

UC Berkeley

Earlier Faculty Research

Title

Modeling Land Use and Transportation: An Interpretive Review for Growth Areas

Permalink

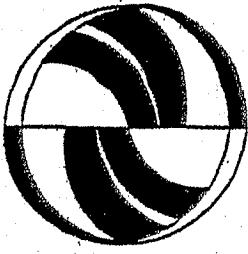
<https://escholarship.org/uc/item/4nw2t7n5>

Authors

Berechman, J.
Small, K. A.

Publication Date

1988



**Modeling Land Use and Transportation: An
Interpretive Review for Growth Areas**

J Berechman
K. A. Small

UCTC
No. 1

The University of California
Transportation Center
University of California
Berkeley, CA 94720

**The University of California
Transportation Center**

The University of California Transportation Center (UCTC) is one of ten regional units mandated by Congress and established in Fall 1988 to support research, education, and training in surface transportation. The UC Center serves federal Region IX and is supported by matching grants from the U.S. Department of Transportation, the California Department of Transportation (Caltrans), and the University.

Based on the Berkeley Campus, UCTC draws upon existing capabilities and resources of the Institutes of Transportation Studies at Berkeley, Davis, Irvine, and Los Angeles; the Institute of Urban and Regional Development at Berkeley; and several academic departments at the Berkeley, Davis, Irvine, and Los Angeles campuses. Faculty and students on other University of California campuses may participate in

Center activities. Researchers at other universities within the region also have opportunities to collaborate with UC faculty on selected studies.

UCTC's educational and research programs are focused on strategic planning for improving metropolitan accessibility, with emphasis on the special conditions in Region IX. Particular attention is directed to strategies for using transportation as an instrument of economic development, while also accommodating to the region's persistent expansion and while maintaining and enhancing the quality of life there.

The Center distributes reports on its research in working papers, monographs, and in reprints of published articles. It also publishes *Access*, a magazine presenting summaries of selected studies. For a list of publications in print, write to the address below.

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.



University of California
Transportation Center

108 Naval Architecture Building
Berkeley, California 94720
Tel: 510/643-7378
FAX: 510/643-5456

The contents of this report reflect the views of the author who is responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California or the U.S. Department of Transportation. This report does not constitute a standard, specification, or regulation.

Research policy and review 25. Modeling land use and transportation: an interpretive review for growth areas

J Berechman†, K A Small#

Institute of Transportation Studies and School of Social Sciences, University of California, Irvine, CA 92717, USA

Received 2 June 1987; in revised form 8 December 1987

Abstract. Urban growth is taking new forms in recently urbanized or formerly suburban areas, characterized by low density, heavy dependence on automobile transportation, and multiple activity centers. In order to understand better such 'contemporary urban areas', researchers need land-use models that realistically capture the key features of such areas and that can handle detailed data sets.

We review the literature on large-scale land-use modeling with this objective in mind. Characterizing the known models along several dimensions describing purpose, conceptual basis, mathematical content, and level of detail, we select models that are representative of the range of approaches taken. Six of these are reviewed in detail, and four others are discussed more briefly.

We find that the existing literature forces one to choose between tractability and suitability for contemporary urban areas. The key omission in the tractable models is economies of agglomeration that would help explain the emergence of subcenters. Most tractable models also lack a dynamic structure suitable for handling rapid disequilibrium growth. Models that contain these two features are suitable for broad-brush computer simulation, but they cannot be calibrated with real disaggregated land-use data. This conclusion leads to some brief suggestions on directions for future work.

1 Introduction

1.1 *Contemporary urban areas*

For many decades, the focus of population and employment growth in the United States has been shifting to its suburbs. Traditionally defined as those parts of metropolitan areas outside central cities, suburbs now account for some 60% of the national population and have greatly increased their share of metropolitan employment.

A somewhat newer development is the increasing economic independence of these areas. Areas formerly thought of as suburban now contain major employment centers, transit systems, high-rise offices, cultural institutions, and other features of city life. The twentieth largest metropolitan area in the United States, consisting of Orange County in California, has no dominant city, yet is no longer a collection of Los Angeles suburbs: its diverse economy provides 84 jobs for every 100 resident labor-force members. The economic independence of Long Island from adjacent New York City has long been recognized by its unique early designation as a metropolitan area lacking even an officially designated central city. The rapid growth of such 'contemporary urban areas' guarantees their increasing importance in the national urban system.

These rapidly growing areas are developing their own patterns of agglomeration and centrality, which are very different from the patterns that inspired the land-use models familiar to urban planners. The newer metropolitan areas are characterized

† Current address: Center for Urban and Regional Studies, Tel-Aviv University, Ramat-Aviv, Tel-Aviv 69978, Israel.

Current address: Department of Economics, University of California, Irvine, CA 92717, USA

by low density, highway orientation, rapid in-migration, many similarly sized activity centers, and significant amounts of undeveloped land close to those centers. This suggests that the underlying determinants of land use are different from those in traditional metropolitan areas. Hence different models will be needed to elicit their nature.

Another characteristic of rapidly growing areas is the prominence of transportation policy as a political issue. In such areas, transportation facilities, mainly highways, are important, expensive, and long lasting. They have a strong and easily observed influence on residents. Furthermore, decisions regarding transportation must often be made well before the shape of the urban area and the nature of its activities are fully known.

It is therefore useful to reconsider the literature on land-use modeling in light of the special needs inherent in studying contemporary urban areas. In the present paper we do so with the ultimate goal of developing models that use empirical data and are capable of analyzing alternative public policies. Hence we are most interested

Table 1. Classification scheme for urban land-use models.

- A Behavioral basis** Most models contain an explicit or implicit theory of what causes land to develop the way it does, such as:
- 1 gravity or entropy,
 - 2 microeconomic—deterministic,
 - 3 microeconomic—stochastic,
 - 4 evolutionary (biological analogy).
- B Time scale** In each model an assumption is made about the length of time over which its solution holds. The time scale may be:
- 1 static (describes a point in time),
 - 2 dynamic (describes change over time),
 - 3 iterative with no real calendric time.
- C Spatial configuration** In the models the actual two-dimensional land surface is simplified in various ways, including:
- 1 one dimensional (a long narrow city),
 - 2 circular (monocentric),
 - 3 rectangular grid,
 - 4 discrete zones.
- D Endogenous sectors** The size, distribution, or properties of one or more of the following sectors may be endogenous, that is, determined within the model itself:
- 1 basic employment,
 - 2 nonbasic employment,
 - 3 housing type,
 - 4 transportation system.
- E Externalities** Of the various externalities that pervade urban areas, two are especially important: congestion on the transportation network, and economies of agglomeration for firms. The model may include either or both of:
- 1 congestion,
 - 2 agglomeration.
- F Solution method** Models are mathematical constructs that mimic reality. In order to produce results, the mathematical equations must be solved. Methods include:
- 1 ad hoc iteration,
 - 2 mathematical programming,
 - 3 stochastic simulation.
- G Applications** The possible responses as to whether the model was applied to a specific urban area are:
- 1 yes,
 - 2 no,
 - 3 partially.
-

in models that meet the following criteria: (a) elucidation of one or more of the characteristics of growing areas noted above; (b) explicit attention to transportation networks, especially highway networks; (c) tractability when confronted with empirical data; and (d) ability to predict the consequences of alternative policies, especially transportation policies.

We do not attempt a comprehensive review of the enormous literature on land-use modeling, such as those by Anas (1988), Batty (1972; 1976), Harris (1985a; 1985b), and Mackett (1985). Instead, our approach takes two steps. First, we briefly discuss some main features of the literature, describing the range of possible approaches by classifying models along several dimensions. Second, we select for detailed review a few specific models that illustrate the wide range of approaches that have been tried. These are selected more for their importance as prototypes than for their historical importance or immediate applicability. Each is described, evaluated in general, then assessed for our particular purpose.

1.2 Varieties of urban land-use models

The literature on urban land-use models is astonishingly heterogeneous. Models are built for purposes as varied as basic scientific understanding, optimization of zoning rules, or detailed forecasting. Their underlying theoretical bases range from entropy to microeconomics. Mathematical tools range from numerical simulation to dynamic control theory. Cities are described as circles, lines, rectangular grids, or matrices of zones. Models may portray a single point in time or changes over time. They may or may not account for employment location, residential location, housing tenure and type, automobile ownership and mode to work, land markets, new development, housing rehabilitation, and highway congestion.

In order to organize ideas, we list in table 1 several dimensions along which urban land-use models may be classified. Of course, most authors have not cooperated with this campaign for neatness, so there is plenty of room for argument. Nevertheless, we find it a useful way to think about the literature, and an aid in selecting a small number of models that are representative.

In table 2 we display our judgments for six models that we have selected for detailed review, plus four that are reviewed more briefly. The basis for our judgments should become apparent in the sections that follow. At this point, we note that nearly every possibility in table 1 is included in at least one of the models reviewed.

Table 2. Classification of selected models.

	A	B	C	D	E	F	G
<i>Gravity</i>							
Garin-Lowry	1	3	4	2	none	1	yes
Putman	1	3	4	2,4	1	1	yes
<i>Deterministic</i>							
Herbert-Stevens	2	3	4	3	none	2	yes
Mills-Kim	2	1	3	1,2,3,4	none	2	no
<i>Economic equilibrium with dispersion</i>							
Anas (1984)	3	1	4	1,2,3,4	1	2	no
CATLAS	3	2	4	3	none	1	yes
<i>Agglomeration economies</i>							
Carruthers	3	2	1	2	2	3	no
Wilson	1	2	4	2	2	1	no
Allen-Sanglier	4	2	3	1,2	2	3	partially
Fujita-Ogawa	2	1	1	1,2	2	3	no

Note: The column headings and numerical entries are those appearing in table 1.

