Review of Existing Land-Use Forecasting Techniques

N. A. IRWIN, Vice President, Traffic Research Corporation, New York, N. Y.

THE BOSTON Regional Planning Project (BRPP) has retained the Traffic Research Corporation (TRC) to develop and apply for preliminary forecasts a mathematical model for estimating future distributions of population, land use and economic activities in the Boston region. BRPP plans to use this model as a device for testing and evaluating the probable future effects of transportation facilities, zoning policies, and possibly other factors under planning control on the distribution and density of development patterns throughout the region.

Most urban transportation planning studies during the past decade have produced or are producing estimates of future land use in all subregions of the areas which have been or are under study, mainly to provide a basis for estimating future traffic demand. (In this context, land use includes population and economic activities as well as structures and land areas.) Methods employed for these land-use estimates have ranged from largely intuitive or judgmental projections to systematic techniques based on a chain of quantitative reasoning which could be reproduced by another group.

The purpose of the present review is primarily to insure that existing techniques are fully utilized, where pertinent, in the development of a land-use forecasting model for the Boston region.

STUDY METHOD

This report is based on a review of the included references and on discussions with researchers concerning various aspects of their predictive techniques. An effort has been made to concentrate on techniques based mainly on explicit formal relationships rather than those which rely mostly on judgment applied subjectively to each subregion. Stress has also been placed on techniques which have been or may be calibrated and/or tested empirically.

The various land-use forecasting techniques described may be grouped into a number of categories according to types of variables, types of restraints and controls, manner of application, degree of operationality, and basic concept employed. Based on the latter criterion only, techniques may be classed as ad hoc, potential, economic, regression, behavioral, etc. Little attempt has been made to group the reviewed techniques along these lines, mainly because most of the techniques are found to fall into several categories simultaneously. The basic characteristics of each technique and some group characteristics are described.

For presentation purposes, however, it is useful to divide the techniques into three groups according to their present development stage. Group 1 comprises techniques which are operational for forecasting purposes or are now being developed to operationality. Group 2 consists of research-oriented studies, aimed primarily at gaining insight into urban processes by empirical testing of certain hypotheses. Group 3 is made up of conceptual studies primarily of interest because of the ideas involved rather than because of any empirically tested relationships which may have resulted. Again, there is some doubt concerning the group to which certain techniques belong; the distinctions are not entirely clear cut, and the choice of group has been fairly arbitrary in some cases.

An attempt has been made to describe the following aspects of each technique: input and output variables, forecasting methods and relationships, methods of calibrating

Paper sponsored by Committee on Origin and Destination.

and/or testing, data requirements and controls, and computational equipment required. Techniques falling into Group 1 have been covered in most detail, although even here some of the techniques have not been developed or described sufficiently to enable coverage of all these points. For techniques falling into Group 2, emphasis has been placed perforce on data and relationships. It has been necessary to describe techniques falling into Group 3 mostly in terms of the concepts and formulations involved.

It is, of course, recognized that the list of 14 land-use forecasting techniques reviewed in this report is not exhaustive. Some techniques, notably those employed in some of the earlier urban transportation studies, have been eliminated because it was felt that their ad hoc nature would contribute little to the present model development project. Others were not included because of their similarity to techniques which have been reviewed; for example, the Pittsburgh Area Transportation Study was developed with reference to the Chicago Area Transportation Study. Still other studies, such as those at Hartford, Conn., Washington, D. C., have been included under the name of the individuals involved. Finally, current land-use forecasting work by a number of groups (the Upper New York State Transportation Studies, Los Angeles Regional Transportation Study, Puget Sound Regional Transportation Study, Twin Cities Area Transportation Study, Southeast Wisconsin Regional Planning Study, Tri-State Transportation Commission, and Ohio-Kentucky-Indiana Transportation Study) has been omitted because of the lack of descriptive detail on their work or plans available at the present time. The land-use models previously developed or currently under development by the Traffic Research Corporation are described in a project description and progress reports which have been submitted previously to the Boston Regional Planning Project and are, therefore, not summarized in the present report.

In summarizing forecasting methods and relationships used in the various techniques, an attempt has been made to present the concepts involved as briefly as possible without oversimplification. In some cases it has been possible to quote directly the relevant mathematical equations. In other cases, however, where presentation of the mathematical formulations in symbolic form and definition of symbols would require a number of pages of quite complex mathematical notation, the mathematical formulations have been presented in words for the sake of brevity. Symbolic presentations of these equations and relationships can be found in the source documents listed in the references.

NOMENCLATURE

A number of technical terms used in describing the various techniques are briefly defined as they are used in this report.

