for the years 1950 1952, as follows·

$$pq = -222\ 65 + 0\ 4357(\pm 0{\cdot}1275)y - 90\ 107(\pm 22{\cdot}703)t$$

$$R = 0\ 69 \qquad S = 375{\cdot}67 \tag{10 106}$$

where

p = price per square foot, in dollars
q = number of square feet per family
y = family income in dollars
t = distance from centre of city in miles

This equation is typical of the genre Many more examples can be found in the book by Muth (1969) and some also in Mills (1969) The main point to make here is that these econometric models seem unduly simple relative to the theory which underpins them For this reason, they will not be pursued any further here, though, of course, this is far from saying that they are without interest The reader is referred to the original sources

The possible exception referred to earlier was the Herbert Stevens (1960) model of residential location We examine this now, and it gives us the opportunity to explore linear programming as a technique for location modelling The discussion fits appropriately into this section as the underlying basis of the model is not very different from Alonso's, and indeed the work was carried out in the same stable in Philadelphia at about the same time.

The Herbert Stevens model, in a notation as consistent as possible with other models in this book, can be written as follows. Maximize

$$Z = \sum_i \sum_k \sum_n T_{i\bullet}^{kn}(b^{kn} - p_i^{kn}) \tag{10.107}$$

subject to

$$\sum_k \sum_n s_{kn} T_{i\bullet}^{kn} \leqslant L_i \tag{10.108}$$

and

$$\sum_i \sum_k T_{i\bullet}^{kn} = P^n \tag{10.109}$$

where

$T_{i\bullet}^{kn}$ = number of type n people locating in zone i in a house of type k;
b^{kn} = the residential budget allocated to the purchase of a residence of type k by a household of type n;
p_i^{kn} = cost to a household of group n of a type k house in i, *exclusive* of site cost;
s_{kn} = area utilized by a household of type n if it has a type k home;
L_i = area of land available for residential use in zone i,
and
P^n = population of type n to be allocated (in households).

No interaction, such as the journey to work, is represented explicitly in this model, but the cost of housing, p_i^{kn}, is to be interpreted as including the costs of all trips made by the household. Thus, the term $b^{kn} - p_i^{kn}$ in the maximand represents the bidding power for site rent, which is clearly equivalent to Alonso's bid price, and so the model maximizes bid prices subject to constraints on land availability and finding everyone a house. An analysis of the dual shows that if bid-rent is maximized, actual rent paid is minimized. (Note that by adjusting L_i after a run of the model, zoning policy type constraints could be handled, and thus carry out the same function as the hybrid spatial interaction model.) Relative to Alonso's micro-analysis, individual households have been grouped by type, and the market clearing mechanism has been replaced by the linear programming action. Thus, as the authors say in their original paper in relating their model to Alonso's: 'His method is difficult to apply directly, but we feel that our linear programming model provides an analogous approach that is both acceptable and workable'. The difficulties with the model in empirical terms are concerned with the estimation of the quantities b^{kn}, p_i^{kn} and s^{kn}. There are also theoretical problems which will be discussed below in Section 10.8. However, for the moment we acknowledge it as an elegant and relatively simple mathematical expression of Alonso's theory.

10.7. OTHER APPROACHES TO RESIDENTIAL LOCATION MODELLING

There are a number of techniques which we have not yet considered, in particular those based on ecological analysis and on the simulation concept. The first of these can be dealt with briefly. No fully formalized mathematical model has been developed by the ecological school. Perhaps the nearest approach to this was the model used to forecast land use in the Chicago Transportation Study (1960). All land uses were considered by the model, but the essence of the residential location mechanism was as follows. An assessment was made of the residential capacity of each zone in the study area. These assessments were then ranked according to distance from the centre and the existing percentage of capacity for each area was plotted, as shown in Figure 10.2 (Hamburg and Creighton, 1959; Wingo, 1961.)

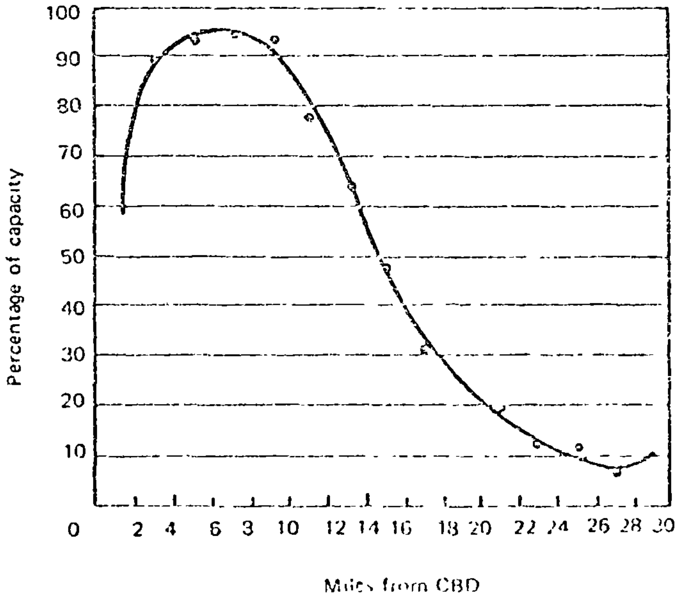


Figure 10.2 Residential capacities in Chicago. (Reproduced from J. R. Hamburg and R. L. Creighton, *AIP Journal*, 25, Figure 3, p. 70, (1959). Reprinted by permission of the *Journal of the American Institute of Planners*.)

An estimate was then made of the shift in this curve caused by future residential development to give a result as depicted in Figure 10.3 (Hamburg and Creighton, 1959, Wingo, 1961.) This gives a proportion of new development in the forecast period which can be applied to each zone, (according to

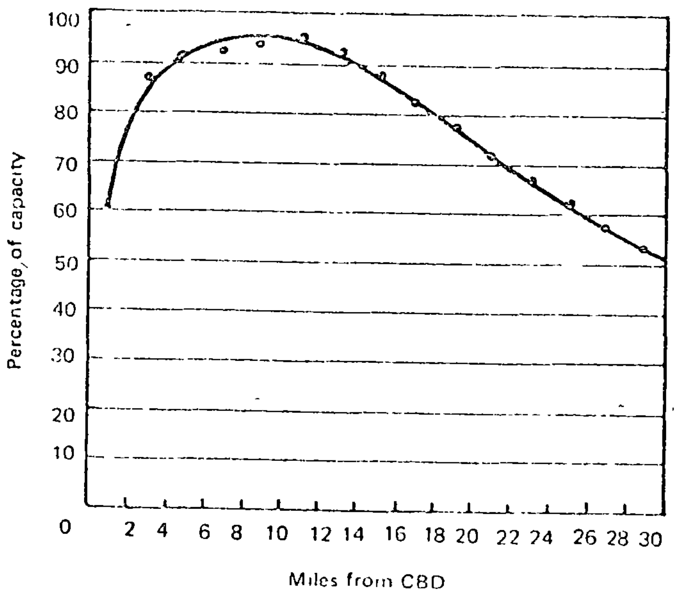


Figure 10.3 Shift of residential capacity curve (Reproduced from J. R. Hamburg and R. L. Creighton, *AIP Journal*, 25, Figure 5, p. 71 (1959). Reprinted by permission of the *Journal of the American Institute of Planners.*)

its distance grouping) These estimates were then modified 'according to staff judgement' and in that sense the model was not a fully formal mathematical one. The model is not strictly an ecological one, but the reader will see that it implies modified concentric ring development, and it was probably no accident that this kind of model was developed in Chicago!

A *simulation* model tackles the problem of estimating residential development by asking 'where, and with what probability, will the next development take place?' It is particularly suitable if the time focus of the model is marginal change. As an example, we will consider the model developed by Chapin and co-workers at the University of North Carolina. This account is based on a paper by Chapin and Weiss (1968) which summarizes six years work on the project.

The model-building exercise itself is preceded by an investigation of the factors which are most important in residential development. Seven strong independent variables emerged from this analysis, three of the 'first rank', four of the 'second'. They were.

First rank:
  marginal land not in use
  accessibility to work areas
  assessed value.
Second rank
  travel distance to nearest major street
  distance to nearest available elementary school
  residential amenity
  availability of sewerage

These factors were identified by a multiple regression analysis, presumably with 'new residential development' as the dependent variable. Particular attention was paid to the time sequence of development. In the complicated decision-making sequence which generates development, certain 'priming' operations were identified which were of particular significance.

For the operation of the model itself the study area is divided into zones; in this case a square grid was used in which zone size was a mere 23 acres. For each zone, an attractiveness index was calculated using the factors described earlier. The total amount of development to be distributed is assumed given for each time period. Development is then allocated to zones (in fact to fractions of zones) using a random number generator biased so as to preserve the attractiveness distribution. The attractiveness indices are then recalculated for the next time period. Attractiveness is correspondingly increased if priming operations have been carried out. The prediction of the likely pattern of residential development in the study area is then made by taking the average of a large number of simulation runs.

10.8. COMPARISON OF THE DIFFERENT APPROACHES TO RESIDENTIAL LOCATION MODELLING: CONCLUDING COMMENTS AND SOME INTEGRATION

We have discussed a range of techniques for residential location modelling and it is now worthwhile to comment and to explore the relations between them.

The spatial interaction models have developed from a simple hypothesis that workers located in a gravity-like way around work-places, to something which incorporates a considerable degree of desirable detail. Indeed, in its most disaggregated form, as discussed above, it is almost misleading to call the model a spatial interaction model: what it actually does is to assign households by type to houses by type, taking into account, *among other things*, the house-workplace interaction. According to the parameter values, the 'other things' may be more important than the interaction; it is a matter of empirical test. It may be better, then, to call the model a *residential assignment model*, though experience with the 'gravity model' shows that the historical name is usually preserved even when the model is perceived differently. Perhaps the most important feature of the model is that it attempts to make the model's predictions maximally consistent with known or assumed information and in this respect it is useful to acknowledge its entropy maximizing origins. This is one of its great strengths relative to the econometric models which typically follow analyses in economic theory. In this respect it is interesting to take the argument a stage further and ask whether the model in any sense contradicts the analyses of economic theorists. There is some indication that it does not. The estimates of the model can be viewed as statistical averages over the behaviour of individual households seeking locations. There seems to be no inherent reason why each one should not be maximizing a utility function. The assignment mechanism of the model can then be interpreted as the model's market-clearing mechanism. In the main model equations, supply-side variables H_i^k and the prices p_i^k appear explicitly. We indicated how the model could be closed by adding an equation which would estimate p_i^k. It would be possible to explore a variety of forms of this equation, including one which reflected measures of 'demand pressure'. We could also develop a supply-side model for the H_i^ks (total or incremental, preferably the latter of course). This all reinforces the interpretation of the model as an assignment model representing a market-clearing operation. It has one substantial advantage over the usual economist's model in this respect. Since we can choose whatever mechanism we like for estimating (or, for example, determining from a planning policy) H_i^k, and indeed p_i^k, machinery other than perfect private market mechanisms can be investigated.

We should also remark that the spatial interaction model is easily extendable. We have seen how the disaggregated model was developed from a very much simpler and less realistic one. With the same ease, effects of mechanisms which were considered important could be incorporated. For example, we might want to build in access to schools and shops, and to incorporate an index of their quality. (Or we might consider that these effects had already been incorporated through the p_i^ks.) Equally, if we can identify household behaviour which is predicted by micro-economic theory, which is inconsistent with the model, then it should be possible to extend or amend it.

At this point it is useful to compare the spatial interaction model with the corresponding linear programming model. The spatial interaction model estimates T_{ij}, the linear programming model T_{i*}^{kn} (where n could conceivably be taken as [illegible]). The first point to note is that in the LP model there will only be as many non-zero T_{i*}^{kn}s as there are constraints. This is a fundamental theorem of linear programming and is presumably what Harris (1962) was referring to when he wrote '... there is no cross-hauling, whereas this is prevalent in [illegible] economic action'. This means that values of k and n (and possibly i)

represent very finely defined groups, the LP model is unlikely to reflect the full variety of actual behaviour. In this sense, in particular, it is possible that the spatial interaction model is more satisfactory. This does lead to the interesting thought that it would be possible to develop an assignment model version (this time it would be misleading to call it a spatial interaction model) of the Herbert–Stevens model which would overcome this difficulty. Suppose we interpreted equations (10.107)–(10.109) as constraints and derived the appropriate entropy maximizing model. This would give (and note the expected sign of μ, as Z is being 'maximized')

$$T_{i*}^{kn} = A^n P^n \mathrm{e}^{-\beta_i s^{kn}} \mathrm{e}^{\mu(b^{kn} - p_i^{kn})} \tag{10.110}$$

and

$$A^n = 1 \Big/ \sum_i \sum_k \mathrm{e}^{-\beta_i s^{kn}} \mathrm{e}^{\mu(b^{kn} - p_i^{kn})} \tag{10.111}$$

to ensure that constraint (10.109) was satisfied, and β_i and μ would be calculated from

$$\sum_k \sum_n s^{kn} A^n P^n \mathrm{e}^{-\beta_i s^{kn}} \mathrm{e}^{\mu(b^{kn} - p_i^{kn})} = L_i \tag{10.112}$$

and

$$\sum_i \sum_k \sum_n A^n P^n \mathrm{e}^{-\beta_i s^{kn}} \mathrm{e}^{\mu(b^{kn} - p_i^{kn})} = Z \tag{10.113}$$

to ensure that constraints (10.108) and (10.107) respectively were satisfied. A^n, β_i and μ would have to be calculated iteratively, and sub-iterations would be needed to solve the equations for β_i and μ. Presumably Z could be increased until its upper limit was reached. As a conjecture, we could suppose that this is achieved by letting μ tend to some limit. Thus, the assignment model principle may provide an alternative to the Herbert–Stevens model which has an appropriately realistic 'blurring' (or Harris's 'cross hauling') in its predictions. This discussion is taken several steps further in Senior and Wilson (1973).

The spatial interaction model also connects to the simulation principle. Essentially, in a simulation model, we need the probability, π_i say, of someone locating in zone i. In the simplest spatial interaction model, estimating T_{ij}, we could take

$$\pi_i = T_{i*}/T_{**} \tag{10.114}$$

where the asterisks as usual represent summation. The definition could be embellished in various ways by disaggregation. Note, however, that particular care must be taken to distinguish between the probabilities of households locating in a zone and of new housing stock being assigned to a zone. The simulation model example given earlier was concerned with the second, and the use of the spatial interaction model in the manner suggested above would be most suitable for the first

In general, then, we favour the spatial interaction model (perhaps more properly called a household assignment model) as the most flexible and general of the models presented, though clearly it should be connected as closely as possible to a basis in micro-analysis The arguments of Senior and Wilson (1973), based on developments of the entropy-maximizing version of the Herbert–Stevens model, suggest that close links are possible We should, however, note one of its disadvantages (particularly relative to the econometric model) It is often very difficult to estimate the parameters of the model, and even when methods can be found, they tend to use much computer time Much work has been done in this field recently, though, and we shall return to the topic in Chapter 12

10.9. UTILIZATION OF SERVICES: BASIC HYPOTHESES

In this section, we consider how people utilize a *given* distribution of services That is, we are making an assumption analogous to that in an earlier discussion of residential location, that we can model what people do apart from the task of modelling housing supply, services and so on

The term 'services' covers a wide variety of activities We shall be concerned with five main categories: retail, personal, educational, health and recreational services Broadly speaking, retail services are concerned with all shopping, personal services with such things as banks and solicitors, education mainly with schools, but also other establishments, health with doctors, health centres and hospitals, recreation with anything from cinemas and dance halls to open space There are, of course, many points of detail which are being neglected do we treat barbers and estate agents as shops or personal services, for example? But this need not concern us for the present and we assume that some suitable classification has been made

Perhaps one of the main distinctions which can be drawn between different types of services is the regulation or otherwise of their use The use of most shopping, personal services and recreational services is not regulated people are free to shop where they choose However, access to schools and health centres is at least partly regulated We shall be concerned mainly with unregulated services (though, as we shall see, we may be constrained by a variety of price and access considerations) The problem of modelling people's behaviour in regulated systems is an important one a problem to which a variety of operational research techniques is often appropriate and we shall return to it from time to time

For unregulated services we may, then, assume that people are interested in satisfaction (or attractiveness) in relation to access and make their utilization decisions accordingly. For regulated services the same considerations apply, but the utilization decision is made by the regulating authority through its regulations Clearly, in the case of unregulated services, we have a spatial interaction problem We consider the spatial interaction model approach in the next section and then a number of alternatives in Section 10.11.

10.10. SPATIAL INTERACTION MODELS OF THE UTILIZATION OF SERVICES

The introduction to this section can be very brief since the reader will already be familiar with the spatial interaction model for shopping, which was used as a first example of a model in Chapter 4. This was a model of the form

$$S_{ij} = A_i(e_iP_i)W_j\,e^{-\beta c_{ij}} \tag{10.114}$$

where

$$A_i = 1\Big/\sum_j W_j\,e^{-\beta c_{ij}} \tag{10.115}$$

to ensure that

$$\sum_j S_{ij} = e_iP_i \tag{10.116}$$

In this model S_{ij} is the flow of retail expenditure from residential zone i to shopping zone j, e_i is mean expenditure per head in zone i, P_i is the population of zone i, W_j is the attractiveness of shops in j, c_{ij} is the cost of travel from i to j, and β is a parameter We need not repeat the justification of this model here but, below, we will set it in its proper historical context and consider various ways of developing it further We can note that, generally for unregulated services, we might expect the flows to satisfy a singly constrained spatial interaction model of the form

$$S_{ij} = A_iO_iW_j\,e^{-\beta c_{ij}} \tag{10.117}$$

where

$$A_i = 1\Big/\sum_j W_j\,e^{-\beta c_{ij}} \tag{10.118}$$

where O_i is the demand for the service generated by residents of zone i, and the other terms are the appropriate analogues of the shopping situation for the service under consideration.

We can now proceed as follows. We will begin with a discussion of the retail model, first setting it in its historical context, then considering some associated measurement problems and some further developments.

A spatial interaction model was first used in the study of retailing by Reilly (1931). To be more precise, spatial interaction hypotheses were used by Reilly to delimit retail market areas. He assumed that a city i attracted people to its shops in relation to a factor P_i/d_i, where d_i was the distance between the person and the city, and P_i the city's population. Then, for a person faced with the choice of two cities (which might be two of our zones), if $P_i/d_i > P_j/d_j$, the person is in the market area of i. Clearly, this process delimits non-overlapping market areas and therefore Reilly's work does not represent a spatial interaction *model* in our sense. If Reilly had interpreted his hypothesis just slightly differently, that P_i/d_i represented the *probability* of the person going to i, then this would lead to a flow model of the form

$$S_i = K\frac{P_i P}{d_i} \tag{10.119}$$

where P is the number of people at the location of the original population. More generally, we could write

$$S_{ij} = K\frac{P_i P_j}{d_{ij}} \tag{10.120}$$

using an obvious notation, as an unconstrained spatial interaction model. So the idea of a spatial interaction model is implicit in Reilly's work, though he used the concepts to delimit non-overlapping market areas. (When there are more than two cities, other problems with Reilly's formula *do* give rise to some overlap in market areas, see Huff, 1964.) Reilly's work was further developed and refined by Converse (1949), though the basic limitations remained.

Although many people worked on spatial interaction models in the next two decades (and the early transport studies of the 1950s in the U.S.A. must have looked at shopping trips), no progress was made with retail models until the work of Huff (1962, 1964). In terms of our earlier notation, Huff proposed that

$$p_{ij} = \frac{W_j c_{ij}^{-n}}{\sum_j W_j c_{ij}^{-n}} \tag{10.121}$$

represented the probability that a resident of i would shop in j, where W_j was measured as shopping centre size, in selling area for a class of goods (thus implying that the model was used separately for different types of good) and

c_{ij} as travel time. Then

$$S_{ij} = C_i p_{ij} \tag{10.122}$$

where C_i is the number of consumers at i. Market areas were then expressed in terms of probability contours using (10.121).

Lakshmanan and Hansen (1965), working independently, used a traffic model modified to make it singly constrained, to produce the model which we have been using as an example and which was presented as equations (10.114) and (10.115) above. The reader can easily check that, with consistent definitions of the variables, the models of Huff and Lakshmanan and Hansen are equivalent. It is an interesting point in relation to the nature of scientific discovery that the same model could be derived from such different viewpoints.

Lowry (1964), again working independently, also used a spatial interaction model. As nearly as possible in our notation, it can be written

$$S_{ij}^k = gP_i c_{ij}^{-n_k} \tag{10.123}$$

$$S_j^k = \sum_i S_{ij}^k + h^k E_j \tag{10.124}$$

where S_{ij}^k is the flow for sector k, P_i is the population of i, E_j employment in j, S_j^k is total activity in j, g and h^k are constants and n_k is a parameter. As with Lowry's residential location model, the first equation represents a very primitive unconstrained spatial interaction model with no attractiveness factor. As Lowry was mainly interested in retail sector employment generated by some population distribution, P_i, his units were 'retail employment in sector k generated'. The important feature of Lowry's model for this discussion (other features will be noted in the section on the Lowry model in the next chapter) arises from the second equation and the term $h^k E_j$, which represents retail activity generated from employment. Since we can casually observe employed people shopping from work in lunch hours and so on, it seems sensible to incorporate such a term, and it is strange that no workers since Lowry have bothered to do so. Lowry assumed that only employees working in the same zone as the shopping centre shopped here. More generally, we would expect a spatial interaction term. We can also note that several authors have run the model separately for different types of good. Henceforth, we will assume that this should be done if possible, but we will not bother to carry the equivalent of Lowry's k subscript explicitly; any model written down is for some 'type of good'.

Although much work has been done on shopping models since the seminal work of the authors referred to above, no new conceptual breakthroughs have been achieved. Different measures have been used of distance and attractiveness, together with a variety of attenuation functions. However, a number of possible further developments can be briefly explored.

Our starting point should perhaps be a model represented by the following equations (using the usual, or an obvious, notation):

$$S_{ij}^1 = A_i^1(e_i^1 P_i)W_j^1\,\mathrm{e}^{-\beta^1 c_{ij}} \tag{10.125}$$

$$S_{ij}^2 = A_i^2(e_i^2 E_i)W_j^2\,\mathrm{e}^{-\beta^2 c_{ij}} \tag{10.126}$$

$$A_i^1 = 1\Big/\sum_j W_j^1\,\mathrm{e}^{-\beta^1 c_{ij}} \tag{10.127}$$

$$A_i^2 = 1\Big/\sum_j W_j^2\,\mathrm{e}^{-\beta^2 c_{ij}} \tag{10.128}$$

and

$$S_{*j} = \sum_i S_{ij}^1 + \sum_i S_{ij}^2 \tag{10.129}$$

The superscript 1 denotes flow from 'home'; 2, flow from 'employment'; e_i^1 and e_i^2 now mean expenditure per head and per employee on the appropriate retail goods. It should not be difficult in principle to test such models, though there may be data difficulties. Usually, we would expect to have $\beta^2 > \beta^1$. β^2 may be so high that Lowry's original assumption could be considered sound, but that is a matter of empirical investigation. Empirical test would also show whether or not W_j^2 could be taken to be the same as W_j^1.

The next obvious development arises from a consideration of behaviour in relation to the distance term. All the considerations of mode and mode availability of Chapter 9 for transport flows in general apply to shopping trips. There are many ways in which we could disaggregate by mode. The simplest is to add a modal superscript, k as in Chapter 9, to equations (10.125)–(10.128) and include a k-summation in equations (10.129). However, this causes a number of problems. We would have to estimate such quantities as e^{1k} and e^{2k}, and the model, like the transport model without a person-type term, would not be able directly to take account of car availability. Indirectly, this would be taken into account through what would be, in effect, a trip-end modal split model. So some kind of person-type superscript is needed also. This could simply be added. However, if we add this, together with the mode superscript, we are then obtaining turnover by aggregating over a large number of modal runs. This may be all right in theory, but in practice, there will rarely be the data to sustain it.

The principle to be used in applying spatial interaction model concepts to other sectors should by now be clear to the reader. A singly constrained spatial interaction model has been used by Morrill and Kelly (1970), for example, to study the flow of patients to hospitals.

It is less likely that spatial interaction modelling methods will be applicable to regulated service systems. For such a system, the mechanism represented by the model will be the regulation mechanism. If a local authority builds its primary schools so that 'no child has to walk more than half a mile to school', and so that the 'market areas' do not overlap, then these rules represent the model. Such cases can be tackled *ad hoc* in a straightforward manner. One particular example of a regulated system is worth noting, however. Suppose O_i is the demand for a service in i, and D_j is the 'supply'. Then T_{ij} may be chosen (i.e. the regulations arranged so that this happens) so that

$$\sum_j T_{ij} = O_i \tag{10.130}$$

$$\sum_i T_{ij} = D_j \tag{10.131}$$

and some objective function is optimized. For example, travel cost

$$C = \sum_i \sum_j T_{ij} c_{ij} \tag{10.132}$$

may be minimized. This is the 'transportation problem' of linear programming. This may be considered a special case of a spatial interaction model—doubly constrained and with parameter $\beta \to \infty$.

We should note that alternative principles of spatial interaction modelling, and especially those using 'intervening opportunities' concepts, could be and have been used for the service sector. Some interesting work has been done along these lines, but no fundamentally new procedures have emerged (see Harris, 1964, Cordey-Hayes and Wilson, 1970).

Finally, we should note that we have contented ourselves with writing about total cross-section models and have not developed marginal dynamic models equivalent to those in the corresponding residential location section. This assumption seems to be justified: in effect, we are arguing that people respond fairly quickly to changes in service sector supply, and therefore an equilibrium approach is justified.

10.11. OTHER APPROACHES TO SERVICE UTILIZATION MODELLING

A number of other approaches to service-sector modelling are possible. Many of them are conveniently reviewed in a National Economic Development Office publication (Distributive Trades E.D.C., 1970). As with residential location, some economists have studied the problem (for example, see Bacon, 1970), but as yet no operational 'economic theory models' have been developed. The main alternative approach to spatial interaction modelling is based on the concept of market area, usually non-overlapping, and some subset of the concepts associated with central place theory. The resulting models are usually unsatisfactory simply because we know that market areas *do* overlap and they will not be considered any further here. However, we should note that some of the concepts of central place theory, notably those associated with the idea of 'hierarchy', are obviously useful and have not yet been successfully incorporated into spatial interaction models, though there seems no reason why this should not be possible. It would be interesting to see central place theory rewritten, as it affects retailing and other services, with spatial interaction model concepts replacing the non-overlapping market area concepts but the other main features being retained. It is tempting to set this as an exercise for the reader, but perhaps it is at least a Ph.D.-size problem! One final qualification must be added to these remarks: there may be some service sectors which would probably be regulated ones, where the market areas do *not* overlap, and some other techniques may then be appropriate, as noted in the previous section.

10.12 THE SUPPLY SIDE

In relation to the discussion of preceding sections on the location of person activities, we must now discuss the location of the buildings and associated facilities which provide the infrastructure of these activities. In particular, we are interested in the supply of housing, services and industry, the last two of which give the spatial distribution of jobs.

Different principles are involved in supply-side modelling. With population activity, both total cross-section and marginal models were appropriate in different circumstances. Much longer time periods are associated with buildings and generally it will only make sense to build marginal models. Furthermore, decisions about supply-side units are made by organizations. Typically, there are far fewer organizations than individuals and therefore the style of model used for population activities, which is based on underlying statistical averaging assumptions, will not usually be appropriate. This is one reason why the relevant theory, for example of industrial location, is usually couched in micro-analytical terms based on the theory of the firm. This often means that the only *comprehensive* models which we can set up from this sort of basis are econometric models which have properties consistent with micro-analytical theory. A possible exception to this general rule is the housing sector, where relatively large numbers of new houses are built. In many cases, however, the decisions involving a large proportion of these houses are made by a small number of organizations, for example, by local government authorities.

The obvious implication of the above discussion is that comprehensive supply-side models are difficult to build. There is often a saving grace, however: the estimates of change in housing supply, services and jobs have to be made and these are precisely the variables which are to a greater or lesser

extent under the planners' potential control, either directly, as with the provision of public authority housing or services, or indirectly through planning permission or zoning controls. Thus, on many occasions, a model as such is not needed, only a specification of a range of alternative plans.

Given the implications of this introductory discussion, and particularly the qualification of the preceding paragraph on the need for supply-side models, we will content ourselves in the rest of this section simply by outlining briefly some of the models which have been developed. The discussion of underlying hypotheses will be postponed.

Firstly, we consider a very simple model of marginal housing supply, due to Hansen (1959). He assumes that in the time period under consideration, there is some given quantity of housing G to be allocated. The model predicts H_i, the quantity allocated to zone i, as a function of vacant land in i, V_i, and accessibility to, say, employment A_i, which we might take as

$$A_i = \sum E_j e^{-\beta c_{ij}} \tag{10.133}$$

using an obvious notation. Then

$$H_i = G\frac{V_i A_i^{\alpha}}{\sum_i V_i A_i^{\alpha}} \tag{10.134}$$

where α is a parameter to be estimated. This simple model has some useful characteristics and in particular points out the dependence of the process on land availability, as well as other desirable properties such as accessibility. It has been discussed as a residential location model (Swerdloff and Stowers, 1966), but we see how useful it is to make the sharp distinction between the location of people and the location of stock. This need not then be considered as one of a number of alternative residential location models, but δH_i from this model could be used as an input to a doubly constrained *person* dynamic residential location model.

A completely different approach to modelling housing supply was attempted by the Community Renewal Project in San Francisco (Robinson *et al.*, 1965). The essence of the model is a calculation of what different types of people would pay for different types of housing (including all existing housing, which is 'aged' in the model), and then new housing supply is generated which provides the highest yield for the developer. The main detail of the model is reserved for the person-type and home-type dimensions and there is relatively little spatial detail. However, it does represent an alternative model-building principle for the housing supply side, though no work on these lines seems to have been done since the San Francisco project.

In the case of services such as retailing, it is possible to argue that supply adjusts rather elastically to meet demand, and that models of the form described in Section 10.10 can thus also be considered to be supply-side models. This is, in effect, what Lowry (1964) assumed in his retail model. To seek alternative modelling approaches, we have to turn to econometric modelling. Since, as already argued, we have to do this for industrial location modelling, and since many modellers have anyway considered both sectors together, henceforth we shall do the same. Further, since some econometric model builders include equations for population or houses as part of their model system, we shall follow suit where appropriate and such equations can be considered as alternative residential or housing location models.

We shall consider three examples of econometric modelling which illustrate the problems of interest. Each piece of work was carried out by a firm of planning consultants in North America. Traffic Research Corporation in relation to Boston (Hill, 1965), Alan M. Voorhees and Associates in relation to Connecticut (Lakshmanan, 1968) and CONSAD in relation to Pittsburgh (Putman, 1967). The models are of the form

$$= \sum a_{ij} y_j + \sum_j b_{ij} \delta y_j \tag{10.135}$$

where x_i represents some measure of activity in zone i, and δx_i is the change in some time period. y_js are 'explanatory' variables (which, as we shall see later, may include some of the x_is). The particular choice of variables depends on the hypotheses to be represented. These may be chosen so as to make the model maximally consistent with some micro-analytical theory, as noted earlier, or simply to reflect empirical relationships discovered through preliminary statistical analysis. In the three examples we are using, a wide variety of hypotheses is represented. We can attempt to put them together eclectically and we obtain the following summary.

Industrial employment locates independently of local consumer markets, but may depend on

(i) access to labour market
(ii) access to services
(iii) access to other similar employment (because of inter-industry linkages)
(iv) land capacity (which depends on land availability, value and development costs)

Service employment may depend on

(i) access to consumers (households and organizations)
(ii) access to related services
(iii) access to the labour market
(iv) land capacity

Population or new housing, may depend on

(i) access to jobs
(ii) access to services
(iii) housing quality and price

(iv) social structure, represented by such things as affinities between income groups
(v) land capacity and associated density constraints

The authors of the Empiric model (Hill, 1965) make a distinction between *located* variables, the measures of economic activity or location, and *locator* variables, the explanatory variables in addition to located variables which are sometimes used as such. Thus, for each located variable R_i, we have an equation of the form

$$\delta R_i = \sum_{j \neq i} a_{ij} \delta R_j + \sum_j b_{ij} (Z_j \text{ or } \delta Z_j) \tag{10.136}$$

where the Z_js are locator variables, δZ_js the change in such variables, and the a_{ij}s and b_{ij}s are coefficients. In the original Toronto study, the variables chosen as located variables were

(i) white-collar population
(ii) blue-collar population
(iii) retail and wholesale employment
(iv) manufacturing employment
(v) other employment

and the locator variables were

(i) densities for land in different uses
(ii) zoning practices
(iii) car accessibility measures

The other two models use a different concept of change, the so-called *differential shift*, and so a preliminary discussion of this concept is appropriate. It originates in the work of Fuchs (1962). We can proceed to a definition in terms of population change. Let $P_i(t)$ be the population of zone i at time t, and consider a period from 0 to T. Then, using an asterisk to denote summation, $P_*(t)$ is the population of the region, the study area, at time t. Then, define

$$\alpha(0, T) = \frac{P_*(T) - P_*(0)}{P_*(0)} \tag{10.137}$$

as the regional rate of population change. We can use DS as an operator to represent differential shift, and

$$DSP_i(0, T) = P_i(T) - (1 + \alpha(0, T))P_i(0) \tag{10.138}$$

Thus the differential shift in period 0 to T is the difference between the zonal population in time T and the population which would have been achieved if the zone had grown at the regional growth rate. Thus, differential shift emphasizes inter-zonal differences in the pattern of change. Note that, from the definition

$$\sum_i DSP_i(0, T) = 0 \tag{10.139}$$

If the time period is clear from the context, the (0, T) suffix will be dropped.

Now, suppose we measure economic activity by employment, and let $E_i^k(t)$ be the employment in sector k in zone i at time t. Then the regional growth rate for sector k is defined as

$$\beta^k(0, T) = \frac{E_*^k(T) - E_*^k(0)}{E_*^k(0)} \qquad (10.140)$$

and the corresponding differential shift by

$$DSE_i^k(0, T) = E_i^k(T) - (1 + \beta^k(0, T))E_i^k(0) \qquad (10.141)$$

Model equations can now be formulated as before, with differential shift variables replacing the gross change variables. However, before we explore examples of such models, a digression is in order.

This digression aims to relate the differential variables defined above to alternate definitions. Although the alternative definitions will not be used again here, they may help the reader with the literature (see, for example, Smith, 1966).

The alternative referred to consists of definitions of *comparative*, *compositional* and *competitive* shifts. Essentially, comparative shift is the differential shift in total employment of a zone, relative to an expected regional growth of total employment; competitive shifts are shifts in total employment relative to expected growth calculated as though each sector grew at its regional rate; the compositional shift is the difference between the two and represents a changing sector-mix. Suppose

$$E_*^*(t) = \sum_k E_*^k(t) \qquad (10.142)$$

is total employment in the region at time t, and

$$E_i^*(t) = \sum_k E_i^k(t) \qquad (10.143)$$

is total employment in zone i at time t. Then

$$\gamma(0, T) = \frac{E_*^*(T) - E_*^*(0)}{E_*^*(0)} \qquad (10.144)$$

is the rate of change of total employment in the region. Let CPAR, CPOS and CPET be the comparative, compositional and competitive shifts in zone i. Then, we can define

$$\text{CPAR } E_i^*(0, T) = E_i^*(T) - (1 + \gamma(0, T))E_i^*(0) \qquad (10.145)$$

$$\text{CPOS } E_i^*(0, T) = \sum_k \frac{E_i^k(0)E_*^k(T)}{E_*^k(0)} - \frac{E_i^*(0)E_*^*(T)}{E_*^*(0)} \qquad (10.146)$$

Using the definition of $\beta^k(0, T)$ from equation (10.140), CPOS $E_i^*(0, T)$ can be written

$$\text{CPOS } E_i^*(0, T) = \sum_k (1 + \beta^k(0, T))E_i^k(0) - (1 + \gamma(0, T))E_i^*(0) \qquad (10.147)$$

Then, we can define

$$\text{CPET } E_i^*(0, T) = \text{CPAR } E_i^*(0, T) - \text{CPOS } E_i^*(0, T) \qquad (10.148)$$

$$= E_i^*(T) - \sum_k (1 + \beta^k(0, T))E_i^k(0) \qquad (10.149)$$

Note that using the earlier definition of differential shift in equation (10.141), if we sum over k we get

$$\sum_k DSE_i^k(T) = E_i^*(T) - \sum_k (1 + \beta^k(0, T))E_i^k(0) \qquad (10.150)$$

$$= \text{CPET } E_i^*(0, T) \qquad (10.151)$$

so competitive shift is the sum of individual differential shifts. This ends the digression and henceforth we shall use differential shift by sector, the definition of equation (10.141).