2 Gravity-type models

2.1 *The modified Garin-Lowry model*

2.1.1 *History* Of all the land-use models developed to date, the Garin-Lowry family of models is the most well known and used. It was first developed by Lowry (1964) in order to simulate spatial patterns of residential and service development in Pittsburgh. The model was revised and extended in many ways, and it was applied to various cities and regions in the USA and elsewhere. It has inspired several quite different models, including those by Putman and Anas that are reviewed later in this paper.

In his original model Lowry assumed as exogenous the level and location of basic employment, that is, of jobs involving production of goods and services sold outside the city or region. The level of population and its dependent service employment were then computed through the use of an economic-base model; and their spatial distributions were obtained through interaction functions from which the name 'gravity model' is derived by analogy. Constraints on zonal activity densities were also included. Thus, policy simulations could be carried out by specifying a new pattern of basic employment, or by changing various parameters.

Subsequent extensions were introduced by Crecine (1964) and Goldner (1968). In Crecine's Time Oriented Metropolitan Model (TOMM), the population was disaggregated by socioeconomic status, and the time elapsed between the base and forecast years was accounted for explicitly. In Goldner's Projective Land Use Model (PLUM), Lowry's gravity functions were replaced by intervening-opportunity functions, and zone-specific activity rates and population-serving ratios were used.

A major reformulation by Garin (1966) improved the original model in several ways. Garin explicitly incorporated interaction submodels (containing the gravity formulae) that distributed *all* activities at each iteration of the calculation. Garin also cast the entire model in matrix notation, thereby simplifying the precise description of the model and exposing the underlying equilibrium inherent in the iterative solution procedure. Garin's formulation, with Lowry's zonal density constraints reintroduced, is known as the 'modified Garin-Lowry model' (Batty, 1972) and is the one we describe below.

2.1.2 *Structure* A flow chart of the Garin-Lowry model, adapted from Batty (1976), is depicted in figure 1 to demonstrate the way in which the economic-base mechanism, the spatial-allocation submodels, and the constraint procedure interact. In a typical application, the model is first 'calibrated' by adjusting its parameters so as to reproduce as closely as possible an existing urban area; it is then used to simulate the impact of new basic-employment forecasts or of policy changes.

The input data include zonal levels of basic employment, interzonal travel-cost matrices for home-to-work and home-to-shopping trips, zonal levels of attractiveness for residential and service location, and control parameters of the economic-base mechanism. Based on these inputs, first, the workers in the basic sector are allocated to residential zones. The incremental residential population and the resulting incremental dependent service employment are then calculated. This increment of employment is distributed to zones of workplace, and a corresponding increment in population is derived and distributed spatially to zones of residential location. This entire iterative process continues until the economic-base mechanism converges.

At each iteration, a check is made to ascertain that zonal densities of service employment and residential population are within preset bounds. If not, an iterative procedure (internal to the economic-base iterations) is used to reallocate the latest increments by changing the zonal attraction parameters.

The output from the model includes vectors of residential population and household-dependent employment, trip tables for work and shopping trips, vectors of residential-attractor and service-attractor weights, and travel parameters.

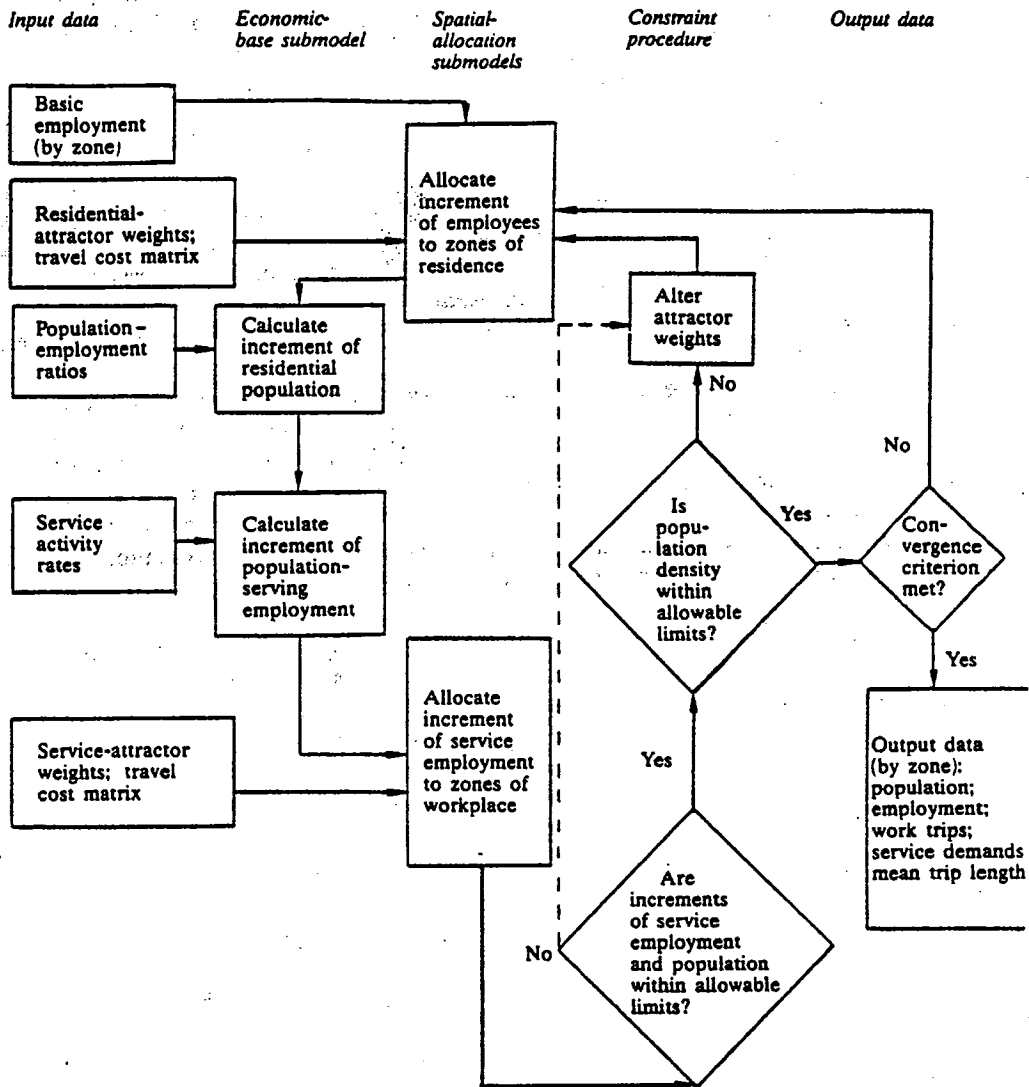


Figure 1. The modified Garin-Lowry model [source: adapted from Batty (1976)].

2.1.3 Evaluation The Garin-Lowry model is operational. It has been used successfully in replicating observed spatial distributions of land-use activities, and in analyzing the impacts of major regional changes. Data requirements are modest. Calibration amounts to adjusting only a few parameters, mainly those of the gravity functions and zonal attractiveness weights. The estimation of these parameters usually involves a nonlinear search procedure for the gravity-model parameters, and an ad hoc search algorithm to meet the requirements of the constraints. The search procedure for the gravity-model parameters has been shown to have a solution (Batty, 1976), but a solution to the algorithm to meet the constraint requirements may or may not exist. Once the parameters have been estimated, simulating the impact of an external change such as a new transportation facility is,

in principle, straightforward; for example, see Foot (1981). Temporal and spatial transferability of the model may, however, require extensive recalibration.

Because the model has a simple structure, activities may be further disaggregated into classes with only modest amounts of additional information. Likewise, with only minor changes in specification, one can include zone-specific characteristics such as commercial floor space, or a more detailed representation of the transportation network (Geraldes et al, 1978; Turner, 1975).

Another strength is that the spatial-allocation mechanism within the model can be shown to mimic the result of randomness in the decisions of individuals (Williams, 1977). Indeed, in several of the papers by Anas, as discussed later, this strength is capitalized upon by formulating stochastic-choice models of individual behavior in order to provide microeconomic interpretations of the gravity functions. Hence the Lowry model is able to represent the dispersion in locational decisions that seems to characterize real cities, and that is lost in the deterministic models discussed in the next section.

The main disadvantage of the Garin-Lowry model is its lack of any underlying economic or behavioral theory. It includes no supply side for urban development, and no equilibrating prices. These omissions mean that the housing industry and the industries supplying industrial and commercial development are entirely ignored, and thereby analyses of such important factors as tax policy, mortgage rates, housing deterioration, vacancies, and abandonment are precluded. Moreover, without prices it is impossible to investigate normative properties such as welfare gains or conditions for optimality.

This lack of economic or behavioral theory presents an especially serious problem in the allocation submodels. The exogenous variables, travel costs and site characteristics, omit such important factors as income, technology, land prices, neighborhood externalities, and agglomeration economies.

A related problem is that the employment multiplier has no calendric time dimension, so it is impossible to predict the pace at which changes will occur. The iterative solution mechanism unfortunately lends itself to being misinterpreted as depicting a dynamic process, but it does not.

Other problems with the model are widely discussed in the literature and need not be elaborated here. These include the definition of basic employment, the use of zones of unequal size or shape, the neglect of population and employment outside the boundaries of the region, the assumed exogeneity of certain coefficients, and the possible nonconvergence of the solution algorithm.

2.1.4 Suitability The Garin-Lowry model has been applied to study a number of policy issues relevant to fast-growing regions. These include new transportation facilities, increased labor-force participation, changes in residential attractiveness, and changes in zoning regulations. However, the analysis of transportation investments is hampered by the inability of the model to account for endogenous congestion and route choice, or to provide a basis for measuring the value of benefits produced.