Model—a systematic method, based on logical or mathematical relationships, for describing, simulating and forecasting real-life processes, in this case the distribution of land-use categories throughout an urban region. The term is used synonymously with technique.

Land-Use Categories—include structures, economic activities, employment, floor areas, population, land areas, and generally any items that can be used to describe urban subregions and regions in quantitative terms.

Input Variables—data categories required to operate a land-use model.

Output Variables—data categories whose subregional values are estimated or forecast by a land-use model.

Relationship—a quantitative statement showing how one variable depends on one or more other variables. It may be in the form of a mathematical function, inequality or equation, or possibly in tabular form.

Parameter—a number which is part of a mathematical relationship and is either constant in value or takes one of a number of specified values; also known variously as a coefficient or a constant.

Calibration—determination of the best values of parameters in the various relationships comprising the model so that the model describes as closely as possible historic situations and events of the type it is intended to simulate.

Independent Variable—a data category which is felt to be the cause rather than the effect

of a particular process, and is, therefore, considered as an input variable rather than an output variable with reference to a relationship describing the process. Conversely, a dependent variable is a data category which is considered to be an effect of a particular process, and is, therefore, considered to be an output variable of the relationship in question.

Reliability Check—a quantitative evaluation of the probable forecasting accuracy of a model. This can be based on a statistical appraisal of the model's relationships and/or on application of the model to simulate a historical situation which was not used in its calibration, followed by a statistical comparison of calculated and observed results.

Regression Analysis—the process of describing by an equation the relationship between a dependent variable (output variable) and one or more independent variables (input variables) so that the equation so derived describes the relationship, as represented by a number of observations, with a minimum of error. A common method of regression analysis is the method of least squares.

Coefficient of Correlation, R—a statistical measure of the degree to which changes in the dependent variable are explained by changes in the independent variable(s). Definitions of this and other commonly used statistical terms may be found in standard statistical texts. Values of R quoted with reference to some of the models described herein should be treated with caution as a means of comparing models; R is influenced by factors such as number of observations, the slope of a relationship, and zone size, so that quoted values of R should not be used to compare one technique with another unless care has been taken to render other things equal. The square of this expression, R^2, is known as the coefficient of determination and is used conventionally as the measure of predictive ability of a relationship or model.

Recursive—consisting of a series of sequential forecasts where output from one forecast is used as input for the next in the series. Recursive in time consists of a series of sequential forecasts, each of a different point in time and each using about the same functional relationships between input and output, where output from the forecast of one point in time is used as input for the forecast of the next point in time.

Monte Carlo Simulation—simulation of the action of a large number of entities by simulating in detail the actions of a sampled subset of the total number of entities. The name derives from the use of random numbers in estimating probabilities required to simulate various decisions made by the sampled entities. The simulation of the activities of one randomly selected group or bundle of the entities is known as a game.

GROUP 1—OPERATIONAL OR QUASI-OPERATIONAL TECHNIQUES

Penn-Jersey Transportation Study

The Penn-Jersey regional growth model represents the largest research effort during the past few years devoted to the development of a land-use forecasting technique. Since Herbert and Stevens presented the basic theoretical structure in 1960 (2, 3), the model has been further elaborated and modified in various papers (4-8, 10, 13). At present, the model is not yet operational but is being intensively developed.

The model operates recursively in time by making a 5-year prediction, using the results of the first prediction as a basis for the next 5-year prediction.

The core of the model is the simulation of residential location based on the economic theory that individual households tend to maximize their locational advantage and that land is allocated to that group of households which can bid the highest price for it. The inputs to the activity distribution phase of the model for one iteration are as follows.

1. Inputs describing the state of the system, i.e., population location and characteristics; activity location by industry or other activity category; existing stock of buildings by area and type; vacant land by characteristics by area; accessibility to opportunities by area (from transportation model); projected income levels; in-migration to be accommodated by type of in-migrant; and growth of economic activity to be accommodated.

2. Data reflecting assumptions and policies, i.e., public open space reserved; public redevelopment; and land development controls and public services.

3. Subsidiary models, parameters and relationships such as calculation of land availability, relationships expressing rate of release of land for new uses, and rent levels from preceding iteration; household and population changes over time, applicable rates, demographic and other; probabilities of households of different types moving from initial location before end of iteration, by household type and area type; division of housing market between renters and owners, and applicable rates; relationships summarizing desirability of areas that determine household budgets to be devoted to location; formula for computing cost of transportation and costs of housing of different types, for households of different types in different areas; formulas for determining locational patterns of consumer-serving industries, including local government, on the basis of the location of population and accessibility measures (transportation conditions); growth and aging characteristics of all other industries; and locational preferences of all other industry based on area characteristics and accessibility (transportation) measures.