The Voorhees model was a set of equations each of the form

$$DSX_i^m(0, T) = \sum a_{mn} Y_i^n \qquad (10.152)$$

for the shift in the variable X_i^m for zone i. The Y_i^ns are explanatory variables and the a_{mn}s are coefficients.

The independent variables, when the shift referred to an *industrial employment* sector, were:

(i) sum of all service sector differential shifts
(ii) total employment in the sector at time 0
(iii) total service employment at time 0
(iv) capacity for industrial employment
(v) access to professional services

The independent variables for service sector shifts were:

(i) sum of all service sector differential shifts
(ii) total differential shift in industrial employment
(iii) total employment at time 0
(iv) access to population

The independent variables for population (by income group) shifts were:

(i) differential shift in population in next highest income group
(ii) total differential shift in industrial employment
(iii) sum of all service sector differential shifts
(iv) sum of population differential shifts
(v) access to population in same income group
(vi) capacity for additional population

It is interesting to compare and contrast the EMPIRIC and Voorhees models in relation to our earlier eclectic list of hypotheses but this is left as an exercise for the reader. However, one important point should be noted at this stage — each model uses variables which are dependent in one equation as independent variables in others. This means that two-stage least-squares, maximum-likelihood, or some other estimation technique must be used to estimate the coefficients rather than the simple regression analysis technique mentioned in Chapter 5 which would lead to biased estimates.

The third example, the CONSAD sub-model for the North East Corridor Transportation Project (NECPT) is mentioned because it shows how simpler models can be postulated which do not involve simultaneous equation estimation techniques (Putman, 1966). We use the notation developed above and, in addition, we introduce an accessibility variable as

$$A_i^E(t) = \sum_j E_j^*(t) f(c_{ij}) \qquad (10.153)$$

as a measure of access to total employment. Then, the assumptions made, taking $k = 1$ as the basic/industrial sector, $k > 1$ as service sectors, were

$$DSE_i^1(0, T) = a_1 E_i^1(0) + a_2 E_i^*(0) + a_3 A_i(T) + a_4 \qquad (10.154)$$

$$DSP_i(0, T) = b_1 P_i(0) + b_2 A_i(T) + b_3 E_i^*(T) + b_4 \qquad (10.155)$$

where $a_1 - a_4$ and $b_1 - b_4$ are coefficients to be estimated. Service employment was not estimated through a shift mechanism, but in total (implying a strong equilibrium assumption) as

$$E_i^k(T) = c_1 E_i^k(0) + c_2 P_i(T) + c_3 E_i^*(T) + c_4 \qquad (10.156)$$

The reader will recognize that this is a very crude retail model indeed, completely neglecting interaction outside the zone being estimated. Some difficulties were discovered in obtaining satisfactory empirical estimates in relation to these equations, and in a later paper Putman (1967), in developing a model for Pittsburgh, was using a different model altogether.

Many more examples of econometric models can be found in the literature (for two more recent pieces of work, see Moses and Williamson, 1963, and Seidman, 1969). Seidman's work, in particular, will be described in greater detail in Chapter 11.

10.13 REFERENCES

W. Alonso (1964) *Location and land use*, Harvard University Press, Cambridge, Mass.

R. P. [illegible] (1970) An approach to the theory of consumer shopping behaviour, *Urban Studies*, 8, pp 55-64.

F. S. Chapin, Jr. and S. F. Weiss (1968) A probabilistic model for residential growth, *Transportation Research*, 2, pp 375-390.

Chicago Area Transportation Study (1960) *Final Report*, Chicago.

P. Converse (1949) New laws of retail gravitation, *Journal of Marketing*, 14, pp 279-304.

* M. Cordey Hayes (1968) Retail location models, Working Paper 16, Centre for Environmental Studies, London.

M. Cordey Hayes and A. G. Wilson (1970) Spatial interaction, *Socio-economic planning sciences*, 5, 73-96.

* E. L. Cripps and D. H. S. Foot (1969) The empirical development of an elementary residential location model for use in sub-regional planning, *Environment and Planning*, 1, pp 81-90.

* E. L. Cripps and D. H. S. Foot (1970) The urbanization effects of a third London airport, *Environment and Planning*, 2, pp. 153-192.

Distributive Trades E.D.C. (1970) *Models for shopping studies*, National Economic Development Office, London.

E. S. Dunn (1954) *The location of agricultural production*, University of Florida Press, Gainsville.

* R. V. Eastin and P. Shapiro (1973) Design of a location experiment, *Transportation Research*, pp 17-29.

* R. V. Eastin and P. Shapiro (1973) The design of a location experiment: a continuous formulation, *Transportation Research*, pp 31-38.

J. W. Forrester (1968) *Principles of systems*, Wright-Allen Press, Cambridge, Mass.

J. W. Forrester (1969) *Urban-dynamics*, M.I.T. Press, Cambridge, Mass.

V. R. Fuchs (1962) The determinants of the redistribution of manufacturing in the United States since 1929, *Review of Economics and Statistics*, 44, pp 167-177.

J. R. Hamburg and R. L. Creighton (1959) Predicting Chicago's land use pattern, *Journal of the American Institute of Planners*, 25, pp 67-72.

F. E. I. Hamilton (1967) Models of industrial location, in R. J. Chorley and P. Haggett (Eds.) *Models in geography*, Methuen, London, pp 361-424.

W. G. Hansen (1959) How accessibility shapes land use, *Journal of American Institute of Planners*, 25, pp 73-76.

B. Harris (1962) Linear programming and the projection of land uses, Paper 20, Penn-Jersey Transportation Study, Philadelphia.

B. Harris (1964) A model of location equilibrium for retail trade, mimeo, Institute for Urban Studies, University of Pennsylvania.

* B. Harris (1968) Quantitative models of urban development, in H. S. Perloff and L. Wingo (Eds.) *Issues in urban economics*, Johns Hopkins, Baltimore, pp 363-412.

J. Herbert and B. H. Stevens (1960) A model for the distribution of residential activity in urban areas, *Journal of Regional Science*, 2, pp 21-36.

D. M. Hill (1965) A growth allocation model for the Boston region, *Journal of the American Institute of Planners*, 31, pp 111-120.

* A. G. Houghton (1971) The simulation and evaluation of housing location, *Environment and Planning*, 3, pp 383-394.

D. L. Huff (1962) A note on the limitations of intra-urban gravity models, *Land Economics*, 38, pp 64-66, 39, pp 81-89.

D. L. Huff (1964) Defining and estimating a trading area, *Journal of Marketing*, 28, pp 37-38, also reprinted in P. Ambrose (Ed.) (1970) *Analytical human geography*, Longman, London, pp 161-171.

W. Isard (1956) *Location and space-economy*, M.I.T. Press, Cambridge, Mass.

T. R. Lakshmanan (1965) A model for allocating urban activities in a state, *Socio-economic planning sciences*, 1, pp 283-295.

T. R. Lakshmanan and W. G. Hansen (1965) A retail market potential model, *Journal of the American Institute of Planners*, 31, pp 134-143.

J W Forrester (1968) *Principles of systems*, Wright Allen Press, Cambridge, Mass

J W Forrester (1969) *Urban Dynamics*, M.I T. Press, Cambridge, Mass

R A Garin (1966) A matrix formulation of the Lowry model for intra-metropolitan activity location, *Journal of the American Institute of Planners*, 32, pp 361-364.

W Goldner (1971) The Lowry model heritage, *Journal of the American Institute of Planners*, 37, pp 100-110.

H. R Hamilton *et al* (1969) *Systems simulation for regional analysis*, M I T. Press, Cambridge, Mass

B Harris (1964) A note on the probability of interaction at a distance, *Journal of Regional Science*, 15, pp 31-35

D M Hill (1965) A growth allocation model for the Boston Region, *Journal of the American Institute of Planners*, 31, pp 111–120

I S Lowry (1964) *A model of metropolis*, RM-4035-RC, RAND Corporation, Santa Monica

• I. S Lowry (1968) Seven models of urban development a structural comparison, *Special Report 97*, Highway Research Board, Washington, D C , pp 121–145.

P. H Rees and A G Wilson (1973) Accounts and models for spatial demographic analysis I: aggregate populations, *Environment and Planning*, 5, 61–90.

K. Schlager (1965) A land use plan design model, *Journal of the American Institute of Planners*, 31, pp 103–111.

D R Seidman (1969) *The construction of an urban growth model*, Plan Report No 1, Technical Supplement, Volume A, Delaware Valley Regional Planning Commission, Philadelphia

J. L Taylor (1972) *Instructional planning systems*, Cambridge University Press, London

A. G Wilson (1967) A statistical theory of spatial distribution models, *Transportation Research*, 1, pp 253-269

A G Wilson (1969-A) Developments of some elementary residential location models. *Journal of Regional Sciences*, 9, pp 377–385

• A. G Wilson (1969-B) The integration of accounting and location theory in urban modelling, in A J Scott (Ed) *Studies in regional science*, Pion, London, pp. 89–104, also reprinted in A G Wilson (1972) *Papers in urban and regional analysis*, Pion, London, pp 154-169.

A G Wilson (1971) Generalizing the Lowry model, in A G Wilson (Ed) *Urban and regional planning*, Pion, London, pp 121–133, also reprinted in A G. Wilson (1972) *Papers in urban and regional analysis*, Pion, London, pp 58 70.

A G Wilson (1973-A) Further developments of entropy maximizing transport models, *Transportation planning and technology*, 1, 183 193

A G Wilson (1973-B) Learning and control mechanisms for urban modelling, Working Paper 35, Department of Geography, University of Leeds.

CHAPTER 12

Practical considerations: data, calibration and testing

12.1. THE NATURE OF THE CALIBRATION AND TESTING TASK

We had one example of calibration in Chapter 4, when the parameter *n* in the model given by equation (4 18) was estimated by maximizing the sum of squares of differences between predicted and observed terms in equation (4 21) Since then, we have concentrated on model-building principles and have neglected the topic of calibration and testing. The principle of calibration of a model is to choose the parameters of the model to optimize one or more goodness-of-fit statistics. In this chapter we outline methods for doing this for a wide range of examples and firstly we detail the range of the calibration problem in relation to the models of Chapters 7–11 Following this, we concentrate on method: in Section 12 2 we discuss the task of defining goodness-of-fit statistics and in Section 12 3 we show how to find parameters which optimize these in different examples In Section 12 4 we add a note on the task of estimating entire unsmoothed empirical curves. Then in Sections 12.5–12 9 we review the calibration and testing problems for the models of Chapters 7–11, in each case paying special attention to likely data availability In Section 12 10, some special procedures and short cuts are described.

We note at the outset that we concentrate more on calibration than on testing, though since calibration almost always involves the calculation of goodness-of fit statistics which can then be evaluated for the best fit, this is little restriction

There are three main types of calibration The models of Chapters 7 and 8, the demographic and economic models, and the category analysis trip-generation model of Chapter 9 all involve the calculation of rates or coefficients for particular sub-groups of the population This process is a kind of model calibration The art is to group sub-sets of the total population into categories in such a way that the rates for each sub-set are stable, and the actual distribution of values for each sub-set about the 'mean' rate is a 'narrow' one This kind of analysis is what Winsten (1967) calls regression analysis without prior hypothesis about the nature of the response surface, more typically a linear assumption is made about this surface, leading to *linear* regression analysis

Next, indeed, we consider parameter estimation in linear models of the form

$$y = a_0 + a_1x_1 + a_2x_2 + \ldots \tag{12.1}$$

or log linear models

$$\log y = a_0 + a_1x_1 + a_2x_2 + \ldots \tag{12.2}$$

in situations where observations are available of $(y, x_1, x_2, \ldots)$ and $a_0, a_1, a_2, \ldots$ are parameters to be estimated. This estimation task is standard and well-known: a sum of squares will be minimized, and some statistic such as R^2 used to measure goodness-of-fit. This technique can only be used when rather strong assumptions are satisfied—the independent variables should not be correlated with each other, they should be normally distributed, and the unwritten error term should also be normally distributed and should have mean zero. Needless to say, this method of parameter estimation is often used in circumstances where these conditions are not satisfied, and biased estimates of the parameters may then be obtained. Examples of the use of linear equations within one of the models discussed in this book are the trip-generation equations (9.1) and (9.2) in Chapter 9.

We should note that some models which do not at first sight appear to be linear can be transformed to take a linear form. For example

$$y = a_0x_1^{a_1}x_2^{a_2}\ldots \tag{12.3}$$

or

$$y = a_0\,e^{a_1x_1}\,e^{a_2x_2}\ldots \tag{12.4}$$

become

$$\log y = a_0 + a_1\log x_1 + a_2\log x_2 + \ldots \tag{12.5}$$

or

$$\log_e y = \log_e a_0 + a_1x_1 + a_2x_2 + \ldots \tag{12.6}$$

(which is equivalent to equation (12.2)) after taking logs. Most of the models which appear in Chapters 9–11 are intrinsically non-linear (a concept which, according to Batty (1971) was first introduced by Draper and Smith (1966)) in that they cannot be transformed this way. A simple example is the shopping model given in equation (4.18) in Chapter 4.

Non-linear models may have one or more parameters and, unsurprisingly, the complexity of the calibration procedure increases rapidly as the number of parameters increases. Where several parameters are involved, an issue may arise as to whether they should be estimated sequentially or [illegible]

In the following sub-sections, we shall assume that methods of estimating rates and parameters in linear models are, on the whole, familiar and we shall concentrate almost entirely on parameter estimation in non-linear models.

12.2. GOODNESS-OF-FIT STATISTICS

Suppose we have a set of observations, x_i^{obs}, and a set of model predictions, x_i, how do we compare the two? The simplest measure is the sum of squares (which we used for this purpose in Chapter 4):

$$S = \sum_i (x_i - x_i^{obs})^2 \tag{12.7}$$

This will take a minimum value for the 'best' fit, though it will be difficult to assess how good the fit is as the value of S depends on the dimensions of the variables and the units chosen.

Perhaps a better procedure of direct comparison is to plot x_i against x_i^{obs}, and to fit a straight line between them, say using the model

$$x_i = a_0 + a_1x_i^{obs} \tag{12.8}$$

For an exact fit, a_0 should be zero and a_1 should be 1. Deviation from these values shows a worsening fit, and the actual values (together with an inspection of the plots) may give an indication of any systematic bias in the model.

The next step is to find measures of association between x_i and x_i^{obs} which are dimensionless. Any standard statistical programme which estimated a_0 and a_1 in equation (12.8) would also estimate the correlation R between the sets of variables and R^2, the amount of variance explained by the model:

$$R^2 = 1 - \frac{\sum_i (x_i - x_i^{obs})^2}{\sum_i [x_i^{obs} - (1/n)\sum_i x_i^{obs}]^2} \tag{12.9}$$

where n is the number of variables. R^2 varies between 1, for an exact correspondence and 0.

The other obvious alternative is to use χ^2 defined as

$$\chi^2 = \sum \frac{(x_i^{obs} - x_i)^2}{x_i} \tag{12.10}$$

The use of these (and other measures) are discussed in the urban modelling context by Batty (1970, 1971) and Evans (1971). In calibration, we would find model parameter values to maximize R^2 and to minimize χ^2.

The variables x_i used to construct these variables may be any model outputs. In a spatial interaction model, they should usually be the interaction matrices, though if such a model is singly constrained and is being used as a location model, then spatial distribution of some interaction-end totals, a distribution of some activity, may be appropriate.

The problem which often arises when goodness-of-fit statistics such as those defined above are used in urban and regional modelling for calibration purposes is that they vary rather slowly as the parameters of the model vary. That is, they are often relatively insensitive to parameter estimation. This problem can be avoided in many cases of parameter estimation in non-linear models by the use of maximum-likelihood procedures (or, in this case analogously, by entropy-maximizing procedures). Such methods for this field are discussed as maximum-likelihood methods by Blackburn (1970), Evans (1971) and Batty and Mackie (1972), and as entropy-maximizing methods by this author (Wilson, 1970). Another equivalent procedure can be derived from Bayesian methods, as in Hyman (1969) The method will be explained here only in the broadest outline, and the reader is referred to the previously cited works for the details Assume that the model can be written so that p_i is the probability of a member of the population being in the i-state, so

$$p_i = \frac{x_i}{x} \tag{12.11}$$

for total population x, and x_i^{obs} is the number observed in state i as before Then the likelihood function is

$$L = \prod_i p_i^{x_i^{\text{obs}}} \tag{12.12}$$

and it can be shown that the best estimates of the model's parameters can be obtained by choosing those which maximize L, or more usually, and equivalently, log L.

This procedure is equivalent to the entropy-maximizing procedure (see Appendix 1 and Wilson, 1970) in the following sense. For each parameter of the model, the procedure produces an equation to be solved for that parameter. This equation turns out to be the constraint equation which would be used to generate the same model as an entropy-maximizing model, and the equivalent parameter is then the Lagrangian multiplier associated with the constraint. Either way, the parameter is obtained by solving the appropriate equation It is really a matter of taste and convenience as to whether maximum likelihood methods or entropy-maximizing methods are used to produce the equation which is to be solved for each parameter. The important point is that, for each parameter, there *is* such an equation, and the statistic associated with it is the *best* goodness-of-fit statistic for that parameter.

The method is best illustrated by an example Consider a one-parameter version of the shopping model given in equation (4.18).

$$S_{ij} = A_i e_i P_i W_j c_{ij}^{-\beta} \tag{12.13}$$

The maximum-likelihood equation for β is

$$\sum_i \sum_j S_{ij} \log c_{ij} = \sum_i \sum_j S_{ij}^{\text{obs}} \log c_{ij} \tag{12.14}$$

using an obvious notation That is, S_{ij} is to be substituted from equation (12 13) into (12 14) and (12 14) solved for β. The entropy-maximizing constraint which gives rise to the β term in the entropy-maximizing derivation of equation (12.13) is

$$\sum_i \sum_j S_{ij} \log c_{ij} = C \tag{12.15}$$

where C is a constant which in this case turns out to be the same as the right-hand side of equation (12 14) and so the resulting estimation procedures are the same (Note also that if *mean* trip expenditure is known, this implies that the right-hand side quantities are known.)

If an exponential function was used, so that the model was

$$S_{ij} = A_i e_i P_i W_j \mathrm{e}^{-\beta c_{ij}} \tag{12.16}$$

then the maximum-likelihood equation would be

$$\sum_i \sum_j S_{ij} c_{ij} = \sum_i \sum_j S_{ij}^{\text{obs}} c_{ij} \tag{12.17}$$

and the entropy-maximizing equation

$$\sum_i \sum_j S_{ij} c_{ij} = C \tag{12.18}$$

which are identical, and which can be solved for β This does remind us, of course, that we still have to choose between alternative forms of function in such models.

If a second parameter is now introduced, as in the equation (4 18) models, a new estimating equation must be found for that parameter (see Batty and Mackie, 1972). For γ in the equation (4 18) model the maximum-likelihood equation is

$$\sum_i \sum_j S_{ij} \log W_j = \sum_i \sum_j S_{ij}^{\text{obs}} \log W_j \tag{12.19}$$

and the equivalent entropy-maximizing constraint would be

$$\sum_i \sum_j S_{ij} \log W_j = \log W \tag{12.20}$$

where log W is in fact given by the right-hand side of (12 19) and is to be interpreted as the total benefit obtained by people choosing centre j where they can obtain benefit log W_j relative to other centres. (In the original derivation of this term, another method was used (Wilson, 1967, 1970) and this argument now suggests that where a new parameter is involved, it is better to introduce a new constraint explicitly.)

It is interesting to reflect that the above discussion records the theoretical respectability of a result which had been known in practice for a long time, at least for the main parameter of a spatial interaction model; that the most sensitive fitting statistic for the parameter in the cost function was something like mean-trip length. The newer results, however, have not only given us theoretical justification, but also more precision: we now know that mean-trip length is appropriate if the trip cost function is $e^{-\beta c_{ij}}$, but for a general function, $f(c_{ij})$, it should be written in the form $e^{-\beta h(c_{ij})}$ and then the mean value of $h(c_{ij})$ will be the appropriate statistic. Thus, for the power function, as we have seen, $c_{ij}^{-\beta}$ can be written $e^{-\beta \log c_{ij}}$ and so the mean value of $\log c_{ij}$ is the best statistic.

We can summarize the overall discussion as follows. For each parameter in a model, ideally the maximum-likelihood estimating equation (or entropy-maximizing constraint equation) should be obtained, and this provides the appropriate goodness-of-fit statistic. However, we still need to know something about the goodness-of-fit of the model as a whole, and for this we will still need something like R^2 or χ^2 tests applied to the model's main variables. This will not only give us the overall indication we need, but may help us to choose between different forms of function: for example, for travel impedance.

12.3. THE MECHANICS OF PARAMETER ESTIMATION

Again, we concentrate mainly on estimation procedures for parameters with non-linear models. In the first instance, to illustrate the methods available for parameter computation, we will consider only single parameter models. The model predictions x_i can then, for these purposes, be considered functions of the parameter, say β, only, and similarly the goodness-of-fit statistics. Thus, our task is to maximize $R^2(\beta)$, or to minimize $\chi^2(\beta)$ or to solve

$$C(\beta) = C^{\text{obs}} \tag{12.21}$$

where C is the 'function' which turns up in the maximum-likelihood equation. (For the single parameter shopping model, equation (12.17) would be the specific form of (12.21) for example.) Figure 12.1 shows diagrammatically what is involved in the three cases mentioned. Almost always, the procedure used to find β when the model is non-linear will be a numerical one within a computer. In each case, the problem is to identify the value of β indicated by the point P on the figure.

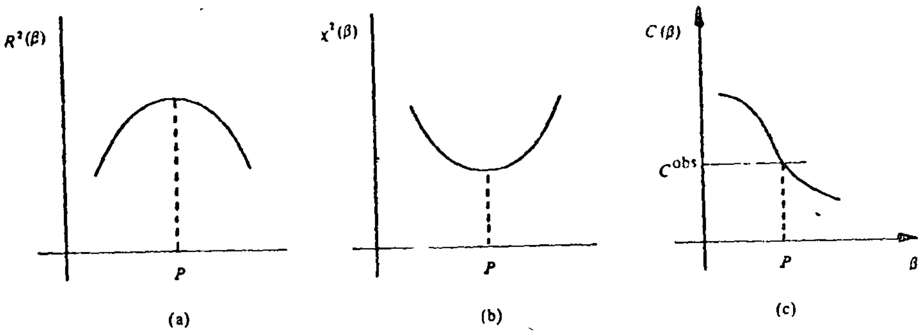


Figure 12.1 Maximum, minimum and equation solving problems in calibration

For the minimization and maximization problems, some sort of search procedure is usually used. The simplest possible procedure is to run the model for a range of values of β, and choose the β which optimizes the statistic, if necessary repeating with a smaller increment connecting the different β-values. However, this is relatively inefficient. Batty (1971) illustrates the use of Fibbonaci and 'golden-section' search routines in this context. Suppose, for definiteness, that we are trying to maximize $R^2(\beta)$ in Figure 12.1(a). At the kth step of the procedure to be described, it has been established that β lies between β_1^k and β_2^k (as in Figure 12.2) and β_3^k and β_4^k are chosen such that

$$\beta_1^k < \beta_3^k < \beta_4^k < \beta_2^k \tag{12.22}$$

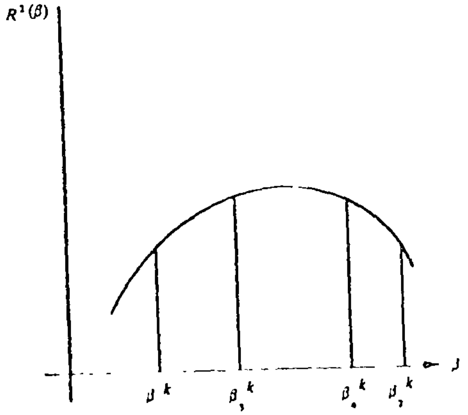


Figure 12.2 Fibbonaci search. (Based with permission on M. Batty in *Environment and Planning* Vol. 3, Pion, 1971, Figure 2, p. 420.)

β_3^k and β_4^k are chosen such that

$$\beta_3^k = \frac{F_{N-1-k}}{F_{N+1-k}}(\beta_2^k - \beta_1^k) + \beta_1^k \tag{12.23}$$

$$\beta_4^k = \frac{F_{N-k}}{F_{N+1-k}}(\beta_2^k - \beta_1^k) + \beta_1^k \tag{12.24}$$

where the Fs are the Fibbonaci numbers defined by

$$\left.\begin{aligned} F_0 &= F_1 = 1 \\ F_n &= F_{n-1} + F_{n-2}, \qquad n \geqslant 2 \end{aligned}\right\} \tag{12.25}$$

if $R^2(\beta_3^k) > R^2(\beta_4^k)$ then β lies in the interval (β_1^k/β_4^k) and if $R^2(\beta_3^k) < R^2(\beta_4^k)$, β lies in the interval (β_3^k/β_2^k). This gives two bounds for the next stage. Batty shows that after N steps the remaining interval, δ^N, is at most

$$\delta^N = \frac{1}{F_N}(\beta_2^1 - \beta_1^1) + \varepsilon \tag{12.26}$$

for some small number ε, and so F_N, and hence N, can be determined from equation (12.26) for any desired value of δ^N.

Golden-section search replaces F_{N-1-k}/F_{N+1-k} in equation (12.23) by the number 0·382 (to which it is approximately equal for large N) and F_{N-k}/F_{N+1-k} in equation (12.24) by 0·618 for similar reasons. Batty (1971) notes that golden-section search is approximately 13% slower than Fibbonaci search, though it is easier to programme for the computer.

These procedures can also be used for the equation-solving problem of Figure 12.1(c) and equation (12.21). An equivalent problem (and again we follow Batty, 1971) is to minimize $|C(\beta) - C^{\text{obs}}|$. However, other methods are also available for equation solving and it is to these that we now turn.

The simplest possible iterative procedure is as follows. Suppose after k steps, our estimate is β^k. Then

$$\beta^{k+1} = \beta^k \frac{C(\beta^k)}{C^{\text{obs}}} \tag{12.27}$$

is a suitable adjustment procedure. However, this may converge rather slowly, and Hyman (1969) has suggested that for $k > 1$, the second-order formula

$$\beta^{k+1} = \frac{(C^{\text{obs}} - C(\beta^{k-1}))\beta^k - (C^{\text{obs}} - C(\beta^k))\beta^{k-1}}{C(\beta^k) - C(\beta^{k-1})} \tag{12.28}$$

should produce more rapid convergence, and several workers have since used this result.

Another possibility is to use a Newton-Raphson procedure. Let

$$\beta^{k+1} = \beta^k + \varepsilon^k \tag{12.29}$$

Then

$$\begin{aligned} C(\beta^{k+1}) &= C(\beta^k + \varepsilon^k) \\ &= C(\beta^k) + \varepsilon^k \left.\frac{\partial C}{\partial \beta}\right|_{\beta^k} \end{aligned} \tag{12.30}$$

to the first-order in ε^k. Since we wish $C(\beta^{k+1})$ to be C^{obs}, ε^k can be obtained from equation (12.30) as

$$\varepsilon^k = \frac{C^{\text{obs}} - C(\beta^k)}{(\partial C/\partial \beta)|_{\beta^k}} \tag{12.31}$$

and equations (12.29) and (12.31) define a suitable iterative procedure. The derivative can be evaluated numerically.

These methods can be extended to the case where two or more parameters are to be estimated. Where many parameters are involved, the model system can usually be decomposed in some obvious way, and the calibration problem correspondingly. Firstly then, let us consider a two-parameter problem: our model is a function of parameters α and β, our general goodness-of-fit statistics are $R^2(\alpha, \beta)$, $\chi^2(\alpha, \beta)$; the maximum-likelihood equation for α is, say

$$W(\alpha, \beta) = W^{\text{obs}} \tag{12.32}$$

and for β, say,

$$C(\alpha, \beta) = C^{\text{obs}} \tag{12.33}$$

The important point to note at the outset is that all the statistics are functions of both parameters. This means that, in principle, the parameters have to be estimated simultaneously. We should also note that Batty (1971) found that, with a two-parameter spatial interaction model, the response surfaces of the goodness-of-fit statistics were such that no *one* of them could be used to determine both parameters simultaneously. However, if, as recommended, maximum-likelihood equations are used, there is one such equation for each parameter, and the difficulty disappears. This was demonstrated for the same model by Batty and Mackie (1972).

This is also a convenient point to include a warning about 'accidental' optimal values of some goodness-of fit statistics for some two parameter models and, in particular, the shopping model shown in equation (4.18). The reader can easily check that if $\alpha = 1$ and $\beta = 0$ in the case where W_j is taken as $S_{\cdot j}^{\text{obs}}$, then $S_{ij} = S_{ij}^{\text{obs}}$ exactly. R^2 for example, is then 1. The response surface of $R^2(\alpha, \beta)$ is then peaked towards the maximum at $\alpha = 1$, $\beta = 0$, and this makes R^2 virtually useless for calibration purposes with this kind of model. This is another good reason for using the maximum-likelihood equations for parameter estimation. The 'bogus calibration' problem was first discussed in a Government Paper on the flows of goods to ports (Ministry of Transport, 1966).

Before discussing how the univariate methods can be extended, it is useful to comment on the separability of the calibration problem. The likely result (which is being stated at the level of conjecture, with some empirical backing, but not as a general *proved* theoretical result) is that the maximum-likelihood statistic for a parameter is likely to be maximally independent of variation of other parameters (though unlikely to be completely independent), but mutual *dependence* is likely to be the order of the day with other statistics.

The simplest calculation scheme for a two-parameter model involves the computation of goodness-of-fit statistics for a 'grid' of parameters of α and β. The values of α and β which optimize the statistics can be chosen, and the computation repeated if necessary for a finer grid around this value. This method, although rarely the most efficient, is simple and robust and has been much used in practice.

The obvious way in which to attempt to improve this method is by using Fibbonacci or golden-section search algorithms. This can be done within an iterative scheme as follows: for given α, calculate β using an efficient search procedure, iterate as necessary.

When the maximum-likelihood equations are being solved, two variable versions of the adjustment schemes given by (12.27) and (12.28) could easily be developed. A two-variable Newton Raphson scheme to solve equations (12.32) and (12.33) can be developed as follows. Suppose

$$\alpha^{k+1} = \alpha^k + \varepsilon_1^k \tag{12.34}$$

and

$$\beta^{k+1} = \beta^k + \varepsilon_2^k \tag{12.35}$$

From equation (12.32)

$$\begin{aligned} W(\alpha^k + \varepsilon_1^k, \beta^k + \varepsilon_2^k) &= W^{\text{obs}} \\ &= W(\alpha^k, \beta^k) + \varepsilon_1^k \frac{\partial W}{\partial \alpha}\bigg|_{\alpha^k, \beta^k} + \varepsilon_2^k \frac{\partial W}{\partial \beta}\bigg|_{\alpha^k, \beta^k} \end{aligned} \tag{12.36}$$

and from equation (12.33)

$$\begin{aligned} C(\alpha^k + \varepsilon_1^k, \beta^k + \varepsilon_2^k) &= C^{\text{obs}} \\ &= C(\alpha^k, \beta^k) + \varepsilon_1^k \frac{\partial C}{\partial \alpha}\bigg|_{\alpha^k, \beta^k} + \varepsilon_2^k \frac{\partial C}{\partial \beta}\bigg|_{\alpha^k, \beta^k} \end{aligned} \tag{12.37}$$

These are simultaneous equations which can be solved for ε_1^k and ε_2^k, so that equations (12.34)–(12.37) form a suitable iterative scheme.

12.4. A NOTE ON ESTIMATING 'UNSMOOTHED' CURVES

Suppose we had a model of the form

$$T_{ij} = KO_iD_jf(c_{ij}) \tag{12.38}$$

in the usual notation and where K is a constant. Then the *form* of $f(c_{ij})$ can be examined by plotting observed values of T_{ij}/O_iD_j against c_{ij}. This is a plot of $f(c_{ij})$ up to the constant, K. More typically, K would be replaced by one or more balancing factors, say

$$T_{ij} = A_iB_jO_iD_jf(c_{ij}) \tag{12.39}$$

again in the usual notation. A plot of T_{ij}/O_iD_j against c_{ij} now gives a possibly biased picture of $f(c_{ij})$ because the A_is and B_js are themselves functions of $f(c_{ij})$ and are not constants. How, in this situation, can we get a reasonable empirical picture of what we might expect to be the function $f(c_{ij})$? The original calibration procedure of the Bureau of Public Roads (1965) for gravity models helps answer this question. We are not recommending this as a *calibration* procedure as such, but only in this context as a way of answering the question raised above. We illustrate the method with the model given in equation (12.39). The aim is to calculate the 'mean' value of $f(c)$ for a range of points 'c'. Suppose c varies, as a label, over such a range, and we use the notation $c_{ij} \in c$ to denote that c_{ij} falls in a certain range, say $a(c) \leqslant c_{ij} \leqslant b(c)$, where $a(c)$ and $b(c)$ are defined for each label c. As a convention, c itself may be taken as the mid-point of the range. Suppose the ranges are exclusive and exhaustive: each c_{ij} is in one and only one. The function $f(c_{ij})$ is then defined either as

$$f(c_{ij}) = f(c), \qquad c_{ij} \in c \tag{12.40}$$

or is obtained by interpolation between function values for adjacent cs; $f(c)$ is obtained as follows. After k iterations, suppose

$$T_{ij}^k = A_i^kB_j^kO_iD_jf^k(c_{ij}) \tag{12.41}$$

where

$$\sum_j T_{ij}^k = O_i \tag{12.42}$$

determines A_i and

$$\sum_i T_{ij}^k = D_j \tag{12.43}$$

determines B_j^k. We adjust $f^k(c_{ij})$ to ensure that the number of trips in the range 'c' reproduce observations. That is, so that

$$\sum_{c_{ij} \in c} T_{ij}^k = \sum_{c_{ij} \in c} T_{ij}^{\text{obs}} \tag{12.44}$$

for each c. This can be achieved by setting

$$f^{(k+1)}(c) = f^{(k)}(c)\frac{\sum_{c_{ij}\in c} T_{ij}^{\text{obs}}}{\sum_{c_{ij}\in c} T_{ij}^{k}}. \tag{12.45}$$

In this way, we obtain an unsmoothed function $f(c)$. This procedure is analogous to that of equation (12.27), and there is presumably a second-order improved version analogous to equation (12.28).

We may then wish to decide whether the resulting function is 'most like' a negative exponential function, or a power function or whatever, and to estimate parameters for a smoothed function by maximum-likelihood methods. In the same way as it is unwise to carry out linear regression analysis without looking at a plot of the observations, it may be argued that it is unwise to estimate the parameters for some smooth function without looking at the unsmoothed function $f(c)$ as calculated above. We shall note another application of this method in relation to housing expenditure in Section 12.8.

12.5. DEMOGRAPHIC MODELS

Unlike many other model sectors, the demographic modelling sector has a wealth of data available to it from census and birth and death registration returns. We emphasized in our discussion in Section 7.7 that the data have to be manipulated very carefully to get them into the form needed for calculating model rates and an accounting framework is essential for this purpose. There is no need to repeat that discussion here. This rate-estimation task is the main calibration task for this set of models. There may, however, be additional tasks, arising from possible sub-models for birth and death rates, and from the use of a spatial interaction model of migration. If a migration model of the form of that given in equations (7.35) and (7.36) is developed, then the standard procedures described above should be used to estimate β^r. There may also be sub-models for migration outflow and inflow (or attractiveness), the O_i^r and W_j^r terms in (7.35), and this may involve either regression analysis or category analysis and a procedure analogous to that used for trip generation. Interested readers should pursue the discussion of simultaneous or sequential estimation in this context in Section 12.7, which could also be applied in principle to these migration models.