The existence of multiple activity centers in contemporary urban areas suggests strong agglomeration economies, which are not accounted for by the model. Indeed, the model takes as given the number and spatial distribution of basic-sector jobs, one of the very features of a rapidly growing area that one would like to explain.

For these reasons, we view the Garin-Lowry model as an unsatisfactory starting point for any study whose primary goal is to elucidate those features most characteristic of contemporary urban areas.

2.2 Putman: Integrated Transportation and Land-use Model Package (ITLUP)

2.2.1 History This model combines two separate components: a land-use model, and a transportation network model. As originally constructed by Putman (1973; 1974), each component was a modification of a previously existing model. The land-use component was IPLUM, a modification of Goldner's version of the Garin-Lowry model of land use; and the network component was a conventional capacity-restrained incremental-assignment model of a transportation network.

In the first version of ITLUP, a preliminary allocation of land-use activities was used to produce trip matrices; the resulting travel times on the (possibly congested) transportation network were then fed back into the land-use model component to produce new activity distributions. The entire process was redone iteratively until it converged.

In a later project, the land-use model component was revised by improving calibration techniques and by modifying the spatial-allocation formulae. The revised land-use component was called Disaggregated Residential Allocation Model (DRAM). The overall package is documented in Putman (1983).

2.2.2 Structure The general scheme is shown in figure 2. Base-year data include the spatial distribution of all employment and residential activities, including characteristics of households by zone, and characteristics of the uncongested base-year highway network. These are entered into the land-use and network models in

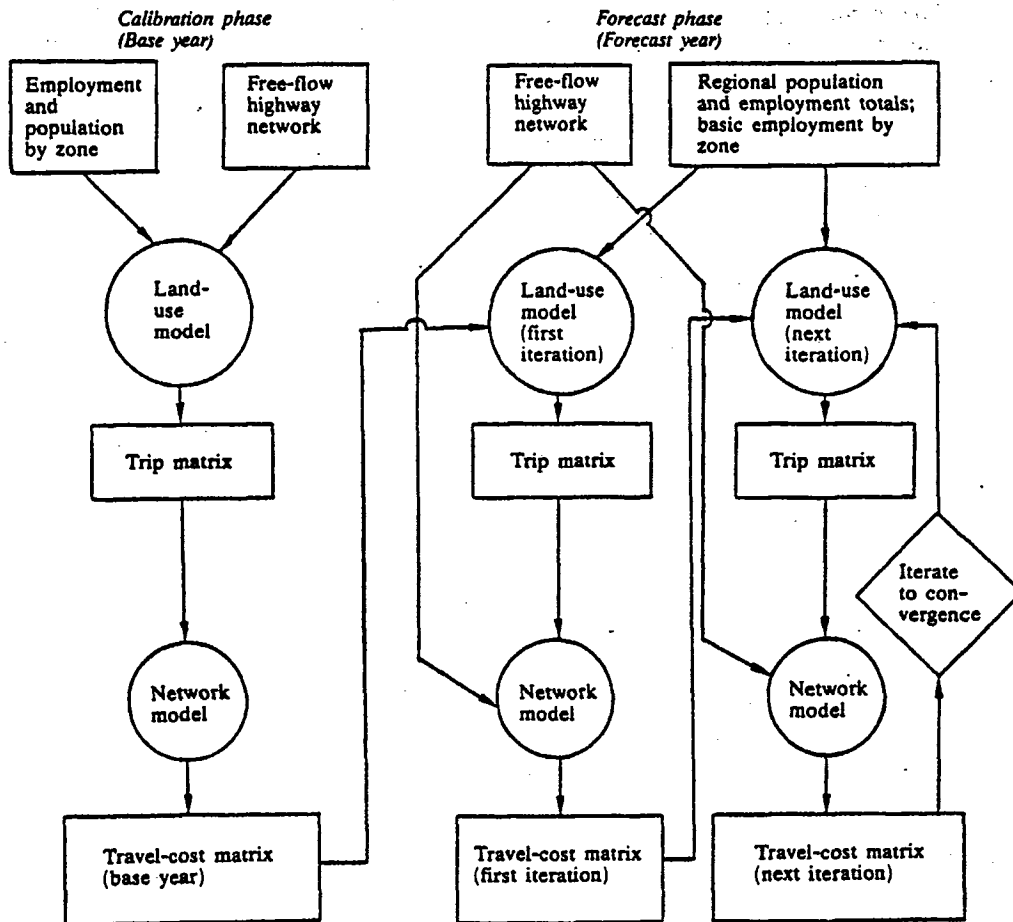


Figure 2. The ITLUP model [source: adapted from Putman (1975)].

order to produce an estimate of the base-year travel costs prevailing on the congested highway network.

For a forecast year, it is only basic employment for which data on spatial distribution must be supplied; for population and other employment, regional totals constitute the only data requirements. Each iteration then includes one iteration of the land-use model (producing a trip matrix), followed by an application of the network-assignment model (producing a travel-cost matrix). In the first iteration, the base-year travel-cost matrix for the congested network serves as the required input to the land-use model. Iterations continue until the distribution of activities stops changing.

2.2.3 Evaluation The greatest advantage of ITLUP is its explicit attention to the transportation network. Simulations take account of highway congestion and the resulting reallocation of activities. This is especially important, of course, in analyzing changes in the highway infrastructure itself; but the direct linkage of the two model components provides a fuller explanation of many urban phenomena. Putman's model is probably the first deserving to be called a transportation/land-use model.

Except for this feature, ITLUP shares most of the disadvantages of the Garin-Lowry model. It contains neither demand nor supply functions, nor a price mechanism for achieving market equilibrium. Nor does it portray changes in real time. In addition, there is some doubt about the convergence properties of the model. Berechman (1981) has shown that the iterative process may converge to something other than the general equilibrium. Berechman and Gordon (1986) show further that this problem is a general one for linked models that do not equilibrate demand and supply for both land and transportation. As a result, the solution may depend upon the particular iterative procedure used. In addition, in the network-assignment procedure all trips are assumed to be by auto, and simultaneous loading of all trips is not permitted, rather a portion of the trips is loaded at each iteration.

2.2.4 Suitability ITLUP shares with other Lowry-derivative models the disadvantages noted earlier for representing rapidly growing areas. However, it is a definite improvement because it permits consistent treatment of traffic congestion.

3 Deterministic economic equilibrium

3.1 *The Herbert-Stevens model*

3.1.1 History The modern economic theory of urban land use began with the work of Wingo (1961) and Alonso (1964). They focused attention on the role of land markets in residential location, asserting that households trade off higher site cost against lower commuting costs; whereas landowners rent (or sell) to the highest bidder. The demand side of the market can be characterized thus: given its (exogenous) workplace location, a particular household has a 'bid-rent function' which describes the most it could pay to live at each possible location and still achieve a given level of satisfaction (utility). The supply side is simply that each location is rented to the highest bidder. Equilibrium occurs when all households of a given type are equally well off, and their levels of utility have adjusted so that every household occupies exactly one site.

Much of the literature since Wingo (1961) and Alonso (1964) is not planning oriented, and such radical assumptions as monocentricity are made. Nevertheless, the Wingo-Alonso theory has also sparked major innovations in empirical land-use models of urban location and structure. The earliest was the model by Herbert and Stevens (1960), developed as part of the Penn-Jersey Transportation Study of metropolitan Philadelphia. Harris (1963) and Wheaton (1974) cleared up a

conceptual confusion about the equilibrium level of household utility underlying the bid-rent functions. Subsequently, the Herbert-Stevens model has influenced quite different modeling efforts including those by Mills, Boyce, Los, Kim, and Anas described later in this paper, as well as the National Bureau of Economic Research model (Ingram et al, 1972).

3.1.2 *Structure* The model is presented here as in Herbert and Stevens (1960). Suppose there are:

- U areas which form an exhaustive subdivision of the region, indicated by the superscripts $K = 1, \dots, U$;
- n household types, indicated by subscripts $i = 1, \dots, n$;
- N_i households of type i ;
- m residential bundles (each described by observable characteristics of a site, house, lot, and set of trips including work trips), indicated by subscripts $h = 1, \dots, m$.

We make the following definitions:

- b_{ih} is the bid-rent by a household of type i for residential bundle h ;
 - c_{ih}^K is the annual cost to a type i household of the residential bundle h in area K , exclusive of site cost, that is, it includes costs of travel and of construction and maintenance of the building;
 - s_h is the lot size included in the residential bundle h ;
 - L^K is the number of acres of land available for residential use in area K .
- The endogenous variables are X_{ih}^K , the number of households of type i choosing residential bundle h in area K .

The programming problem is to maximize aggregate site rents paid by households, given total land availability and the need to accommodate the entire population:

$$\text{maximize } \sum_{K=1}^U \sum_{i=1}^n \sum_{h=1}^m X_{ih}^K (b_{ih} - c_{ih}^K),$$

subject to

$$\sum_{i=1}^n \sum_{h=1}^m s_h X_{ih}^K \leq L^K, \quad K = 1, \dots, U,$$

$$\sum_{K=1}^U \sum_{h=1}^m X_{ih}^K = N_i, \quad i = 1, \dots, n,$$

$$X_{ih}^K \geq 0, \quad K = 1, \dots, U, \quad i = 1, \dots, n, \quad h = 1, \dots, m.$$

The solution generates Lagrangian multipliers that can be interpreted as shadow prices:

- r^K the annual rent per unit of land in area K ;
- v_i an annual 'subsidy' for each household of type i .

The first-order conditions are then readily interpreted as equilibrium conditions for households and landlords. Wheaton (1974) subsequently clarified the interpretation of v_i by noting that in an area with fixed population, utility levels (and hence bid-rents) would adjust until, in equilibrium, all the v_i values were equal to zero. He therefore proposed an additional iterative loop seeking that condition.