First, the model makes all calculations related to land availability. Only certain amounts of land are released in each zone for possible development in each use. This is contingent on zoning and redevelopment policy and on speculative holding. The latter is taken into account by withholding land in zones where rent has risen rapidly during the previous projection period. The model then proceeds to locate households by aging the households in each zone. Each household member ages 5 years or dies. Some households break up, others have children, others are newly formed, others emigrate from the region. The households remaining in the region, plus the projected in-migration, form the population in the future year. Each household will have changed character by this aging and changing of its income level and will have a certain probability of changing residence during the projection period. This information and information concerning newly formed and in-migrant households are used in calculating a pool of locators of each household type.

Each locating household type is assigned a locational budget for each zone equal to the annual amount of money it is prepared to spend therefor transportation and housing. This depends on the income of the household and its preference for different types of zones and possible housing types in the zone. The real annual cost of housing (not including land rent) and the transportation cost in each zone are then deducted from the locational budget of each household type for each zone. The transportation cost is related inversely to the accessibility of the household to destinations of the collection of trips it is likely to make during the year. The residual of these deductions from budgets is called rent-paying ability. Each household type has a unit rent-paying ability for each zone equal to the amount per square foot of land it is willing to pay to locate in its preferred housing type in that zone. It is assumed that each household type seeks to maximize its rent-paying ability and the landowners seek to maximize the returns on their property. The end result of the model is that households locate in such a way as to maximize aggregate rent-paying ability. The model simulates this process through a linear program (2, 3, 10) which maximizes aggregate rent-paying ability subject to two types of constraints: that all households must be located and that the total land used in each zone must not exceed the land made available.

The following notation is used in the formulation of the linear program.

U = areas forming an exhaustive subdivision of the region, indicated by superscripts $K = 1, 2, \ldots, U$.

n = socio-economic groups indicated by subscripts $i = 1, 2, \ldots, n$. For each group in each area, the subscript i also refers to a specific locational bundle, chosen from the set of available bundles, which affords the highest unit rent-paying ability for that group in that area.

b_i^K = annual location that a household of group i will use if it locates in area K $(i = 1, 2, \ldots, n)$.

c_i^K = annual cost to a household of group i of locational bundle it purchases in area K $(i = 1, 2, \ldots, n)$, $(K = 1, 2, \ldots, U)$.

s_i^K = number of units of land in locational bundle purchased by household of type i in area K (i = 1, 2, ..., n), (K = 1, 2, ..., U).

L^K = number of units of land made available in an iterative period for residential development in area K (K = 1, 2, ..., U).

N_i = projected number of households of socio-economic group i to be located in region during same iterative period (i = 1, 2, ..., n).

X_i^K = number of households of socio-economic group i located (by model) in area K in that iterative period (i = 1, 2, ..., n), (K = 1, 2, ..., U).

Allocation model.—The linear programming model for allocating households to land has the following form:

$$\text{Maximize } Z = \sum_{i=1}^{n} \sum_{K=1}^{U} X_i^K \left(b_i^K - c_i^K \right) \tag{1}$$

Subject to

$$\sum_{i=1}^{n} s_i^K X_i^K \leq L^K \qquad (K = 1, \ldots U),$$

$$\sum_{K=1}^{U} - X_i^K = -N_i \qquad (i = 1, \ldots n), \text{ and}$$

$$\text{all } X_i^K \geq 0 \qquad (K = 1, \ldots U), \ (i = 1, \ldots n)$$

The model proceeds to locate consumer-oriented and government activities mostly in relation to population of each zone and nearby zones. Some attempts will be made to develop a hierarchy of such activities with respect to the size of market in order to determine the degree of orientation to residential areas.

A procedure somewhat similar to the residential model is proposed for locating industry, wholesaling, warehousing, etc., using an aging process for firms, their locational budgets, transportation costs, structure cost, and consequent rent-paying ability. Land cost in each zone will also be included, determined by the residential model. Certain types of heavy industry and specialized activities will be located by hand. Alternatively, industries may be distributed according to probabilities to all land areas where the balance between cost and desirability is suitable. Penn-Jersey staff are also considering the adaptation of Karl Dieter's polimetric model, currently under development by TRC on behalf of the BRPP, for possible use as a submodel for locating industries in the Penn-Jersey region.