12.6. ECONOMIC MODELS

The calibration problems with economic models are similar in principle to those of demographic models, mostly involving rate calculations for the input–output tables which can then give the technical coefficients needed for the model, but the data problems are much worse. It is difficult to assemble the data to build input–output models at the national scale in the U.K., for example, let alone at the urban and regional scale. Almost always, an urban or regional model can only be built if a special survey is carried out, as was the case with Artle's (1959) work on Stockholm and Morrison's (1972) work on Peterborough to cite but two examples. Faced with this situation, others have attempted to devise methods for estimating rates for local models by 'adjusting' in some way rates obtained from national models (see Hewings, 1971, Round, 1972).

Separate econometric sub-models may be built and estimated to provide final-demand variables in the input–output models, for example, see Stone (1967). These can be calibrated using standard regression techniques.

Unsurprisingly, multi-regional input–output models are even more rare than single-region ones. When, ultimately, they are developed, then there will be spatial interaction model calibration problems. The nearest point to this being reached in the U.K. is perhaps Chisholm's (1971) and O'Sullivan's (1971) work on freight flows.

12.7. TRANSPORT MODELS

There are two main sources of transport data for model calibration: special surveys, which have been common in large cities in the last 10–15 years and, for particular trip purposes, notably the journey to work, census data. In the U.K., census work-trip data are available from the 1961, 1966 and 1971 censuses, and for some areas in 1966 and 1971 they are available on a reasonably fine-zone system. The special surveys are usually available for one point in time only, but in some cities, such as London, second surveys are now being carried out—the second in 1972 following the first in 1962. In general, then, we can start from the assumption that some reasonably good data will be available, in the case of the special surveys collected specifically for model calibration, and so we can proceed to discuss their use.

We first illustrate the calibration problem using the model described in Section 9.4. For convenience we repeat the main equations of the model here. In equation (9.7), we refer to household category h, and later we define this to be the triple (I, n, p), where I is income group, n is car-ownership (in the associated household) and p is the family-structure index. Thus, using this explicitly, and assuming that 'n' in O_i^n refers to non-car-owning ($n = 1$, say) and car-owning ($n = 2$, say), equation (9.7) can be written

$$O_i^n = \sum_{I=1}^{6}\sum_{n'\in H(n)}\sum_{p=1}^{6} a_i(I, n', p)\,\Gamma(I, n', p) \tag{12.46}$$

where $n' \in H(n)$ means that the summation is for $n' = 0$ only for $n = 1$, and $n' = 1, 2$ or more, for $n = 2$. Equation (9.27), coupled with the number of households, if i, in zone, gives $a_i(I, n, p)$ as

$$a_i(I, n, p) = H_i f(p) \int_{a_I}^{a_{I+1}} P(n|x)\phi(x)\,dx \qquad (12.47)$$

The trip attractions were given in equation (9.15) which is repeated directly here:

$$D_j = \sum_l b_j(l) t(l) \qquad (12.48)$$

The trip-distribution and modal-choice equations are (9.40)–(9.44) and are also repeated directly:

$$e^{-\beta^n C^n_{ij}} = \sum_{k \in \gamma(n)} e^{-\beta^n c^k_{ij}} \qquad (12.49)$$

$$T^{*n}_{ij} = A^n_i B_j O^n_i D_j e^{-\beta^n C^n_{ij}} \qquad (12.50)$$

$$A^n_i = 1 \Big/ \sum_j B_j D_j e^{-\beta^n C^n_{ij}} \qquad (12.51)$$

$$B_j = 1 \Big/ \sum_i \sum_n A^n_i O^n_i e^{-\beta^n C^n_{ij}} \qquad (12.52)$$

$$T^{kn}_{ij} = T^{*n}_{ij} \frac{e^{-\lambda^n c^k_{ij}}}{\sum_{k \in \gamma(n)} e^{-\lambda^n c^k_{ij}}} \qquad (12.53)$$

Equations (12.46)–(12.53) are cycled through separately for each trip purpose and these trips are then added together, combined with commercial vehicle trips (Section 9.5) and bus trips, and are then input as vehicle trips to the assignment procedure.

With the assignment procedure described in Section 9.6 no calibration is involved, except adjustment of network representation and link speed flow characteristics to obtain the best possible fit to write flow data. But this is to adjust the components of the model as simplification of the real world to best fit the real world, it is not *parameter* adjustment, and hence is not calibration in the usual sense. Suppose then, that our task is to estimate the parameters in equations (12.46)–(12.53) (for each trip purpose, except in equation (12.47) which is common to all purposes).

In the trip-generation equations, the rates $T(I, n', p)$ in (12.46) and $t(l)$ in (12.48) have to be obtained. The distributions of households, H_i and economic activities, $b_j(l)$ can be obtained from location models or as planning inputs. The distributions $f(p)$, $P(n|x)$ and $\phi(x)$ each have parameters which have to be estimated by comparing predicted distribution of these quantities with observed distributions. (The only possible complication in this process arises when $\phi(x)$ has to be estimated from zonal car-ownership data, and this procedure is described in Chapter 9.)

β^n in equation (12.50) can be estimated by the methods described in the previous two sub-sections, with mean-trip length as goodness of fit statistic. That is, β^n is chosen to ensure that

$$\sum_i \sum_j T^{*n}_{ij} C^n_{ij} = \sum_i \sum_j T^{*n\text{obs}}_{ij} C^n_{ij} \qquad (12.54)$$

There is one unusual feature of this particular model. C^n_{ij} is constructed from the modal costs, c^k_{ij}, by equation (12.49) and *is itself a function of* β^n. So, if the right-hand side of equation (12.54) is considered as the 'observed' value, then this also varies with the parameter. However, the right-hand side varies much more slowly with β^n than the left-hand side, and no difficulties arise in practice.

We now proceed to the modal split equation (12.53). Define, for convenience;

$$M^{kn}_{ij} = \frac{T^{kn}_{ij}}{T^{*n}_{ij}} \qquad (12.55)$$

Then the maximizing-likelihood equation, or entropy-maximizing constraint equation, (cf. Wilson, 1973) is

$$\sum_i \sum_j \sum_{k \in \gamma(n)} M^{kn}_{ij} c^k_{ij} = \sum_i \sum_j \sum_{k \in \gamma(n)} M^{kn\text{obs}}_{ij} c^k_{ij} \qquad (12.56)$$

and so λ^n can be found by solving this equation using the methods of section 12.3.

This completes our discussion of the main structure of the transport model calibration process, but a number of other points must now be noted. First, we have assumed above that in both equations (12.50) and (12.53), the negative exponential function is best. Other forms of function should also be tested, and the methods of Section 12.4 should be used to produce unsmoothed, but correctly normalized, empirical functions for inspection. Thus, we could rewrite equation (12.50) as

$$T^{*n}_{ij} = A^n_i B_j O^n_i D_j f^{nD}(C^n_{ij}) \qquad (12.57)$$

and obtain $f^{nD}(C^n_{ij})$ as follows. Let $f^{nD(m)}(C)$ be the estimate of f^{nD} after m steps in some iterative procedure. Then

$$f^{nD(m+1)}(C) = f^{nD(m)}(C) \frac{\sum_{C^n_{ij} \in C} T^{*n\text{obs}}_{ij}}{\sum_{C^n_{ij} \in C} T^{*n(m)}_{ij}} \qquad (12.58)$$

where $T^{*n(m)}_{ij}$ is the model estimate after m steps.

In such a procedure, the equation for C_{ij}^n, (12.49) would have to be replaced by one which was independent of the travel function chosen, for example

$$C_{ij}^n = \min_k (c_{ij}^k) \tag{12.59}$$

or the obvious generalization of equation (12.49), which would be

$$f^{nD}(C_{ij}^n) = \sum_{k\in j(n)} f^{nD}(c_{ij}^k) \tag{12.60}$$

After n steps, we would have

$$f^{nD(m)}(C_{ij}^{n(m)}) = \sum_{k\in j(n)} f^{nD(m)}(c_{ij}^k) \tag{12.61}$$

and C_{ij}^n in equation (12.58) would have to be replaced by $C_{ij}^{n(m)}$.

When a function has been chosen by this procedure, then in order to estimate the parameters of the corresponding smoothed function, the maximum-likelihood equation is

$$\sum_i\sum_j T_{ij}^{*n} h^n(C_{ij}^n) = \sum_i\sum_j T_{ij}^{*n\,\mathrm{obs}} h^n(C_{ij}^n) \tag{12.62}$$

where h is defined by

$$e^{h^n(C_{ij}^n)} = f^{nD}(C_{ij}^n) \tag{12.63}$$

A similar procedure could be developed for the modal split function. In the distribution model above, we choose $f^{nD}(C)$ to ensure that trips for which $C_{ij}^n \in C$ equalled observed trips in the same class. The equivalent modal split criteria, based on the maximum-likelihood equation (12.56) would be: estimate $f^{nM}(c)$ such that

$$M_{ij}^{kn} = \frac{f^{nM}(c_{ij}^k)}{\sum_{k\in j(n)} f^{nM}(c_{ij}^k)} \tag{12.64}$$

with observed modal share equal to model modal share for each cost group $c_{ij}^k \in c$. That is, at the mth step of an iterative procedure

$$f^{nM(m+1)}(c) = f^{nM(m)}(c)\,\frac{\sum_{c_{ij}^k\in c} M_{ij}^{kn\,\mathrm{obs}}}{\sum_{c_{ij}^k\in c} M_{ij}^{kn(m)}} \tag{12.65}$$

using an obvious notation.

This procedure emphasizes the role of $f^{nM}(c_{ij}^k)$ in the modal split model, or $e^{-\lambda^n c_{ij}^k}$ in equation (12.53). It is more customary, as in the discussions of equation (9.38) and (9.39), to divide numerator and denominator by $e^{-\lambda^n c_{ij}^k}$ or $f^{nM}(c_{ij}^k)$ and to see modal split as a function of cost differences, in the case of the negative exponential function, or whatever combination of costs from several modes, in the case of the general function. This creates a feeling that the model, and the resulting maximum-likelihood fitting procedure, should perhaps be stated directly in terms of such differences. However, it seems probable that an equation of the form (12.62) is in any case the only way in which this can be done and all the modes treated equally symmetrically, so perhaps in terms of model *form* we are being sufficiently general. This does, however, suggest an alternative to the fitting procedure. We might emphasize modal choice between each pair of modes, so that we write (12.62) as

$$N_{ij}^{kk'n} = \frac{T_{ij}^{kn}}{T_{ij}^{k'n}} = \frac{M_{ij}^{kn}}{M_{ij}^{k'n}} = \frac{f^{nM}(c_{ij}^k)}{f^{nM}(c_{ij}^{k'})} \tag{12.66}$$

and choose $f^{nM}(c)$ so that relative modal shares are equal in the model and the survey. This only produces a neat result if the form of f is such that (12.66) can be rewritten in terms of cost differences or cost ratios. Suppose, either

$$\frac{f^{nM}(c_{ij}^k)}{f^{nM}(c_{ij}^{k'})} = F^{nM}(c_{ij}^k - c_{ij}^{k'}) \tag{12.67}$$

or

$$\frac{f^{nM}(c_{ij}^k)}{f^{nM}(c_{ij}^{k'})} = F^{nM}(c_{ij}^k/c_{ij}^{k'}) \tag{12.68}$$

The first case arises when f is an exponential function, and the second when f is a power function. The reader can easily check that, apart from replacing e by another constant, or changing the base of logarithms when C_{ij} is replaced by $\log C_{ij}$ to get the power function, these are the only functions of cost differences or cost ratios respectively for which relationships (12.66) and (12.67) hold in conjunction with the modal split equations (12.62) and (12.66). Other forms of function could be used in (12.62), but this would not lead to a function of cost differences or ratios in equation (12.66). The converse result is slightly different: a wide range of functions of cost differences or ratios could be used, but unless it was an exponential or power function, separability in the form (12.66) would not be possible.

This suggests that we could also investigate the other kind of separability: we could seek functional forms such that

$$N_{ij}^{kk'n} = f^{nM}(c_{ij}^k) - f^{nM}(c_{ij}^{k'}) \tag{12.69}$$

and

$$N_{ij}^{kk'n} = F^{nM}(c_{ij}^k - c_{ij}^{k'}) \tag{12.70}$$

or

$$N_{ij}^{kk'n} = F^{nM}(c_{ij}^k/c_{ij}^{k'}) \tag{12.71}$$

Equations (12.69) and (12.70) could only be compatible if

$$N_{ij}^{kk'n} = a(c_{ij}^{k} - c_{ij}^{k'}) \tag{12.72}$$

for some constant a, and equation (12.69) and (12.70) only if

$$N_{ij}^{kk'n} = a(\log c_{ij}^{k} - \log c_{ij}^{k'}) \tag{12.73}$$

neither of which we shall consider further.

Returning to the main line of the argument, the converse case is to hypothesize that

$$N_{ij}^{kk'n} = F^{kk'nM}(c_{ij}^{k} - c_{ij}^{k'}) \tag{12.74}$$

for the cost difference case (which is the only one we will consider); the reader can repeat the argument which follows for the power function case without difficulty. Notice that labels kk' have been added, since in principle a different function could be chosen for each k, k' pair. The function need only be defined for $k \neq k'$ and $k > k'$ (say). In fact, it is useful to assume that for $k = k'$

$$F^{kknM}(0) = 1 \tag{12.75}$$

and that

$$F^{kk'nM}(c_{ij}^{k} - c_{ij}^{k'}) = \frac{1}{F^{k'knM}(c_{ij}^{k'} - c_{ij}^{k})} \tag{12.76}$$

to extend the definition to all (k, k') combinations.

If we sum the first equation in (12.64) over k, we get

$$\sum_{k \in_i(n)} N_{ij}^{kk'n} = \frac{1}{M_{ij}^{k'n}} \tag{12.77}$$

so (replacing k' by k and k by k'),

$$M_{ij}^{kn} = \frac{1}{\sum_{k' \in_i(n)} N_{ij}^{k'kn}} \tag{12.78}$$

So that in terms of the arbitrary function $F^{kk'nM}(c_{ij}^{k} - c_{ij}^{k'})$,

$$M_{ij}^{kn} = \frac{1}{\sum_{k' \in_i(n)} F^{k'knM}(c_{ij}^{k'} - c_{ij}^{k})} \tag{12.79}$$

Thus for any F, there is always a simple formula for M_{ij}^{kn}. Recalling equation (12.75) it can also be written

$$M_{ij}^{kn} = \frac{1}{1 + \sum_{k' \in_i(n),\, k' \neq k} F^{k'knM}(c_{ij}^{k'} - c_{ij}^{k})} \tag{12.80}$$

The [illegible] result was that this can only be translated into the separable form (12.64) if F is the exponential function.

From the calibration viewpoint, we can now take (12.74) as the model. The procedure for estimating an unsmoothed function would be, using a by now obvious argument,

$$F^{kk'nM(m+1)}(\Delta c) = F^{kk'nM(m)}(\Delta c)\frac{\sum_{c_{ij}^{k} - c_{ij}^{k'} \in \Delta c} N_{ij}^{kk'n\,\mathrm{obs}}}{\sum_{c_{ij}^{k} - c_{ij}^{k'} \in \Delta c} N_{ij}^{kk'n(m)}} \tag{12.81}$$

in order to get $N_{ij}^{kk'n}$ right within each cost differential bracket, Δc.

We can now summarize the position reached on modal split calibration as follows. For the Chapter 9 model, restated as equation (12.53) we gave the maximum-likelihood equation to be solved as (12.56). We have since shown how to use the Section 12.4 method to investigate alternative forms of function. If we choose to work with the model given in equation (12.64), we use the procedure given in (12.65). If we chose a function $f^{nM}(c)$ on this basis and then wished to estimate the parameters of the corresponding smoothed function, the maximum-likelihood equation to solve for the parameters would be

$$\sum_i \sum_j \sum_{k \in_i(n)} M_{ij}^{kn} h^n(c_{ij}^{k}) = \sum_i \sum_j \sum_{k \in_i(n)} N_{ij}^{kn\,\mathrm{obs}} h^n(c_{ij}^{k}) \tag{12.82}$$

where

$$e^{h^n(c_{ij}^{k})} = f^{nM}(c_{ij}^{k}) \tag{12.83}$$

If, however, we choose to work in terms of cost differences, then the model equation is (12.74) (with modal split being given by (12.78)) and we estimate an unsmoothed function $F^{kk'nM}(\Delta c)$ by the procedure (12.81). To estimate the parameters of the corresponding smoothed function, the maximum-likelihood equation is

$$\sum_i \sum_j N_{ij}^{kk'n} h^{kk'n}(c_{ij}^{k} - c_{ij}^{k'}) = \sum_i \sum_j N_{ij}^{kk'n\,\mathrm{obs}} h^{kk'n}(c_{ij}^{k} - c_{ij}^{k'}) \tag{12.84}$$

where

$$e^{h^{kk'n}(c_{ij}^{k} - c_{ij}^{k'})} = F^{kk'nM}(c_{ij}^{k} - c_{ij}^{k'}) \tag{12.85}$$

If we additionally decided that $F^{kk'nM} = F^{nM}$, that is, the functional form was independent of k and k', the summations in equations (12.82) and (12.84) would be extended to cover k and k'.

This concludes our discussion of how to apply the methods of Section 12.4 to the distribution and modal split models to explore the construction of alternative functions for these models.

The second major point to discuss concerns the relationship of C_{ij}^{n} to the c_{ij}^{k}s. The exponential form given in equation (12.49) is a natural and convenient form for the model as described in Chapter 9. We have already indicated in the discussion of equation (12.61) that if the distribution model function is changed, there is a case for replacing equation (12.49) with (12.61) It is, however, made clear in an earlier book, (Wilson, 1970, p. 34) that any form of function could be explored for this task. In general, then, we might assume that

$$C^n_{ij} = \phi^n(c^1_{ij}, c^2_{ij}, \ldots) \tag{12.86}$$

In the absence of further knowledge, we might replace equation (12.49) by (12.61), but there is no reason why we should not try to investigate the form of ϕ^n by other means However, the only available way to do this is to run the model for a range of φ^n's and see which one gives the best overall goodness-of-fit statistics for the model If the goodness-of-fit statistics are relatively insensitive to the choice of ϕ^n, then there would be a case for retaining either equation (12.49) or equation (12.61) for convenience

A third major point to be taken up concerns other parameters which are implicit in the model as stated so far, and in particular the weights internal to the generalized costs c^k_{ij}. We show in equation (9.20), that

$$c^k_{ij} = a_1 t^k_{ij} + a_2 e^k_{ij} + a^k_3 d^k_{ij} + p^k_j + \delta^k_j \tag{12.87}$$

where the terms on the right-hand side relate to travel time, excess time, operating costs, parking costs and a modal penalty respectively. Detailed definitions are given in Chapter 9 following equation (9.30); a_1, a_2 and a^k_3 are coefficients to be estimated. Several questions arise at more or less fundamental levels. For example, other functional forms could be tested in which different independent variables were used The essential calibration question, though, is: should the parameters a_1, a_2 and a^k_3 be estimated simultaneously with other parameters in the distribution and modal split models, or outside the model? It seems reasonable to obtain at least the relative weights in this way, even though the overall sensitivity of cost is trip length and modal split is determined within the model by other parameters.

Quarmby (1967) has shown how the weights a_1, a_2 and a^k_3 can be estimated using discriminant analysis. The terms a_1 and a_2, which value different kinds of time, were expressed as a percentage of income, and in this form have been utilized (as relative weights, λ^n still being determined by calibration within the model) in other studies (cf the SELNEC Study, Wilson, Hawkins, Hill and Wagon, 1969) In the SELNEC Study, *ad hoc* adjustments were made to improve the fit in the distribution and modal split models It was found, for example, that for the c^k_{ij}s in the distribution model, the terms $p^k_j + \delta^k$ were best omitted, but they were retained in the modal split model In effect, then, slightly different generalized costs were used in two of the sub-models

In the case where there are only two modes, we should note that the modal split model, and hence the weights in the generalized cost expression, can be estimated by logit regression analysis (Theil, 1965; pp 77 87). The model, for persons of type n, can be written

$$p^{kn}_{ij} = \frac{1}{1 + e^{-\lambda^n(c^k_{ij} - c^{k'}_{ij})}} \tag{12.88}$$

where p^{kn}_{ij} is the probability that a person will travel from i to j by mode k (the other mode being k') Algebraic manipulation gives

$$\frac{p^{kn}_{ij}}{1 - p^{kn}_{ij}} = e^{\lambda^n(c^{k'}_{ij} - c^k_{ij})} \tag{12.89}$$

so that, using equation (12.86)

$$\log \frac{p^{kn}_{ij}}{1 - p^{kn}_{ij}} = \lambda^n a_1(t^{k'}_{ij} - t^k_{ij}) + \lambda^n a_2(e^{k'}_{ij} - e^k_{ij}) + \lambda^n a^k_3 d^k_{ij} - \lambda^n a^{k'}_3 d^{k'}_{ij} + a_0 \tag{12.90}$$

where

$$a_0 = p^k_j - p^{k'}_j + \delta^k - \delta^k \tag{12.91}$$

The coefficients in this equation can now be obtained by regression analysis.

It remains only to discuss briefly the calibration issues which arise from other kinds of transport models, such as those outlined in the 'further developments' section of Chapter 9. Perhaps the dominant feature of the American models discussed around equations (9.81) to (9.96) in Chapter 9 — the Quandt–Baumol abstract mode model, the Charles Rivers Associates sequential disaggregate model and Ben-Akiva's simultaneous disaggregate model — is that they are all designed so that their parameters can be estimated by standard econometric techniques In the case of the three examples cited, this can be achieved simply by taking logs in the main model equations. It is almost as though the desire to use such estimation procedures forms a major component of model design criteria In other words, these models arise from a much more inductive (as opposed to deductive style) than has been the policy in most of this book If a more deductive style is adopted, this throws more onus on the model builder to *understand* his parameters, and eventually to develop sub-models which predict parameter values as part of his overall theory One of the benefits of the maximum-likelihood or entropy-maximizing approaches, for example, is that the parameter p^n which turns up in equation (12.50) is obtained by solving the maximum-likelihood equation (12.51), and its behaviour can be predicted if the future value of

$$C^n = \sum_i \sum_j T_{ij} C^n_{ij} \tag{12.92}$$

can be predicted This emphasizes a new model building problem rather than a parameter-estimation problem

Apart from these observations, there is nothing to be added to the remarks on the calibration of these models which were made in Chapter 9. There *are* real issues of sequential versus simultaneous estimation, but we would argue, as in the discussion on the generalized cost weights above, that some separability is not a bad thing. Some authors associate sequential estimation with the hypothesis that travellers make sequential choices, and simultaneous estimation with the hypothesis that travellers make simultaneous multiple choices (cf. Brand, 1972, Ruiter, 1972, for example), but it can be argued that such an identification of behavioural hypotheses with methods for estimating more aggregative models cannot be made.

We should perhaps note finally, in relation to the three examples under discussion, that two of the models are disaggregate models and would be estimated by using household rather than zone-aggregated data. Then, if the various probabilities on the left-hand side of these equations are interpreted as frequencies, they will usually be (0.1) variables.

It only remains to comment on the calibration issues raised by the suggested modifications to the 'mainstream' Chapter 9 model in equations (9.108) through to (9.128). The main issue raised in equation (9.109), for example, which is

$$T_{ij}^{pn} = A_i^{pn} O_i^{pn} X_j^p \, e^{-\beta_i^{pn} c_{ij}^{pn}} \tag{12.93}$$

is that the parameter is now a function of i. The maximum-likelihood equation would be

$$\sum_j T_{ij}^{pn} c_{ij}^{pn} = \sum_j T_{ij}^{pn\,\mathrm{obs}} c_{ij}^{pn} \tag{12.94}$$

On the face of it, a large problem has been created because each equation contains all the β_i^{pn}s. However, this should not be too serious. We could devise an iterative scheme as follows: take some set of starting values for the β_i^{pn}s (say $\beta_i^{pn} = \beta^{pn}$ obtained in the usual way without i-dependence). Then solve each equation (12.94) for β_i^{pn} keeping the others fixed. Then repeat with the new estimates and so on. Convergence should be fairly rapid. An analogous problem and solution would arise with λ_i^{pn} in equation (9.11) if it was decided to keep the i on the λ_i^{pn}.

The disaggregate model given by equation (9.124) (9.128) presents no special problems. However, if h still referred to an individual, then it would be customary to drop h from β_i^{pn} and λ_{ij}^{pn} and to estimate β_i^p, λ_{ij}^p by something like logit regression analysis on household data, possibly carrying out the analysis separately for a number of groups.

12.8 LOCATION MODELS

We now turn to the location models of Chapter 10, beginning with residential location models. The data requirements for calibrating the simpler

spatial interaction models are very similar to those of journey-to-work transport models, for obvious reasons, and are usually easily met. Data is a more difficult problem for the disaggregated models whether of the spatial interaction/assignment type or the economic theory type. It then becomes necessary to have data on houses by location by type, and price, and jobs by location and wage as independent variables. Since the dependent variable is T_{ij}^{kw}, observations of this are also needed for calibration purposes. These sorts of data are not readily available and, ideally, special surveys are needed. This is often not possible, and more makeshift methods have to be used; 'estimating' observed values of T_{ij}^{kw} by making conditional probability assumptions, for instance (and for an example of this, see Senior and Wilson, 1973).

When the data is available, the calibration problems are relatively straightforward. For the simpler models, the procedures outlined in Sections 12.2–12.3 for one- and two-parameter spatial interaction models can be used. Methods analogous to those of Section 12.7 for transport models could be used to investigate the possibilities of using alternative functional forms for attractiveness factors and cost functions. There is also scope here for investigating a variety of possible measures of attractiveness as well as functional transformations of whatever measure is used.

In the case of the disaggregated spatial interaction model, the principles are similar though now, of course, there are more parameters. The simplest example of such a model was that given in equation (10.40):

$$T_{ij}^{kw} = A_i^k B_j^w H_i^k E_j^w \, e^{-\beta^w c_{ij}} \, e^{-\mu^w [p_i^k - q^w (w - c'_{ij})]^2} \tag{12.95}$$

The maximum-likelihood equation for β^w is

$$\sum_i \sum_j \sum_k T_{ij}^{kw} c_{ij} = \sum_i \sum_j \sum_k T_{ij}^{kw\,\mathrm{obs}} c_{ij} = C^w \tag{12.96}$$

in the usual way, and for μ^w is

$$\sum_i \sum_j \sum_k T_{ij}^{kw} [p_i^k - q^w (w - c'_{ij})]^2 = \sum_i \sum_j \sum_k T_{ij}^{kw\,\mathrm{obs}} [p_i^k - q^w (w - c'_{ij})]^2 = \sigma_w^2 \tag{12.97}$$

The usual calibration procedure is to take a set of starting values for all the β^w and μ^w, and then, given these, to find β^w and λ^w for each w in turn using the two-variable Newton–Raphson procedure of equations (12.34)–(12.37), and then to repeat the cycle with the new values until a suitable convergence is achieved. (For an example, see Cripps and Cater (1972).)

Again, alternative cost functions could be estimated using the methods of Section 12.4, and in this case, an alternative functional form for the budget price term, $e^{-\mu^w [p_i^k - q^w (w - c_{ij})]^2}$ could be investigated.

Define

$$v_i^{kw} = p_i^k - q^w(y^w - c_{ij}^k) \qquad (12.98)$$

We could then replace $e^{-\mu^w v_i^{kw^2}}$ in equation (12 95) by $f^w(v_i^{kw})$ where $f^w(v)$ was determined by an iterative process in which, after the mth iteration

$$f^{w(m+1)}(v) = f^{w(m)}(v)\frac{\sum_{v_i^{kw}\in v}T_{ij}^{kw\text{obs}}}{\sum_{v_{ij}^{kw}\in v}T_{ij}^{kn(m)}} \qquad (12\,99)$$

The econometric models of residential location mentioned in Section 10 6 have parameters which, by definition, can be estimated by standard procedures In the case of the economic theory approach, as represented by the Herbert–Stevens' model in equations (10.107)-(10 109), there are no visible parameters in the model, but some are implicit in the definition of the bid rents b_i^{kn} It is common, for example, to assume a linear form for b_i^{kn} and then to estimate the coefficients by regression analysis using actual rent data Such a procedure would also have to be used to estimate b_i^{kn} in the integrated model presented in equations (10 110)-(10.113). But also the parameter μ would have to be estimated, with

$$\sum_i\sum_k\sum_n T_{i*}^{kn}b_i^{kn} = \sum_i\sum_k\sum_n T_{i*}^{kn\text{obs}}b_i^{kn} = Z^{\text{obs}} \qquad (12.100)$$

as maximum-likelihood equation (cf equation (10 113)).

We can now turn to models of the utilization of services. The most commonly-used model in this field is the shopping model and, even for shopping, the data are often poor The only systematic data in the U K are provided by the Census of Distribution, which takes place spasmodically, and in such a way that successive censuses are rarely comparable Information on housing expenditure or shopping goods is available from the Family Expenditure Survey Information on shopping trips is only available from special surveys, and usually in person-trip units rather than shopping-expenditure units. Some interaction data is vital to shopping model calibration as, without it, the only goodness-of-fit statistics which can be used are those which suffer from the $\alpha = 1$, $\beta = 0$ 'bogus calibration' problem Interaction and other necessary data probably could be obtained for a range of public services, libraries, health services, welfare services and so on, from tickets and records of various kinds, but there is little experience of modelling in these fields Data are even more difficult to come by for other private services: banking, legal services and so on

If data can be assembled, and a spatial interaction model is to be built, the calibration task is straightforward and has, in fact, been discussed as the basic example in Sections 12 2 and 12 3

It only remains to mention the models of the location of economic activity in Section 10 7. Since these are all econometric models, they can be estimated by standard procedures

12.9. A NOTE ON GENERAL MODELS

Most general models are likely to be assemblies of partial models whose calibration problems we have been discussing above If the models are linked, however, as are the residential-location model and the service-sector model within Lowry's general model, then in theory there is a degree of interdependence of the parameters in the different models This is no different in principle to the interdependence of parameters in a multi-parameter partial model and can be handled in the same way, using an iterative procedure This has been done for the Lowry model by Batty (1970), who found that although there was some relationship between the parameters of the two main models, this was not particularly strong, and the calibration task was straightforward

12.10. SOME SPECIAL PROCEDURES

There are many special and alternative procedures which can be adopted for model calibration and testing within the general framework presented in this section Many of the procedures outlined above involve nests of iterative cycles, both in calibration, and within the model itself (for example calculating A_is and B_js). These iterations can be very time-consuming on the computer, and there is much scope for investigating alternative ways of structuring the iteration in an attempt to save time Two examples are presented here first, an approximation used in the calibration of the distribution model in the SELNEC study, and second, a non-iterative calculation for solving the maximum-likelihood equation for β in a distribution model These methods will only be outlined here, and the reader is referred to the original papers for the details by Wilson, Hawkins, Hill and Wagon (1969), and Evans (1971) respectively.

In the SELNEC case, the problem was the time taken up during the iterative calculation for A_i^n and B_j in a model of the form of equation (12 50) with two-person types, so that there were two parameters to be estimated, β^1 and β^2. The essence of the procedure adopted to relieve the situation was to calibrate (that is, to run the model for a range of β^1 and β^2) using a singly constrained model, with B_js fixed at 1, but with D_j replaced by $\hat{B}_jD_j$, where $\hat{B}_j$ was obtained from one doubly constrained run The hope is that the B_js are relatively insensitive to variations in β^1 and β^2 Then, of course, final values of β^1 and β^2 can be checked and corrected with more doubly constrained runs This turned out to be a useful time saving procedure

Evans (1971) is concerned with a maximum-likelihood equation of the form of (12 54), but stated as follows for one-person type:

$$C(\beta) = \sum_i \sum_j T_{ij} c_{ij} = y \qquad (12.101)$$

where y is the observed value and the equation is to be solved for β The essence of the method is that $C(\beta)$ can be calculated from $C(O)$ and the derivatives of C evaluated at $\beta = 0$ in a Taylor expansion:

$$C(\beta) = C(0) + \beta \frac{dC(0)}{d\beta} + \frac{\beta^2}{2!} \frac{d^2C(0)}{d\beta^2} + \ldots \qquad (12\ 102)$$

His paper contains explicit formulae for the derivatives up to the tenth order If $C(\beta)$ in equation (12 102) is replaced by the observed value y from equation (12 101), then this gives a polynominal equation in β which can be solved by any of the standard methods The number of terms needed in the Taylor expansion can be estimated by looking at the first-order approximation for β.

12.11. REFERENCES

* W Alonso (1968) Predicting best with imperfect data, *Journal of the Institute of American Institute of Planners*, 34, pp 248-255.

R Artle (19[illegible]) *[illegible] structure of the Stockholm* [illegible] Institute of the [illegible] School of Economics, rep[illegible] Cornell University Press, Ithaca, New York.

M. Batty (1970) Some problems of calibrating the Lowry model, *Environment and Planning*, 2, pp 95-114

M Batty (1971) Exploratory calibration of a retail location model using search by golden section, *Environment and Planning*, 3, pp 411 432

M Batty and S Mackie (1972) The calibration of gravity, entropy and related models of spatial interaction, *Environment and Planning*, 4, pp 131-250

A. J Blackburn (1970) A non linear model of the demand for travel and An alternative approach to aggregation and estimation in the non-linear model, in R E Quandt (Ed) (1970) *The demand for travel theory and measurement*, Heath Lexington Books, Lexington, Mass , pp 163 196

D Brand (1972) The state of the art of travel demand forecasting, a critical review, mimeo Harvard University Cambridge, Mass

Bureau of Public Roads (1965) *Calibrating and testing a gravity model for any size urban area*, Government Printing Office Washington, D C

M Chisholm (1971) Forecasting the generation of freight traffic in Great Britain, in M Chisholm, A E Frey and P Haggett (Eds) *Regional forecasting*, Butterworths, London pp 431 442

* E L Cripps (1970) A comparative study of information systems for urban and regional planning I Scandinavia, Working Paper 10, Urban Systems Research Unit University of Reading

* E L Cripps and A F Cater (1972) The empirical development of a disaggregated residential location model some preliminary results, in A G Wilson (Ed) *Patterns and processes in urban and regional systems* Pion, London, pp 114 145

* E L Cripps and P Hall (1969) An introduction to the study of information for urban and regional planning, Information Paper 8, pp 22 81, Centre for Environmental Studies, London

N R Draper and H Smith (1966) *Applied regression analysis*, John Wiley, New York

A W Evans (1971) The calibration of trip distribution models with exponential or similar functions, *Transportation Research*, 5, pp 15-38

G. J D Hewings (1971) Regional input-output models in the U K some problems and prospects for the use of non-survey techniques, *Regional Studies*, 5, pp 11-22

G M Hyman (1969) The calibration of trip distribution models, *Environment and Planning*, 1, pp 105-112

Ministry of Transport (1966) *Portbury*, H M S O , London

* G H Orcutt, H W Watts and J B Edwards (1968) Data aggregation and information loss, *American Economic Review*, 58, pp 773-787

P O'Sullivan (1971) Forecasting interregional freight flows in Great Britain, in M. Chisholm, A E Frey and P Haggett (Eds), *Regional Forecasting*, Butterworths, London, pp 443 450

W. I Morrison (1972) An input-output model of Peterborough, unpublished Ph D Thesis, Department of Geography, University College, London

D A Quarmby (1967) Choice of travel mode for the journey to work some findings, *Journal of Transport Economics and Policy*, 1, pp 273-314, also reprinted in R E Quandt (Ed) (1970) *The demand for travel theory and measurement*, Heath Lexington Books, Lexington, Mass , pp 235-295

J I Round (1972) Regional input-output models in the U K : a re-appraisal of some techniques, *Regional Studies*, 6, pp 1-9

. E Ruiter (1972) Analytical structures, paper prepared for H R B Conference on Travel demand forecasting, Williamsburg, Virginia, to be published

M L Senior and A G Wilson (1973) Disaggregated residential location models. some tests and further theoretical development, in E L Cripps (Ed) *Proceedings*, Fifth Annual Conference, British Section, Regional Science Association, Pion, London, to be published.