3.1.3 *Evaluation* The Herbert-Stevens model was the first based on economic principles to be applied to data for a real metropolitan area. This theoretical foundation of the model is a major strength, as it is through this that residential price determination is explained and results with economic interpretations are given. Another important advantage is the use of linear programming, a well-understood technique with computational algorithms available even for very large problems. However, only the residential sector is determined. Furthermore, the

representation of an urban residential land market as a linear problem implies that no spatial or other types of externalities exist; as emphasized by Harris (1985b), this greatly detracts from its realism. Another drawback is that in the model the supply of land is assumed to be completely inelastic; hence speculatively held vacant land cannot be accounted for. In addition, the required data on housing characteristics, shapes of bid-rent curves, lot sizes, and construction costs are hard to obtain.

The nature of time is somewhat unclear. The model seems most naturally suited to a description of a long-run static equilibrium in the housing market. However, Herbert and Stevens describe it as applying to one iterative period in which a prespecified amount of population is to be allocated to a prespecified allotment of newly available land (Herbert and Stevens, 1960, page 22). This seems to require embedding the model in some unspecified larger model of urban growth. Another input that must be derived from some external model is the set of destinations in the 'trip set' that constitutes part of a residential bundle.

Two additional limitations have been noted, and their modifications have been proposed. Both are aptly described by Los (1979), who presents a comprehensive extended model. First, in the deterministic framework of Herbert-Stevens,

"no account is taken of the possible dispersion of preferences among households of a given type. The linear-programming formulation can generate an allocation of households to dwellings where some types of households can be completely absent from some zones or some housing types. This is too extreme a representation of actual household behavior." (Los, 1979, page 1248.)

The remedy, as indicated by Senior and Wilson (1974), is to add to the objective function a term that produces dispersion in the allocation of households of a given type:

$$-\left(\frac{1}{\mu}\right) \sum_{ihk} X_{ih}^k \ln X_{ih}^k,$$

where \ln is the natural logarithm, and μ is a parameter related to the amount of dispersion in household preferences. They interpret this term as the negative of entropy, though, as we shall see below in reviewing Anas's work, it has a behavioral counterpart in a stochastic model of individual household preferences.

Second, in the Herbert-Stevens model the transportation sector is taken as exogenous. This is remedied by Boyce (1978) and Los (1979), who add a transportation network to Senior and Wilson's version of the Herbert-Stevens model. It is fully integrated into the land-use portion of the model by including as separate variables the link-specific traffic flows, and by adding to the objective function terms that involve generalized travel costs as functions of these variables. The Kuhn-Tucker conditions for a solution then include Wardrop's condition for user equilibrium on a congested network.

These additions result in a nonlinear programming problem for which there exist practical algorithms assuring convergence, such as that by Evans (1976). In recent work by Boyce et al (1983), Chon et al (1983), and Boyce and Lundqvist (1986) the model has been applied to Chicago and Stockholm.

3.1.4 Suitability On the basis of the above discussion, it seems doubtful that the Herbert-Stevens model, in its original or extended form, is suitable for the analysis of a rapidly growing region. In such a region the housing market is strongly influenced by expectations of capital gains, speculation, credit availability, and physical or zoning constraints on new development. With significant growth occurring within a time span far shorter than the life of buildings, one cannot

expect such an area to be in long-run equilibrium. Finally, as noted earlier, externalities such as economies of agglomeration are particularly important for growing areas and cannot be included within the linear cost structure of the model.

3.2 *Mills-Kim: linear programming models of efficient cities*

3.2.1 *History* During the 1970s, Mills (1972; 1974a; 1974b; 1976) published a series of papers in which he attempted to operationalize the pioneering work by Wingo, Alonso, Muth (1969), and himself on what has come to be known as the 'new urban economics'. His goal was to show "how models can be constructed that not only are solvable when formulated in realistic detail, but also have both market and planning interpretations" (Mills, 1972, page 101). To this end, he developed a linear programming framework for representing an urban area with a single shipping point for export goods. As in the case of the new urban economics, this work gives an important role to land prices and to capital-land substitutability in the production of buildings. It also imposes a high degree of geographical symmetry, namely that the city consists of four quadrants that are mirror images of each other. The linear programming approach attempts to bypass the severe restrictions typically imposed on continuous-space models in order to achieve analytic tractability.

Mills also wanted "to shed light on ... one of the most important sources of inefficiency in urban areas, namely improper pricing and resource allocation in ... transportation" (1972, page 101). However, his model is fully elaborated only for efficient cities. Mills does use the model to argue theoretically that equilibrium cities would be economically efficient were transportation priced at marginal social cost (Mills, 1976); but a fully fledged model of an equilibrium city in the presence of underpriced transportation is only hinted at.

Ripper and Varaiya (1974) modified the model to allow export through the periphery and also worked out a dynamic version. Through numerical simulation, they found that the optimal solution involved considerably higher rents and more congestion in the dynamic version than in the static version.

Hartwick and Hartwick (1974) extended the model of Mills to remove the requirements of monocentricity and symmetry. By expanding the number of flows explicitly accounted for in the equations, they allowed an arbitrary number of export nodes to be specified in advance. Subcenters may then appear around these nodes in the solution. Hartwick and Hartwick also allowed for intermediate goods in production and accounted for their transportation within the city.

Kim (1979) extended the model still further by allowing for several transportation modes, as characterized by different combinations of fixed and variable cost. In Kim's solution procedure, the optimal transportation network for one of the modes (called 'subway') is determined endogenously along with optimal land use and production technology. The subway network is constrained to be continuous, but otherwise it can take on any shape; there seems no reason why it could not be renamed 'intraurban expressways network' and its parameters set so as to generate an optimal expressway system, superimposed on the regular street grid that constitutes the ubiquitous transportation mode.

3.2.2 *Structure* We describe the version by Kim (1979). Space is divided into squares of a specified width, usually taken to be one mile. Each square is identified by two integer coordinates (i, j) . Kim, like Mills, conserves on numbers of variables and equations by requiring the four quadrants of the city to be mirror images of each other. Square $(1, 1)$ of each quadrant is designated as the primary export node, which in the solution becomes part of the central business district (CBD).

Other export nodes can be specified exogenously, along with the unit cost of shipping a given export commodity from that node. Hence, for example, cost parameters could be chosen so that the CBD became the primary node for exporting services, and a suburban node (representing, perhaps, an access point to an interstate highway network) became a node for exporting manufactures. Each square has an exogenously specified amount of land available for development.

In the following, Kim's notation is adapted slightly to correct some ambiguities. There are \bar{F} produced commodities, which may be consumed, exported, or used as intermediate goods. Commodity \bar{F} is housing. These commodities are produced from each other and from two primary inputs, designated $\bar{F}+1$ (land), and $\bar{F}+2$ (capital). Each of the \bar{F} commodities can be produced by a variety of linear processes or 'activities', designated $s = 1, \dots, \bar{s}$, in which inputs are combined according to fixed input-output coefficients. Different activities have different coefficients, hence land and capital are used in different proportions; for ease of discussion, Kim follows Mills's protocol of interpreting each activity as production in a building of a particular height.

Available land, as already noted, is specified exogenously for each square; it has opportunity cost (presumably in agriculture) of R_s per acre per year. Capital is supplied elastically at a fixed rental rate R .

There are three modes of transportation. Local streets ($m = 1$) are available for transporting all commodities including commuters, whereas express bus ($m = 2$) and subway ($m = 3$) are available only for commuting. Commuting is interpreted as the commodity flow associated with the use of produced good \bar{F} (housing) as an input in the production of other goods. Nonwork trips are not represented.

Other exogenous variables include the amount of transportation service required to ship a unit of each commodity; transportation user-cost coefficients; and total exports of each commodity. Endogenous variables, computed as part of the solution, include output of commodity r by activity s in square (i, j) ; exports of commodity r from each designated export node; commodity flows across each boundary between squares; and units of transportation services produced within each square, by mode and commodity.

The solution is obtained through the minimization of an objective function with three components: cost of production, cost of transportation, and cost of exporting. The first is the opportunity cost of the land and capital used in production, added over commodities r , activities s , and squares (i, j) . The second is the opportunity cost of land and capital, plus the user costs, incurred in producing transportation services. The third is proportional to the volume of exports at each export node.

This objective function is minimized subject to five types of constraints. First, exports of each commodity from the various export nodes must add to the exogenously given total exports of that commodity. Second, commodity flows across the boundaries of each square must balance, taking into account any production within the square or export from that square. Third, the transportation services supplied must be sufficient to handle the commodity flows. Fourth, land used in various production processes (including transportation and housing) within each square must not exceed the amount available. Fifth, the subway system must be continuous.

A feature of the original formulation of Mills, which was dropped by later authors, was a series of discrete congestion levels and associated cost parameters that allowed the efficient solution to involve varying degrees of congestion. This is the only feature that would cause the efficient city (the solution to the problem just described) to differ from a long-run equilibrium city in which all products, including transportation services, are produced by competitive enterprises. Kim does not

include congestion in the model reviewed here, although he does so in a later and somewhat different model (Kim, 1983).

3.2.3 Evaluation A major advantage of this series of models is the attention to opportunity costs and the corresponding interpretation of dual parameters, which appear in the solution, as shadow prices. Kim's version allows flexible geographic shapes, and direct planning of efficient commuter-transportation systems.

The major difference between this and other cost-minimization models, such as Herbert-Stevens, is that the simulations are numerical examples rather than representations of actual urban areas. Kim's model is not intended to be calibrated in a precise way with detailed data from an actual city; instead, the investigator has great latitude in defining exogenous parameters so as to produce a city that looks 'realistic'. This has its dangers: given enough patience and cleverness, one may produce a simulation bearing uncanny resemblance to a real city, yet having poor predictive power and no normative significance. Furthermore, as Hartwick and Hartwick discovered, solutions may not always be unique. Nevertheless, simulation does offer a way of increasing our understanding of complex systems. For example, by generating such cities with Kim's model, one might learn a good deal about the merits of subways connecting subcenters.