The output of the model will be a spatial distribution of different types of industrial and commercial activity with the corresponding amount of land used and a spatial distribution of household types, housing types, and land rents. A transportation model for distributing and assigning trips, etc., is part of the regional growth model and uses input from and supplies output to the activity distribution phase. A flow chart of one iteration of the regional growth and transportation model is reproduced in Figure 1, (5).

The model disaggregates the independent variables to a greater extent than most other land-use models. Such disaggregation (for example, into different size households of different income levels) is designed to identify the basis of changing behavioral characteristics of aggregate groups so that these changes can be predicted more reliably when the changed composition of the groups is known. A technique called latent class analysis will be used to stratify households into a small number of homogeneous groups of different locational characteristics in order to avoid stratification on the multitudinous independent variables available from the survey data. No results have yet been reported on calibration or testing of the model. Considerable data are avail-

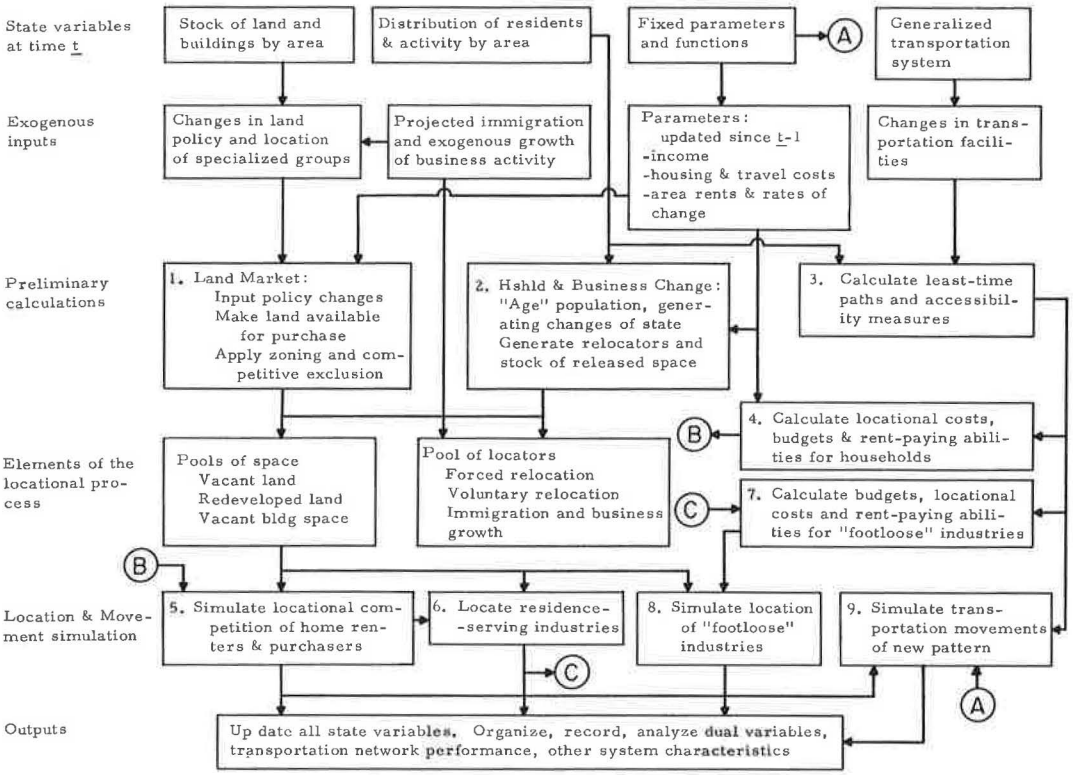

Figure 1. Flow chart: one iteration of regional growth and transportation model.

able for this purpose, including a 1960 survey of 65,000 households with information on household and trip making characteristics. The Penn-Jersey study has available an IBM 1401 data processing system and 7090 or 7094 computers. The computers will be required for application of the regional growth model.

An alternate system, the simplified distribution model (14), is also being developed to distribute land uses to Penn-Jersey districts (comprising about four zones each). This will first distribute industrial development (using the same models as the regional growth model), then households, and then household-oriented activity. Households will be distributed by specifying incremental population by household types for each county, predicting the socio-economic characteristics of each district and assigning the households accordingly. The changes in each district will be established by the initial state (socio-economic class) of the district and the probabilities of transition to other states.

Chicago Area Transportation Study

The Chicago method (29-32) cannot properly be called a mathematical model since much of the procedure involves elements of judgment rather than explicit mathematical relationships. Thus, the results of the forecast are probably not reproducible unless done by the same study team.

The region is divided into zones according to a grid system. These zones are grouped by sector and by ring. A district is a group of zones common to a particular sector and a particular ring. The object of the method is to predict the future year (1980) population and manufacturing employment by zone and the number of acres of each zone to be devoted to major land-use classes—residential, commercial, public open space, manufacturing, transportation (such as airports), streets and alleys, and others.