R Stone (1967) *Mathematics in the social sciences*, Chapman and Hall, London.

H Theil (1965) *Economics and information theory*, North Holland, Amsterdam

A G Wilson (1967) A statistical theory of spatial distribution models, *Transportation Research*, 1, pp 253-269.

A G Wilson (1970) *Entropy in urban and regional modelling* Pion, London

A G Wilson (1973) Further developments of entropy maximizing transport models, *Transportation planning and technology*, 1, pp 183 193

A G Wilson, A F Hawkins, G J Hill and D J Wagon (1969) Calibrating and testing the SELNEC transport model, *Regional Studies* 3, pp 337 350, also reprinted in A G Wilson (1972) *Papers in urban and regional analysis*, Pion, London, pp 202 215

C B Winsten (1967) Regression analysis vs category analysis, pp 17 19, *Seminar Proceedings*, Trip end estimation, P T R C , London

APPENDIX 1

Entropy-maximizing methods

The principles and uses of entropy-maximizing methods are described in another book by the author (Wilson, 1970) and only a brief presentation is given here. The methods will be outlined in relation to spatial interaction models, though they could be applied more widely.

The family of spatial interaction models developed in Chapter 6, and frequently used in the rest of the book, were based on the three assumptions (6.1)–(6.3). The basic situation is illustrated in Figure A1.1. Some interaction between zones i and j is taken as proportional to each of a mass at i, a mass at j, and inversely proportional to some function of the distance (or travel cost) between them. The model is based on a Newtonian analogy.

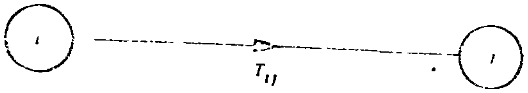


Figure A1.1 Basic interaction

The entropy-maximizing method changes the basis of the analogy and deals directly with the basic components of the system of interest. Suppose we are concerned with people and, say, the journey-to-work. The Newtonian analogy takes 'residential population' and 'jobs' as masses. The entropy-maximizing method works with individuals, assesses their probability of making a particular journey to work and, essentially, obtains the interaction as a statistical average. We illustrate the method for one case only: the production-attraction constrained model of Chapter 6. The reader should then be able to see that an analogous method can be used for other members of the family of spatial interaction models, or other 'probability assignment' models, and he can find the details in the book cited above (Wilson, 1970).

Suppose elements of the matrix $\{T_{ij}\}$ satisfy

$$\sum_j T_{ij} = O_i \qquad \text{(A1.1)}$$

$$\sum_i T_{ij} = D_j \qquad \text{(A1.2)}$$

$$\sum_i \sum_j T_{ij} c_{ij} = C \qquad \text{(A1.3)}$$

Consider three levels of resolution at which the system of interest can be described as shown in Figure A1.2 (and, to fix ideas, let us again take the journey to work as an example).

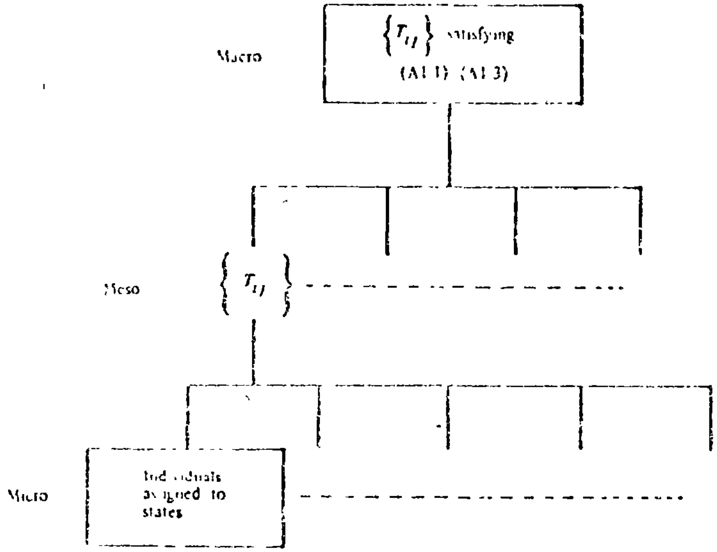


Figure A1.2 Levels of resolution in the analysis of interaction

At the most detailed, or micro, level, we could assign individuals to particular work trip categories, as indicated in Figure A1.3.

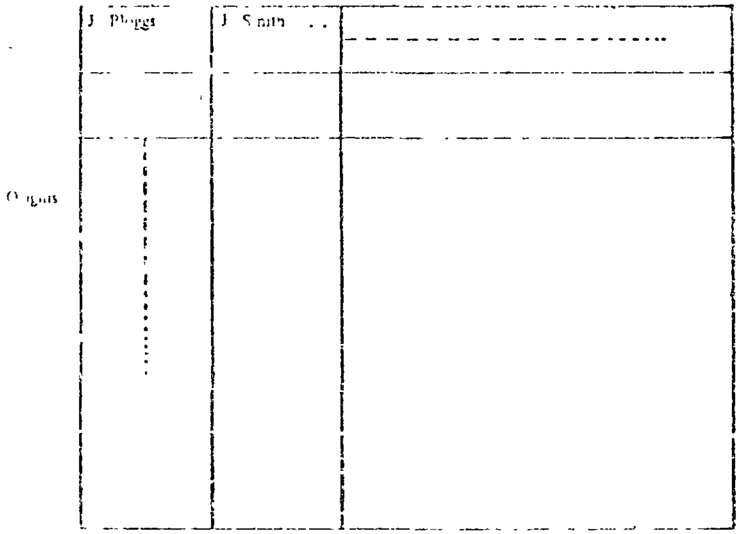


Figure A1.3 A micro state

At the intermediate, or meso level, we are only interested in the total interaction counting individuals, as shown in Figure A1.4.

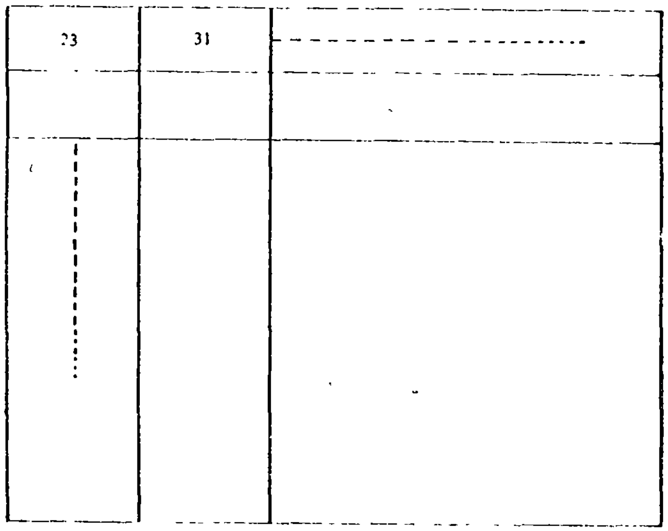


Figure A1.4 A meso state

At the macro level, our only restriction is that $\{T_{ij}\}$ should satisfy the given constraints (A1.1)–(A1.3). Clearly, there are many micro-states, assignments of individuals, which give rise to a given meso-state, and many meso-states which give rise to the given macro-state, as indicated in Figure A1.2. We now make the assumption *that all micro-states (within the restriction of the overall macro-constraints) are equally probable*. Then, the most probable meso-state is that with the greatest number of micro states associated with it.

The number of ways of assigning individuals to the boxes of Figure A1.3 and giving rise to $\{T_{ij}\}$ is $W(\{T_{ij}\})$ where

$$W(\{T_{ij}\}) = \frac{T!}{\prod_{ij} T_{ij}!} \tag{A1.4}$$

It can be shown (Wilson, 1970), that if $\log W(\{T_{ij}\})$ is maximized subject to the constraints (A1.1)–(A1.3), then the resulting $\{T_{ij}\}$ has elements given by

$$T_{ij} = A_i B_j O_i D_j \mathrm{e}^{-\beta c_{ij}} \tag{A1.5}$$

where

$$A_i = 1\Big/\sum_j B_j D_j \mathrm{e}^{-\beta c_{ij}} \tag{A1.6}$$

and

$$B_j = 1\Big/\sum_i A_i O_i \mathrm{e}^{-\beta c_{ij}} \tag{A1.7}$$

ensure that constraints (A1.1) and (A1.2) are satisfied, and β is the Lagrangian multiplier associated with (A1.3). This is the doubly constrained model presented in equations (6.16)–(6.18) except that the general function $f(c_{ij})$ has been replaced by the negative exponential function, $\mathrm{e}^{-\beta c_{ij}}$.

A1. REFERENCES

* P. Gould (1972) Pedagogic review: Entropy in urban and regional modelling, *Annals, Association of American Geographers*, 62, pp. 689–700.
A. G. Wilson (1970) *Entropy in urban and regional modelling*, Pion, London

Transpn Res., Vol 11, pp 117-126 Pergamon Press 1977 Printed in Great Britain

USE OF A LOWRY-TYPE SPATIAL ALLOCATION MODEL IN AN URBAN TRANSPORTATION ENERGY STUDY

JERRY L. EDWARDS
Department of Civil and Mineral Engineering, University of Minnesota, Minneapolis, MN 55455, U.S.A

(*Received 3 June* 1975, *in revised form* 21 *April* 1976)

Abstract—This paper describes the structure of a Lowry-type spatial allocation model and the results of using the model to evaluate several different hypothetical urban forms with respect to the level of transportation energy consumed in urban passenger travel The Lowry-type model was used together with simple models of trip generation, mode choice and trip assignment to produce realistic land-use patterns and characteristic travel behavior in each city Interzonal trip matrices were computed for each of five trip purposes (work, two types of shopping or service, non-home based and social) and, following the assignment of trips to the network, energy models for automobile as well as transit were utilized to assess the transportation energy consumed over all trip purposes in each of several different urban spatial structures

INTRODUCTION

The urban transportation planning process in the U.S. has traditionally responded to changes in urban and regional goals and to shifts in the characteristics of the policy environments in which these processes have operated. For example, current issues related to community disruption and environmental/ecological damage wrought by the construction and operation of transportation facilities are now treated as very important elements in the development of an urban transportation plan One need only mention as evidence the number of environmental impact statements which have been written, pursuant to the requirements established in the National Environmental Policy Act of 1969, and dealing with federally financed highway projects having potentially deleterious environmental effects.

The consideration of environmental issues has been incorporated in the planning process and will no doubt remain important in the future. However, the oil boycott of the United States in late 1973 and early 1974 suddenly brought to light a problem which threatens to overshadow in importance all other transport planning issues. That problem is how to deal with a constraint imposed on all transportation systems from the shortage and dwindling supplies of fossil fuels Although fuel availability is not currently a major issue, it is clear that U.S dependence on crude oil from the Middle East for much of its fuel needs could, in the event of a renewed boycott, once again create serious energy shortages.

Essentially all of the policy options considered in the U.S. during the last oil embargo were short-run in their outlook. Clearly, a more effective mix of policies and actions would address the long-run nature of the problem as well. If public policy is to respond to what appears to be a long-run problem with an appropriately long-run solution, then there is strong logic in attempting to understand the fundamental relationships between urban spatial structure and energy consumed in urban passenger travel. This is because the demand for travel is a derived demand, and the spatial arrangement of urban activities is a fundamental determinant of that demand

The research to be described in the subsequent sections was founded on the premise that the structure of urban land-use and transport networks can have a very significant effect on the consumption of energy in urban passenger travel If the magnitude and direction of these effects could be discovered and documented, some strong guidance might be provided for future long-term policy development Toward that end, a major study was undertaken to examine the travel requirements and the transportation energy characteristics of many alternative urban spatial structures (Edwards, 1975, Edwards and Schofer, 1975) The purpose of this paper is to describe the structure of the spatial allocation model and to present a statistical analysis of the results of the experiments

MODELING STRATEGY

At the outset of this study, two feasible strategies for conducting a policy-oriented investigation were considered. One way was to gather existing data on land-use patterns and travel behavior from current metropolitan transportation planning studies However, such an effort would entail great expenditures of time and effort in gathering and manipulating such data sets. Another way was to treat the problem in an entirely abstract fashion, in which mathematical programming techniques are used to allocate persons to residences and employees to work places in a way that optimizes a given travel-related objective. Such normative approaches typify the investigations of Hemmens (1967) and Dantzig and Saaty (1973) which rest upon the assumption of aggregate optimizing behavior; yet, it is unlikely that persons behave in such a manner as to achieve a social optimum

The difficulties inherent in the totally empirical approach and the uncertainties associated with the normative approach made it necessary to attack the problem from a different perspective A compromise was chosen, in which data from an existing city could be utilized together with spatial allocation models to simulate travel behavior in a series of hypothetical cities Such an approach might also be more flexible than the other two approaches, in the sense that one could vary many elements in the urban spatial structure to examine their effects upon transportation energy consumption

CONCEPTUAL FRAMEWORK OF PRESENT STUDY

Conceptually, this study attempted to do the following. Choose a representative American city in which aggregate travel behavior has been observed and documented.

Resettle the residents of that city into each of several different hypothetical cities and analyze the travel patterns, transportation energy requirements and accessibility characteristics arising from changing the spatial variables (shape, form, density patterns, etc.) By utilizing appropriate trade-off analyses, identify factors of fundamental importance in leading to higher transport energy requirements and those which bring on higher levels of regional accessibility to activities of interest

In the experimental design process, values of the activity variables (population, employment, etc) must be consistent across all designs—providing a basis for comparing the travel required to connect activities within each design—in order that we may assess the effects of changing the interaction variables (e g the highway network) through which these activity variables interact Activity variables, therefore, are fixed in quantity but not by location for all designs, while interaction variables are allowed to take on values differing across the designs.

In this study, the fixed set of city attributes (activity variables) was taken from an existing city (Wilbur Smith and Associates, 1965), in order to ensure that the results of the effort would be well grounded in fact. These attributes, given along with the interaction variables in Table 1 include total population, total employment by category, labor force participation rate, interzonal impedance (friction) factors by trip type, and trip rates per capita by trip type Impedance factors represent increasing propensity of travelers to make trips of various lengths, and vary for the same trip purpose between cities Therefore, these factors are more correctly viewed as belonging to the set of spatial variables and should not be considered fixed across different urban spatial patterns. The same is true of trip rates per capita by trip type, for which there is evidence to suggest that trip frequency depends upon the spatial arrangement of activities and upon the transport system (Jordan, 1972) However, while recognizing that these attributes might vary across urban designs, there is no theory which might be utilized in accounting for the variation. Therefore, in lieu of impedance factors which are functions of the spatial pattern, a single set of factors was selected to apply for each trip purpose across all urban designs. Similarly, trip generation was assumed to depend solely upon distance from the city center, a surrogate variable selected to represent the combined effects of auto ownership, income and family size.

Table 1 Urban structural variables

Activity Variables (Fixed Attributes of Case Study City)	Interaction Variables (Varied Over the Series of Urban Designs)
Total Population	Urban Form
Employment, by Category	Transportation Network, for Each Mode
Labor Force Participation Rate	Network Modes (Technologies)
Population Serving Ratios[*]	Network Levels of Service
Floorspace Requirements for Employee, by Type	Mode Choice, by Trip Type
Interzonal Friction[**] Factors, by Trip Type	
Auto Occupancy, by Trip Type	
Trip Rates Per Capita, by Trip Type	

[*] Defined to be the number of service workers by type per capita.

[**] Defined in the context of the gravity trip distribution model

MODELING PROCEDURE

The detailed land-use patterns emerging from the selected combinations of activity and interaction variables were created by utilizing a simple spatial allocation model of the Lowry type (Lowry, 1964) The choice of the Lowry-type model was largely a function of three requirements of the study· (a) the data needs of all sub-models would have to be relatively small, (b) the models should be computationally efficient, and (c) the land-use patterns designed by the spatial modeling procedure ought to be realistic

Recent experience in Great Britain (Batty, 1969) suggested the use of a Lowry-type land-use model in the context of a New Town or a hypothetical city would satisfy these requirements, and it would offer significant advantages over more conventional land-use models since it has been widely used in Britain during recent years and appeared to replicate land-uses in realistic ways. In addition, data requirements were few, the standard travel models (trip generation, distribution and assignment) were easily incorporated with the framework of the model, and its relatively simple structure permitted experimentation over a wide range of urban designs in a relatively efficient manner

The modeling procedures used in the study can be briefly described as follows· To begin with, the values of the activity variables (attributes) are taken from an existing city Next, the interaction variables (media) through which the activity variables interact are specified exogenously, and both sets of variables are utilized by the spatial allocation model to design a land-use pattern and to estimate the person travel emerging from that pattern In turn, these trips are assigned by mode to routes of the network Estimates of the total transportation energy required for nearly all daily person travel are obtained by use of separate models of automobile, bus, and rail rapid transit energy consumption Finally, measures of accessibility to various activities of interest, including to population and shopping are obtained. The sequence of models is given in the following flow chart, Fig 1

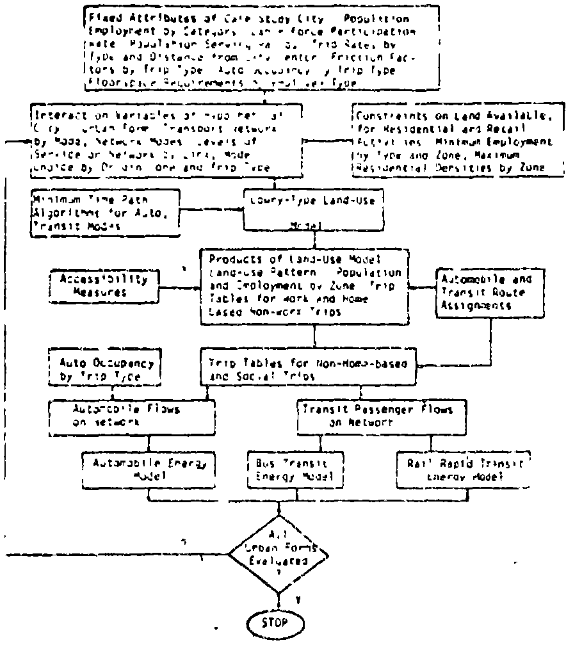


Fig. 1. Modeling sequence.

THE SPATIAL ALLOCATION MODEL

The Lowry model and its descendants (Batty, 1969, 1970, 1971, 1972; Crecine, 1964; Garin, 1966; Goldner, 1971; Goldner and Graybeal 1965; and Wilson, 1970) of which the spatial allocation model in this study can be considered a member are based upon the assumption that the spatial structure can be organized into activities (population, basic and service employment) and land-uses. Furthermore, it is assumed that service employment and population are derived from basic employment. Operationally, the non-residential activities can be defined in the following way: basic employment consists of all workers employed by firms which are site-oriented, i.e. which tend to be space-intensive, or which tend to locate independently of the local populace; and the service employees are all those who work for firms which are client-oriented, i.e. whose location behavior is dependent upon the residential pattern

The model to be described in the subsequent paragraphs can be characterized as a sequel to the work of Batty (1971). That being the case, a number of important steps in its derivation are not repeated here. Instead, the model is displayed in skeletal form, except for those parts which are original extensions of the earlier work Before presenting the mathematical structure of this model, however, a brief verbal description of the allocation process seems appropriate.

From an exogenous distribution of basic employment, the model allocates these employees to residences by a work-trip spatial interaction model. The population associated with these employees is found by applying the reciprocal of the labor force participation rate. This population demands a certain number of service employees (non-basic) of a particular type, and that number is found by multiplying the client population by a population-serving ratio. The service employees are allocated to their shop sites by a service-trip spatial interaction model. From the shop sites, these employees find residences by way of the same work-trip spatial interaction model used to allocate the basic employees to homes. The number of families of the service employees is found, which in turn demand services and who require additional service employment. The process of deriving additional residential and employment activity eventually ceases (converges) and once convergence is attained, the population-holding capacities of each zone are checked against the allocated values of population as well as the minimum service employment values. If any of these constraints is violated, the parameters used to drive the spatial interaction models are changed to prevent this from recurring and the allocation procedure is repeated until all of the constraints are satisfied.

The spatial interaction model, whose required inputs are given in Table 2 and whose variables are defined in Table 3, is described mathematically in the following paragraphs.

The index m denotes the iteration number in which additional activities are derived, and the index n denotes the interaction in which the locational constraints are tested.

The work-trip spatial interaction model is given as follows:

$$a_{ij}(n) = \frac{U_i(n)G_iF_{ij}}{\sum_i U_i(n)G_iF_{ij}} \tag{1}$$

$$T_{ij}(m,n) = E_i(m,n)a_{ij}(n) \tag{2}$$

$$P_j(m,n) = \alpha \sum_i T_{ij}(m,n) \tag{3}$$

from these one obtains the procedure of allocating employees to residences:

$$P(m,n) = E(m,n)A_i(n). \tag{4}$$

The service-trip spatial interaction model is specified in the following equations.

$$b_{ijk}(n) = \frac{V_{jk}(n)H_{jk}FF_{ijk}}{\sum_j V_{jk}(n)H_{jk}FF_{ijk}} \tag{5}$$

$$S_{ijk}(m,n) = \beta_k P_i(m,n)b_{ijk}(n) \tag{6}$$

$$NT_{ijk}(m,n) = \gamma_{ik}P_i(m,n)b_{ijk}(n) \tag{7}$$

Table 2. Inputs to spatial allocation model

Basic employment, by Zone
Network Interzonal Travel Times, by Mode
Labor Force Participation Rate
Population Serving Ratios, by Service Category (e g , personal business, recreational, convenience, shopping, etc)
Friction Factors (Interzonal Impedance Factors), by Trip Type
Estimates of the Zonal Attraction for Receiving Population, by Zone
Estimates of the Zonal Attraction for Receiving Service Activity by Type, by Zone
Available Land for Development, by Zone
Maximum Residential Density Constraints by Zone
Minimum Service Employment by Type and by Zone
Floorspace Per Employee by Service Type and by Zone
Daily Trips Per Person by Trip Type and by Origin Zone
Mode Choice by Origin Zone and Trip Type

Table 3 Definitions of variables in spatial model

α	reciprocal of the labor force participation rate
$a_{ij}(n)$	probability of traveling from work place in zone i to residence in zone j
$A(n)$	N-dimensional square matrix whose i, j element is $a_{ij}(n)$
$A_1(n)$	N-dimensional square matrix whose i, j element is $\alpha a_{ij}(n)$
β_k	population-serving ratio for service category k
$b_{ijk}(n)$	probability of a resident in zone i traveling to zone j for service purpose k
$B_k(n)$	N-dimensional square matrix whose i, j element is $b_{ijk}(n)$
$B_{1k}(n)$	N-dimensional square matrix whose i, j element is $\beta_k b_{ijk}(n)$
γ_{ik}	number of daily trips per person from zone of residence i for service purpose k
$E_{jk}(m,n)$	number of service workers of type k employed in zone j, iterations m and n
$E_i(m,n)$	number of workers of all service types employed in zone i, iterations m and n (When $m = 1$, it represents the number of basic employes allocated exogenously to zone i)
$E(m,n)$	N-dimensional row vector whose ith element is $F_i(m,n)$
$\Gamma_k(n)$	vector of total employment associated with service establishments of type k, outer iteration n
$E(n)$	total employment vector
F_{ij}	friction or impedance factor associated with traveling from workplace in zone i to residence in zone j
FF_{ijk}	friction or impedance factor associated with traveling from zone of residence i to zone of service j for service purpose k
G_j	utility or locational attraction of zone j for receiving population
H_{jk}	utility or locational attraction of zone j for receiving service employees of type k
i, j	zonal indices
k	index representing service employment type
K	total number of service employment types
N	number of zones
$NT_{ijk}(m,n)$	expected number of trips for non-work purpose k from zone of residence i to zone of service j, iterations m and n
$NT_{ijk}(n)$	total expected number of non-work trips for purpose k from zone i to zone j
$P_i(m,n)$	population allocated to zone i, iterations m and n
$P(m,n)$	N-dimensional row vector, whose ith element is $P_i(m,n)$, iterations m and n
$P(n)$	total population vector
$S_{ijk}(m,n)$	expected number of service workers of type k allocated to zone j, serving residents in zone i, iterations m and n
$T_{ij}(m,n)$	expected number of workers traveling from work in zone i to home in zone j, iterations m and n
$T_{ij}(n)$	total expected number of work trips from zone i to zone j
$U_j(n)$	balance factor during iteration n to prevent over-allocation of workers to residences in zone j
$V_{jk}(n)$	balance factor during iteration n to insure against too few employees of type k being allocated to zone j

The procedure of allocating service employees to service establishments is found by defining

$$E_{jk}(m+1,n) = \sum_i S_{ijk}(m,n) \tag{8}$$

and

$$B_1(n) = \sum_{k=1}^{K} B_{1k}(n) \tag{9}$$

from which one obtains

$$E(m+1,n) = E(1,n)[A_1(n)B_1(n)]^m \tag{10}$$

For a fixed value of n, the process of iterating over m converges (Edwards, 1975) and one obtains

$$E_k(n) = \sum_{m=2}^{\infty} E_k(m,n) \tag{11}$$

from which it is true that

$$E_k(n) = E(1,n)A_1(n)[I - B_1(n)A_1(n)]^{-1}B_{1k}(n). \tag{12}$$

Now, if the total employment vector $E(n)$ is defined to be

$$E(n) = E(1,n) + \sum_{k=1}^{K} E_k(n) \tag{13}$$

and if total population is defined to be

$$P(n) = \sum_{m=1}^{\infty} P(m,n) \tag{14}$$

then total employment and population are found as follows

$$E(n) = E(1,n)[I - A_1(n)B_1(n)]^{-1} \tag{15}$$

$$P(n) = E(n)A_1(n) \tag{16}$$

Once the allocations have been made in iteration n, each zone is examined to see if the constraints on the minimum number of retail employees are satisfied, as well as checking to make sure overcrowding of population has not taken place If for any zone the minimum retail employment constraint is violated another iteration is required and the attraction index for receiving service employment on the next iteration is set to zero Similarly, if any zone is found to have violated the residential density constraint, the zones which have additional holding capacities are made proportionately more attractive, while the attraction is proportionately diminished for those zones which are overcrowded. Once the constraints are all satisfied, the interzonal work trips are estimated as follows

$$T_{ij}(n) = E_i(n)a_{ij}(n) \tag{17}$$

In a similar fashion, the non-work trip matrices are found, for each purpose k as follows:

$$NT_{ijk}(n) = \gamma_{ik}P_i(n)b_{ijk}(n). \tag{18}$$

The final products of the spatial allocation model are given in Table 4

The model utilized in the study differs from the work of Batty and others by allowing for the definition of disaggregate service employment types In fact, each type of service employment (e g insurance workers) corresponds uniquely with a particular non-work or service trip

Table 4 Products of the spatial allocation model

Total Employment for Each Service Category, by Zone

Total Population, by Zone

Total Land Occupied by Non-residential Activities, by Zone

Work Trip Table and Characteristics by Mode
(Number of Person and Vehicle Trips, Total Vehicle Miles, Average Trip Length, Time and Speed, Trip Length Distribution)

Non-work (Service) Trip Tables and Characteristics
(Same as for Work Trip)

made by urban residents (e.g. personal business) Consequently, there is a one-to-one relationship between the types of service workers and the types of service (shopping) trips made by residents Early in the study, an attempt was made to disaggregate the existing employment data to conform precisely with the following non-work (service) trip purposes personal business, recreational, school, shopping G A F (goods, apparel and furnishings) and shopping for convenience purposes. However, there were considerable problems encountered in trying to categorize each employment type into a service category corresponding exactly to one of the non-work types Furthermore, except for the convenience shopping purpose, all trip types had relatively long average trip lengths (well over six minutes vs about five minutes for convenience shopping trips) Therefore, in this study the shop-to-client interaction was considered to be significantly different in the case of the other types from that of the convenience type In the first case, it can be said, because of the relatively long trip lengths, that the trip purpose comprising personal business, recreational, school, and G A F. shopping trips is accomplished by being made to locations which are not in close proximity to their clients In contrast, trips for convenience shopping purposes are made to sites in relatively close proximity to home. Therefore, it was decided that the non-work trip purposes excluding convenience shopping trips would be collectively considered to be one purpose and named type *N* (insensitive to location of client), and convenience shopping would be another purpose, type *S* (sensitive to location of client). Accordingly, the employees are either of the basic type or of the non-basic types, type *N* or *S* Recall that Lowry's categorizations of non-basic types were metropolitan, local, and neighborhood (Lowry, 1964)

EXPERIMENTAL DESIGN

The basic attributes (activity variables) of the selected case study city are those of Sioux Falls, South Dakota (population 95,000) found in a 1965 metropolitan transportation study (Wilbur Smith and Associates, 1965) and augmented with data from Baton Rouge, Louisiana (Wilbur Smith and Associates, 1967) The population, employment and travel data are representative of those found in small-to-medium sized American cities and were selected for several reasons First of all, it was felt that there would be too many attendant computational difficulties resulting from a study of a large metropolitan area For example, the number of traffic assignment zones in an urban transportation study can be considered to be proportional to the population, and computational times increase essentially proportional to the square of the number of zones. Secondly, by selecting a smaller city, one should be able to comment on the efficiency of New Town forms and conurbations in view of the fact that several British and American New Towns are projecting comparable population and employment levels

Three basic urban shapes were selected for study and they are depicted in Figs. 2–4. The first, in Fig. 2, is a concentric ring shape enclosing roughly the same land area (147.25 mi^2) as the Sioux Falls area This basic ring shape with superimposed grid transport network is typical of most midwestern cities in the U.S. The other two forms are more radical, one being a pure linear shape depicted in Fig. 3, and a polynucleated shape shown in Fig. 4. The linear shape was selected to be an urban shape which represents city forms having low transportation capital costs (in terms of roadways and transit facilities), good proximity to activities in terms of distance, and a relatively compact (9.75 mi^2) land-use pattern. The polynucleated shape, on the other hand, is a component of form which is attractive from the point of view of accessible open space, but represents a land-use pattern incorporating nuclei of fairly high density (a total developed area of only 6.1 mi^2) having neighborhood and community facilities within walking distance. A total of 35 experiments were conducted utilizing these three basic shapes, and the detailed experimental design is set out in Edwards (1975). Two additional experiments were conducted using a pure cruciform design, depicted in Fig 5. The cruciform, it was felt, might combine the best features of the linear and polynucleated shapes, with the basic design offering physical separation of neighborhoods from the commercial and industrial areas, yet the relatively compact land use (10.25 mi^2 of developed land) is spread out in such a manner as to offer good accessibility to open space.

Selection of zone size for each shape was done in such a manner as to capture as much vehicular traffic as possible It was assumed in the concentric ring shape of Fig 2 that a good deal of the urban activity would be

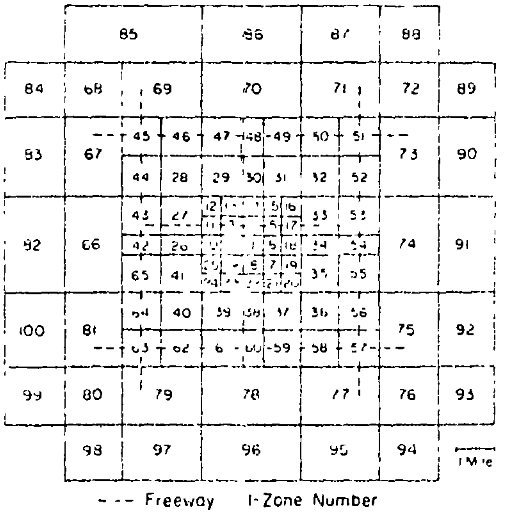


Fig. 2. Concentric ring shape

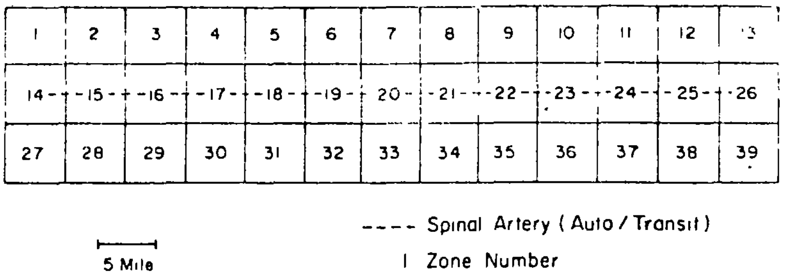


Fig. 3. Pure linear shape.

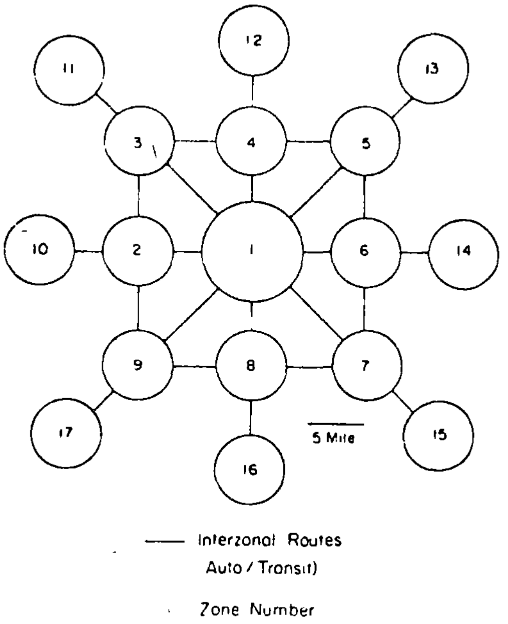


Fig 4 Polynucleated shape

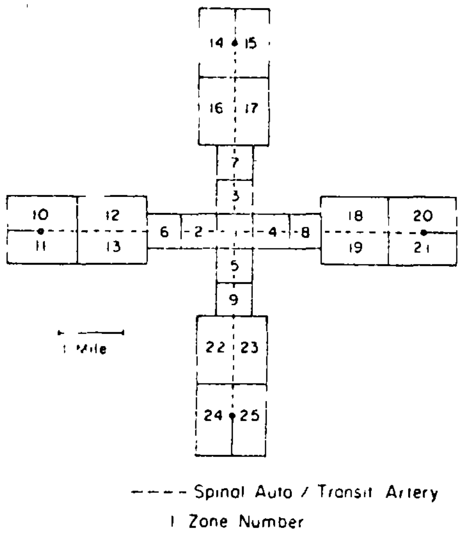


Fig 5 Pure cruciform

transacted in the central 25 zones Consequently, zonal size was chosen to be 0.25 mi² (0.5 mile on a side), and it was assumed that intrazonal trips in this area would be accomplished on foot The area encompassed by zones 1–81 (20.25 mi²) is roughly comparable to the area enclosed by the City of Sioux Falls Consequently, zones 26–65 were designed to be areally larger (1 mi²), owing to the assumption that zonal activities in this ring would not be as numerous as those in the central area, and most automobile trips would likely cross the zonal boundaries. The zones on the periphery are, in turn, even larger (in excess of 2.25 mi²) than zones 1–65.

Selection of the urban form in each experiment is incomplete without specifying density patterns to accompany the chosen shapes As previously mentioned, the patterns were essentially determined by specifying the attraction utilities G_i (eqn 1) and H_{jk} (eqn 5) In those experiments where uniform (sprawled) densities were desired, the attraction values were directly proportional to zonal area Where peaked density patterns were desired, the attraction utilities were proportional to desired levels of population (or employment) densities

A priori specification of the network modes, levels of service and mode choices by zone are given in detail in Appendix A to Edwards (1975) along with the specific values of the attractor variables for each experiment.