On the negative side, the model is of an efficient city, not an actual one. Even if modified along the lines suggested by Mills, the model is strictly long run; no attention is given to longevity of capital and hence no insight is offered into the dynamics of urban change. Furthermore, this model shares with Herbert-Stevens and most of the new urban economics a cost-minimizing determinism that neglects important sources of dispersion. For example, using the model, one would never predict that two objectively identical commuters would travel in opposite directions on the same highway segment; yet Hamilton (1982) shows that such cross-hauling is the dominant feature of real commuting patterns.

The model appears difficult, though not impossible, to adapt to realistic geographical features such as harbors, mountains, and rivers. There is no conceptual barrier, but breaking the symmetry would add greatly to the number and complexity of variables, equations, and constraints. A similar comment applies to putting the treatment of congestion back into the model.

Another limitation is that export nodes must be specified in advance, so the model cannot predict the degree of multicentricity that will (or should) develop.

3.2.4 Suitability The disadvantages noted above seem particularly damaging to the use of these models for the analysis of new and rapidly growing urban areas. Such areas are characterized by changes on a time scale much shorter than typical lifetimes of buildings; hence the assumption of long-run equilibrium is severely limiting. Such areas are also characterized by great dispersion of origins and destinations, long commutes, and cross-hauling; these features cannot be explained by deterministic cost-minimization.

Kim's focus on multicentric development is certainly attractive for rapidly growing areas. But, because the export nodes must be specified in advance, the model cannot be used to illuminate the process by which the pattern of subcenters is determined.

The ability to build an arbitrary transportation network could be extremely useful if, as suggested above, it could represent a system of expressways. (This might require changing the rules to allow this mode to serve commodities as well as commuters.) It should not be difficult to add constraints representing expressway links that already exist. This would offer an opportunity not only for planning future highway systems, but for studying whether recent growth of the network was efficient.

4 Economic equilibrium with dispersion

4.1 *Anas: a random-utility general-equilibrium model*

4.1.1 *History* Anas (1984) describes a complete general-equilibrium model including land use, building type, employment and residential location, and traffic flow on a congestible highway network. The model unites no less than five major strains in the literature on urban form and transportation, several of which we have already discussed.

First is the spatial-interaction framework of Lowry, in which travel between pairs of zones is represented explicitly, and in which location decisions depend, among other things, on the generalized cost of such travel between a given zone and all other zones. Anas considers this aspect so central that the phrase "Lowry-type model" appears in his title. Second and equally important is the general economic equilibrium framework, attributed to Mills (1967; 1972), in which land and building rents adjust to ensure equilibrium, and equilibrium conditions are generated either from behavioral principles or from the solution to a mathematical programming problem.

The other three strains of literature may be viewed as techniques for generating specific parts of the model. The microeconomic representation of diversity in the choices of an observationally homogeneous group of firms or households is that of econometric models of discrete choice, as expounded by McFadden (1973). The macroscopic representation of the same phenomenon, appearing as terms in the objective function whose maximization yields the equilibrium conditions, is interpreted as entropy, as expounded by Wilson (1967). Finally, the term in the objective function used to generate a travel-network equilibrium comes from Beckmann et al (1956), who in turn were applying a principle of Wardrop (1952).

Several authors had already synthesized smaller combinations of these five approaches. In combined trip-distribution and traffic-assignment models, such as by Evans (1976) and Florian et al (1975), both the entropy and the network-equilibrium paradigms are used. Stochastic traffic assignment, as in Daganzo and Sheffi (1977), combines discrete choice and network equilibrium. Williams (1977) demonstrated the equivalence between logit models of discrete choice and entropy maximization. Anas himself (1980; 1982; 1983) developed models of housing markets that combined discrete choice with general economic equilibrium. Curiously, Anas fails to mention the triple combination of general economic equilibrium, entropy, and network equilibrium that was developed by Los (1979), and in a rather different form by Kim (1983).

The full Anas model is extremely ambitious and has never been applied empirically.

4.1.2 *Structure* The model consists of a fixed number of firms in each of two industries (basic and service), and a fixed number of employed residents. These actors make six choices, each denoted by a different subscript:

- i is the spatial zone of employment ($i = 1, \dots, I$),
 - j is the spatial zone of residence ($j = 1, \dots, I$),
 - k is the spatial zone of shopping ($k = 1, \dots, I$),
 - n is building type (those used by firms are $n = 1, \dots, N$; those used by residents are $n = N + 1, \dots, M$),
 - ρ_1 is the travel route between zones i and j for work trips (peak period),
 - ρ_2 is the travel route between zones i and k for shopping trips (off-peak).
- The travel network and times of day are exogenous:
- h is highways link ($h = 1, \dots, H$),
 - t is time period ($t = 1$ peak, $t = 2$ off-peak).

The price of the export good, produced by the basic industry, is fixed. All other prices are endogenously determined as part of the solution. These prices are:

- w_{li} wage in industry l in zone i ($2I$ prices),
- p_{2i} price of service commodity produced in zone i (I prices),
- R_{ni} rental price for building of type n in zone i (MI prices),
- c_{th} travel cost on link h during time period t ($2H$ prices),
- τ_{th} travel time on link h during time period t ($2H$ prices).

Hence the total number of unknowns is $(M+3)I+4H$. An equal number of equations provides the equilibrium conditions:

(a) Equilibrium in labor market, by industry l and zone i ($2I$ equations):

Demand is stochastic, based on the probability that a firm of type l chooses zone i and building type n , summed over n .

Supply is stochastic, based on the probability that a household chooses job in zone i , residence in zone j in building type n , shopping in zone k , and routes ρ_1 and ρ_2 , summed over all except i .

(b) Equilibrium in service-commodity market, by zone k (I equations):

Demand is stochastic, based on the probability of household choice as above, summed over all except k .

Supply is stochastic, based on the probability that a firm of industry type 2 chooses location in zone k and building type n , summed over n , and multiplied by an exogenous output coefficient.

(c) Equilibrium in building rental market, by building type and zone (MI equations):

Demand is stochastic, by firms ($n = 1, \dots, N$) or by households ($n = N+1, \dots, M$).

Supply is stochastic, based on developers (that is, landowners) in each zone maximizing profits over the possible types of buildings, including none.

(d) Equilibrium on network links ($4H$ equations):

Link flows arise from household demands for routes (see above), allocated to the links comprising those routes.

Link costs and travel times are determined as a function of link flows by a specified congestion technology.

Anas goes on to prove the existence and the uniqueness of a solution under quite mild conditions, namely that there be enough land to accommodate the exogenously given numbers of firms and households, and that the generalized cost of travel on a link (defined as a linear combination of cost and travel time) be a positive, strictly increasing, and strictly convex function of traffic flow.

The proof involves the setting up of a nonlinear programming problem in which the equilibrium equations are generated as first-order conditions. The objective function includes three entropy terms, each summing a term of the form $v \ln v$ over the relevant indices on v . The first, in which v is number of firms in zone i using building type n , generates stochastic variation in location choices and building-type choices of firms. The second, in which v is number of households choosing zones i, j, k , employment industry l , residential building type n , and routes ρ_1 and ρ_2 , generates stochastic variation along all these choice dimensions. The third, in which v is the amount of land in zone i devoted to building type n , generates stochastic variation in land use. It is these three terms that account for the nondeterministic nature of the solution and also provide a useful interpretation as a measure of consumers' surplus (Williams, 1977). The other terms in the objective function are deterministic. Two of them generate maximizing behavior on the part of firms and households, as in Herbert and Stevens (1960) or Mills (1972): these consist of the exogenous portions of aggregate firm profits and of aggregate household utility. The final term is the negative of the integral of average generalized cost of travel as link traffic flow is raised from zero; this is the term introduced

by Beckmann et al (1956), and used by Los (1979), to generate a 'user equilibrium' on the travel network. This objective function is maximized subject to a series of resource and accounting constraints that yield, as Lagrangian multipliers, all the endogenous prices in the system, plus a shadow value of land in each zone and a shadow value of a firm in each industry.

4.1.3 Evaluation The model of Anas is the most theoretically complete of any we have reviewed. Its solution always exists, its equations can be written fairly compactly, and it brings together important insights from several branches of the literature. Its formulation as the solution to a mathematical programming problem lends elegance and connects the theory to well-understood computational techniques. Presumably, because congestion is the only externality included, the model could easily be converted to a normative one depicting an efficient city, just by changing the traffic-equilibrium term to an integral of marginal rather than average generalized travel cost (so that it represents total generalized travel cost). This would produce a so-called 'system equilibrium' (that is, a social optimum) rather than a user equilibrium.

Virtually all quantities of interest for intrametropolitan study are determined endogenously, with the exception of the configuration of the transportation network. Total employment by industry is fixed in advance, but its location is determined as part of the solution; so is the location of vacant land, types of development, and building vacancies.

A major disadvantage is the enormous number of variables and equations involved. Suppose, for example, that we represent the seven-county San Francisco Bay Area with the zones and highway network used by its Metropolitan Transportation Commission in the 1970s. There are 440 zones and 11 036 links. Suppose also that we need 15 building types, as in Mills (1972). We then have 52 064 variables $[(18 \times 440) + (4 \times 11\,036)]$ to solve for in a nonlinear program. Clearly, implementation of such a model would have to be done in a sketch planning framework in which a highly aggregated representation of a metropolitan area is analyzed for the broad outlines of urban development.