The procedure followed may be described as land-use accounting where predictions are made on a large area basis and the area figures are used as controls for the totals of the component subareas. The hierarchy is typically region, ring, district, zone. Numerous checks and balances are involved in the procedure and specific account is taken of discontinuities such as large regional shopping centers. Considerable emphasis is placed on internal consistency and the stability and reasonableness of results.

The procedure depends on the following data, which can be considered as input:

1. Base year land occupancy patterns—population, employment, and land use;
2. Existing zoning ordinances and community plans affecting the use of vacant land;
3. Plans for redeveloping the central built-up area; and
4. A forecast of future year total population for the city, the study area, and the metropolitan area and a forecast of total employment by industry type for the study area and the metropolitan area.

In the Chicago study, unlike in Boston, the study area lies within the metropolitan area.

As a first step, the stability of already built-up land over time was examined. It was found that the land already built-up in 1940 was used in much the same way in 1956. On this basis it was decided to use the same population density and proportional use of land in 1980 for all zones built-up in 1956, with the exception of the central area of the city, where the present conditions were substituted by a redevelopment plan of the Chicago Department of City Planning.

Second, all vacant land was classified into the various land-use classes. Public open space was designated according to present conditions and expected open space standards for the future population. Manufacturing land was designated as all land presently zoned for industry, plus tracts listed by the Commonwealth Edison Co., as favorable for industrial sites, plus other sites considered suitable. Area for streets was designated as present street area plus a percentage of usable vacant land. Railroad land, airport land, and trucking warehouse and other nonmanufacturing-industrial land was designated mainly on the basis of trends in the demand for such uses. The remaining usable land was designated as residential and commercial. Local commercial land was designated according to a per capita rate and based on a preliminary estimate of population density. Non-local commercial land was mainly regional shopping centers located according to present plans and estimates of regional requirements.

Population was distributed to the residential land by calculating holding capacity of each zone. This was done by specifying the net residential density for each zone and multiplying by residential acres. A stable pattern of density as a function of distance from the central business district (CBD) was observed between 1940 and 1956 and projected to 1980. Considerable study was made of percent capacity as a function of distance from the CBD at different points in time and in different sectors. A downward sloping curve was observed for all cases and a curve was conjectured for 1980 which would contain the expected population within the study area. Adjustments were made according to characteristics peculiar to each zone.

Manufacturing workers were distributed by a similar procedure of calculating capacities according to available land and densities, and relating percent capacity to distance from the CBD. An attempt was also made to relate workers in each ring to the population in each ring and this was adjusted according to a downward sloping percent capacity relationship with distance from the CBD.

The Chicago method was not tested for accuracy by predicting known values of a present activity using only data on past activity. One of the weak points is the percent capacity curve which, though measurable in 1956, appears to be largely arbitrary in 1980. The parameters were quantified typically by plotting survey data and in some cases data for more than one time period and observing the stability of the relationships over time; a minimal amount of statistical technique such as regression analysis was utilized and much of the adjustment was based on judgment. It appears that most of the computation could be done utilizing accounting-type data processing equipment except possibly for a few curve-fitting procedures where a computer might be of some help.

The accuracy of the procedure depends a great deal on the judgment of the study team. Considerable emphasis has been given to internal checks and alternate ways of making particular predictions as well as to maintaining control totals in aggregate areas.

The procedure of land-use accounting could be used in conjunction with more explicit mathematical prediction techniques in apportioning the land-use distribution to traffic zones from the larger subregions which might reasonably be the basic areas of the mathematical model.

Study by Voorhees, Barnes, and Hansen

This section summarizes the techniques (33-39) used to predict land-use variables in connection with transportation studies carried out in Washington, D. C., and Hartford, Conn.

Traffic prediction models require as input the values of various land-use variables for each zone. In the Hartford Area Transportation Study (35, 36, 38), the following variables had to be predicted: (a) manufacturing employment, (b) service employment, (c) retail trade employment, (d) population, and, as a by-product, (e) number of cars registered.