RESULTS

The total transportation energy required in each experiment is exhibited together with the corresponding regional accessibility to population in Fig 6 Total energy refers to that energy required to accomplish all daily person travel from home to work, from home to service

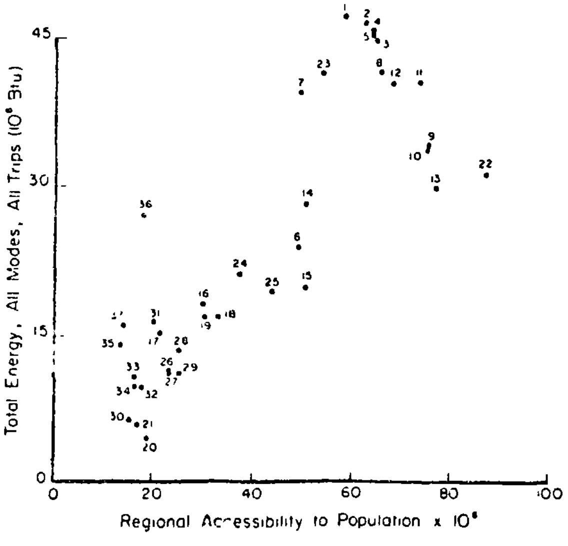


Fig 6 Total energy and regional accessibility to population for each experiment

establishments type N, from home to service establishments type S, from home for social purposes, and one-half of the total daily non-home based travel. Assuming that the amount of daily person travel to home will match that amount from home, the total energy required of all travel (except for those trips of type serve-passenger and change-mode) can be estimated by doubling the amount displayed in Fig 6. Regional accessibility to population is defined in eqn (19)

$$\text{Regional Accessibility} = \sum_{\substack{\text{all zones} \\ i,j}} P_j F_{ij} \qquad (19)$$

when P_j is the population residing in zone j, and F_{ij} is the work-trip friction (impedance) factor incurred between zones i and j. It was found that this regional measure is very highly correlated with measures of accessibility to services of types N and S (Edwards, 1975) and was therefore chosen to represent those measures of accessibility.

The first observation which can be made from Fig. 6 is that there exists wide variation in transportation energy requirements for differing urban structures Generally speaking, those urban structures which are characterized by sprawling land-use patterns have energy requirements which are much larger than those which are relatively compact. For example, the first five experiments have total energy requirements anywhere from 9 to 10 times the energy required in the least energy-intensive structure (experiment 20) At the same time, however, those five urban structures have the greatest dispersion of population and employment (as measured by the second moments of population and employment about zone 1) of all cities examined, while experiment 20 represents a much more compact land-use pattern.

Those cities with compact land-use patterns occupy the more energy-efficient locations in the space of feasible structures For example, the pure linear forms (9 75 square miles) are represented in experiments 24 through 29, the cruciform structures in experiments 30 and 31, and the polynucleated forms are evaluated in experiments 32 through 37 With the exception of experiment 36, which is a polynucleated structure with an auto-only transportation network, these structures occupy the lower left-hand portion of the trade-off space, and represent cities with low energy costs, but concomitant low levels of regional accessibility to population

Those cities which utilize only the automobile as the means of transport have energy requirements which are generally much larger than those cities which employ public transportation For example, all those cities examined in the first 13 experiments utilize only the automobile, as well as experiments numbered 22 through 25 and number 36 If one were to examine the energy value of 20 0 × 10^8 Btu, one finds that only one city (experiment 14) employing a combination auto-transit transport network exceeds that value for total transport energy, and only one city (experiment 25) utilizing the auto as its sole means of transport has a lower energy requirement Hence the value of 20 0 × 10^8 Btu could be taken in this sequence of experiments to be the energy threshold above which nearly all of the automobile-oriented energy costs lie, and below which almost all of the energy requirements of auto-transit cities are found.

Those urban structures with the same basic urban shape (concentric ring) vary in energy requirements, as well as accessibility, according to whether the density pattern of activities (population, employment of each type) is sprawled or concentrated; and depending upon the relative importance of the automobile in the overall transportation system For example, each experiment in numbers 1 through 5 and 9 through 13 represents a city with relatively large land requirements (147 25 mi^2), but the first five are sprawled patterns and experiments 9 through 13 represent cities where density patterns have activities concentrated more in the central area At the same time, experiments 1–5 and 9–13 might be visualized as lying on or near a line segment running from upper left to lower right, with movement downward and to the right being accompanied by increasing concentrations of activities. Once transit is introduced to augment the automobile, the energy requirements fall, but the level of accessibility is decreased as well. For example, experiments 14 through 21 represent basically the same city as in experiment 13, but with different levels of transit and modal choices. Generally speaking, for the same relative transit service (ubiquitous service, and a frequency of 10 buses per hour), an increase in modal split for transit service over the automobile brings about a drop in both total energy required for accomplishing that travel, and regional accessibility to population Accessibility is decreased largely because of the time penalties which transit passengers pay in incurring longer out-of-vehicle and line-haul travel times

A close examination of the results of all 36 experiments reveals several interesting facts, most of which will be discussed in some detail as the discussion proceeds To begin with, total energy consumed by all modes over all trip purposes is very highly correlated with any variable capturing variation in work travel, the most prominent being (a) auto vehicle-miles traveled to work ($\rho = 0.991$), (b) average trip length by auto to work ($\rho = 0.943$) and (c) total person-miles traveled to work on all modes ($\rho = 0.943$). This is not surprising in view of the facts known about the energy intensity of low-occupancy automobiles and of the way in which the spatial allocation model operates. More specifically, the model designs a land-use pattern based upon the journey from work to home. From Fig. 6 one observes that land-use patterns which are relatively expansive (and characterized by long work trips) are also energy-intensive Hence, any variable capturing change in work travel will likely explain a good deal of the variation in energy.

The results of three separate linear regressions of energy upon auto vehicle miles traveled to work are given in Fig 7. The solid line represents the results for all experiments, the dashed line for the auto/transit cities, and the dashed-dotted line for the auto-only cities All were significant at the 1% level. The importance of the work trip is even more graphically illustrated by the relationship between average work trip length by auto and total energy (all trip purposes) for those cities in which the

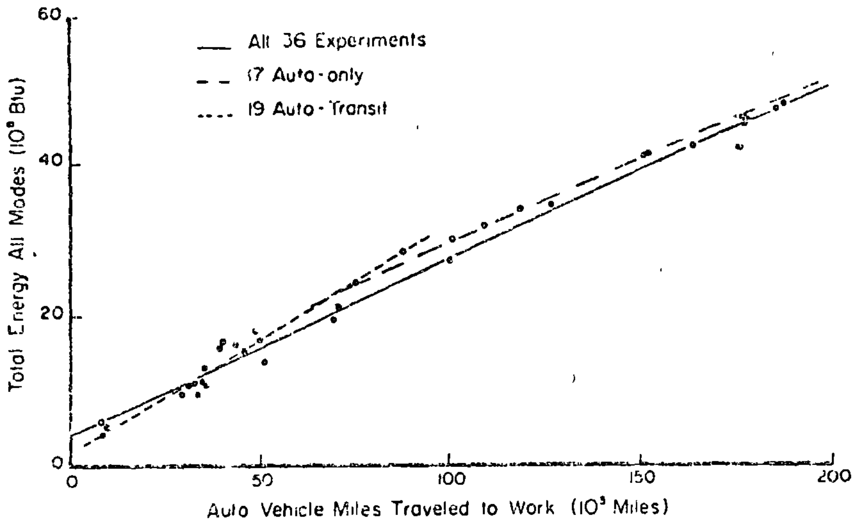


Fig 7. Total energy, all modes vs auto vehicle miles to work for 36 experiments

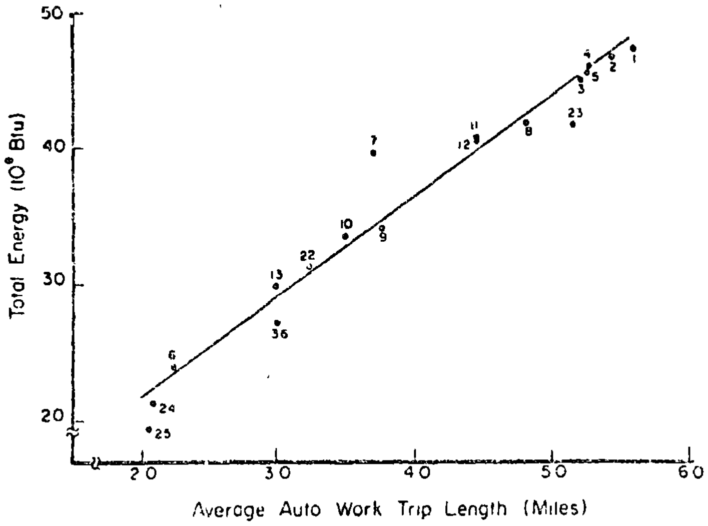


Fig 8. Total energy as a function of average work trip length for 18 auto-only experiments

auto is the sole mode. These results are illustrated in Fig 8. The equation of the line segment drawn is given in eqn (20) as follows

$$\underset{(19.1)}{\mathrm{TEA} = 7.34}\ \mathrm{WAL} + \underset{(4.35)}{7.09} \tag{20}$$

$R^2 = 0.96$, $F_{1,16} = 363.0$ (t statistics beneath coefficients)

where TEA is the total energy (by automobile) for all trip purposes in 10^9 Btu, and WAL is the average work trip length by auto in miles. The regression is significant at the 1% level, and with the exception of experiment 7, the straight line captures nearly all of the observations. That experiment represents a spread city with a large expansive low density land-use pattern. There is a relatively short average work trip for the developed area because employment is spread uniformly over the entire area. However, intrazonal travel (e.g. service trips of type 5) outside the central zones is accomplished entirely by auto and there is extensive stop and delay movement, making this structure relatively energy-intensive.

Even though the average length of auto work trips correlates very highly with energy in the auto-only cities, this variable does not correlate as well ($\rho = 0.752$) in those cities where both modes are present. Indeed, not even the average work trip length over all modes will capture this variation, as seen by Fig. 9. This is because a work trip in a multi-modal city is no longer the sole determining factor in explaining variation in total energy. It is, rather, one of several important factors which are found to influence total energy consumed in transportation. Included among those factors are urban shape and the ratio of transit person trips to all person trips (mode split). More specifically, Fig. 10 exhibits apparently different sensitivities of energy, to changes in average work trip length depending upon the shape. Experiments 14 and 16–21 represent concentrated ring cities, while numbers 32–35 the polynucleated types. However, numbers 32–35 are characterized by the same mode split (50% transit) while

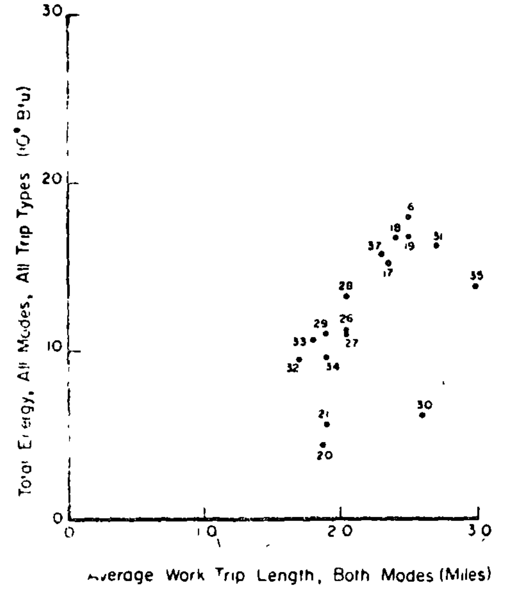


Fig 9 Total energy, all modes, all trip types, as a function of average work trip length over all modes

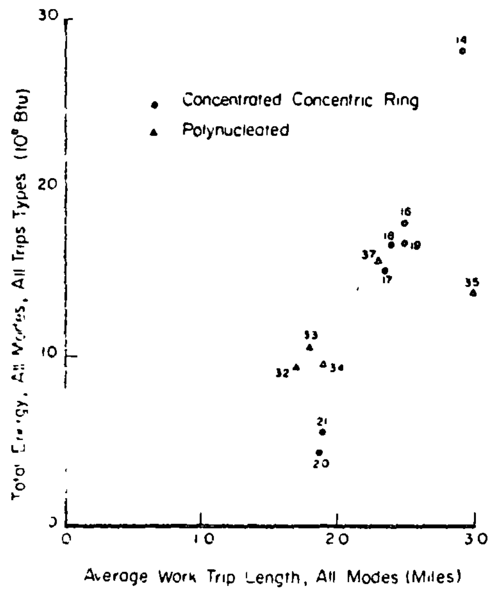


Fig 10 Total energy, all modes, all trip types as a function of average work trip length, all modes for different urban forms

number 14 differs from numbers 16–19 (roughly 10% transit as opposed to about 37% transit) which in turn differ from numbers 20 and 21 (roughly 65% transit).

This suggests, among cities of the same shape, that the energy sensitivity to average work trip length is also a function of mode split. This functional relationship, as yet untested, is displayed in Fig 11. Experiments 1–13, as well as numbers 22 and 23 were conducted with only the auto mode, while number 14, as mentioned previously, employed transit as well as auto. By constructing checked lines parallel to the solid line, Fig 11 implies that variation in energy, for cities of the same shape and mode split, can be predicted by variation in average work trip length.

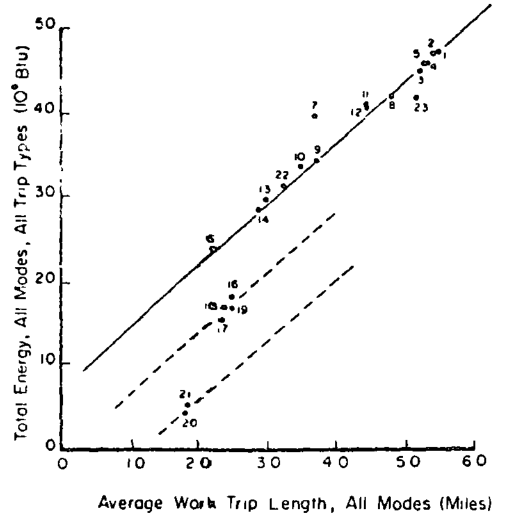


Fig 11 Total energy, all modes, all trip types as a function of average work trip length, all modes for 20 concentric ring cities

URBAN STRUCTURAL COMPONENTS IMPORTANT IN DETERMINING TRANSPORTATION ENERGY

From one experiment to the next, there were several factors varying simultaneously, making it very difficult in many cases to separate the important factors from those which are only secondary. However, in most cases, there emerged a clear correspondence among sets of experiments between variation in total energy consumption and the factor under consideration. Detailed analyses of these factors (or structural components) are given in Edwards (1975) and Edwards and Schofer (1975). The principal components found to be important in explaining differences in energy were (1) the urban form, (2) the overall level of service provided by the transportation system (as measured by overall travel speed), and (3) the extent to which public transit plays a role in the transport system (the mode split). With respect to the first component, four distinct dimensions were separately identified as contributing subcomponents. These were (a) urban shape, whether concentric ring, linear, polynucleated or cruciform; (b) the geographic extent, (developed land area in square miles), (c) the degree of concentration of population about the city centroid (measured by the second moment), and (d) the degree of employment concentration about the city centroid (also measured by

Table 5 Urban development factors to be considered in formulating energy conserving policies

Contributory Causes leading to	
Relatively High Energy	Relatively High Accessibility
Expansive land-use	Compact land-use
Spread of population, employment	Concentration about central zones of residential and retail activities
Stop and delay automobile traffic	Free flow on street and freeway network
High level of service	High level of service
Long work trips	Short work trips
Predominance of automobile in transportation system	Predominance of automobile in transportation system

the second moment) Most of the salient factors associated with higher energy levels are succinctly summarized in Table 5.

CONCLUSIONS

In summary, the spatial allocation model attempts to simulate an urban system whose components adjust themselves immediately to some static equilibrium. This rather idealized view of an overwhelmingly complex urban setting is in many ways unrealistic since the model structure does not include the dynamic nature of cities, and location behavior is simulated at only one point in time. This does not mean that the model is unsuitable for spatial planning or for study of the urban spatial structure of cities. It does mean, as Massey (1973) and Batty (1970) have pointed out, that the model can only be used for conditional impact analysis involving the proper interpretation of the equilibrium state of the system. For example, conditional upon the assumptions relating to preferences of workers for various trip lengths which are inherent in the specification of work trip friction factors, the model allocates workers to homes until equilibrium is attained. Should the assumption about travel behavior be in error, the resulting allocation may or may not be an accurate sketch of the residential location process.

REFERENCES

Batty M. (1969) The impact of a new town--An application of the Garin-Lowry model. *J Town Planning Inst.* 55. 428-435.

Batty M. (1970) An activity allocation model for the Nottinghamshire-Derbyshire subregion. *Regional Studies*, 4. 307-332.

Batty M. (1971) Design and construction of a subregional land-use model. *Socio-Economic Planning Sciences*, 5, 97-124.

Batty M. (1972) Recent developments in land-use modelling: A review of British research *Urban Studies*, 9, 151-177

Cinlar E (1969) *Stochastic Processes*, p. 80. Northwestern University.

Crecine J. P. (1964) TOMM (Time-Oriented Metropolitan Model) Technical Bulletin No 6 Community Renewal Program, Pittsburgh.

Dantzig G. and Saaty T. (1973) *Compact City* W H Freeman and Co., San Francisco.

Edwards J. L. (1975) Relationships between transportation energy consumption and urban spatial structure. Unpublished Ph D. dissertation, Department of Civil Engineering, Northwestern University, Evanston, Illinois.

Edwards J. L. and Schofer J. L. (1975) Relationships between transportation energy consumption and urban structure: results of simulation studies. Paper presented at *54th Ann. Meet. Trans Res. Board*, Washington, D.C. To be published in future issue of *Trans Res Record*.

Garin R. A. (1966) A matrix formulation of the Lowry model for intrametropolitan activity allocation. *J. Am. Inst Planners* 32, 361-364.

Goldner W. (1971) The Lowry model heritage. *J Am. Inst Planners* 37, 100-110.

Goldner W and Graybeal R. S. (1965) *The Bay Area Simulation Study: Pilot Model of Santa Clara County and Some Applications*. Center for Real Estate and Urban Economics, University of California, Berkeley.

Hemmens G. (1967) *Experiments in Urban Form and Structure* pp. 32-41. Highway Research Record No. 207, Highway Research Board, National Research Council, Washington, D C

Jordan J. David (1972) *Distribution Patterns of Automobile Travel in the New York Metropolitan Area.* pp 159-160 Highway Research Record No. 392, Highway Research Board, National Research Council, Washington, D.C

Lowry I S (1964) *A Model of Metropolis.* RM-4035-RC, Rand Corporation, Santa Monica

Massey D. B. (1973) The basic: service categorization in planning. *Regional Studies* 7, 1-15

Wilbur Smith and Associates, Inc. (1965) *Sioux Falls Area Metropolitan Transportation Study.*

Wilbur Smith and Associates, Inc. (1967) *Baton Rough Area Metropolitan Transportation Study.*

Wilson A. G. (1970) *Entropy in Urban and Regional Modelling* Pion, London

Biométrie Praximétrie
XII 1-4, 1971

What are mathematical models and what should they be ?

Frank HARARY

Research Center for Group Dynamics
Institute for Social Research
The University of Michigan
Ann Arbor, Michigan 48016, USA

No matter what the area of scientific research, whether social or physical, mathematical thinking is involved, explicitly or implicitly. At the least, the precise formulation of a problem entails some aspect of set theory and logic. Generally speaking, the working scientist uses the term 'mathematical model' for whatever branch of mathematics he may be applying to his present problem. On the other hand, the purist mathematician-logician insists strictly on the use of 'model' to mean a certain interpretation of an abstract axion system in the real world.

We begin with a self-contained development of the concepts needed for the discussion of research processes. This leads to the distinction between the real and abstract world, and the interaction between them by interpretation and abtraction. A similar, but conceptually different bifurcation is proposed for the two levels of research : digging into the foundations versus extending the horizons of knowledge. These considerations are assembled into a comprehensive Research Schema which enables a concise analysis of scientific discovery. Classical illustrations are provided, including true stories about Newton, Darwin, Freud, and Einstein. We conclude with some subjective evaluations of acceptability of mathematical models.

1. — What Are They ?

We have just noted that the word 'model' has different meanings for e mathematician and the scientist. When a mathematician uses the word, is referring to the physical or social realization of his theory. On the other hand, when a scientist speaks of a mathematical model, he means the

area of mathematics which applies to his work. Thus one[*] could say as a mnemonic aid that a model is always the other fellow's system Contrariwise it also appears to be customary by usage to refer to « research » as what goes on in your own domain.

In order to define a model rigorously, it is convenient to develop (as in Wilder [4] or in a more elementary presentation, Richardson [3]) several notions in the foundations of mathematics. Recall from high school geometry that Euclid's axioms are about as follows (depending on which book you read). The words « point » and « line » are undefined terms.

A_1 (Axiom 1) Every line is a collection of points.

A_2 There exist at least two points.

A_3 If u and v are points, then there exists one and only one line containing u and v

A_4 If L is a line, then there exists a point not on L.

A_5 If L is a line, and v is a point not on L, then there exists one and only one line containing v which is parallel to L.

Axiom 5 is the celebrated « Parallel Postulate » of Euclid.

An *axiom system* $\sum = (P, A)$ consists of two sets : a set P of primitives and a set A of axioms. *Primitives* are the deliberately undefined terms upon which all definitions in the system are based. *Axioms* are statement which are assumed to be true, and from which other statements called *theorems*, can be derived. Primitives and axioms serve to avoid so-called circular definitions and circular reasoning Each axiom in the system is an assertion about the primitives.

Euclid's axiom system consists of two primitives, 'point' and 'line', and five axioms. When Euclid developed geometry, he made a distinction between axioms and postulates. Both were statements whose truth was assumed, but axioms were considered self-evident while postulates were not ! This distinction eventually proved unnecessary and even undesirable, and today axiom and postulate are synonyms.

We shall denote by T or $T(\sum)$ the set of all theorems derivable from an axiom system $\sum$. Then a *mathematical system* (P, A, T) is an axiom system together with all theorems derivable from it.

An *independent axiom* A of $\sum$ is one which cannot be derived from the remaining axioms. An *axion system* is *independent* if every axiom is independent. It is called *consistent* if there are no two contradictory statements in $T(\sum)$.

---

[*] Following Abraham Kaplan (oral communication).

One of the classical problems in 19th Century mathematics was to determine whether or not Euclid's Parallel Postulate, A_5, was independent. The consensus of opinion was that A_5 was dependent, that is, it could be derived from A_1 -- A_4. Unsuccessful attemps to derive A_5 led to the discovery instead of non-euclidean geometry. The two types of non-euclidean geometry are now respectively called *hyperbolic geometry* (Bolyai-Lobachewski independently) in which there can be many parallels through a point to a line, and *elliptic geometry* (Riemann) in which there can be no such parallel.

An *interpretation* of an axiom system is an assignment of meanings to its primitives which makes the axioms become true statements. The results of an interpretation of Σ is called a *model* for Σ. This is the strict use of 'model' mentioned earlier.

An axiom system is called *satisfiable* if it has at least one model. Two models, M_1 and M_2 of Σ are *isomorphic* if there is a 1-1 correspondence between the elements of M_1 and those of M_2 which preserves every Σ-statement. In a *categorical* axiom system, any two models are isomorphic.

To illustrate, consider an axiom system with primitives $P = \{S, \circ\}$, where S is a set of integers, and ° is a binary operation. This little circle, ° is chosen as an undefined term for a binary operation denoted $a \circ b$, in order to avoid preconceived notions that a familiar symbol like $a + b$ would bring to mind. The following statement A_1 — A_4 are called *group axioms*, and any set S on which they hold is called a *group* under the operation °.

A_1 (Closure Law) S is closed under °, that is, if a and b are in S, $a \circ b$ is in S.

A_2 (Associative Law) Operation ° is associative, that is, $a \circ (b \circ c) = (a \circ b) \circ c$ for all a, b, and c in S.

A_3 (Identity Law) There is a unique element i in S, called the identity element, such that $a \circ i = i \circ a = a$ for all a in S.

A_4 (Inverse Law) For every a in S, there is a unique element, written a^{-1} and called the *inverse* of a, such that $a \circ a^{-1} = a^{-1} \circ a = i$. Each of the four group axioms is independent, and so this is an independent axiom system. To verify that this axiom system is satisfiable, we now display a model.

One model for this system is the set $S_1 = \{1, -1\}$ under multiplication. Thus this is called a group of *order* 2, i.e., having just two elements. The identity element is 1, each element has itself as an inverse, and S is obviously closed and associative, as can be seen from the following multiplication table :

×	1	—1
1	1	—1
1	—1	1

Another model for this axiom system is the set $S_2 = \{0,1\}$ under addition modulo 2. We define the sum of a and b mod 2 to be the remainder of $a + b$ after division by 2. Under this operation, we see at once from the next table that S_2 is closed and associative, 0 is the identity, and each element is again its own inverse. Thus S_2 is also a group of order 2.

+ mod 2	0	1
0	0	1
1	1	0

More generally, one can take S to be the set {0, 1, 2, . ., n — 1} and $a \circ b$ to mean $a + b$ (mod n) Then for each positive integer n, we get a distinct group of order n Thus the above axiom system for groups is not categorical, since it has many non-isomorphic models.

These two groups, S_1 and S_2, are isomorphic since we can let operation × correspond to + and mod 2 and set {1, — 1} to {0,1}, and all statements derivable from the axioms still hold. That the two models are isomorphic is also shown in the fact that their tables both have the following form :

°	a	b
a	a	b
b	b	a

In fact, any pair of groups with two elements are isomorphic, so it is customary to speak of « the group of order two »

The study of group theory was originally motivated by properties which are possessed by the symmetries of a configuration, whether it be geometric, algebraic, architectural, physical, or chemical. It is readily verified that symmetries satisfy the four group axioms. For example, the inverse of a symmetry of a configuration is the corresponding symmetry mapping done in reverse.

2. — Two Worlds : Abstract and Empirical

The realm of research activity is naturally divided into two worlds : the abstract and the empirical The abstract world is generally regarded as the domain of the mathematician, logician, or purely theoretical physicist, while the empirical world is inhabited by experimental scientists of many varieties : physical, social, and others * There is a growing tendency, however, for people to live in both worlds in these interdisciplinary times.

Those who work entirely in the abstract world are engaged in deriving new theorems either from axioms or from an existing theory or coherent body of theorems. Such results are usually expressed in symbols rather than numbers, and rarely touch upon the real world.

On the other hand, the inhabitants of the empirical world « work for a living ». Some live in laboratories and perform experiments in order to collect meaningful data leading to a scientific theory.

The two worlds are shown in Figure 1. The two loops, called theory builing and experimentation, represent purely theoretical and purely experimental research

Figure 1 exhibits a symmetric pair of directed links between the worlds, the first of which can be called interpretation in accordance with the use of this word in the preceding section. In a confrontation between these two worlds, the mathematician's theorems become predictions about the real world, which can be tested by the scientist. If a prediction is verified by an appropriate experiment, the scientist feels that the theorem really works, and the mathematician has found a *realization*.

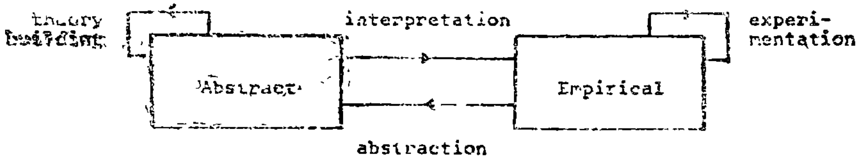


Fig 1. — Two worlds

If the predictions are entirely incorrect, the model cannot be used, However, in cases where the predictions are not verified, yet are « rather close » to correct, further abstraction is in order to construct a working model This abstraction in the light of the experiment may suggest alternate hypotheses which should result in new theorems. These theorems hopefully will lead to better predictions than previously, and to a working model.

---

* It has been established empirically that the less scientific a subject, the more likely it is that its practitioners call it a science Outstanding examples include (in alphabetical order): Christian science, library science. military science, political science, and [illegible] science.

3. — Two Worlds : Two Levels

Each of our two worlds may be divided into two levels As we have indicated, the upper level of the abstract world deals with the development of mathematical systems by the derivation of theorems. We have discussed interaction between worlds at this level by means of interpretation and abstraction. In this section we shall observe that this same type of interaction can occur at the lower level.

The lower level of the abstract world deals with the foundations of mathematics, axioms, and logic. The research activities might involve trying to prove consistency or independence of an axiom system.

A rather esoteric and dramatic recent example of an important discovery at this level is given by the definitive work of Paul Cohen [1]. It is known (see Wilder [4], for example) that a 1-1 correspondence can be constructed between the natural numbers, 1, 2, ..., and all the integers, ... —3, —2 —1, 0, 1, 2, ..., and between the integers and the rational numbers. These three sets of numbers are all said to have the same (infinite) *cardinality* which is conventionally denoted $\aleph_0$.

It is also known that there are more real numbers than integers The real line is sometimes called the *continuum*, and so *c* is written for the number of reals The *continuum hypothesis* states that there is no infinite set with cardinality between $\aleph_0$ and *c*.

Cohen proved that the continuum hypothesis (as well as its negation) is consistent with the usual axioms of set theory. As a consequence, it is independent and can neither be proved nor disproved in that axiom system. Analogous to the development of non-euclidean geometry, two entirely different axiom systems have been created; one by assuming the continuum hypothesis, and the other by taking its negation. Cohen also proved the independence of the « axiom of choice ».

On the other hand, the lower level of the empirical world also deals with foundations, but in the form of the basic laws of science, Kepler's Laws of Planetary Motion, Darwin's Law of Natural Selection, Newton's Laws of Motion, Kirchhoff's Laws of Electricity, and Einstein's Law of Special Relativity are all there.

The link between the two worlds at this lower level is quite analogous to that at the upper level. Thus interpretation of an axiom leads to a basic law about the real world, while on abstraction, a coherent set of scientific laws becomes an axiom system. The schematic representation of interaction between the two worlds is shown in Figure 2.

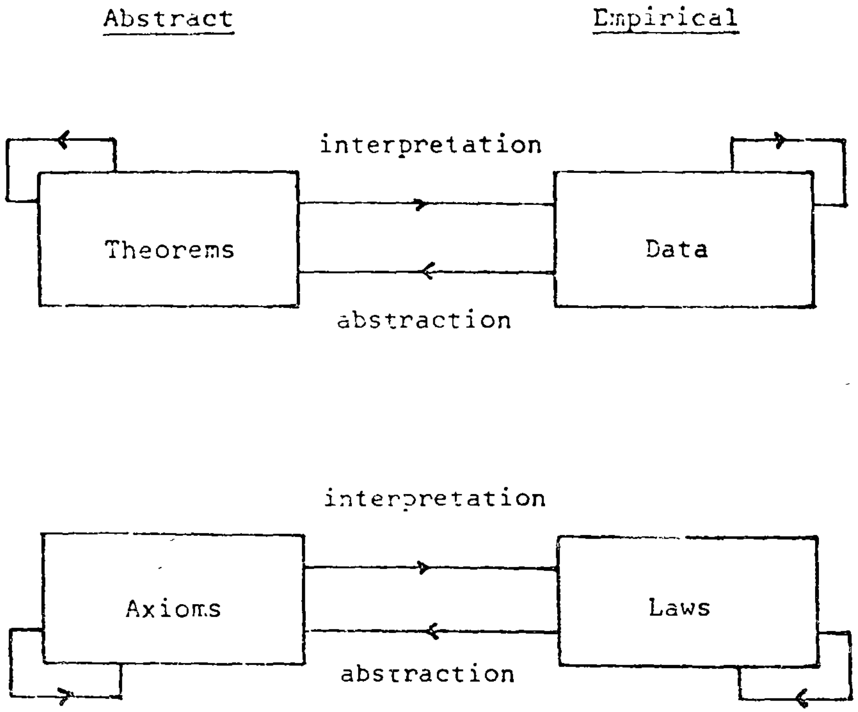


Fig 2 — Interaction between the two worlds

4. — Two Levels : Derivation and Selection

Having discussed interaction between the two worlds, we shall now establish links between their upper and lower levels. The process of climbing from the lower level to the upper in the abstract world can be regarded as *derivation*. For we begin with an axiom system and then, sometimes painfully, derive progressively complicated theorems to obtain a mathematical system.

Now consider how one goes from the upper level to the lower. From an existing body of theorems, an axiom system is to be built. To accomplish this, we select a body of particularly appropriate and fruitful theorems to use as axioms. This process of *selection* yields a small, more manageable and often more powerful system. which is conductive to the derivation of new theorems.

Selection in the empirical world involves collecting and studying vast amounts of data, and observing a pattern which may suggest a general law. Thus it is actually the *induction* process

There appears to be no direct link in the empirical world from the lower level to the upper. Derivation does occur, and in fact uses the deduction process, but again and again we find that it takes the « long way around », as shown in Figure 3. One begins with several scientific laws (lower right), and abstracts them to formulas (lower left) from which theorems can be derived (upper left) which make predictions about the real world (upper right) It is convenient, however, to draw the link representing derivation directly as well, as we do later.

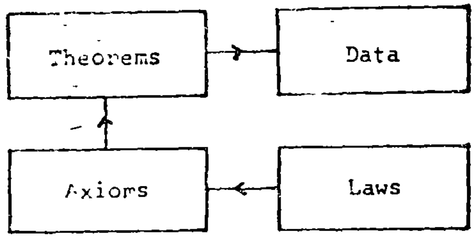


Fig 3 -- Derivation in the empirical world

In general, innovative research is initiated in the upper level, and particularly in the upper right quadrant This is due to the fact that the great majority of natural and fundamental questions arise from an attempt to observe or explain empirical phenomena In fact, most research is done at the upper level, both right and left, while almost no one continuously remains at the lower level.

For example, in ancient Egypt, the discovery of geometric formulas was necessitated by the search for improved techniques in measuring and surveying. Problems in geometry were solved long before Euclid organized the subject in an axiomatic formulation.

5. — Research Schema

We contend that the above Research Schema represents all the types of interaction between the abstract and empirical worlds during the processes of research and discovery. Its two diagonal links are shortcuts which represent research processes that go directly to « opposite » quadrants. There do not seem to be any directly ascending diagonal links

It is rarely but definitely possible to predict scientific laws from a body of theorems without actually working with experimental data. This is represented by the diagonal from upper left to lower right in the Research Schema. We shall see that Einstein took this route in his formulation of the theory of special relativity

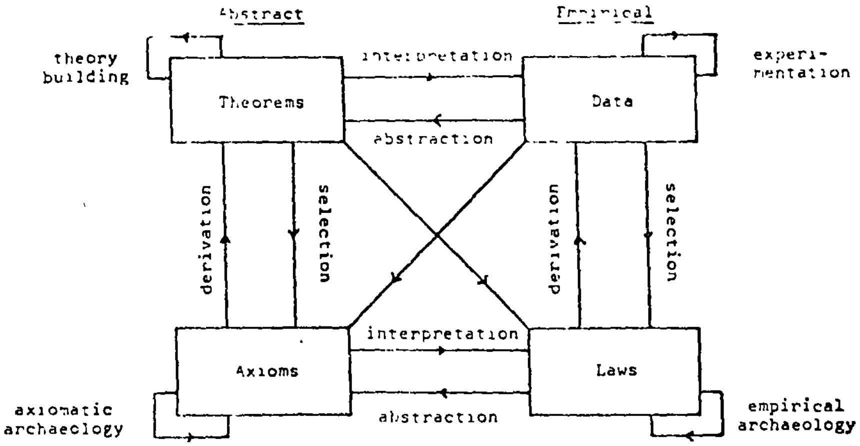


Fig 4. — Research schema

The shortcut from experimental data to axioms, skipping the formulation of laws, occasionally occurs in the social sciences when a careful analysis of data patterns produces a set of formulas that can be taken as axioms. These are then interpreted, and hopefully suggest an empirical law, without the selection process

When considering routes between the two worlds, one must also allow for traversing loops at any quadrant one or more times. The upper right loop, for example, when traversed several times, indicates repeated efforts in observation and collection of data, before attempting to select corresponding laws.