A corresponding difficulty is the data requirements of the model. The following quantities, among others, must be specified: length and capacity of each link; possible routes between each pair of zones; cost, land requirements, number of employees, and floor space associated with each building type; value of land in nonurban uses; price of the export commodity; and congestion technology. Furthermore, a discrete-choice model must be estimated for a firm in either industry choosing among I locations and N building types; another is needed for a household simultaneously choosing its job location, job type (basic or service), location of home and shopping, residential building type, and routes for work trips and shopping trips. Alternative-specific dummy variables may be included in many of these choice functions to account for special characteristics of each zone. This is well beyond the existing capabilities of demand models.

Alternatively, the researcher could view many of these data requirements as opportunities for parametric specification. In that case, it becomes a simulation model rather than a strictly empirical one, with the same advantages and disadvantages of modeling flexibility as those in the model of Kim, except that many more parameters must be specified.

Three other disadvantages are shared with the linear programming models. No economies of scale or externalities other than congestion are represented. The solution is a long-run equilibrium, not a dynamic growth path. Finally, the nature of capital-land substitution, to which Mills (1967) attributes central

importance, is buried within the numerical values of various exogenous parameters, making its role somewhat difficult to isolate from other forces.

4.1.4 Suitability Because of its size, this model is best suited to sketch planning and broad-brush policy analysis—for example, tracing the effects of a major highway improvement on the location of large business centers. It would be very difficult to calibrate the model on actual historical data for a given metropolitan area, but plausible choice functions might be pieced together from past studies, and other parameters could then be set in the spirit of a simulation.

However, its long-run nature and the absence of any economies of agglomeration make the model poorly suited for the study of rapid growth, or of areas where the spontaneous development of major industrial or commercial centers is of crucial interest. Hence, we do not regard it as a prime candidate for modeling rapidly growing suburban areas.

4.2 *Anas: Chicago Area Transportation/Land-use Analysis System (CATLAS)*

4.2.1 History Anas (1982; 1983) has developed an empirically implementable model incorporating some of the desirable features of the general-equilibrium model reviewed above. Stochastic discrete-choice models are featured to represent dispersion in choices, and prices to portray market clearing. Empirical tractability is achieved by taking all employment location and transportation system characteristics as given, concentrating on a detailed representation of supply and demand in housing.

The model is explicitly dynamic, with a recursive structure in which short-run market clearing with a given housing stock occurs each year, and long-run changes in housing stock occur in response to the resulting rents. It was designed especially to predict property-value changes resulting from transportation policies. It has been calibrated with data from Chicago and used to simulate the impact of a fixed-rail extension there.

4.2.2 Structure Housing demand in each of I residential zones is defined as the expected number of households that will choose to live in that zone, as a function of housing rent. This is determined by a choice function for households, $P_{imh}(R_i, X_i, Y_{ih}, \alpha)$, which gives the probability that a worker in zone h will choose to commute using mode m and to live in zone i with average rent R_i , residential characteristics X_i , and transportation characteristics Y_{ih} ; α is a set of empirically estimated parameters. Aggregate demand for housing in zone i is therefore:

$$\sum_{hm} N_h P_{imh}(R_i, X_i, Y_{ih}, \alpha),$$

where N_h is the number of workers in zone h . In practice, Anas distinguished only two workplace zones, CBD and non-CBD.

Housing supply in each zone is defined as the expected number of housing units there that will be offered for rent or that will be owner-occupied. This is determined by another choice function representing the behavior of owners of housing. The function $Q_i(R_i, X_i, \beta)$ gives the probability that the owner of a dwelling unit in zone i will offer it for occupancy, given average rent and residential characteristics; β is another set of empirically estimated parameters. Aggregate supply of housing in zone i is therefore $S_i Q_i(R_i, X_i, \beta)$, where S_i is the size of the existing housing stock in zone i .

Short-run equilibrium is depicted by I simultaneous equations in the I unknowns, R_i , each equation setting aggregate demand equal to aggregate supply in one zone. Long-run adjustments in housing stock are specified through other equations that take into account the profitability of renting or occupying a housing unit.

4.2.3 *Evaluation* CATLAS is the most theoretically satisfactory model that has been actually applied to a large data set for a real city. Economic markets are represented explicitly, and thereby welfare analysis of the results is permitted. Its use allows for more realistic dispersion than do the economic models deriving from Herbert-Stevens and Mills, while providing a microbehavioral basis for that dispersion through the stochastic demand formulation.

The model also has shortcomings. First, only the residential sector is depicted, and hence the model cannot be used to address questions about employment location. (Nevertheless, residential location accounts for the bulk of all urban land use and is therefore an important topic in its own right.) Second, housing markets are assumed to be perfectly competitive, although we know that housing markets are strongly affected by neighborhood effects such as ethnic attraction, by environmental externalities such as pollution, by urban amenities such as lakeshores or public services, and by economies of scale in residential land development. Anas had good reason to exclude such phenomena, because, as shown by Werczberger and Berechman (1988), they often destroy the uniqueness of equilibrium, which is one of the key goals of this modeling effort.

Third, with CATLAS one cannot account for congestion on the transportation network, an understandable limitation in view of its initial purpose of permitting analysis of transit improvements. Finally, large numbers of estimated choice functions are required, as are large amounts of data on housing characteristics, transportation characteristics, and rents.

4.2.4 *Suitability* Applicability of the model to rapidly growing areas is limited by the focus only on the residential sector, lack of economies of scale, and neglect of those urban amenities that are of prime policy interest in a new urban area. Although the housing stock does have a long-run dynamic element, it is doubtful that the simple adjustment mechanism used could adequately represent the strategic decisions of large developers or owners of large tracts of vacant land.

5 Agglomeration economies

5.1 Introduction

The models discussed so far generate dense clusters of activity only because of some location—an exogenous employment center or export node—that many people or firms want to be close to. This can help explain the existence of the traditional port city or railroad town, but it cannot help in the explanation of the employment centers in many newer sprawling metropolitan areas.

The reason an urban center or subcenter can develop in a nondescript spot is economies of agglomeration: benefits that one economic actor receives through proximity to another. At the aggregate level, agglomeration economies are a type of economy of scale, by virtue of which a large cluster of activities produces more efficiently than a small cluster. At the 'micro' level, agglomeration economies are a type of positive externality, in which the activity of one party benefits another more than is reflected in any monetary payments. To emphasize this dual interpretation, economies of agglomeration are also called 'external economies'. They were dramatically illustrated in case studies by Vernon (1960) and Chinitz (1961), and they have become a cornerstone in the standard explanation for the existence of cities.

The impact of agglomeration economies on clustering of industrial location has never been included in full-scale models of land use, despite this early recognition of its importance. Indeed, such models nearly always require that the location of basic employment be specified in advance. This greatly limits their value for forecasting

because of the difficulty and uncertainty in supplying the needed forecasts of basic-employment distribution. It also limits their value for policy analysis because of the inability to account for feedback from other sectors to basic employment.

Agglomeration economies may be classified by the types of interaction among economic actors. 'Production economies' arise from interactions among firms that reduce production costs: for example, shipments of intermediate goods, flows of information, or coordination of activities. This type explains concentrations of production activities, particularly those heavily dependent on the kinds of personal meetings that are fostered by location within a single business district. 'Shopping economies', by contrast, arise from the advantage to customers of their being able to combine several tasks into one trip, causing them to favor commercial establishments that are clustered. Shopping economies were explored systematically by Harris and Wilson (1978), who discovered sudden dramatic shifts in the whole pattern of retail clustering as the scale-economy parameter is changed. There are also agglomeration economies known as 'neighborhood effects' which affect residential location, for example, the desire to live near members of a particular economic or ethnic group; we do not discuss these effects here.

Two trends have greatly altered the kinds of clustering typifying newer urban development. First, cheaper transportation has reduced the importance of shipment costs in industrial location, leading to a more sprawling overall development pattern. At the same time, the increased importance of service and technical industries requiring frequent personal interaction has led to thriving centers of office activity. Often these centers sprout like mushrooms at apparently random spots throughout an otherwise low-density urban area. Economies of agglomeration, especially of the 'production' type involving information and coordination, are a key to understanding this phenomenon. Hence, in searching for land-use models that can help clarify recent urban growth, we should seek models that incorporate such economies explicitly.

Of course, activity clusters could occur merely from the existence, within a limited area, of locations sharing particular traits desired in common by many firms or households. Whether clustering is mainly due to this or to agglomeration economies is an empirical question that can be answered only within models allowing for both phenomena.

5.2 Office location within a central business district

In one class of models, a need for contacts between firms is postulated explicitly, and it is on this basis that a pattern of location within a business district is derived. Two such models that rely on economic concepts, including an endogenous distribution of land rents, are by O'Hara (1977), and Tauchen and Witte (1983). Although the research reported in these articles establishes important theoretical principles for analyzing what goes on inside a business district, they have not been developed far enough to portray an actual city. Nor do they predict the spontaneous emergence of business subcenters.

Empirical information about contacts among office workers is scarce. An exception is the study by Goddard (1973), which provides survey data about personal and telephone contacts among office workers in central London. One may expect that the rising interest in telecommunications and its locational effects (Salomon, 1986) will lead to expanding empirical knowledge in this area.

5.3 Central place theory

Another class of models is built on central place theory. This theory was developed to explain the sizes and functional distributions of cities within a region, but it is equally applicable to the sizes and functional distributions of subcenters within a

metropolitan area. Because central place theory is based on market areas, the concept is applicable only to the 'shopping' type of agglomeration economy.

The original works in this area, Losch (1954) and Christaller (1966), are based on rigidly deterministic behavior by consumers, leading to a hierarchical structure in which each center performs certain functions plus all the functions of centers lower on the hierarchy. However, some recent extensions have incorporated more realistic behavioral assumptions and some randomness, leading to a quite varied set of possible outcomes. What ties these extensions to the central place literature is a continuing emphasis on a hierarchy of functions that cities can perform, on the simultaneous determination of business and residential location, and on the tension between the need of producers to pursue economies of scale and the desire of customers to reduce transportation costs.