The percentage of total growth in manufacturing employment occurring in each zone was related to different variables by multiple correlation analysis (linear multiple regression). Ultimately nine were used: (a) highway accessibility to the labor force in the base year, (b) vacant industrial land, (c) tax rate, (d) sewer facilities index (related to system capacity), (e) rail service (subjective rating), (f) water facilities index, (g) travel time to airport, (h) promotion (activity of town in promoting industrial activity) and (i) industrial land close to expressway. It was found that (b) and (d) were of prime importance and (a), (e), and (g) were second in order of importance, though all nine variables were used in the final equation:

$$\text{Growth index} = 12X_1 + 37X_2 + 5X_3 + 34X_4 + 12X_5 + 2X_6 + 19X_7 + X_8 + 5X_9 + 120 \quad (2)$$

$$\text{Zonal growth} = \text{total growth} \times \frac{\text{zonal growth index}}{\sum(\text{zonal growth indices in all zones})}$$

The coefficient and significance levels of the three best equations are shown in Figure 2 (36); R^2 is not stated. Note that the equation of analysis 3 was used even though four of the variables had no significance.

Distribution of growth of service employment was predicted using an arbitrary factor equal to the product of base year highway accessibility and retail employment. Distribution of retail employment growth was made proportional to distribution of population growth. Population growth was distributed according to future year employment accessibility and holding capacity for new development. (The actual form of the relationship is not stated, but is evidently similar to the linear dependence in the manufacturing land relationship.) Adjustments were made on the basis of other factors such as prestige locations and building codes from results of a questionnaire sent to building contractors. Number of cars was made proportional to population times the car ownership. The latter was predicted to vary with income for different residential density groups according to data from studies in other areas of the country.

The growth in the various land-use categories was distributed in a predetermined order using the results of the previous distribution as input to the next one. The order was established on the grounds that certain activities were more alert to changes and should be settled first. The order was industrial employment, population, retail employment (service employment was not mentioned but presumably was predicted after retail). To make a prediction for 1990, two intermediate predictions were made for 1965 and 1975. This accounted for feedback effects, where the results of the future year projection over the first time period are used to describe the base year of the next time period. The zones for which the predictions were made were the 41 cities and

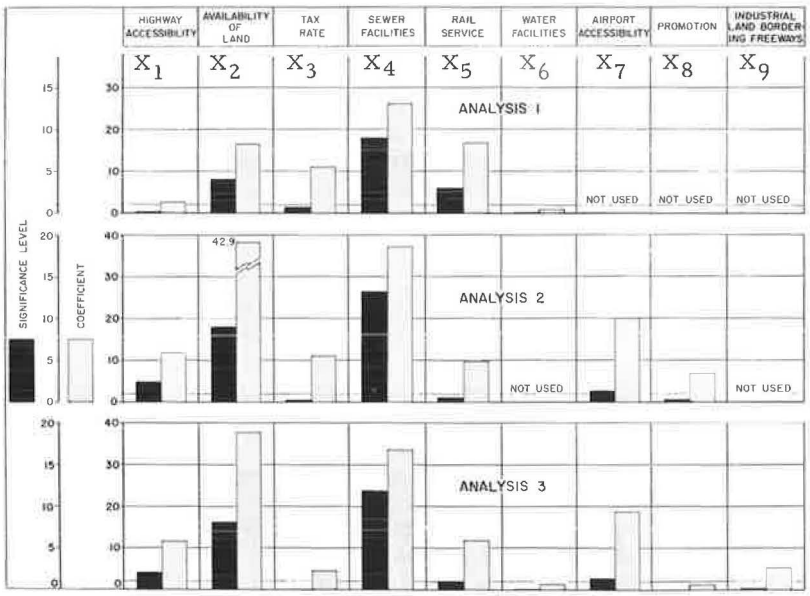

Figure 2. Bar chart of coefficients and levels of significance used in industrial land equations, Hartford area traffic study multiple correlation analysis.

towns in the Hartford area. Most of the calculations involved in applying the model were apparently carried out manually, although it is understood that U. S. Bureau of Public Roads least time path programs were used in determining accessibilities. A number of computer types, including the IBM 1401 and 7090, could be used for this purpose.

The relationships were based largely on the data describing the Hartford area in 1950 and 1958. In the predictions, considerable adjustments were made in the relationships developed, since the mathematical equations did not adequately describe the distribution of growth between 1950 and 1958. For example, it was found that much industrial development not predicted by the model occurred in a certain traffic corridor. Population change was not adequately described by the model; in particular it tended to underestimate the suburbs' share of growth. Retail employment growth did not correspond to the pro rata assignment to population growth because of the development of large shopping centers. Comparisons of estimated vs observed values of population growth distribution and retail employment distribution were presented using known values of future year (1958) industrial employment distribution and population distribution, respectively, as a basis for calculating predicted values. The comparisons were made in the form of maps showing for each year the observed value and predicted value side by side by means of bar charts. No values for R are stated, though they could presumably be calculated from the maps (Figs. 3 and 4, 36). The comparisons would probably show even poorer correspondence if predicted values of industrial employment distribution and population distribution were used respectively in the estimates.