One must also note that the most direct route is not often taken in research. This will become evident in the next section when we take a closer look at particular cases of discovery.

6. — Sketches of Discovery

We shall illustrate the Research Schema with the work of several men who represent varied branches of science and mathematics. We begin with Euclid, whose work in the axiomatization and derivation of what we now call euclidean * geometry is represented schematically in Figure 5.

Euclid : Although Euclid is the acknowledged father of geometry, his main contribution was to its organization rather than to its derivation. The early Egyptians already knew the rudiments of geometry, including a form

---

* It has been said that the ultimate recognition of a man's contribution is conferred when his name is made an adjective and not capitalized.

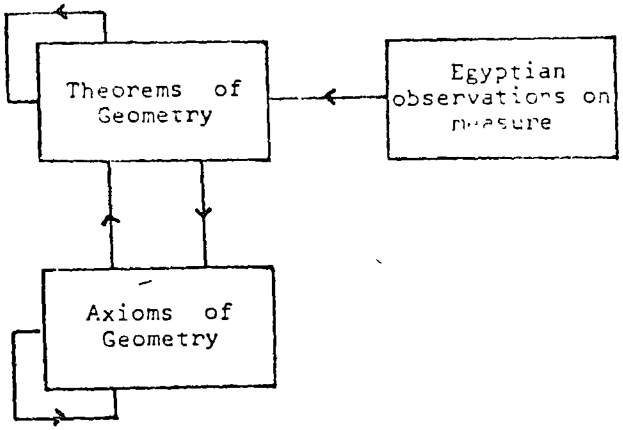


Fig. 5. — Euclid's Research Schema

of the pythagorean theorem, and formulas for the area and volume of many geometric figures. Thus we attribute the upper right quadrant in Figure 5 to the Egyptians. The emphasis on proof, however, was introduced by the early Greeks and Euclid's contemporaries developed many of the theorems of geometry. Euclid selected the five axioms above from existing results. He then proved from these all the tehorems of geometry then known and a few new ones, and presented a logical organization of the material in an exhaustive text. By today's standards, Eulcid's axiomatic work is not rigorous, but it was an outstanding accomplishment for its time

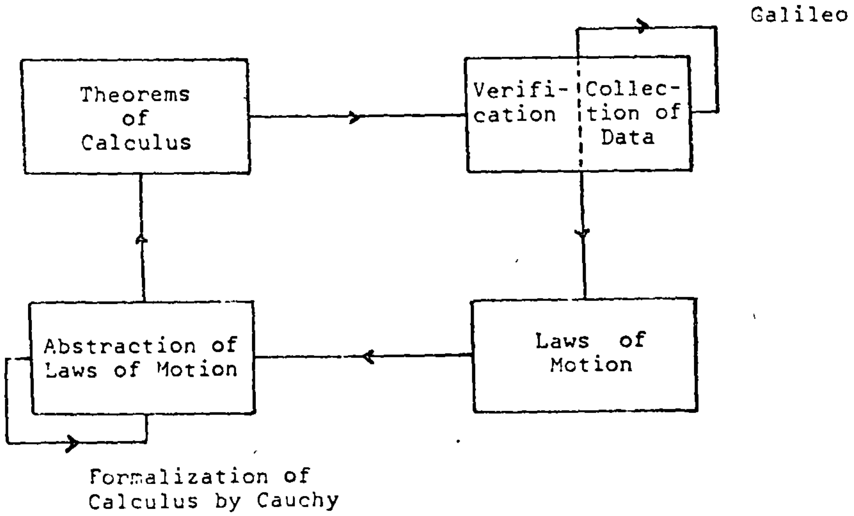


Fig 6 — Newton's Research Schema

Newton Unlike Euclid, Newton occupied every quadrant of the Research Schema His first work was on the upper level of the empirical world, where he experimented in chemistry and optics while still a student Newton's most important results, however, were not derived from his own data; but from the work of those before him. His formulation of the Laws of Motion was induced from Galileo's extensive experimentation. Hence, we credit the upper right loop in Newton's Research Schema to Galileo Newton's Laws of Motion have been stated as follows :

1. Every body will continue in its state of rest or uniform motion in a straight line unless it is compelled to change that state by impressed force

2. The rate of change of momentum is proportional to the impressed force and takes place in the line in which the force acts.

3 For every action, there is an equal and opposite reaction.

Newton left the empirical world and entered the abstract by expressing his laws symbolically as equations His work with these resulted in the discovery of both differential and integral calculus. Others independently discovered these concepts, but it is believed that only Newton and Leibnitz (who discovered calculus independently) realized that differentiation and integration were inverse processes

Calculus did not become mathematically precise until the next century when Cauchy introduced the necessary concepts of limit and infinite sequence. We draw a loop in the lower left quadrant of Figure 6 to represent Cauchy's work in the foundations of calculus

This new branch of mathematics readily produced an abundant supply of theorems. The predictions which resulted were tested in the laboratory, and found to be entirely correct within the range of current measuring instruments.

Einstein Eventually, more accurate measuring devices revealed that Newton's Laws of Motion could not explain the behavior of light on either the microscopic or astronomical level Furthermore, the Michelson-Morley experiment proved conclusively that « ether » did not exist These discoveries led to a period of great activity in physics pioneered by Albert Einstein.

Like Newton, Einstein's major work resulted from data collected by scientists before him. Einstein was a purely theoretical physicist, and never worked in the upper right quadrant of the Research Schema himself. But he certainly stimulated an enormous number of experiments there. He proposed the following empirical axiom system as laws of light motion .

1. No physical object can travel faster than the speed of light.

2. The speed of light depends not at all on the relative positions of the source of light and the observer, or their relative speeds.

3. The mass at a velocity v of a particle equals its mass at velocity 0 divided by $\sqrt{1 - v^2/c^2}$, where c is the speed of light.

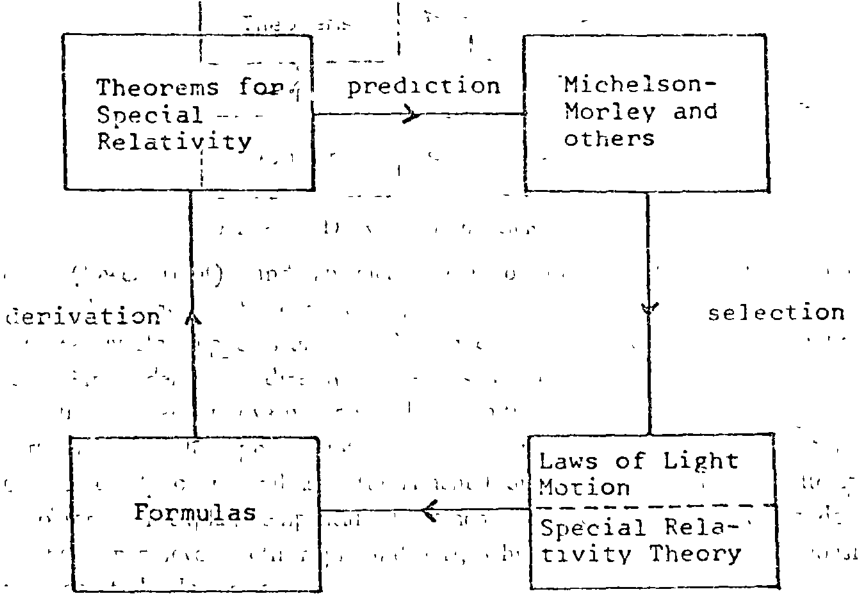


Fig 7 — Einstein's Research Schema

Einstein abstracted these three laws to an axiom system, from which he derived the body of theorems interpreted as the theory of special relativity. He found that in particular, his distance formulas for relativity theory were related to those of hyperbolic noneuclidean geometry, thus relativity theory provides a physical model for hyperbolic geometry. The Research Schema for this discovery is shown in Figure 7. We begin in the upper right with the Michelson-Morley experiment, and then go to the Laws of Motion of Light in the lower right, and their abstractions in the lower left From there we go the theorems of special relativity in the upper left, and finally to the experimental verification in the upper right where this cycle started Einstein then went around this cycle again with his more refined theory of general relativity, which led to more precise predictions of physical measurements

Darwin Charles Darwin spent most of his life doing research in only one quadrant of the Research Schema, the upper right. His research career began when he became the official naturalist on the good ship Beagle, and embarked upon a five-year voyage He made observations on all species of animals he could find, and took voluminous notes During the remainder of his life, Darwin analyzed and classified these notes and all other available information. The climax of his work was the formulation of his Law of Natural Selection and his Theory of Evolution

Darwin's theory asserts that all animal species have descended from a common origin. The variety of species results from «natural selection», in which those animals which are best adapted to their environment survive. Due to occasional mutations, certain animals in a species are better able to survive than others. These mutations may be passed on to their offspring who in turn will tend to survive and reproduce, eventually resulting in a new species which has been naturally selected.

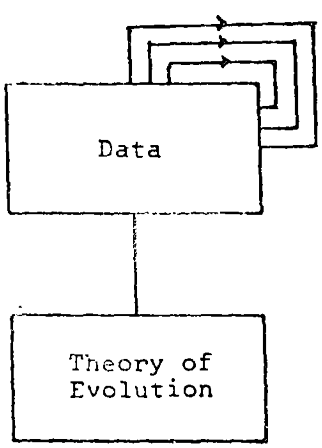


Fig 8 — Darwin's Research Schema

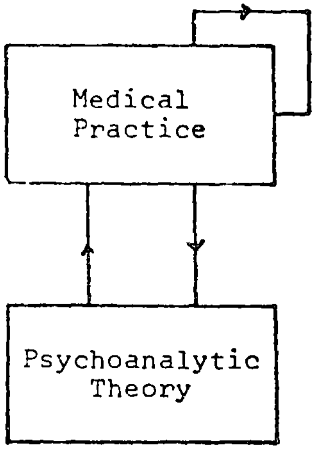


Fig 9 — Freud's Research Schema

Freud : Sigmund Freud, like Darwin, stayed in the empirical world. In fact, their Research Schemata are quite alike, as seen in Figures 8 and 9. He began with a medical degree and turned from general practice to specialization. Freud (in collaboration with J. Breuer initially) did research in the treatment of «hysterical» patients who had physical symptoms for which no physical cause could be found. He inferred from the study of many cases that the symptoms could be traced back to some repressed childhood trauma, and went on to develop the concept of the subconscious together with the id, ego, superego. First through hypnosis, and later through «free association», Freud was able to induce himself and his patients to recall these forgotten experiences, and alleviate their symptoms.

Much of the psychoanalytic theory which Freud developed is still highly controversial today, although it has made a lasting impact on the development of many modern theories in psychology.

There has been a highly publicized report of the proof of a deep and important theorem by a mathematician while boarding a bus in Paris It may be just as true as the anecdote about Newton's finding his law of gravitational attraction when an apple fell off its tree and landed on his head This sort of phenomenon does occur, but fortunately is not an intrinsic part of the discovery procedure. In the words of Hans Zinsser,

> «It is an erroneous impression, fostered by sensational popular biography, that scientific discovery is often made by inspiration... This is rarely the case. Even Archimedes' sudden inspiration in the bathtub; Descartes' geometrical discoveries in his bed; Darwin's flash of lucidity on reading a passage in Malthus; Kekule's vision of the closed carbon ring came to him on top of a London bus; and Einstein's brilliant solution of the Michelson puzzle in the patent office in Berne, were not messages out of the blue. They were the final co-ordinations, by minds of genius, of innumerable accumulated facts and impressions which lesser men could grasp only in their uncorrelated isolation, which — by them — were seen in entirety and integrated into general principles. The scientist takes off from the manifold observations of predecessors, and shows his intelligence, if any, by his ability to discriminate between the important and the negligible, by selecting here and there the significant steppingstones that will lead across the difficulties to new understanding. The one who places the last stone and steps across to the *terra firma* of accomplished discovery gets all the credit. Only the initiated know and honor those whose patient integrity and devotion to exact observation have made the last step possible.»

When a researcher has become sufficiently steeped in his problem, he has amassed enough meaningful data (mathematicians also accumulate data via «thought-experiments») to perceive the proper pattern and conceive the correct conjecture. This is a necessary but not sufficient step toward establishing a theorem A proof, which is valid, must be supplied; otherwise, the assertion remains a conjecture The two talents of conjecture and proof appear to be quite separable

7. — What Should They Be ?

It is becoming more fashionable to use mathematical models as a powerful analytic device for advancing scientific research in a remarkable variety of disciplines. This usage is certainly not unwarranted, since models when used with care and discretion can and should be of great use in the clarification of existing problems and the formulation of important new ones. Unfortunately, it seems that models are misused all too often. The word 'model' is sometimes bandied about by people with little conception of its real meaning simply because it is *au courant* They don't even define 'model', but use the word to suit their own purposes.

A model need not be impressively confusing in order to be valuable. In fact, one of the main contributions of a model lies in its ability to simplify a problem, and so it should be no more complicated than necessary.

Neither should a model be symbol-rich but idea-poor. Models which hide miniscule content behind a mass of symbolic formulas tend to look impressive, but add nothing. « Mystery is no criterion of knowledge » For example, a recent paper in a leading psychological journal had only one abstract idea : the number of elements in the union of two sets is the sum of the number of elements in each minus the number they have in common Alas, the author apparently did not recognize it as the simplest special case of the Principle of Inclusion and Exclusion

Another unfortunate use of mathematical models occurred in a published paper in sociology in which there were ten axioms and zero theorems. However, an interpretation was then given which resulted in ten empirical theorems, one for each axiom Thus 1-1 correspondence between axioms and empirical theorems simply involves the preparation of axioms which will yield desired empirical assertions

Furthermore, an axiom system should not be constructed for the artificial purpose of deriving just one theorem which has already been verified statistically. Clearly such a model only clutters the literature and does not involve genuine derivation.

We do not wish to lay all the blame for the misuse of mathematical models on scholars in the empirical world; it occurs in the abstract world as well. The following passage by the eminent linguist Gustave Herdan [2] shows the dual roles the two worlds can play in the misuse of models.

« Without going into details, I will only mention a certain quantitative relation known to linguists as the 'Zipf law'. Mathematicians believe in it as a law, because they think that linguists have established it as a relation of linguistic facts, and linguists believe in it because they, on their part, think that mathematicians have established it to be a mathematical law. As can be shown in five minutes, it is not a law at all in the sense in which we speak of natural laws ».

Loosely stated, this law of Zipf proposes a high correlation between the frequency of use of words and their brevity

Another typical superficial use of mathematical models involves the bland assumption that the most elementary parts of an existing branch of mathematics apply unchanged to a problem in social science. Typical examples include high school algebra, coordinate geometry, matrix manipulation, graph theory, and the probabilistic theory of Markov chains. In such models, the typical procedure is to assign empirical terms to the mathematical variables by way of interpretation at the lower level. Then the existing theorems and methods of calculation are translated at the upper level into statements which are claimed to be new empirical findings.

What, then, should mathematical models be ? We have suggested that they should lead to new theorems, but this is not always necessary. The precise thinking involved in the careful formulation of an axiom system will lead to an improved conceptualization of the empirical phenomena at hand. This in turn can suggest the proper variables to measure, and perhaps an approach to the measurement problem.

Sometimes an existing area of mathematics can be quite useful as a mathematical model provided it is augmented by one or more new axioms suggested by the real world. The most productive models, however, have involved derivation. For it is only after the derivation of new theorems that unexpected and far-reaching predictions can be made From a mathematician's viewpoint, it is best if derivation leads to nontrivial theorms, which actually qualify for publication in the mathematical literature. To summarize, it is our personal and perhaps controversial contention that mathematical models will lead to significant and natural growth in both the abstract and empirical worlds.

Aknowledgement.

Research supported in part by Grant MH22743 from the National Institute of Mental Health.

REFERENCES

[1] P J COHEN, *Set Theory and the Continuum Hypothesis*, Benjamin, New York, 1966.

[2] G HERDAN *Sonderdruck aus Zeichen und System der Sprache*, Vol 2, Akademie-Verlag, Berlin, 1962, p. 108

[3] M RICHARDSON. *Fundamentals of Mathematics* 3rd ed , Macmillan, New York, 1966

[4] R L WILDER, *Introduction to the Foundations of Mathematics*, 2nd ed , Wiley. New York, 1965.

Technical Review

THE LOWRY MODEL HERITAGE

William Goldner

The Lowry model constituted a breakthrough of a particularly intractable conceptual barrier. Notwithstanding its attractive features, including the promise of meaningful operationality, the model has been implemented in only a few domestic applications; its adapters abroad have had more success. In this process, the model has gradually been transformed to more satisfactorily meet the requirements of the regional and transportation planning agencies that depend on allocation modeling. As newly developing conceptual embellishments are made operational, the Lowry model should become an instrument that will enlarge our knowledge of the metropolis and lay a solid foundation for the next generation of models.

The contributions of Dr. Ira S. Lowry to the development of land use modeling have been much greater than the sum of the reports of his work. Somehow, the conceptual framework of what has come to be known as the "Lowry model" has stimulated a population explosion of successors, each with meaningful elaborations. The development of the original model has had this effect because: the promise of meaningful operationality was a prime stimulus; the simplicity of the causal structure had substantial appeal; and the opportunity to enlarge and embellish the framework encouraged further work.

This short review of Lowry-type models was initiated primarily to identify the increments to the Lowry conceptual framework that other model developers and users have made, either as experiments in design, or as operationally successful versions. Some results, particularly in international cases, appear to counter, at least to some extent, the recent statement that "... those of us ... who like to think that the use of sophisticated analytic techniques in public decision-making has been spreading, might be surprised and disappointed" (Hemmens, 1970).

The Lowry Model

The Lowry model was developed in 1962-63 and the report of its framework published by the RAND Corporation in 1964 (Lowry, 1964). The model was developed as part of a modeling system to generate alternatives and aid decision-making in the Pittsburgh Comprehensive Renewal Program (CRP). This program was complex, embodying an articulated set of models in addition to the Lowry model. An extremely forthright evaluative report covering the destinies of the Pittsburgh CRP models by a study staff member mentions that:

> "... it was too ambitious. In 1963, when the model (system) was designed, there seemed to be some possibility that the detailed problems in constructing the model could be worked out. At the present time, the task of building such a model would correctly be regarded as monumental, and the event is still some time in the future. Efforts can be better directed to limited purpose models, for investigating the behavior of the models themselves and for testing hypotheses which are potentially useful in models... Adequate data with which to run the model have never been assembled, and around Pittsburgh the interest in the model is generally low. The City decided, after several years of work had gone by, not to continue the simulation effort, so most of the models have not been tested or used" (Lee, 1968).

Notwithstanding this discouraging prognosis for the Pittsburgh CRP, the Lowry model as reported in the RAND publication, was widely distributed and thoroughly reviewed (Fleisher, 1965). Experimental and operational versions were developed by other modelers, and the wide spectrum of modeling efforts that grew out of this innovative stimulus is reflected in numerous reports that only now are beginning to come into view.

CAUSAL STRUCTURE

One of the appealing characteristics of the Lowry model is its straight

William Goldner is currently Research Economist at the Institute of Transportation and Traffic Engineering, University of California (Berkeley). He has served as Consultant on Regional and Locational Models for the Bay Area Transportation Study Commission (BATSC), and as a Lecturer in Urban Economics at the University of California (Berkeley). Dr Goldner received his bachelor's degree from Stanford University and his doctorate from the University of California, both in economics.

forward and easily understood causal structure. First, an ingenious multiplier concept is used to generate the spatial distribution of total employment from the exogenously provided location of basic employment. Then, the model generates households by reversing the journey-to-work into a work-to-residence trip so that workplaces can be used to locate the residential stock. In addition, these allocations are subject to constraints: an upper bound capacity constraint for households, a minimum threshold for size of service employment. This structure is diagrammed in Figure 1

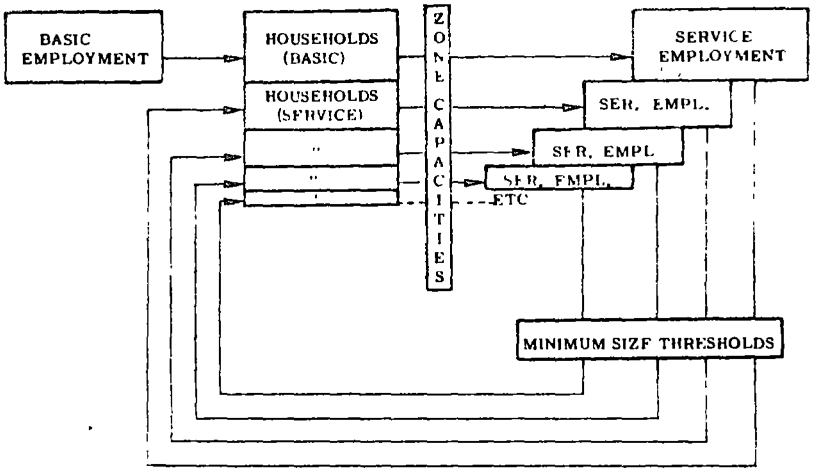


FIGURE 1 Causal Structure of the Lowry Model

MAJOR CONCEPTUAL ELEMENTS

In addition to the causal structure, the Lowry model incorporates several concepts that are fundamental to the class of models it precedes. These include the concepts of "basic" and "population [illegible]" employment, the [illegible] of the allocation [illegible] the treatment [illegible].

The Basic-Retail Dichotomy Because of the resemblance of terms to the economic base concept, there has been some misunderstanding regarding the nature of the fundamental structural split in the Lowry model. The starting point is that a group of activities in the model are located in relationship to markets made up of households, thus, these activities are implicitly related to populations and purchasing power. All activities for which a local market or service area can be identified for final products or services are in the category of "retail." The criterion is locational, flowing from the existence of a *local market or service area.*

All other industries, including those producing intermediate products for industries serving local markets, are located in relation to *other influences*, such as unique site features, interindustry linkages, agglomeration economies, resource availabilities, and interregional transportation routes. These industries, referred to as "basic," are in fact a residual remaining after the "retail" industries have been identified. Lowry clearly states that this nomenclature is used for convenience, not for descriptive taxonomy (Lowry, 1964), and implies that "site-oriented" and "residence-oriented" might be more exact labels. The regionwide inputs which the model requires for "basic" employment and for calibration of multipliers to allocate "retail" employment are supplied exogenously by a regional economic model. The design of this regional aggregate model may properly be subject to economic base aggregate model criticisms. However, these critiques do not carry over to the allocational framework embodied in the Lowry model.

In Lowry's notation, the basic-retail dichotomy is expressed as follows:

$$E = E^B + E^R \qquad (1)$$

Lowry disaggregates retail employment into three classes so that

$$E = E^B + E^{R1} + E^{R2} + E^{R3} \qquad (2)$$

These aggregated totals for the study area plus the exogenous allocation of E^B to individual zones provide estimates of parameters that generate allocation of retail categories. Let

$$a^{RK} = \frac{E^{RK}}{E} \quad (K = 1,2,3) \qquad (3)$$

then,

$$E = E^B + a^{R1}E + a^{R2}E + a^{R3}E \qquad (4)$$

The Multiplier The iterative solution used by Lowry and the decomposition of employment in (4) provide the basis for a multiplier effect in the working of the model. From (4),

$$E^B = E - a^{R1}E - a^{R2}E - a^{R3}E$$

$$E^B = E\,(1 - \sum_{K=1}^{3} a^{RK}) \text{ and} \qquad (5)$$

$$E = \frac{E^B}{1 - \sum_{K=1}^{3} a^{RK}}$$

To generate dwelling units, the employment is factored by the dwelling unit/employment ratio, f.

$$N = fE \qquad (6)$$

All of the above concepts deal with aggregates and have their analogs in contemporary macroeconomic theory. The allocation of these quantities to individual zones is the particularly innovative contribution of Lowry. To develop these allocational concepts, we must consider the zonal system, the network of interzonal distances, and the allocation functions that were used.

The Zonal System The sensitivity of the allocation process to the zonal system is often not mentioned. However, the configuration and number of zones are crucial to the operational development of an allocation model. The number of zones is constrained by the computer capacities to which the program is to be fitted, and also by the importance of locational exactness for which the output is to be used. The configuration of zones is controlled by data collection processes, and particularly by the need for spatial congruence for data on several alternative variables. Census tracts are often used in large regions because of the availability of maps, boundaries, and data on residential household characteristics. However, a more desirable configuration is the system of grid cells, usually square. These require sophisticated coding from primary source data.

The Lowry model used one-mile square grid cells as zones. The city blocks falling within these cells were assembled so that block boundaries approximated the one-mile square. Land use, employment at workplace, and residential population were allocated to city blocks or to these larger cells to constitute the data base.

The Interzonal Network The grid cell configuration has the special advantage that simple geometric relationships define the interzonal centroid-to-centroid distances. Distances calculated in this way are much more economical to produce than the complex and elegant interzonal times (skim-trees) which are computed using minimum time-path procedures based on extensive definition and inventorying of transportation facilities. Lowry used the grid cell schema, having tested to his satisfaction that air distance is highly correlated with time-distance measured by more sophisticated methods

The Allocation Functions The Lowry model contains three groups that are subject to spatial allocation: (1) employees to residences; (2) households to services; and (3) employees to services. For (1), Lowry used a gravity function for the work-to-residence allocation, calibrated on the basis of trip indices calculated for residence to-work relationships. The actual functional form is:

$$\frac{dp}{dr} = ar^{-X}$$

In the calibration of the model, the actual values used are

$$\frac{dp}{dr} = r^{-1.33}$$

with dependence on normalization procedures to readjust the function for bypassing the coefficient a.

The residence-to-service allocation (2) was fitted by reciprocals of quadratic functions of the form:

$$\frac{dp}{dr} = (a - br - cr^2)^{-1}$$

The actual values used (Lowry, 1964) are:

	a	b	c
Neighborhood shopping	.5107	–.7400	–.2699
Local shopping	.0116	–.0012	–.0202
Metropolitan shopping	.0664	–.0442	–.0156

Allocation (3) from workplace to shopping is assumed to take place within the zone in which the workplace is located, therefore, no allocation function is required.

Allocation weights for locating services are required to allow each influence, work-to-shop and residence-to-shop, to be combined. These proportions (Lowry, 1964) are:

	Residence	Work
Neighborhood	.90	10
Local	.70	.30
Metropolitan	.50	.50

Aggregation and Disaggregation The Lowry model is a macromodel aggregating within its structure many variables that would influence the outputs if they were explicit. The basic-retail dichotomy is one form of disaggregation, and the retail category, in turn, is broken down into three types of shopping clusters, metropolitan, local, and neighborhood. Although Lowry's experiments indicate he considered other variables, such as occupation, and also a finer breakdown of industries within retail as possible bases for disaggregation, the constraints of computer capacity and limitations of data did not allow for further disaggregation.

The Treatment of Time The model generates an equilibrium allocation of activities consistent with network distances and basic employment locations, modified by lower and upper bound constraints. Lowry suggests that the iterative process may represent a time oriented buildup of trade activities and the associated increments to the housing stock but does not press this interpretation very hard. His more felicitous reference to the "instant metropolis" is also more persuasive for its time implications.

Constraints Lowry anticipated that constraints would have to be incorporated into residential allocations and retail activities. For residential development, he spurned zoning ordinance density limits as well as projection of existing densities in favor of a maximum density limit of approximately sixty-five dwelling units per acre. This high level does not appear to have been constraining in the model, with the possible exception of one zone.

Similarly, he anticipated that the model algorithm did not have the property of bringing together retail activities to reflect the external economies of scale that are visible in strip retail development and in shopping centers of various sizes. His solution was to identify the lower employment threshold of such clustered activities and exclude retail development below the limit. The thresholds which were used in the model (Lowry, 1964) distinguished between centers of differing sizes as follows:

Type of cluster	Minimum number of employees within zone
Neighborhood facilities	30
Local facilities	250
Metropolitan facilities	20,000

Descendants of the Lowry Model

In general, Lowry-type models display certain characteristics in common: (1) partitioning of employment into a market-oriented category called population-serving or "retail," and a residual termed "basic," or site-oriented; (2) the causal system leads from "basic" employment to residential population to population-serving employment; (3) the population-serving allocation grows out of a multiplier relationship applied to basic employment.[12] However, each descendant makes certain fund mental additions to the Lowry framework.

TOMM (TIME ORIENTED METROPOLITAN MODEL), 1964

The earliest revision of the Lowry model was generated in Pittsburgh where the stimulus to make the model operational in the CRP was immediately apparent. The responsibility was undertaken by John P. Crecine, under the aegis of the Consad Research Corporation. The published technical report (Crecine, 1964) indicates that three major revisions were incorporated:

1. Conversion to a "marginal allocation model that allows only a portion of the establishments and households to move in a certain period of time, rather than the aggregate, allocative model of Lowry's";

2 Household disaggregation ". . . by income, housing characteristics, social characteristics, or all three":

3. Limitation of "the simulation study to locational characteristics within the City's Boundaries. . . ."

In addition, the report mentions that "the City will be sectored into *n* census tracts," shifting the zonal system from the mile-square grid adopted by Lowry.

The implementation of the "marginal allocation" process was made to depend on a partitioning of activities between "stable" residential and commercial activities (those that do not have the opportunity to relocate during the time period) and locatable activities. The determination of this partitioning was to be made exogenously but is not described in the monograph.

A comparable partitioning carries over to land uses which are categorized as unusable, "basic," stable retail and service, public, and stable residential land (as a residual). This breakdown of land use information enlarges considerably on the bare-bones requirements in the Lowry model.

TOMM also disaggregates households by type (unspecified) in the allocation process, which mediates between the propensity to reside with comparable households and the propensity to commute. Parameters which weight the relative importance of these two propensities vary for each household type and are unspecified.

With regard to the operationality of TOMM, Lee reports that "a trickle of resources have gone into the model in the past few years, from the Department of City Planning and from Consad, but not enough to make the model usable" (Lee, 1968). Crecine states, "Like other land use, population, and commercial location models, the degree to which T.O.M.M. can be validated is a function of data availability. Because of incompleteness, incompatibility, and non-existence of most small area data and the lack of time-series data, T.O.M.M. remains in the development stage" (1968).

BASS I (BAY AREA SIMULATION STUDY), 1965

The Lowry model was also the stimulus for parallel effort by Goldner and Graybeal at the Center for Real Estate and Urban Economics at the University of California (Berkeley). They had the measurement of the impact of industrial location as a primary purpose (Goldner and Graybeal, 1965). Their model was designed to test the sensitivity of the commercial and residential allocation system to the exogenous emplacement of a large plant at a specific location.

BASS I had several design features that differed from the parent model. First, it used census tracts rather than grid squares as zones. Second, it generated employees, population, and households (Lowry uses households as a surrogate for population). Third, and most important, many of the parameters which are applicable to the whole system in the Lowry model were disaggregated to individual tract-specific form. Labor force participation rates and land absorption coefficients for residential and commercial allocation reflected the base-year relationships, rather than a systemwide average. These tract-specific parameters were closely related to the layers of development to which the urban region has been subject over time. Thus, close to the center, residential densities are high and family size small, contrasted with the development margin where the low-density suburb with larger families is characteristic.

A fourth change from Lowry was the abandonment of the disaggregation of population-serving employment. In effect, the model traded off this disaggregation in favor of the spatial disaggregation mentioned above.

BASS I was a pilot version and was used for testing the effects of the emplacement of several industrial parks. Out of the experience with this model, a program for improvement and redesign was developed. Graybeal chose to pursue a redesign strategy that generated a composite system of separate models. These are included in parts of BASS III, a non-Lowry system of models (Center for Real Estate and Urban Economics, 1968), and reported in other places (Graybeal 1966a; Graybeal, 1966b). Goldner organized a revision more consistent with the Lowry framework, which became PLUM (Goldner, 1968).

THE GARIN-ROGERS CONTRIBUTIONS, 1966

A serendipitous contribution to the stream of developments originated with a graduate student, Robert A. Garin, in the planning workshop of Professor Andrei Rogers, during the work on BASS I (Garin, 1966). Garin expressed the fundamental Lowry algorithm in vector and matrix format. Using this notation, he demonstrated that the iterative process used by Lowry to generate population serving employment could be replaced by elementary matrix operations to obtain an exact rather than an approximate solution.

Professor Rogers (1966) provided one additional fillip to this development by adding a time dimension to the input vector. This replaced the static equilibrium solution of Garin with an equilibrium displaying *stationarity,* that is, a distribution that arises after repeated application of an unchanging recursive growth process.

Both of these developments were demonstrated with experimental ten-zone allocations. However, for operational use, there are several problems. First, the Garin formulation does not comprehend the constraints which Lowry imposed. Neither the minimum size constraint for population-serving employment nor the maximum density constraint for residential development was included in the matrix operations. In addition, the inversion of a matrix for a large size zonal system presented problems of computer storage and time-cost that might be bypassed by useful approximations. Finally, Rogers' requirement of constant recursive growth is not always consistent with exogenous growth forecasts reflecting shifting age composition of the population and drifting labor force participation

CLUG (THE CORNELL LAND USE GAME), 1966

A parallel development was also taking place in the form of a heuristic game designed to teach planning principles to public officials and students. Professor Allen G. Feldt of the Cornell University Department of City and Regional Planning saw the need for an instrument to bridge the gap between the complexities of planning expressed in sophisticated mathematical terms and the decision-influencing comprehension needed by senior planning officials and local legislators. Working independently, he devised a game that is analogous to TOMM (Feldt, 1966).

The model embedded in the game has several elaborations which extend the conceptual framework of the Lowry model. Residential allocation can occur at four densities corresponding to the scaling of densities from single family upward to large multiple unit developments. Instead of a network, there is emphasis on infrastructure (utilities, local government services, and the like) to guide the configuration of development. There is also a governmental revenue and outlay process which relates to the value of land and improvements. In fact, there is a flow of money passing through the system, starting with bids for vacant land, provisions for buying and selling land, and for other transactions involving wage payments, retail trade, and provision of goods.

The Lowry causal sequence—starting with infrastructure to basic industry to residences to retail—is clearly embedded in the framework. The hand version of the game was essentially a pedagogical tool and, in a later version, was supplemented by a short computer program to eliminate hand computations and accounting processes that were necessary for the game to proceed.

A DYNAMIC MODEL OF URBAN STRUCTURE, TOMM II, 1968

Still using the same model name, TOMM, Crecine presented a more completely documented version of his earlier model in 1968. The model enlarges upon the Lowry model and the earlier version of TOMM in several ways: the "variables . . . are of a much more disaggregate nature, the concept of site amenities is introduced, and zoning constraints are explicit" (Crecine, 1968). Among the elaborations incorporated into the revised version are: (1) white collar and blue collar workers; (2) several household types; (3) inclusion of the effect of site amenities and economic externalities in determining site valuation; (4) incorporation of effects of density, zoning restrictions, and market imperfections on rents; and (5) recognition of the inertia in the urban locational system that results from the durability of physical property and infrastructure.

These changes are backed by careful theorizing and explicit attention to computer applications. Crecine summarizes, "T.O.M.M. is still in a development stage. This version, however, approaches the limits of this particular approach to urban locational phenomena. Further efforts on T.O.M.M. should focus on developing an appropriate data base and on the considerable parameter estimation problems. Additional theoretical refinements would appear to have only marginal payoffs" (Crecine, 1968).

This enlarged version of TOMM is ". . . serving as the spatial-location device in the METRO project at the University of Michigan" (Crecine, 1968). The project is attempting to develop more sophisticated operational gaming techniques through which several models incorporating technical formulation of plans and decisions by policymakers are matched and evaluated for disparities (Duke and Burkhalter, 1966).

PLUM (PROJECTIVE LAND USE MODEL), 1968

The line of development that grew out of the experiments with BASS I led to the operational version of PLUM (Goldner, 1968). PLUM was implemented to provide the land use allocations and small zone forecasts of population, dwelling units, and employment used by the Bay Area Transportation Study Commission (BATSC). In addition to the modifications mentioned with regard to BASS I, PLUM incorporated several additional concepts:

1. Network times were created by careful generation of minimum time-paths (skim trees) with alternatives for free-flow and peak-hour versions, augmented in both cases by terminal times.