Carruthers (1981) simulates a one-dimensional city over time. Firms make sophisticated profit-maximizing decisions about location and scale, taking into account likely actions of future firms: each firm wants a large initial market share, but also wants to forestall future entry by close competitors. Production in each firm involves a predetermined degree of scale economies, the magnitude of which determines the position of the industry of that firm in the hierarchy of urban functions. Workers choose residences according to a predetermined spatial distribution around the firm. Demand by each household for produced goods is based on price plus transportation cost, with parameters varying among households. Randomness is inserted into the locational decisions of firms, and the resulting patterns are generated numerically with the help of a random-number generator. These results typically show a mixture of monopolies, oligopolies, and competitive firms. Carruthers finds that the size distributions of urban centers generated by his model conform to the general predictions of the traditional central place theory.

Wilson (1981, pages 179-195) combines a market-area approach to retail shopping-center supply with a Lowry-type spatial-interaction model of residential location and retail demand. Household expenditures on housing, land, and retail services are governed by exogenous parameters, but the specific locations are governed by a spatial-interaction function. The model is made dynamic by a simple adjustment mechanism in which the change in supply of land, housing, or retail floor space is proportional to the discrepancy between supply and demand. However, the model contains no prices, and therefore cannot account for any tendency of price adjustments to alleviate disequilibrium. Wilson is particularly interested in exploring the potential for such a model to show 'bifurcations': ranges in which a small change in parameters will result in a qualitatively distinct growth path. No actual simulations are presented.

Allen and Sanglier (1981) provide computer simulations of a stochastic two-dimensional system of urban centers over time, within the central place framework. The geography is based on a square grid. Centers arise spontaneously on the basis of specified random fluctuations in demand, and then grow according to equations that are reminiscent of population biology: rates of change are proportional to the gap between a predetermined 'natural carrying capacity' (supplemented by jobs) and the actual amount of activity, but modified by crowding effects. New industries appear whenever a predetermined size threshold is reached. Shopping economies of agglomeration appear through a rather complex function describing the attractiveness of a given center to the surrounding population. This function includes prices, which seem to be exogenous, and a crowding effect.

Allen and Sanglier find that, in most simulations, the urban systems reach a period of stability of form. Interestingly, though, there may be several qualitatively

distinct stable forms for a given set of parameters; and constraints on the early history of the system can influence which stable form will ultimately develop. This is an attractive feature, because it means that one can incorporate specific historical accidents into a simulation of the future. No suggestions are given for calibrating the model from data for a real city. Efforts by Pumain et al (1984) to do so for Rouen, France, met with considerable difficulties including multiple solutions, counter-intuitive results, and a consequent reliance on ad hoc methods.

To date, none of the central place models have been used for serious empirical prediction or for testing of detailed land-use plans. Indeed, their potential for such applications appears limited by their rather crude underlying behavioral postulates. Nevertheless, they are of interest because of their ability to flesh out the theoretical notions of external economies in the form of detailed simulations, and because there are so few approaches to quantitative modeling of agglomeration economies.

5.4 Production economies of agglomeration

Several authors have recently tried to model economies of agglomeration of the 'production' type, thereby explaining centers of industrial or office activity.

Clapp (1984) starts from first principles, postulating that certain agents must contact every other agent, as in the case of office-location models discussed earlier. Even one such agent is enough to cause the spontaneous emergence of a single business district, and bilateral contacts among several agents can result in subcenters. The model is for a linear city, and no simulations are attempted.

Fujita and Ogawa (1982) study a linear city in considerably greater detail. Agglomeration economies are postulated rather than derived, being described by a 'locational potential function' that is similar to accessibility measures in spatial-interaction models. There are two sectors: business, and residential. The authors find that several quite distinct equilibrium patterns may emerge, depending on the parameters of the potential function and on the ratio of transport cost to production price. These patterns include: monocentric city with an all-business or a mixed-use center surrounded by a residential district; monocentric city with mixed-use throughout; duocentric 'twin city' pattern; tricentric pattern with or without some outbound commuting toward the satellite business districts. Multiple equilibria are possible with some parameter values. Dynamic growth paths are not considered, but the pattern of equilibrium land rents is fully described.

5.5 Suitability

The models reviewed in this section are not intended for detailed planning or forecasting in a real urban area, but rather for numerical simulation. Given the rudimentary state of our knowledge of agglomeration economies, this is probably an advantage: rather than force a more well-developed model into a situation where it just does not fit, one can use these flexible models to explore a variety of explanations for observed patterns.

The approaches of central place theory and of the production externality offer possibilities, though each specific model reviewed has severe drawbacks. Only the three central place models are dynamic, and of those only the model of Carruthers contains any economic behavior as reflected in prices; but his is a one-dimensional model. The two-dimensional model of Allen and Sanglier, though overly mechanical, can simulate a wide variety of realistic growth patterns.

The stochastic element in the models of Carruthers and Allen-Sanglier is a major advantage if one wishes to confront a model with the actual history of a single urban area. Starting at an initial configuration, such a model can provide a statistical description of the range of likely possible growth histories that can be compared with the actual growth history. Knowledge of important local decisions

and events can provide hints as to the nature of the random events that actually occurred, and this information can be used to assess the plausibility of the model as the generator of the actual history. Given a plausible model and a knowledge of the random events up to a given point, one can generate forecasts of possible futures. Not only that, one can study the 'counterfactual': What if old man Bixby had not been tricked into selling his farm to Shark Development Corporation for a tenth of its market value, thereby triggering development there instead of on plots owned by more stubborn neighbors?

The only models reviewed here that permit study of production externalities are one-dimensional and static. Nevertheless, Fujita and Ogawa's work provides a good deal of insight into the ranges of conditions that lead to different patterns. Further development of this kind of model might well lead to useful two-dimensional dynamic simulations of patterns typical of industrial or research and development parks.

6 Conclusions

Our goal has been to determine what land-use models are most suitable for realistic empirical application in a 'contemporary urban area'. In particular, we want to highlight the features that tend to distinguish such rapidly growing, formerly suburban, areas from older metropolitan areas with dominant central cities. If further progress is to be made in understanding the forces driving late twentieth-century urban growth, a suitable framework is needed to guide data collection in such areas.

Our review suggests that among existing land-use models, it is possible to find models that are empirically tractable on real data *or* that highlight at least some features of contemporary urban areas; but not both. Those models such as Garin-Lowry, Herbert-Stevens, or CATLAS that have been designed to accommodate the complexity and limitations of real data are either static in nature, unrealistically deterministic, or limited to the housing sector. Most oversimplify the transportation sector and none include agglomeration economies, which affect the clustering of firms. On the other hand, those models that are more complete theoretically, such as Anas's general-equilibrium model, require too many data and have too many equations to estimate; whereas models such as those of Carruthers or Allen and Sanglier that focus on dynamic growth and agglomeration economies have been used only for simulations of prototype cities.

One implication of this dilemma is obvious, but can never be repeated too often: the best model depends upon the purpose. If you want realistic forecasts of the development of a particular city, you have no choice but to use an operational model with reasonable data requirements and, probably, an inadequate theoretical structure. If you want to explore efficiency in land-use allocation, it is logical to choose one of the programming models. If you want a better theoretical understanding of land speculation or of the role of historical accident, you need a model with sophisticated dynamics and long-run disequilibrium, and will have to give up on realistic detail. If you want to understand the formation of subcenters, your model should include agglomeration economies.

Researchers can afford to feel less constrained by solution algorithms because computational costs are falling rapidly. For example, some disadvantages of the Lowry and Putman models could be eliminated simply by further iterating the model, and by reallocating *all* activities and traffic at each iteration. Complexity will continue to place limits; but those limits will be more closely tied to the capacity of the researcher to supply data and check for errors than to the capacity of the computer to do the calculations. This suggests, for example, the disaggregation of

decisionmakers into many classes, all of which behave according to a single parametric family of choice functions that can be estimated all at once.

In the end, we are unable to describe the best approach for modeling contemporary urban areas. We have argued that dynamics, randomness, and agglomeration economies each play a key role; and we have seen that these features are lacking in operational models with realistic detail. From this point, one could pursue any of three directions. One could start with a working empirical model and add agglomeration economies and dynamic adjustment; this would greatly alter the character of the model and would therefore require a new solution mechanism. Or one could start with a dynamic model with agglomeration economies, and add greater realism and detail; this would add enormous flexibility to an already flexible simulation model, therefore requiring new techniques for the reliable calibration of its parameters. A third approach would be to build an entirely new model from scratch.

We are confident that through one or more of these approaches, a deeper understanding of the forces acting in contemporary urban areas can be attained. For better or worse, the degree of success will depend, as always, on the creativity and common sense of those researchers who make the attempt.

Acknowledgements. This paper was prepared in conjunction with a study of Activity and Transportation Systems Development, funded by the Institute of Transportation Studies, University of California. We are grateful to Richard Arnott, David Boyce, Genevieve Giuliano, and Britton Harris for comments on an earlier draft.