Hansen (34) describes a relationship for predicting the distribution of metropolitan population growth as follows:

$$\frac{P_i}{P_t} = \frac{A_i^{2.7}\, O_i}{\sum_{j=1,n} (A_j^{2.7}\, O_j)} \tag{3}$$

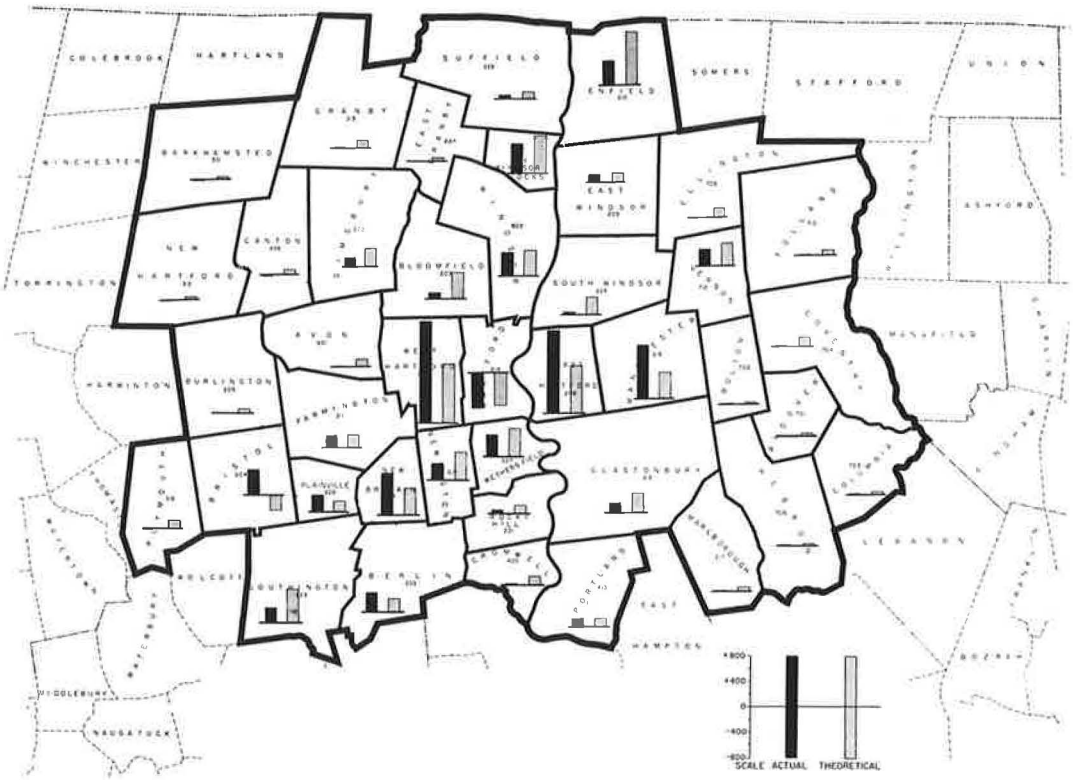

Figure 3. Retail employment change, 1947-1955, by towns, Hartford area traffic study.

where

P_i = increase in residential population zone i;
P_t = total increase in metropolitan population;
O_i = developable land in zone i; and
A_i = accessibility to (future year) employment.

The exponent 2.7 was derived from changes between 1948 and 1955 in the Washington metropolitan area. It is reported that comparison of actual and calculated subregional changes in residential population between 1948 and 1955 revealed that 40 percent of the zones showed calculated values of the changes to be within ± 30 percent of observed changes and 70 percent of the zones showed calculated changes which were within ± 60 percent of observed changes.

A similar procedure was used on behalf of the National Capital Transportation Agency to make predictions of 1980 population throughout the Washington, D. C., metropolitan region (39). The relationship used was the following:

$$P_i = P_t \frac{A_i H_i Q_i}{\sum_{j=1,n} (A_j H_j Q_j)} \tag{4}$$

where

H_j = holding capacity of a zone, based on developable land times probable future net population density; and
Q_j = a special factor taking into account features peculiar to each zone which are felt to encourage or retard growth.