2. The gravity allocation function, which has biases strongly influenced by the sizes of zones in the zonal system and is also deficient for its treatment of short trips, was replaced by a more satisfactory function, the reciprocal transformation in logarithmic form. These allocation functions were disaggregated by three types of trips, work-to-home, work-to-shop, and home-to-shop, and were disaggregated spatially by county, requiring calibration of twenty-seven parameters.

3. An additional variable at place of residence allowed population, employed residents, and dwelling units to form a consistent triad that is linked by three vectors of parameters. Parameters, including population per household, workers per household, and population per worker, were zone-specific (spatially disaggregated) and adjusted to drift through time in conformity with exogenous forecasts of employment and population.

4. The model simulates trips rather than estimating them in correspondence to an accessibility index.

5. Land use accounting includes zone-specific acreage for residential and vacant land, rather than the residual treatment characteristic of the Lowry model and TOMM.

6. Constraints of residential development accommodate to land capacity and simultaneously to in-

creased density as the zone is filled. Excess demand exerts pressure on density in relation to density transformation coefficients which are calibrated to cross-section data for each county.

7. Comparative statics is the basis for the generation of time increments which are added to initial conditions data to generate target year forecasts, either on a one-step or a five-year recursive basis.

8. The equation system is solved as a causal chain which is made possible by the substitution of approximations of the multiplier process for the matrix inversion suggested by Garin.

A. G. WILSON'S ENLARGED CONCEPTUAL FRAMEWORK, 1969

Interest in the Lowry model is international in scope. One nucleus of both conceptual development and experimental work is the Center for Environmental Studies in London. There, Alan G. Wilson and his colleagues have established a beachhead on which the Lowry flag may legitimately be emplaced.

In an early paper, Wilson evaluated the residential model of Lowry as an example of an "elementary" model, that is, a model which did not require detailed "knowledge of preference structures and individual utility functions in various forms" (Wilson, 1969a). Wilson posed five issues which called for solution:

1. The explicit nature of the allocation function used,

2. The absence of attractiveness characteristics in the residential zones,

3. The problem of placing population limits as planning constraints on specific zones;

4. The corresponding problem of placing capacity constraints on selected zones,

5. The special features of the intervening opportunity allocation contrasted with the standard travel cost function.

In a terse, precise exposition, including the use of the concept of constrained entropy maximization to [illegible] Wilson offered his analytical solutions to these five problem areas. This work focuses on the potential use of more satisfying functions than the gravity function in the allocation process and on the systematic handling of constraints.

A second paper (Wilson, 1969b) goes much farther, exploring the role that disaggregation might play in improving elementary models and offering provocative suggestions for elaboration. The first and most significant proposal relates to partitioning the residential locators on the basis of four contrasting behavioral patterns. These reflect the behavior of workers who are:

1. Locationally unconstrained (locational choices involve both work-place and residence);

2. Constrained to fixed residences (locational choice involves workplace);

3. Constrained to fixed jobs (locational choice involves place of residence as in the Lowry model);

4. Constrained by fixed residences and fixed jobs (during some specified time period, no locational choices are to be made).

This treatment of behavior reflects important issues in development modeling. In-migrants to the region are represented in category 1—locationally unconstrained. Category 2—fixed residence—begins to account for sociological issues that have been missing from development models. Workers pinned to their residences include occupants of ghettos and public housing, such as the poor, the aged, and racial minorities. Explicit identification of this category allows the work-locating process to be modeled differently for these groups. Category 3 encompasses the Lowry model's residential allocation process when fixed job locations are assumed. And, category 4 covers those households that remain stable during any given time.

Wilson then formulates the strategy for solution of this enlarged framework, including a specification of the data requirements and some simplifying procedures to make the problem more tractable.

He also covers the need to disaggregate elementary models, for at least

1. different income groups;
2. different wage levels by location,
3. different types of house; and
4. variations in the price of houses by location.

Further disaggregation is deemed desirable for:

5. type of locational behavior;
6. number and type of workers per household;
7. social class/occupation stratification; and
8. number of cars per household.

The supply side of the modeling process has been virtually neglected in elementary models, and Wilson has contributions to make in this direction also. He generates the vacant houses made available by movers and new houses coming on to the market, allowing this supply to be mediated with demand by means of a housing price variable. Alternative approaches to the supply of new houses are postulated, deriving impetus from public planning configurations or from private profit maximizing incentives. Filtering of families through the housing stock is included in the process.

This summary of Wilson's work hardly does justice to its elegance, comprehensiveness, and sophistication. Nor does it include any mention of the equally valuable contributions on subjects not so closely related to the Lowry model lineage. The research that is being carried on at the Center for Environmental Studies in London constitutes a conceptual umbrella held high over derivative work being carried on in the United Kingdom, particularly in the transformation of these ideas into operationality.

Empirical Work With Lowry-Type Models

A host of examples of Lowry-type model implementations exists in England. The documentation of these covers at least six subregions, while problems evaluated include the impact of new towns, the location of major airports, and the planning of urban structure within the framework of a prospective reorganization

FIGURE 2 Areas Using Lowry-type Models

of local government. (See Figure 2). These "sub-regions" are parts of Britain's eight national, economic planning regions. In American terms, these subregions would most closely correspond to medium sized Standard Metropolitan Statistical Areas (SMSA's) with multiple central cities, such as Albany-Schenectady-Troy, or Allentown-Bethlehem-Easton.

A SUBREGIONAL MODEL, 1969

Cripps and Foot, working at the University of Reading and supported in part with funds from the Center for Environmental Studies, have devised an operational version of the Lowry model, augmented by several of the variants developed by Wilson. These include the elegantly modeled use of population limits for zones which are presently filled, those that have room for population expansion, and those for which "planned" limits were set by local decisionmakers. In addition, the allocation functions include selected measures of attractiveness both for residential and service activities (Cripps and Foot, 1969a).

The original objective of the model was to prepare a plan for the County of Bedford, particularly for the new town of Milton Keynes. Cripps and Foot describe the calibration process in great detail, performing a service of documentation that is generally absent in operational versions. In addition, the evaluation of their own methodology and of the planning and development policies they were testing sets high standards for the literature.

AN EXPANDED SUBREGIONAL MODEL, 1969

Building on their experience with the model mentioned above, Cripps and Foot (1969b) enlarged the zonal system in a revised version:

> The main and original objective for developing the model was to prepare a structure [illegible] for the County of Bedford. [illegible] the core of the study area was obviously the County and [illegible] system of interest need only include additional areas which had major interaction with the County. This original objective has had to be modified to some extent to account for possible major developmental impacts other than Milton Keynes and this accounts for the further extension of the study area (Cripps and Foot, 1969b).

This second paper is noteworthy for its more detailed documentation of the model's concepts, data requirements, and [illegible]. In addition, several approaches to evaluation of the model's output are described involving comparisons of mean times spent commuting to work and traveling to shop. Further, there are appraisals of the effect of the constraints imposed by the model on spatial interaction.

On the agenda for further model development, Cripps and Foot express a preference for attempting to make Wilson's partitioning of behavioral types operational. This preference is implemented in a comprehensive proposal (Cripps and Batty, 1969) incorporating model development, surveying data availability, and testing in the unitary areas 52 (Reading and Berkshire) and 53 (West Surrey) that are proposed local units to be set up upon implementation of the report of the Royal Commission on Local Government (1969).

THE URBANIZATION EFFECTS OF A THIRD LONDON AIRPORT

The significance of the expanded area and zonal system comes into view in a third paper by Cripps and Foot (1970). The "major developmental impacts other than Milton Keynes" mentioned directly above turn out to be two potential sites for a third London Airport located in the study area. The ready availability of this operational model is brought to bear on another type of major planning problem, and with their methodology tested, Cripps and Foot are able to provide decisionmakers with a substantial catalog of measurements of impact. These include pop[illegible]tion, employment increase and distribution, land use relationships, economic structure, noise impact, and spatial interaction differentials.

NORTHWEST ENGLAND, 1969

Working in a parallel course of development to Cripps and Foot, Michael Batty applied the model to an area covering Central and Northeast Lancashire (Batty, 1970a). The version Batty used is noteworthy for direct use of the Garin matrix inversion process. In this application, Batty omitted the constraints usually included, but this is not illogical because the study area is a static and almost declining subregion. The purpose of the model implementation is to test the potentialities of locating a new [illegible] and these implications are pursued with great interpretative skill. The model was run recursively using a five-year module to cover a twenty-five-year time span, thereby forecasting the development time-paths for each zone. Furthermore, variations in parameters were used to achieve [illegible] the calibration process.

Of particular interest are the evaluative innovations that were tested. Batty calculated population potentials for the base and target years to generate a visual perspective of development impact. He also used correlation and variance analysis to test the [illegible] and calculated [illegible] work-trip length and mean service-trip length.

NOTTINGHAM-DERBY, 1969

Further development of the Batty version was applied to a subregion that incorporated a much larger population than those previously reported. With the larger population, the grain of development was more continuous and raised the issue of differences in model treatment between separated multi-nodal regions and continuous uni-nodal ones. Most significantly, Batty incorporated capacity constraints into the Garin-style matrix operations in this version and looked to the results for clues differentiating multi-nodal and uni-nodal development (Batty, 1969).

Batty's concern with calibration problems became the basis for a particularly cogent summary in an article synthesizing both of the above-mentioned experiences (Batty, 1970b).

A MODEL OF A TOWN (READING AND STEVENAGE), 1968

The value of parallel research efforts focused on the same problems and supported by the same source of funds, is exhibited by the concurrent development of an experimental model of the Lowry type by the team of Echenique, Crowther, and Lindsay, working at the University of Cambridge (Echenique *et al.*, 1969).

The contrasts with the Cripps-Foot effort start with the emphasis on the single town as a node, as compared with the multi-nodal region (subregion). Then, the zonal system is made up of a uniform [illegible] of one-kilometer cells, a la Lowry. Third, and most important, structures on the land are represented by a newly modeled variable - available floor space. Fourth, the creation of floor space in anticipation of its occupancy by residences and commercial enterprises is incorporated in a crude investment goods submodel which provides an enlarged emphasis on the supply side. Fifth, the tests of the model in a more or less freely developing town, Reading, and in one which is almost exclusively planned, Stevenage, suggest that this version of the model family can bridge a wide range of alternatives when applied in a planning mode.

The monograph is an interim report and promises further work in conceptual development, particularly in the direction of disaggregation of the type suggested by Wilson, as well as additional testing of these elaborated systems.

THE LJUBLJANA MODEL, 1970

The most intensive testing and comprehensive uses of the Lowry model in its original form are included in the technical reports of the American-Yugoslav Project. The study centers on the city of Ljubljana and includes the area surrounding it. The project was devised as a demonstration in urban and regional planning "basically to establish a frame of reference for [illegible] tion of a wide [illegible] forces, [illegible] plex interaction [illegible] of the urban region as a whole" (Music and Barber, 1970).

Stubbs and Barber (1970) applied the Lowry model (with only minor changes. Census tracts rather than grid zones were used. The interzonal distances were carefully composed of combinations of transit and vehicle travel times instead of airline distances. The model was modified slightly to reflect the importance of agricultural employment in the hinterland surrounding the city. The demand for population-serving activities from place-of-work was not maintained within the model structure.

With these few changes to adapt it to the conditions in Ljubljana, the model was run iteratively for five year periods from [illegible] to [illegible]. Careful study and judgment were used to tune the adjustable parameters and constraints so that a series of development alternatives were produced for evaluation.

It is remarkable that this most complete test of the Lowry model should have been conducted in Yugoslavia, in a foreign tongue, and partially under foreign auspices. Equally noteworthy is the applicability of the model under governmentally controlled development.

Stages of Development

The models included in this review were carried to differing degrees of completion, reflecting the varied interests of their creators, as well as the resources and data that were brought to bear upon their development. Three cumulative stages can be distinguished: *conceptual*, *experimental*, and *operational*. The *conceptual* stage is characterized by plans and strategies to extend the Lowry model, including embellishments of causal structure, variables, and theory and methodology. In the *experimental* stage, conceptual modifications are confronted with data, and successive trials and refinements, usually involving the use of computer programs, are carried out. The *operational* stage signifies that the experimental work is carried through to

TABLE 1 *Lowry Models By Stages of Development*

Development stage and name of model	Number of zones
Conceptual	
Garin-Rogers	not applicable
A. G. Wilson	not applicable
Experimental-Conceptual	
Lowry	456 (650 with dummy zones)
TOMM	189 (300 with urban perimeter)
BASS I	127 (260 with urban perimeter)
CLUG	196 (14 x 14 grid squares)
Crecine (TOMM II)	not documented
Echenique, *et al.* (Reading)	130
Echenique, *et al.* (Stevenage)	49
Operational-Experimental-Conceptual	
PLUM	291
Cripps-Foot (Bedford)	70
Cripps-Foot (Bedford and environs)	130
Batty (Central and N. E. Lancaster)	51
Batty (Nottingham-Derby)	62
Ljubljana	123

forecasts, evaluation, and applied decision-making.

Table 1 attempts to classify individual models, based on documented reports. For experimental models, this classification may reflect a transient status toward full operationality that is not yet documented.

Changes to the Lowry Model

Having reviewed the model variants in chronological order, it is appropriate to summarize the major changes and additions.

DYNAMICS AND QUASI-DYNAMICS

The models attack the handling of the *time dimension* in several ways. The most ambitious is the Wilson partitioning in which behavior of non-movers is contrasted with those who move their residences, others who move their employment locations, and still a third group who move both. These flows of movers are obviously measured in time units, as well as implying other dynamic influences.

TOMM I, earlier than Wilson, distinguished between "stable" and "relocatable" entities, a more elementary breakdown that leads to the Wilson partitioning. PLUM uses comparative statics, the comparison of equilibrium allocations at two points in [illegible] the growth vector. The British models generate static equilibrium future states, in the [illegible] Lowry model, but are moving in the direction of the Wilson approach in future versions. Finally, there is Rogers' intriguing concept of "stationarity," attained by successively iterating a matrix until it attains a stable allocation and an unchanging growth rate.

In evaluating the handling of time in these Lowry-type models, the consensus for the static form among the operational versions stands out with overwhelming clarity—even PLUM uses comparative statics. In contrast, much of the conceptual and experimental work attempts to incorporate some form of dynamics. There is much to be said for the long-run static equilibrium versions in terms of optimality, transferability of models among metropolitan areas, and economy of [illegible] requirements. On the other hand, the clear definition of the time-path of change, including shorter [illegible] points on that path, allows present policies to be more closely adapted to the long-run state (Harris, 1970).

DISAGGREGATION

Virtually all the models look on the prospect of further disaggregation with high expectations. The suggestions and examples depend heavily on data availability and expansions of computer capacity in addition to the particular modeling processes that are to be [illegible] to the data.

At the work[illegible] several forms of disaggregation are anticipated. TOMM II experiments with a white-collar/blue collar-break. In addition, Wilson presents wage stratification. CLUG is unique in distinguishing office activities from basic and population-serving.

Stratification at place of residence includes income levels, [illegible] types (Wilson, TOMM I, TOMM II, CLUG), and site amenities, usually in the form of surrogate variables (Wilson, TOMM II, Cripps and Foot).

Stratification differentiating the division [illegible] and [illegible] the region characterizes [illegible] of BASS I and PLUM. This spatial stratification implicitly reco[illegible] the historically varying influences on urban form and structure as one moves from the center to the outer perimeter of the region.

Several operational models (Cripps and Foot, Batty, and PLUM) abandon the Lowry stratification of population-serving [illegible] with the minimum size thresholds for these activities in favor of a single homogeneous category.

It is significant that although the conceptual and experimental modelers call for expanded disaggregation, the operational versions have not pushed far in this direction. It may be that we have some kind of grand experiment here that provides data for Alonso's cogent discussion of specification error and imperfect data as constraints on the complexity of modeling (Alonso, 1968). He states:

> . . . it is perfectly conceivable that we can devise predictive models which are beyond the capacity of the [illegible] in the sense that, although they are more "accurate" in their specification, the quality of the data results in a deterioration of prediction. I raise the question of whether in the field of land use and traffic models we have not gone beyond the best predictors. I must stress that I do not know whether we have or have not: but we must try to find out (p. 251).

CONSTRAINTS ON DEVELOPMENT

Several alternative approaches are presented to the problem of establishing upper and lower bounds on development. In some cases these are to keep the modeling process under control; in others, to allow planning units to be used. There is frequent dependence on the normalization ad[illegible] sometimes rationalized as a reflection of interzonal competition, particularly when the gravity function is used for allocation.

Wilson's most elegant treatment imposes economic constraints by confronting household income with costs of housing and transport. Conceptual and data problems intervene to hinder an easy solution, although Wilson does push the problem to its complex limit. CLUG contains a similar mechanism in simpler form for gaming objectives.

PLUM includes a zone-specific constraint process on residential development involving the holding capacity of each zone at existing density, and then modifying the density upward to respond to the pressure of excess demand at the zonal level.

Cripps and Foot impose constraints established by planners on some zonal capacities, leaving others [illegible]

AREAL UNITS, ZONES, AND NETWORKS

All of the models start with a defined study area, partitioned into a set of zones with [illegible] linked to other [illegible] by a network. These seemingly elementary inputs have substantial influence on the models' processes and their treatments differ among the models analyzed here.

As has already been mentioned, the study area may be a town (Echenique), a subregion (Batty, and Cripps and Foot), or a region. The regional concept used here is the [illegible] place (hierarchical) concept emphasizing urban development, rather than the region in the sectional sense, such as the South.

The study area is partitioned into zones which may be grid squares (Lowry, Echenique), census tracts (TOMM I, Ljubljana), or [illegible] tracts or of enumeration districts (PLUM, Cripps and Foot).

The critical variable used from the net[illegible] the [illegible] costs, or cost in [illegible] between [illegible] centroids (air-distance as per Lowry, BASS I), or road distance can be used (Echenique). [illegible] be estimated for a given network or for [illegible] by [illegible] and [illegible] and the application of free-flow or peak-hour speeds plus terminal times for each link (PLUM, Ljubljana).

VARIABLES

In addition to the variables implied by the suggestions for disaggregation, several of the models incorporate new variables to more completely comprise the development process. Land rent (TOMM II) and house prices and wages (Wilson) remedy an obvious omission in the urban economic mechanism.

CLUG emphasizes the whole package of infrastructure services, rather than transportation alone, as a necessary condition for development. Floor space is uniquely present in the Echenique model of a town.

MONEY FLOWS

A unique challenge for modeling reality is contained in the Cornell Land Use Game. Money passes through the system, encompassing wage payments, consumer spending, transportation costs, rental costs of business and residence sectors, local property taxes, and profit accumulation. Although the game context provides simplifying conditions which allow this to happen, the analogs in the model context should generate some experiments to comprehend regional accounts.

CALIBRATION AND EVALUATION TECHNIQUES

Although the authors are modest in their claims, the extensive work on calibration of the models by Batty and by Cripps and Foot constitutes a step toward optimization that is not visible elsewhere. This extensive experimentation clearly has its roots in the entropy maximization framework developed by Wilson.

Associated with this optimizing approach is the detailed variety of tests used for evaluation of the model outputs. Eschewing coefficients of determination (R^2) as totally inadequate, Batty and Cripps and Foot imaginatively create and test a whole battery of evaluative indices that will surely find their way into planning practice.

Conclusions

This review of Lowry-type models generates some significant conclusions. First, the model's innovative character is confirmed by the large number of prototypes designed to use the same conceptual framework. Second, a wide range of variants are documented, reflecting attempts to use the model for differing purposes, with varying data resources, as well as for [illegible] styles. Third, a remarkably high proportion of the efforts are confined to experimental objectives which did not reach the operational stage. Finally, the exceptional performance of British modelers is noteworthy at the conceptual, experimental, and operational levels.

This last conclusion merits some amplification. Wilson, in one of his earlier papers, distinguished between "elementary" models and "more advanced model concepts which demand knowledge of preference structures and individual utility functions in various forms." He continues, "Such a limited approach remains useful as more empirical work is possible in the short run using elementary models, and so they are worth developing until the two streams of work merge in the long run" (Wilson, 1969a). This commitment to a need for, and a use of, models in which elegance was traded off against "more empirical work" is one clue to the British achievements. Also, the organizational and financial coordination, embodied in the establishment of the Center for Environmental Studies, enhanced with funding capability, seems to have been an effective device. Finally, the variety of uses of urban activity models which are coming into view within the British planning framework provides sufficient demand for, and support of, operation and experiment. If, indeed, the streams of work are to merge, this review suggests there are several potential convergence points already visible.

Author's Note. *This review is part of a monograph being prepared by the author under a research contract with the U. S. Department of Transportation, Federal Highway Administration. The opinions, findings, and conclusions expressed herein are those of the author and not necessarily those of the Federal Highway Administration.*

Eric Cripps of the University of Reading was particularly helpful in providing materials and perspective on relevant British experience.

NOTES

[1] These criteria account for the omission from this discussion of the BASS III models described in *Jobs, People, and Land* (Berkeley: Center for Real Estate and Urban Economics, Institute of Urban and Regional Development, University of California, 1968), and other urban development models.

REFERENCES

Alonso, W. (1968) "Predicting Best With Imperfect Data," *Journal of the American Institute of Planners*, 34 (July): 218-255.

Batty, M. J. (1969) "The Development of an Activity Allocation Model for the Notts, Derbys Sub-Region" (Manchester, England: Department of Town and Country Planning, University of Manchester).

Batty, M. J. (1970a) "Models and Projections of the Space Economy," *Town Planning Review*, 41(2): 121-147.

Batty, M. J. (1970b) "Some Problems of Calibrating the Lowry Model," *Environment and Planning*, 2: 95-114.

Center for Real Estate and Urban Economics (1968) *Jobs, People, and Land: Bay Area Simulation Study (BASS)* (Berkeley: Institute of Urban and Regional Development, University of California).

Crecine, J. P. (1964) *TOMM (Time Oriented Metropolitan Model)*, Technical Bulletin Number 6 (Pittsburgh Community Renewal Program).

Crecine, J. P. (1968) *A Dynamic Model of Urban Structure* (Santa Monica: Rand Corporation).

Cripps, E. L. and M. J. Batty. (1969) "Outline of Research in the Urban Systems Research Unit" (Reading, England: Urban Systems Research Unit, University of Reading).

Cripps, E. L. and D. H. S. Foot. (1969a) "The Empirical Development of an Elementary Residential Location Model for Use in Sub-Regional Planning," *Environment and Planning*, 1: 81-90.

Cripps, E. L. and D. H. S. Foot. (1969b) "A Land-Use Model for Sub-Regional Planning," *Regional Studies*, 3: 243-68.

Cripps, E. L. and D. H. S. Foot. (1970) "The Urbanisation Effects of a Third London Airport," *Environment and Planning*, 2: 153-92.

Duke, R. D. and B. R. Burkhalter. (1966) "The Application of Heuristic Gaming to Urban Problems," Technical Bulletin B-52 (January) (East Lansing: Institute for Community Development, Michigan State University).

Echenique, M. *et al* (1969) "Development of a Model of a Town," Working Paper 26 (Cambridge, England: Center for Land Use and Built Form Studies), 59 pp.

Feldt, A. G. (1966) *The Cornell Land Use Game* (Ithaca: Center for Housing and Environmental Studies, Cornell University), p. 4.

Fleisher, A. (1965) "Review of a Model of Metropolis by Ira S. Lowry," *Journal of the American Institute of Planners*, 31 (May): 175-6.

Garin, R. A. (1966) "A Matrix Formulation of the Lowry Model for Intrametropolitan Activity Allocation," *Journal of the American Institute of Planners*, 32 (November): 361-4.

Goldner, W. (1968) *Projective Land Use Model (PLUM): A Model for the Spatial Allocation of Activities and Land Uses in a Metropolitan Region*, BATSC Technical Report 219 (Berkeley: Bay Area Transportation Study Commission).

Goldner, W. and R. S. Graybeal. (1965) *The Bay Area Simulation Study: Pilot Model of Santa Clara County and Some Applications* (Berkeley: Center for Real Estate and Urban Economics, University of California).

Graybeal, R. S. (1966a) "A Model for Forecasting Residential Development," *The Regional Science Association Western Section Papers*, 134-9.

Graybeal, R. S. (1966b) *Land Use Report* (Honolulu: Oahu Transportation Study).

Harris, B. (1970) "Change and Equilibrium in the Urban System" (Philadelphia: Institute for Environmental Studies, University of Pennsylvania).

Hemmens, G. C. (1970) *Urban Development Modeling*, Monograph Number 6 (Washington, D.C.: Program of Policy Studies in Science and Technology, George Washington University).

Lee, D. B., Jr. (1968) *Models and Techniques for Urban Planning* (Buffalo: Cornell Aeronautical Laboratory, Inc.)

Lowry, I. S. (1964) *A Model of Metropolis* (Santa Monica: Rand Corporation).

Music, B. and B. Barber. (1970) *Spatial Policies for Regional Development*, Technical Report 11, "Description of Alternative Patterns and Comparison of Lowry and "Hand" Forecasts" (Ljubljana: American Yugoslav Project).

Rogers, A. (1966) "A Note on the Garin-Lowry Model," *Journal of the American Institute of Planners*, 32: 364-6.

Royal Commission on Local Government in England (1969) *Report, Volume I* (London: Her Majesty's Stationery Office).

Stubbs, J. R. and B. Barber. (1970) *Spatial Policies for Regional Development*, Technical Report 10, "The Lowry Model" (Ljubljana: American-Yugoslav Project).

Wilson, A. G. (1969a) "Developments of Some Elementary Residential Location Models," *Journal of Regional Science*, 9: 1.

Wilson, A. G. (1969b) "Disaggregating Elementary Residential Location Models," Working Paper Number 37 (London: Center for Environmental Studies).

The Uses of Theory in the Simulation of Urban Phenomena

BY BRITTON HARRIS

Most of us who are engaged in one form or another of transportation and land use research have focused a very large proportion of our efforts on simulation This means that we have devoted our efforts to reproducing in some recognizable form certain aspects of human behavior or of the performance of mechanical systems or of a combination of these two We have done this generally in order to make predictions, and we have been interested in the [illegible]cy of predictions in order to assist our agencies or other policy-makers in making decisions It is the modest aim of this paper to provide a brief review of some of the ways in which theory can be of assistance in improving the similitude of simulations, and consequently, the accuracy of predictions and the wiseness of decisions There is, I believe, very little which is novel in what I have to say about the relationships which I imagine to exist between theory and practice Consequently, it might be wise to apologize in advance to the philosophers whose ideas I may abuse and to my readers for serving a warmed over menu.

Since there is a good deal of popular jargon which tends to imply that practical activities are useful while activities dealing with theory tend to be nonproductive, I intend to devote a part of this discussion to what might be called paradoxically a down-to-earth defense to these [illegible]ctical activities—and to some extent I shall oppose what I would call crackpot realism with what might be termed realistic idealism. As Bertrand Russell has said, "Nothing is as practical as a good theory."

THEORY AND PRACTICE

n very simple terms, theory is a general statement about the real world In these simple terms, for example, the Pythagorean theorem s one of many consequences of the theory of Euclidean geometry As such it makes its own general statement about the properties of rightngle triangles on plane surfaces, and has had remendous practical influence in surveying and ngineering. This theorem provides the basis for all the well worn formulas of elementary trigonometry, for example There are two ways in which, however, we need to qualify this simpleminded definition of theory, and it is these qualifications which may tend to give the notion of theory some of its otherworldly character. First, when we say that theory talks about the real world, we have to include in that real world the minds and ideas of men Thus, theory may deal to some extent with technology and concrete things on the one hand, and on the other hand with mental constructs which are seldom or never encountered in the physical world outside of men's minds until they have been written down. The real world of mental constructs is a very important one, and in the end has many practical applications. The extension of the trigonometry of measurement into trigonometric functions, for example, is the basis for other large parts of engineering The second qualification is that a theoretic statement about the real world may not be, to the layman at least, a recognizable mapping of the real world, and the nature of the correspondence between the theory and the world and the consequences of the theory may not be readily expressible in everyday language. This sometimes makes it difficult for the layman to conclude at first glance that the theory is in any sense realistic or has any practical consequences.

There is of course an intimate relationship between theory and science or between the verification of theory and the scientific method Since theories consist of statements about the real world, their degree of correspondence with this reality can be tested Where the "real" world in question is one of mental constructs, as in logic and mathematics, the testing may be of a special and somewhat different nature, based on internal relations between constructs It is not in general a requirement of the development of conceptual systems and then theory that any direct correspondence with mate-

Reprinted by permission of the *Journal of the American Institute of Planners*, XXXII, No 5, September 1966, 258 278.

real phenomena should be established, but it has frequently proved to be the case that, after short or long periods of development, such concepts have found important and unforeseen applications to theories of phenomena. This course of events is analogous to, but not the same as the laboratory development of methods and devices which for a long while remain mere curiosities, but which ultimately become technologically important.

We live in an age of rapidly expanding scientific endeavor. Science and the scientific method are being newly applied in the first instance to old systems of human thought such as ethics and philosophy, and in these areas, the boundaries of untestable contention are constantly being narrowed. We now not only know that because of the atomic nature of matter only a finite number of angels can dance on the point of a pin; we also feel greater confidence in the rigor and cogency of philosophy. At the same time, new groups of phenomena are becoming the subject of science. Testable rather than speculative hypotheses are developed about these phenomena, and these hypotheses are organized in increasingly unified systems, frequently of a quantitative nature. Some of this movement towards new applications of the scientific method is occurring in the social sciences, and in this field the two tendencies to reduce the area of philosophical and ethical speculation and to systematize our understanding of objective phenomena go hand in hand.

It is hardly necessary to review the practical ways in which the advances of science during the last two centuries have greatly increased man's control over his natural environment through the application of science to technology. It is more useful to point out, first, that not only is science successful in an objective sense, but also that it is widely accepted publicly and politically, as may be judged by the governmental and private resources which are devoted to it, and, second, that the growth and prestige of science have not simplified but have complicated the distinction between practical and theoretical endeavors.

The "customers" of the scientific establishment are basically interested in results and frequently have shorter time horizons than the scientists. This dichotomy expresses itself in the distinction between science and technology as disciplines, and organizationally in the distinction between mission-oriented research and theoretically oriented research. As Alvin Weinberg has recently emphasized, the objectives of mission-oriented research are externally imposed upon the scientific community by the real or imagined needs of society and by society's control over expenditures, while the objectives of theoretical research are largely generated within the scientific community.

It appears to me that these distinctions, while valid and useful for analyzing and discussing the problems of science and technology, can be unduly overemphasized. Many factors tend to blur the differences. Technology and engineering themselves are becoming more scientific in their basic methods, and consequently engineers are becoming scientists. Mobility between the professions tends to inject mission-orientation and social responsibility into the scientific community, which was in any case never detached from these values. The tremendously accelerated pace of science tends to shorten the scientist's time horizon and bring it more into accord with that of the decision-maker. Finally, the complexities of real life which face decision-makers are driving them away from simplistic commonsense judgments and in the direction of a more comprehensive and quantitative approach to the problems which they face.

It is in fact the magnitude of the problems of societal control in a period of rapid change and development which are providing the impetus for the truly scientific development of the social sciences. Problems such as maintaining peace, feeding billions of people, reducing racial discrimination, and organizing great cities require powerful instruments of control over men and machines. These problems of control can no longer be resolved by an engineering approach which is overwhelmingly oriented towards physical, inanimate, machine systems. Engineers working in transportation planning must pay increasing attention to problems of human behavior, and it is rapidly becoming evident that the relevant behaviors are not only in the fields of transportation demand and transportation system utilization, but also in the field of land use development and locational choice. In a sense, therefore, and almost willy-nilly, the planning-engineering professions find themselves working on a frontier of science. This is the area of social behavior and social control, in which the application of the scientific method has been unduly retarded. In order to explore what implications this situation will have for their work, we must therefore take a closer look at some of the elements of this method.

SCIENTIFIC METHOD

We are used to the idea that man and the other higher primates are endowed with an innate curiosity which leads them to explore their [illegible] environment in an apparently insatiable and not entirely purposive way. There is usually no a priori identifiable useful payoff in some of the exploratory activities of monkeys and children, and one is tempted to make an analogy with the data-collection propensities of social science research and transportation studies. It is also perfectly clear, however, that in man at least, curiosity extends beyond the accumulation of data about the environment. First, even the childlike exercise of curiosity involves the exploration of cause and effect. The "experimenter" will employ some of the simpler ploys of the scientific method to discover what worked when and where. And second, there is frequently an effort to *generalize;* there seems to be a tendency to seek out analogies and similar situations in which earlier findings and elementary 'theories' can be tested. Thus we have in a primitive form the four [illegible] steps in some classic descriptions of scientific method: First, induction; the collection of information and its organization into patterns. Second, generalization, a restatement of the cause and effect relations behind the patterns, or a redefinition of the patterns themselves in a more abstract form which includes the observations as a special case. Third, deduction: the search for new special cases previously unstudied, as suggested by the more general statement, or theory. And finally, testing, a check to see whether the new cases perform as predicted—if not, the theory must be revised. This schema, while useful for analytic purposes, does not correspond in its rigid division of steps with the way in which scientific investigations actually proceed. We will use these categories as a basis for discussion, specifically maintaining, however, that the classification is artificial and if pressed too far, actually harmful.

The testing of theories about the world of phenomena raises special problems in the social sciences which should be generally understood before we take up other aspects of the scientific method. In this discussion I use the term "testing" in preference to the more usual "verification" because in principle no theory can be established, but only disestablished. A theory does not have verity, but verisimilitude. There are of course many theories which are outstandingly successful and for which there have been an almost unlimited number of successful tests, while the unsuccessful tests are nonexistent or occur only under well-defined special conditions. An especially significant case in which a theory may be regarded as firmly established because no counter-example exists is the correspondence between the counting numbers of everyday experience and the invented set of positive integers as defined in modern algebra by the use of the Peano postulates or otherwise. The theory states that these two systems, one from real life and one from mathematics, have the same form, and no exception to this theoretical statement is known or is likely to be discovered. A second example is the Newtonian statement based on the laws of motion, the law of universal gravitation, and Euclidean geometry. This theory relates the mechanics of the real world to a mathematical system of differential equations (a field of study which Newton, in fact, was forced to invent). Up to the point where by relativistic considerations the Euclidean geometry no longer applies in real space, there are no exceptions to this theoretic statement, and it too may be regarded as firmly established. In this latter case, it is important to note that the correctness of the theory has only been established by eliminating or controlling extraneous factors such as air resistance which affect the theoretically defined unimpeded motion of observed bodies.

THE SOCIAL SCIENCES

The special problems of the social sciences arise, as is well known, out of the difficulties of pursuing this experimental method in which most variables are held constant (the *ceteris paribus* assumption of economics) while a lim ited subset of possible variables is manipulated. Where a large number of variables is involved, this difficulty can sometimes be overcome if a large number of diverse observations is available to the researcher, but this is unfortunately not the case with regard to the study of the development, and the manipulation, of large urban areas. Here the case material is limited in extent, and experimental manipulation is both extremely slow and vastly expensive with regard to the aggregate phenomena. Experimental *cum* statistical methods are only possible with respect to smaller elements of the total system. In these regards, science as applied to total development of the function of the urban system is in most respects analogous with its use in astronomy which has a few cases of major interest, subject matter which is inaccessible to experimental manipulation, and the capacity for studying in the physics laboratory elements which do not aggregate by simple addition into the whole.

It may however be argued that the disadvantages of the social sciences in establishing an experimental method have been greatly exaggerated. This argument is advanced on the grounds that the most important tests of theories concern their ability to predict new phenomena or phenomena not previously studied in detail and to extrapolate the effects of causes beyond the ranges in which the causes were originally observed. It is curious to note that the literature of engineering and the social sciences abounds with warnings as to the dangers of extrapolation. If we wish to use, as we are almost forced to do, the power of extrapolation of a theory as a test of its credibility, then this cautionary advice is a frank confession that the relationships being extrapolated have no theoretical basis whatever. From the point of view of testing theories, the social scientists should welcome rather than shun opportunities for extrapolation, since this will provide his main basis for justifying a theory or for designing improvements in it.