References

- Allen P M, Sanglier M, 1981, "A dynamic model of a central place system-II" *Geographical Analysis* 13 149-164
- Alonso W, 1964 *Location and Land Use* (Harvard University Press, Cambridge, MA)
- Anas A, 1980, "A probabilistic approach to the structure of the rental housing markets" *Journal of Urban Economics* 7 225-247
- Anas A, 1982 *Residential Location Markets and Urban Transportation: Economic Theory, Econometrics and Policy Analysis with Discrete Choice Models* (Academic Press, New York)
- Anas A, 1983, "The Chicago Area Transportation-Land Use Analysis System", final report to US Department of Transportation (contract 92028); Department of Civil Engineering, Northwestern University, Evanston, IL 60201
- Anas A, 1984, "Discrete choice theory and the general equilibrium of employment, housing, and travel networks in a Lowry-type model of the urban economy" *Environment and Planning A* 16 1489-1502
- Anas A, 1988 *Modeling in Urban and Regional Economics* to appear as a volume in *Fundamentals of Pure and Applied Economics* (Harwood Academic Publishers, New York) forthcoming
- Batty M, 1972, "Recent development in land-use modeling: a review of British research" *Urban Studies* 9 151-177
- Batty M, 1976 *Urban Modeling: Algorithms, Calibrations, Predictions* (Cambridge University Press, Cambridge)
- Beckmann M J, McGuire C B, Winsten C B, 1956 *Studies in the Economics of Transportation* (Yale University Press, New Haven, CT)
- Berechman J, 1981, "Analytical problems in linking an activity model with a transportation network model" *Environment and Planning A* 13 449-462
- Berechman J, Gordon P, 1986, "Linked models of land use-transport interactions: a review" in *Advances in Urban Systems Modelling* Eds B Hutchinson, M Batty (North-Holland, Amsterdam) pp 109-131
- Boyce D E, 1978, "Equilibrium solutions to combined urban residential location, modal choice and trip assignment models" in *Competition Among Small Regions* Eds W Burh, P Friedrich (Nomos, Baden-Baden) pp 246-264
- Boyce D E, Chon K S, Lee Y J, Lin K T, Le Blanc L J, 1983, "Implementation and computational issues for combined models of location, destination, mode, and route choice" *Environment and Planning A* 15 1219-1230

- Boyce D E, Lundqvist L, 1986, "Network equilibrium models of urban location and travel choices: alternative formulations for the Stockholm region", unpublished report, Department of Civil Engineering, University of Illinois, Urbana, IL 61801
- Carruthers N, 1981, "Central place theory and the problem of aggregating individual location choices" *Journal of Regional Science* 21 243-261
- Chinitz B, 1961, "Contrasts in agglomeration: New York and Pittsburgh" *American Economic Review, Papers and Proceedings* 51 279-289
- Chon K S, Eash R W, Boyce D E, 1983, "Testing of combined urban location and travel choice models", unpublished report, Department of Civil Engineering, University of Illinois, Urbana, IL 61801
- Christaller W, 1966 *Central Places in Southern Germany* (Prentice-Hall, Englewood Cliffs, NJ)
- Clapp J M, 1984, "Endogenous centers: a simple departure from the NUE model" *Papers of the Regional Science Association* 54 13-24
- Crecine J P, 1964 *TOMM: Time Oriented Metropolitan Model* Technical Bulletin 6, Community Renewal Program, Department of City Planning, Pittsburgh, PA
- Daganzo C F, Sheffi Y, 1977, "On stochastic models of traffic assignment" *Transportation Science* 11 253-274
- Evans S P, 1976, "Derivation and analysis of some models for combining trip distribution and assignment" *Transportation Research* 10 37-57
- Florian M, Nguyen S, Ferland J, 1975, "On the combined distribution assignment of traffic" *Transportation Science* 9 43-53
- Foot D, 1981 *Operational Urban Models* (Methuen, New York)
- Fujita M, Ogawa H, 1982, "Multiple equilibria and structural transition of non-monocentric urban configurations" *Regional Science and Urban Economics* 12 161-196
- Garin R A, 1966, "A matrix formulation of the Lowry model for intrametropolitan activity location" *Journal of the American Institute of Planners* 32 361-364
- Geraldes P, Echnique M H, Williams I N, 1978, "A spatial economic model for Bilbao" *Proceedings of the PTRC Summer Annual Meeting*; PTRC Education and Research Services Ltd, 110 Strand, London WC2
- Goddard J B, 1973, "Office linkages and location: a study of communications and spatial patterns in Central London", in *Progress in Planning, Volume 1* Eds P R Diamond, J B McLaughlin (Pergamon Press, Oxford) part 2
- Goldner W, 1968 *Projective Land Use Model (PLUM): A Model for the Spatial Allocation of Activities and Land Uses in a Metropolitan Region* TR-219, Bay Area Transportation Study Commission, copy from Metropolitan Transportation Commission, Berkeley, CA
- Hamilton B W, 1982, "Wasteful commuting" *Journal of Political Economy* 90 1035-1053
- Harris B, 1963, "Linear programming and the projection of land uses", Penn-Jersey Paper 20, Department of City and Regional Planning, University of Pennsylvania, Philadelphia, PA 19104
- Harris B, 1985a, "Synthetic geography: the nature of our understanding of cities" *Environment and Planning A* 17 443-464
- Harris B, 1985b, "Urban simulation models in regional science" *Journal of Regional Science* 25 545-567
- Harris B, Wilson A, 1978, "Equilibrium values and dynamics of attractiveness terms in production-constrained spatial-interaction models" *Environment and Planning A* 10 371-388
- Hartwick P G, Hartwick J M, 1974, "Efficient resource allocation in a multinucleated city with intermediate goods" *Quarterly Journal of Economics* 88 340-352
- Herbert J S, Stevens B H, 1960, "A model of the distribution of residential activity in urban areas" *Journal of Regional Science* 2 21-36
- Ingram G, Kain J, Ginn J R, 1972 *The Detroit Prototype of the NBER Urban Simulation Model* (Columbia University Press and the National Bureau of Economic Research, New York)
- Kim T J, 1979, "Alternative transportation modes in an urban land use model: a general equilibrium approach" *Journal of Urban Economics* 6 197-215
- Kim T J, 1983, "A combined land use-transportation model when zonal travel demand is endogenously determined" *Transportation Research B* 17 449-462
- Los M, 1979, "Combined residential-location and transportation models" *Environment and Planning A* 11 1241-1265
- Losch A, 1954 *The Economics of Location* (Yale University Press, New Haven, CT)
- Lowry I S, 1964 *A Model of Metropolis* RM-4035-RC, The Rand Corporation, 1700 Main Street, Santa Monica, CA 90406

- McFadden D, 1973, "Conditional logit analysis and qualitative choice behavior", in *Frontiers in Econometrics* Ed. P Zarembka (Academic Press, New York) pp 105-142
- Mackett R L, 1985, "Integrated land use-transport models" *Transport Reviews* 5 324-343
- Mills E S, 1967, "An aggregative model of resource allocation in a metropolitan area" *American Economic Review* 57 197-210
- Mills E S, 1972, "Markets and efficient resource allocation in urban areas" *The Swedish Journal of Economics* 74 100-113
- Mills E S, 1974a, "Mathematical models for urban planning, in *Urban and Social Economics in Market and Planned Economies* Eds A Brown, J Licari, E Neuberger (Praeger, New York) pp 113-129
- Mills E S, 1974b, "Sensitivity analysis of congestion and structure in an efficient urban area", in *Transport and the Urban Environment* Eds J G Rothenberg, I G Heggie (John Wiley, New York) pp 192-208
- Mills E S, 1976, "Planning and market processes in urban areas", in *Public and Urban Economics* Ed. R E Grieson (D C Heath, Lexington, MA), pp 313-329
- Muth R F, 1969 *Cities and Housing: The Spatial Pattern of Urban Residential Land Use* (University of Chicago Press, Chicago, IL)
- O'Hara D J, 1977, "Location of firms within a square central business district" *Journal of Political Economy* 85 1189-1207
- Pumain D, Saint Julien T, Sanders L, 1984, "Dynamics of spatial structure in French urban agglomerations" *Papers of the Regional Science Association* 55 71-82
- Putman S H, 1973 *The Interrelationships of Transportation Development and Land Development* (US Department of Transportation, Washington, DC)
- Putman S H, 1974, "Preliminary results from an integrated transportation and land use models package" *Transportation* 3 193-224
- Putman S H, 1975, "Urban land use and transportation models: a state-of-the-art summary" *Transportation Research* 9 187-202
- Putman S H, 1983 *Integrated Urban Models: Policy Analysis of Transportation and Land Use* (Pion, London)
- Ripper M, Varaiya P, 1974, "An optimizing model of urban development" *Environment and Planning A* 6 149-168
- Salomon I, 1986, "Telecommunications and travel relationships: a review" *Transportation Research A* 20 223-238
- Senior M L, Wilson A G, 1974, "Explorations and syntheses of linear programming and spatial interaction models of residential location" *Geographical Analysis* 6 209-238
- Tauchen H, Witte A D, 1983, "An equilibrium model of office location and contact patterns" *Environment and Planning A* 15 1311-1326
- Turner C G, 1975, "The design of urban growth models for strategic land use transportation studies" *Regional Studies* 9 251-264
- Vernon R, 1960 *Metropolis 1985* (Harvard University Press, Cambridge, MA)
- Wardrop J G, 1952, "Some theoretical aspects of road traffic research" *Proceedings, Institute of Civil Engineering, Part II* 1 325-378
- Werczberger E, Berechman J, 1988, "Incorporating neighborhood effects into spatial allocation models" *Environment and Planning A* 20 595-607
- Wheaton W C, 1974, "Linear programming and locational equilibrium: the Herbert-Stevens model revisited" *Journal of Urban Economics* 1 278-287
- Williams H C W L, 1977, "On the formation of travel demand models and economic evaluation measures of user benefit" *Environment and Planning A* 9 285-344
- Wilson A G, 1967, "A statistical theory of spatial trip distribution models" *Transportation Research* 1 253-269
- Wilson A G, 1981 *Catastrophe Theory and Bifurcation: Applications to Urban and Regional Systems* (Croom Helm, Beckenham, Kent)
- Wingo L, 1961 *Transportation and Urban Land Use* (The Johns Hopkins University Press, Baltimore, MD)