INDUCTION AND DEDUCTION

Any acceptance of this criterion for testing theory tends to indicate the ultimate futility of a complete reliance on induction for generating theories. Even in the physical sciences, a fairly thorough knowledge of a particular range of joint variation of phenomena does not guarantee any adequate knowledge of cause and effect or even any complete description of relationships outside the range of observation, and this is even more true of the social sciences. Most of us are thoroughly familiar with the situation which arises when we get a good fit of a polynomial to a set of observed data, only to find that it behaves extremely erratically outside the range of observation. Poincaré has pointed out that if we had a complete knowledge of a portion, however small, of a continuous function, we would have a knowledge of the behavior of the function over its entire range. We could attempt to reach this happy state by developing the function as an infinite series and fitting all its coefficients statistically. Unfortunately, this procedure requires an infinity of observations (and a larger infinity than $\aleph$). Even more important, our observations must be free of errors of measurement, and our function must include all relevant variables, each of which must be measured. It is quite clear that even the process of induction from observed phenomena must be guided by theoretical concepts based on previous experience which will suggest the ranges and objects of observation and the character of the functions to be fitted. A simple engineering example might be found in the difference between the parabolic curve and the catenary. These curves arise under different circumstances in the construction of suspension bridges and lie extremely close together over a certain range of the variables. To distinguish between them by induction would be a hopeless task, especially since the formula for a catenary would not intuitively occur to a statistician, yet in extreme cases the distinction is important for engineering design. The differences are well defined a priori on the basis of a theory which may have been suggested by observation, but which does not spring automatically from it. The difficulties of induction are conclusively delineated, in fact, by the difficulties which arise in social science research in selecting functions for curve fitting and interpreting the results. Linear models are most frequently used because of their simplicity perhaps with the justification that the linear approximation to some unknown function is not unreasonable over the range of the observations. The function being unknown mean that theory is out the window. Perhaps a polynomial is used as some sort of an approximation to a Taylor expansion of a function. In this case, the catenary is defined as a parabola. Where the choice of function is deduced from a priori considerations and not merely to satisfy goodness of fit, we are suddenly in the realm of deduction rather than induction.

Deductive thinking is of very high value in science. Examined closely, the antinomy between induction and deduction is somewhat artificial; on the one hand, as we have just suggested, induction is almost never initiated without some kind of prior theory, however naive, which suggests areas of investigation, relevant variables, and the form which functions might take—while on the other hand, if deduction is unsuccessful and does not result in a confirming instance of the theory on which it is based, then the contradictory evidence may be the basis for a new round of induction. But the importance of deduction as a part of the scientific method remains in spite of the partial artificiality of its separation from induction. At the start of the process of deduction, the investigator is forced to make a statement of a general nature about the real world, in other words, he must formulate a theory. The motive for this formulation frequently comes from psychological forces very closely related to induction and to the search for generality. To follow the processes of deduction suggested by the theory, the investigator must search out new areas or new modes of application of the theory. It is useful to him in defining supposed cause and effect or functional relationships of variables to be investigated. In considering, therefore, the nature and power of the deductive process, we are led naturally to the final and perhaps the essential part of our discussion of theory construction, that of generalization, or the actual formulation of the theory. We must consider this in the light of all of the processes and problems which have been discussed up to this point.

GENERALIZATION

Generalization is the bridge by which the scientist or theoretician crosses over from induction, or the observation of reality, to deduction, or the testing of theories and their application to new phenomena. For this reason, I rather like the name "transduction," which is sometimes applied to it. No matter how often this bridge is crossed in the course of a scientific investigation, the act of transduction always involves some invention on the part of the investigator. The psychology of invention in this field is intricate and fascinating, but a discussion of it is out of place here. The sources of this invention may, however, be better understood through a consideration of its inherent nature.

The construction or invention of a theory involves in essence a precise statement regarding formal relationships, usually including relationships of cause and effect. There is an infinity of possible formal statements of relationships which may be made in their most abstract form in the language of mathematics or logic. Such statements regarding relationships are purely formal and have no reference to the real world. Within the sciences dealing with concepts, they may in fact be developed quite independently of the real world. The problem of theory construction or invention is, then, to make the correct identification between a real phenomenon and a mathematical or logical statement regarding relationships. There are three possible ways in which this may be done, two of which are merely suggestive and one of which tends to satisfy rigorous scientific requirements.

First, an analogy may be recognized at the level of phenomena. Thus, for example, a city may be compared with an organism—say a jellyfish. This analogy is scientifically useless unless two conditions are met. First, the comparative object (the jellyfish) must have a form which has been clearly and logically defined and, second, the object compared (the city) must be unequivocally said to be theoretically identical. In this case, we have identified a correspondence of the third type below, but otherwise we have merely made a statement which is useful for heuristic purposes.

Second, an analogy may be recognized as between a phenomenon and a mathematical or logical construct, but may indeed be very loosely defined. Thus, for example, the gravity formula recognizes an analogy between the decay of trip frequency with distance and the power function X^{a}. This analogy is extremely crude, seizing as it does on the most obvious and easily manipulated of a host of monotonically decreasing non-negative functions. No statement of the gravity model, to my knowledge, states any causal relationships which would generate this particular function in preference to others. I think that we may designate an analogy between phenomena and a logical function as a *homomorphism*, meaning a similarity of form.

Third, an important qualitative change is introduced if a scientist identifies a particular phenomenon as having a clearly defined logical form. The definition of form may have already been made either in the development of logic and mathematics and unrelated to phenomena, or in connection with the development of theory dealing with some other and perhaps completely unrelated phenomena. The use of formal statements pertaining to other phenomena is indeed often suggested by analogies between the phenomena themselves. On occasion, the study of phenomena themselves and the formulation of ideas about cause and effect necessitates the invention of a new relational calculus. This has been the case in Newtonian mechanics and quantum mechanics. In any event the essence of a theoretical statement is to identify an *isomorphism* (identity of form) between a set of phenomena and a logical or mathematical relational system. Thus, the Schneider model for trip distribution, in contradistinction to the gravity model, makes a rigorous statement that the decay of trip frequency is isomorphic to the negative exponential function and consequently also to the radioactive decay of fissionable elements, and Schneider identifies the precise cause and effect relationships on which the isomorphism is based.

It should of course be clear that the borderline between homomorphisms and isomorphisms is blurred, partly because it refers to the motivations and psychology of the scientists. A theory which is in fact generated as an analogy must be presented as an isomorphism, and a badly conceived isomorphism may turn out to be only an analogy. The appropriate distinction can be made only upon close examination of the theory and of its results.

In formulating a theory to serve as a bridge between induction and deduction, the analyst has a number of guides as to desirable features of his formulation. Some of the most significant criteria tend to conflict with one another and others to reinforce each other, depending on circumstances. All arise out of the general characteristics of the scientific process as we have outlined it.

The outstanding criterion, of course, is that the theory should be "correct," that is, that the theory if testable should pass its tests successfully. This is the basis for the essentially practical nature of science, that is, that it says "true" things about the real world. We have seen above that such truth is unpermanent, always awaiting contradictions. If these arise, it frequently is retained circumscribing the generality of the theory, limiting the circumstances in which it applies, and creating new and more general theories to apply to other combinations of circumstances. This criterion of correctness is in general overriding. However, practical considerations frequently lead to the generally indefensible practice of applying theories which have been inadequately tested or which have known errors.

Thus a theory, if untestable, becomes a non-theory; like Milton, science has "no use for a fugitive and cloistered virtue. . . ." There are of course examples of very important scientific theories, such as Einstein's theory of relativity, which when published appeared very difficult if not impossible to test. These difficulties in relation to a reputable and exciting proposal often serve as a spur to experimental work. In the field of social sciences, however, this particular type of non-theory has two other manifestations. One of these is a normative admonitory description of how things should be done; the other is a literary or pseudo-statistical description of the real world. The fact that these things are merely masquerading as theory can easily be exposed by searching for critical tests which could deny their validity. If such tests do not exist, the so-called theory is in fact a non-theory. Occasionally tests acceptable to the authors of the theory will be so circumscribed with restrictive conditions and assumptions as to expose the fact that the theory is of extremely limited application and has dubious stature.

The testability of a theory is a special case of a more general property of useful theories—productiveness or fruitfulness. Important the-

[illegible] in the development of science not only answer the problems posed in the initial stages of induction, but are pregnant with consequences which are only dimly seen by their inventors and which lay the basis for a wide variety and a great number of deductive experiments. Axiomatic systems in geometry, algebra, and logic exhibit this property, for example. The tremendous accomplishments of modern mathematics follow (although not effortlessly) from a very limited set of carefully considered initial assumptions. Similar examples exist outside of conceptual systems. The quantum theory, which was invented to explain anomalies in black-body radiation, has found innumerable applications to phenomena as diverse as photoelectricity and solar spectrography, and is in fact a key element in all modern physics. The social sciences and the planning-engineering professions are somewhat lacking in such key theories, but some nominations could be made. These might include, for example, marginal substitution and general equilibrium concepts from economics, and applications of general systems theory. In any event, while it is somewhat difficult to [illegible] the process by which a theorist comes upon a fruitful theory with many applications while attempting to solve a more limited and more particular problem, it is apparent that solutions of this type are unusually desirable. At the least, theory builders should have this objective in view, especially since this state of mind leads to stripping any problem to its most essential elements, and thus may simplify as well as lead to greater generalities.

Simplicity is in fact an ancient and honored criterion for choosing between otherwise equipotent theories. Occam's Razor, named after a fourteenth century English philosopher, dictates that theories should contain the minimum possible number of hypotheses, and many of the more durable theories elegantly exhibit this characteristic. Because of the large number of conditions, relations, and variables which occur in social science research, this condition is difficult to meet here and frequently conflicts with requirements of realism and testability. It is nevertheless a desirable characteristic, not only for reasons of elegance, economy, and generality, but also for practical reasons which will be discussed in the next paragraphs. Here there is a special pitfall which social science researchers can dig for themselves by the use of modern computational techniques. It has been suggested that, had computers been available at the time of Copernicus, the ease of computation of epicycles might have removed the practical difficulties which led to the construction of the elegant and economical heliocentric theory and the Newtonian theory of celestial mechanics. By the use of computers in the descriptive system of Ptolemy, navigational tables could have been constructed to any required degree of accuracy, and the practical impetus for the Copernican and subsequent revolutions would have been removed. I feel that we fall into the same trap when, as with the introduction of K-factors into the gravity model, we constantly patch up a nonexistent or inadequate theory with computational amendments.

MODELS

To be testable a theory must be manipulable. The experimental method in the social sciences is, as we have said, forced to rely on paper experiments, and [illegible] our professions commonly talk about the use of models. To quote Harmer Davis, "A model is a smaller copy of the real thing, as the woman said about a model husband." This pointed definition does not permit us, however, to distinguish between a mathematical model and a simulation model on the one hand, nor between a simulation model and a theory on the other hand. We shall devote brief attention to these two distinctions.

As we have tried to emphasize above, there are in principle distinct sources of an understanding of cause and effect in the real world and of formal representations of relationships in the world of mathematics and logic. Science is in many respects an effort to establish isomorphisms between these distinct realms. If we refer to a linear programming model of warehouse location, we are referring to just such an assertion about an isomorphism. We might then be correct in speaking of a *mathematical model of warehouse location*. Frequently, however, people speak of the linear programming model, and more generally of *mathematical models* in the abstract without relation to any particular real world phenomena. I would submit that this application of the word 'model' is incorrect, though lamentably ineradicable, because the mathematical linear programming model is not a model or a smaller copy *of* anything.

The distinction between a theory and a simulation model is somewhat more subtle and difficult. A theory in fact could also be said to be a logical or mathematical model of the phenomena to which it refers. It is smaller; it is a copy; and it is *of* the real thing. Yet this identification of a theory with a model somehow goes against the grain. On the basis of very serious consideration, I have redefined models as they are used in the simulation of social and economic events in a way which tends to provoke outraged reactions, but which I believe withstands serious examination and criticism. *a model is an experimental design based on a theory*. Let us examine the implications of this definition somewhat more carefully.

As is well known to workers in our fields, there are many theories which are testable in the sense that a critical experiment can be designed, but which remain untestable in the sense that the data requirements are for practical purposes excessive, that they involve presently unobservable variables, or perhaps most important, that they cannot be cast in a form which will fit into a computer and run economically. These practical considerations do indeed provide a spur to all kinds of experimental ingenuity, and they should by no means dominate the process of theory construction.

In the process of the development of a theory there are many applications of experimental design in which, however, the theorist must invoke models. First, in exploratory or inductive investigations, he is quite apt to use a severely truncated or patently inadequate experimental design such as a multiple regression model to explore relationships and to provide information as to the direction of his next transductive steps. In a more developed form he will use a model more closely corresponding to theory inductively to establish the parameters of relationships. Second, he will use a model for testing in the deductive sense in order to determine the applicability of his theory under a wider range of conditions. Third, used scientifically in a context of projections, the model will provide experimental evidence as to the consistency of the theory and possible inductive evidence as to the sensitivity of the real world to changes in conditions. It may be a matter of scientific but not practical indifference to the scientist that the projective use of models also is important to decision-makers.

One may choose to make a distinction between the value of theory building and the value of experimental work with models, imputing a higher value to the first of these activities. However, in the tradition of British and American experimental science, the theorist usually has some responsibility for making feasible the tests of his ideas, and it is only the boldest and most brilliant innovator in pure theory who can expect others to accept a division of labor in which they will devise feasible tests for his "impractical" formulations. It is this experimental difficulty which often leads to emphasis on the false dichotomy between theory and practice, which can only be overcome by a long-run view of the value of theory and by a nice sense of the potential contributions of new theories whose testing and application may appear outrageously difficult.

There are other criteria which may or may not be useful for the construction and selection of theories but which are frequently in the minds both of scientists and lay people. We have mentioned above the fact that for a variety of reasons a theory may not correspond directly with intuitive and popular ideas about the nature of reality. In this case, the theorist or scientist may be accused of being unrealistic and may feel a social obligation to change or alter the tone of his theory in the direction of popularly understood "realism." Such a compulsion grossly distorts the role of the scientist, which is to identify a genuine isomorphism, most frequently not obvious, between the behavior of the real world and a set of mental constructs. Frequently he has to invent the mental constructs in order to disclose the isomorphism. Many of the most pregnant ideas of the physical and biological sciences, such as the quantum theory, the theory of relativity, or the independence of heredity from environment run counter to widely held and deeply rooted popular ideas, yet without the discovery of these theories and their application to everyday life, the world would have given up a great deal of progress. A search for naive realism is counterproductive in science.

Frequently even though a naive demand for realism may be abandoned, the critics of science will continue to take refuge in an unthinking insistence on comprehensibility. In the field of social sciences, this insistence is based on two circumstances. First, every critic is a member of society, a user of cities, and a participant in the political process. He hence feels intuitively that by virtue of this special status he and most other informed people ought to be able to understand directly all of the theories which purport to define the operations of society, of cities, and of politics. In my view, it would be equally ridiculous to say that because we are all made of protein, we should all understand at a glance the theories of molecular biology. A second circumstance resides in the fact that a great deal of social science research is conducted in such a way that the scientists are close to the administrators, the administrators are close to the decision-makers, and the decision-makers are close to the voters, all with no clear separation of function. Because of the personal and normative nature of the communication between these groups, each link in the chain feels that he ought to know all about what the adjacent link is doing. We may contrast this with the somewhat more impersonal relationships which govern research and development in industry. The laboratory scientist may understand solid state physics in detail. The corporation executive will understand the main directions of this research and its potentialities. The sales department understands the capability of the resultant product, and the customer chooses in the market place between the products of competing technologies and competing companies. The man in the street could not care less about the crucial role of, say, quantum mechanics in the production of his transistor radio. Probably when social science theories produce as effective results as quantum mechanics, the administrators, policy makers, and voters will be less inclined to ask questions and more inclined to judge by results.

A possible requirement for theory which requires brief mention is more likely to be generated by the scientist than by the layman. As a result of the complexity of social phenomena which require holding other things constant, and as a result of the drive for generalization which is inherent in theory building, there is a considerable drive to create theories which are "comprehensive." This drive encounters resistance on two fronts. First, a comprehensive theory may in certain cases become so general as to say nothing about everything. Even if this is not the case, the more comprehensive theories may be the most difficult to manipulate for purposes of testing and application. An important part of theory building is therefore a nice sense of discrimination as to when comprehensive theories are necessary and when they may be appropriately avoided by discretion in the subdivision of the problem into manageable parts. In policy related sciences improper subdivision of the policy-making problems may result in suboptimization, but a subdivision of the problems of the real world and its functioning for purposes of study need not entail this danger.

LAND USE SYSTEMS AND TRANSPORTATION SYSTEMS

In the preceding sections of this discussion, I have developed my ideas with regard to the scientific construction of theory, mainly with respect to the problems of simulating events in the real world of mass behavior in the use of transportation facilities and in the choice of locations, even though this concern has been in the main implicit rather than explicit. There are two other areas related to public decision-making in which theories of a different kind will have to be developed. These deserve brief mention. Transportation and planning literature already recognizes the need for the development of more general theories of decision-making. In crude terms, the questions to be answered by such theories concern what we are planning for, what trade-offs are involved in the public decision process, and what values our plans will satisfy. In more sophisticated terms we turn out to be dealing with difficult problems of public discount rates, collective consumption, spillovers and externalities, the aggregation of utilities, and the reconciliation of conflicting interests. It is to be hoped that the improvement of theories and models in this general area may be expected to reflect backward into the planmaking process so that sketch planning procedures are replaced by optimizing procedures, and so that optimizing is not limited to narrow engineering criteria but is extended to the most general of social objectives. I think it is also predictable that as we explore the problems of decision making, planning and optimizing more thoroughly, we will discover that there are ferocious computational problems which arise in the design process as a result of the huge combinatorial variety which exists in the possible combinations of policies and future conditions. Our fraternal theorists in the field of mathematical programming may be able to make contributions of a theoretical nature with practical applications which are related to the needs of decision making. It is also probable that a clearer formulation of these needs will influence this development of what are essentially design models.

We have now reached the vantage point of a somewhat shaky and perhaps imperfect understanding of some of the processes of science, from which we may view the needs and accomplishments of experimental simulation of transportation systems and land use systems and the behavior of their users. I will not here belabor the point which is now becoming widely accepted in principle—that in many policy making contexts we are dealing with these systems not independently, but as a part of the larger urban metropolitan system or regional system. I will emphasize the fact that most theories of locational behavior contain ideas about transportation costs and convenience, and consequently that locational models must contain as submodels some replication of the transportation aspects of the system. It will also prove useful in the discussion which follows to consider the salient features of all these problems together from the point of view of theory construction, drawing freely upon examples from any field wherever they may be appropriate.

The range of our interest in these phenomena covers a wide span from very large and complex total systems through subsystems which may be defined in engineering terms, in social and economic terms, or in spatial terms, down to the smallest elements of the system themselves. These last may be mechanical components. But the greater interest attaches to decision units—a man driving a car, a family looking for a home, or a corporation deciding to build a new establishment. At each of these levels, different problems arise regarding the appropriate form and content of research.

The broadest view of the over-all system is probably not in itself highly productive, but it is a starting point for certain applications of general systems theory which later affect our view of the components and the elements. General systems theory with respect to the total urban, metropolitan, or regional system will ultimately play a direct role in decision models. Meanwhile, it can be particularly useful in defining the appropriate limits of a system and in guiding the structuring of the problem in such a way that its decomposition into subproblems dealing with subsystems will entail a minimum of distortion. Up to now in both transportation and land use analysis these two problems have been approached largely by intuition and induction. I do not feel that the results have been seriously wrong, but a systematic and better informed approach might provide some surprises and prove a useful guide to research design.

With respect to subsystems properly defined and considered as systems in their own right, general systems theory may very well contribute powerful methods for dealing with system stability as a planning objective and with homeostatic or equilibrating tendencies within systems as handles for both planning and analysis. I have my own intuitive feeling that concepts of equilibrium animate a great deal of research and theory in land use and transportation analysis, but that these concepts are inadequately explored and not sufficiently explicit. For example, transportation analysis and the assignment of traffic to networks with capacity restraints imply a whole pattern of equilibrating behavior on the part of individuals which may or may not lead to system equilibrium and may or may not be related to various forms of optimization. These questions have been very lightly explored by brute force iterative methods in modeling experiments, and their full implications remain to be seriously examined. In land use growth model simulations based upon trend data, there is also a set of unexplored assumptions about tendencies to equilibrium. Whether such an equilibrium exists or ought to exist has in fact been slightly examined in theory. Needless to say, one shot sketch planning or design models and "instant cities" such as Ira S. Lowry's *Model of a Metropolis* are constrained to use either simultaneous determination or optimization, and it seems likely that the former method contains some optimizing assumptions in its behavioral parameters. More generally, I feel that land use behavior as well as land use system performance can hardly be explained without a consideration of land market equilibrium and simultaneous determination—all of which pose major problems for system theory.

COMMUNICATING MECHANISMS

There are a number of interesting problems which arise out of the communication among subsystems and between elements and subsystems and out of the mechanisms by which equilibrating, disequilibrating, and determining forces are transmitted to and from decision units. The organs of the body communicate information leading to action by nerve impulses and those maintaining homeostasis by chemical messengers. What are the messengers in a large city or region? Many of these questions will arise again in the discussion of the behavior of decision units below, but there is some advantage in taking an overview in the context of systems. It is quite apparent that the generic name for these messengers will be information, and it seems quite likely that some gains in theoretical clarity will be achieved if a systematic application of communications theory can be made to the diffusion of information through and about the systems under study. The applicability of this concept is already apparent in the most elementary consideration of the stability of traffic flow systems, and these ideas can probably be extended to land use systems and larger transportation systems. Considered in the communications context there is some merit in combining the study of decision units with a priori considerations from different disciplines as to what information is likely to be important and available. At one extreme this type of merger leads to a consideration of the individual's reaction to the visual environment as developed in studies by Lynch and others. At a different extreme, economics suggests that prices are the messengers by which important economic information regarding, say, the housing market is transmitted. Between these extremes lie many combinations of phenomena which are observable, influential in behavior, and to some extent predictable as consequences of other developments.

The importance of prices as a messenger and of the allocation of money to different purposes (that is, of economic behavior) in private decision-making is so great that it deserves special attention. It is a curious fact that in spite of this a priori importance of monetary phenomena they have really received relatively little emphasis in transportation and land use planning and analysis. For somewhat understandable reasons, transportation planners have been reluctant to explore the importance of pricing policies in alternative transportation systems. Surely, however, this reluctance should not extend, as it frequently does, to the omission of cost factors and the exclusive emphasis on time-distance which is frequently found in network analysis, trip distribution, and even modal split. Fortunately, this default is not universal. In land use analysis, the problem is perhaps even more severe. Housing rents and values are the medium through which consumers communicate with each other their willingness or unwillingness to compete for space, and more commonly land prices are the medium of communication in the competition of residential, industrial, and public uses for land. Yet in the research field, housing value and land prices very seldom appear as variables. So damaging is this omission that expensive and otherwise useful surveys of locational, social, economic, and housing variables by the Penn Jersey Transportation Study and the Tri-State Transportation Committee are partly vitiated by the failure to inquire as to housing value or rent. It must be admitted that the collection of these data and especially of land value information in a research study is fraught with difficulty, but I believe that there is a more serious reason why these values have been neglected in spite of strong theoretical reasons for their inclusion.

If values (prices) are made explanatory variables leading to changes in the behavior of decision units, then future applications of the same theory and its derivative models require that these values be projected under new circumstances. The theorist then faces an ugly dilemma. If he chooses to predict future prices by means of proxy variables, he must build a purely descriptive model for this purpose which contains no ideas about cause and effect; and this being the case, he might just as well have left prices out of the original analysis and included the same proxy variables, admitting from the outset that his theory was in part purely descriptive. If on the other hand he takes the importance of these economic variables seriously, he must face the difficulty of reconstructing a complete market through some form of simulation. This reconstruction is complicated by the existence of submarkets, institutional stickiness, imperfect dissemination of information, and probable lags in equilibrium. If economic considerations were largely peripheral to the theory of land use and transportation systems, there would be less objection to taking the easy way out of this dilemma. I believe, however, that these considerations are so central that economic models must in the future be added to the implementation of transportation and location theory at full scale. This approach will involve much deeper consideration of equilibrium tendencies than was suggested above, and perhaps a much more serious look at some aspects of the behavior of decision units.

BEHAVIOR OF DECISION UNITS

Before turning to a discussion of the theory of the behavior of decision units, I must emphasize a vital distinction between the study of that theory and its application. To a very large extent, the study of the behavior of decision units can be undertaken independently of the simulation of system and subsystem performance which has been the subject of the prior discussion. This is true because at the moment when we examine the actions of decision units, the systems in which they are embedded have already performed their functions, interacted with each other, and thereby generated the environmental conditions and information of which the decision unit has knowledge and on which it acts. In these areas, the experimental approach consists in seeking out instances in which the environment and its informational content differ significantly from other environments, or the decision units differ significantly from other decision units, so that the general application and fruitfulness of the theory may be examined. When, however, the behavior of decision units as understood on the basis of such an analysis is to be explored experimentally under changed assumptions as to policies and technology, an entirely new situation arises. We can no longer assume that various sets of decision-makers are independent of each other. Each reacts with the environment and creates changes which result in messages reaching other decision-makers. This interaction, which is irrelevant to some analyses, becomes critical in system simulation. I thus assume that system simulation and decision analysis interact strongly with each other and that each is necessary for the other. But as a matter of research emphasis, I would give short-term priority to system simulation on the grounds that relevant experiments to test our understanding of the behavior of decision units can probably not be performed without it.

Engineers and planners are vitally concerned with the behavior of households and business establishments in making use of the transportation system and in making locational decisions. Such behavior is the source of transportation demand. Private decisions in respect of automobile ownership, location, and new construction in the aggregate greatly influence the development of cities and regions. Finally, I am sure that if we understood thoroughly the whole constellation of decisions made by individuals and firms, we could understand at the same time the extent to which various urban arrangements satisfy their needs. Such an understanding is a vital key to producing plans and policies which will best serve the public interest.

Some of the differences between practicing planners and engineers can be traced to their different approaches to decision-makers' needs and preferences. The planner typically approaches the problem from the viewpoint of normative standards of behavior and social welfare. This is in part based upon notions of minimum socially acceptable levels of welfare and in part upon an emphasis on the externalities of individual behavior—that is, upon the effects of one's behavior on others. These notions are linked with strong ideas of social control. The engineering approach tends to be more adaptive. Individual behavior is regarded as being largely self motivated and not widely amenable to control. In dealing with supposed patterns of behavior as necessary inputs to engineering estimates, the engineer approaches the problem with inadequate concepts of motivation and of measurement. Neither planners nor engineers are in general well trained in the intricate issues of choice behavior, and present-day economics, sociology, and psychology offer little which is of general applicability to the problems which they face. The following remarks are therefore observations on a dilemma which will ultimately be resolved only by training people and developing methods which embody a combination of all of these disciplines in a new format.

The basic theory of choices by individual decision units deals in terms of alternatives and trade offs, yet if we examine transportation and locational theories or models, we find that these trade offs are remotely reflected, if indeed they may be presumed to have been considered seriously at all. Since the same thing is true of econometric models in related fields, this is not a particularly telling criticism in terms of past performance, but it is clear that it may constitute a barrier which will have to be removed before a great deal of progress can be made.

Much of the difficulty concerns observation and measurement, and perhaps this may best be illustrated with reference to the theory of industrial and commercial location. Industrial location in particular has long been very carefully studied by locational economists and regional scientists, and interregional locational theory is a particularly well developed field. In this location theory three factors are particularly important: internal economies of scale which depend on the size of establishment, external economies of scale, or agglomeration economies which depend on the sizes of the geographical assemblages of activities in which the establishment is located, and locational costs which depend mainly on the cost of land and the costs of interaction. In the complicated urban metropolitan scene these economic variables turn out to be very difficult to define, more difficult to measure, and still more difficult to value. While it may be well known, for example, that the garment industry has large agglomeration economies and is sensitive to its accessibility to a particular labor force and to the cost of industrial space, these variables and their relationships are not well defined. The interaction requirements of offices and the agglomeration economies of retail trade establishments are also imperfectly understood. While these ideas enlighten a good deal of research design, anyone who has tried to set up an industrial or commercial survey knows that it is very difficult to tie them down specifically. Because of this situation and for allied reasons, it is beginning to appear that in spite of the much more sophisticated work over many decades in industrial location, the problems of residential location are more tractable and amenable to sound solution.

In the area of consumer behavior, some difficulty is introduced by the fact that certain decisions are made by individuals, others by households, and still others by individuals in a household context. These difficulties must be faced in research design, but they are relatively minor compared with other more obvious problems. One which has been both recognized and ignored (often simultaneously) is that of aggregation. Some researchers, perhaps moved by data difficulties, are inclined to deal with the means and medians of areal aggregations of data. This method of work is almost enforced by the form of the availability of published census data in certain cases. There is clearly here a latent conception that area averages represent some sort of aggregation of behaviors, but the implicit rules of this aggregation are not explored and frequently the assumed behaviors are not fully defined. The gravity model of trip distribution clearly takes this approach at a descriptive level, while multiple regression models of modal [illegible] may be but a step closer to postulating real cause and effect. The Schneider model of trip distribution postulates more explicit behavioral patterns and works with areal aggregate data. In practice, however, this model [illegible] unexplained, variations in decision behavior because it requires an area-specific determination of the proportions of long and short trips. This specification amounts to a statement that the behavior in the model is incompletely defined.

Those who avoid the implicit assumptions of working with grouped area data by using individual or household observations encounter another level of difficulty which helps to elucidate the problem of aggregation. Behavioral models of individual and household choice invariably produce tests in which only a small part of choice behavior is adequately explained. Typical coefficients of determination are in the range of .15 to .30 While this may mean in some cases that the models and theories employed are inadequate, it is more likely to imply that behavior is influenced by more or less unobservable cultural and psychological factors which (at least at any one time) may be statistically distributed in the population The delicate problem of research design is to know when to stop trying to identify these factors and when to introduce assumptions about the statistical nature of their distribution in the population After the first recourse is exhausted or while it is being further developed, it is apparent that the nature of the assumptions about the statistical distributions of behavior around their observed statistical means may strongly affect the characteristics of their aggregation As a simple example, I have demonstrated elsewhere that if Schneider's L parameter is assumed to have a certain statistical distribution rather than being fixed, his model converts readily to a modified gravity model or to a combined model Certainly considerable statistical expertise will be required to explore this problem further

One of the more subtle and neglected aspects of the analysis of decision units is the role of the history of the unit in its behavior. To some extent, the history of certain units is implicit in their state description—for example, a family head aged 20 is probably recently married. But other and more subtle historical aspects may be overlooked. It is quite clear, for instance, that the history of industrial establishments is related to their tendencies to relocate and the ethnic background of many population groups is related to their choice of residence. It has even been reasonably suggested that consumer choice of mode of travel is related to the individual's history in learning to drive. These historical aspects of the behavior of decision units have two very important relationships with more general systems analysis The historical aspects of decisions are closely tied to the extent of the lags in movement toward system equilibrium, and only systems in which the history of decision units is unimportant will rapidly achieve equilibrium At the same time, the introduction of these histories is a means of dealing quite explicitly with trend data, without at the same time building into the theory an assumption that trends will indefinitely continue. It should be apparent that this historical approach does not lend itself to easy application to aggregate data, at least in analysis And if the histories to be considered become very complex, then Monte Carlo methods are almost required for any projection simulations

In the light of this necessarily incomplete review, we may justifiably conclude, I feel, that a theoretically sound and scientific approach to systems simulation of transportation and land use will require a great deal of rethinking of our theory of decision-units' behavior.

IMPLICATIONS FOR PLANNING

Let us now take a brief final view of the workaday implications of the type of program which I have sketched above The essential elements of this approach are six in number. First, since sound theory has so much to offer for practical progress, the work should be organized on a scientific rather than a mission-oriented or technological basis. We would thus also avoid the dangers implicit in harnessing these activities to suboptimal policies. We would rely on the social and policy motivations of the scientists to maintain a well directed drive toward ultimate application. Second, we would view these problems as related to certain real world systems and would deepen our efforts to achieve successful theories of the operation of those systems. Third, we would give appropriate recognition to the need for the study of the behavior of decision units in the context of larger systems which create their environment Fourth, we would give explicit recognition to the theoretical problem of communication between the systems, subsystems, and decision [illegible]. Fifth, we would recognize that the scope of these investigations will require the unification of parts of different disciplines in institutions and in individuals. Sixth, we would recognize that in a specially defined sense this work is experimental in the best traditions of experimental science and that the experimental method will require special conditions for success.

It would seem to me that the scale of these problems and their importance for long-term policy development tend to argue against scattered research in connection with specific projects. Such projects in any case tend to attempt to base their mission orientation on individual researchers. The resulting tension between the desire of the researcher to satisfy his scientific conscience and the desire of the management to get the job done sometimes borders on the tragic, or the comic. In any event, the problems are of general and national significance, and if worthy of consideration should not be charged to local or to special-purpose studies. It is also apparent that the variety of ability and knowledge required for an assault on these problems can rarely be assembled even in a large study of an ad hoc nature. Consequently, many such studies are repeating the work and perpetuating the errors of other studies for lack of resources to go further and try new methods. Finally, there are serious difficulties of communication within this scientific community which result from the excessive fragmentation of effort.

[illegible] attention needs to be devoted to the requirements of the experimental method in this field. Consider, for example, that we are designing a laboratory for social, engineering, and planning research. Our experimental material, instead of white mice, is extensive data about metropolitan areas and regions. These data must meet certain rigorous standards and be well organized and accessible. Our main experimental tool is probably the computer, but this will include the software or operating programs which embody many or most of the elementary processes of simulation and analysis which we have discovered. Our experimental design is a model or group of models based on theory and using our experimental material (data) and our experimental tools (computers and software). In any good experimental design, we are apt to discover that some special-purpose tools will have to be made—in this case, that new programs will have to be written and in some cases new data collected. The essential aim of an experiment will be to make critical tests of theories by good experimental design and thus to decide, for example, on a clear definition of the relative merits of the gravity model, the Schneider model, the Tomazinis model, and the Harris model of trip distribution. The essential ingredient for progress in addition to all the matters which I have so far discussed is quick turnaround so that experiments may be rapidly executed once they are designed. I would estimate that under current conditions, with practically no standing stock of data and widely diversified programs, the turnaround time on experimental work of this type is roughly three to five years. This time should be reduced by a factor of three or more, and the content of the experiments should be far more conclusive than it is today.

I believe that these standards, both of theoretical excellence and [illegible] performance, are achievable, and that if achieved they will have tremendous payoff in improved planning.

REFERENCES

Gerald A. P. Carrothers, "An Historical Review of the Gravity and Potential Concepts of Human Interaction," *Journal of the American Institute of Planners*, Vol. XXII (Spring, 1956).

Britton Harris, "A Note on the Probability of Interaction at a Distance," *Journal of Regional Science*, Vol. V, No. 2 ([illegible]).

Britton Harris (ed.), "Urban Development Models: New Tools for Planning," *Journal of the American Institute of Planners*, Special Issue, Vol. XXXI (May, 1965).

Ira S. Lowry, *A Model of Metropolis*, Memorandum RM-4035-RC (Santa Monica, California: The RAND Corporation, August, 1964).

Kevin Lynch, *The Image of the City* (Cambridge, Massachusetts: M.I.T. Press, 1960).

Henri Poincaré, *Science and Hypothesis* (New York: Dover Press, 1952).

Morton Schneider, "Gravity Models and Trip Distribution Theory," *Papers and Proceedings of the Regional Science Association*, Vol. V, 1959.

Anthony R. Tomazinis, "A New Trip Distribution Model," Paper #347 (Washington, D.C.: Highway Research Board, [illegible]).

Alvin Weinberg, "But is the Teacher Also a Citizen?" *Science*, Vol. 149 [illegible] (August 6, 1965).