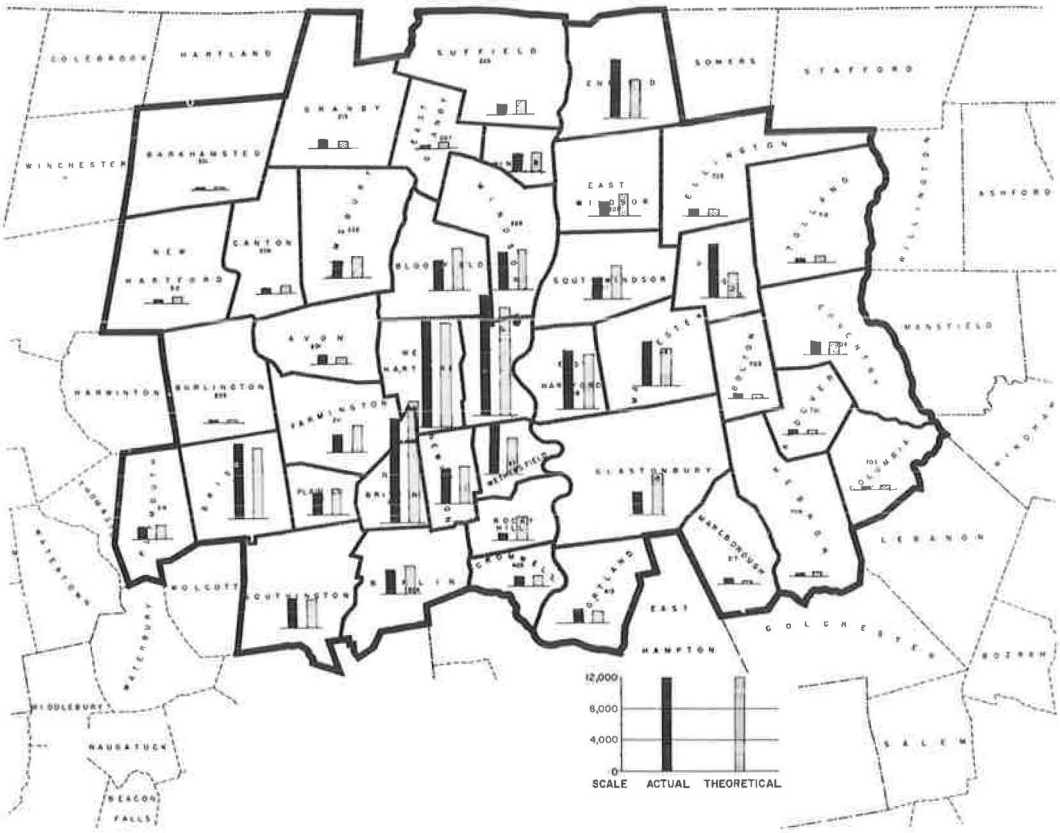

Figure 4. Population change by weighted comparison, 1950-1958, by towns, Hartford area traffic study.

As far as is known, no special calibration or testing steps were carried out as part of this procedure.

Study by Chapin and Weiss

The Chapin-Weiss model for predicting land use has only been formulated (in published form) in nonmathematical terms (40). Tests have been done on different relationships which by linear multiple regression relate two dependent residential land-use variables in zones to various independent variables characteristic of the zonal environment at the same time period and/or at a previous time period.

The linear regression models describe the following two dependent variables: (a) Y_1 = total land in urban use in each zone (in 1960); and (b) Y_2 = dwelling density (or simply "dwellings" since zones are of constant size) in each zone (in 1960). Various combinations of independent variables for two different time periods were tested. The final equations utilizing eight variables were capable of explaining 60.5 percent of the variance in Y_1 but only 44.8 percent of the variance in Y_2. These variables, listed in order of significance, are

1st rank: X_1, marginal land (bad land) not in urban use (1948 and 1960)
X_2, accessibility to work areas (1960)
X_3, assessed value of land (1948)

2nd rank: X_4, travel distance to nearest major street (1960)
X_5, distance to nearest available elementary school (1960)

X_6, residential amenity (1960)
X_7, availability of sewerage (1948)
X_8, zoning protection (not significant 1948)

The mathematical relationships are of the form

$$\begin{aligned} Y_1 &= B_{11} X_1 + B_{12} X_2 \ldots + B_{18} X_8 \\ Y_2 &= B_{21} X_1 + B_{22} X_2 \ldots + B_{28} X_8 \end{aligned} \tag{5}$$

The variables X were chosen by running about 1,400 regressions with different mixtures of 37 variables of two time periods (1948 and 1960) on Y_1 and Y_2 for 1960 in Greensboro, N. C., a city of 200,000 population. The area was divided into 3,980 equal square zones, each 1,000 by 1,000 ft, and values of the variables were specified for each square as numbers from 1 to 9 according to a one-digit coding system. Accessibility to work areas was measured both in terms of travel time and travel distance. Comparative analyses showed no significant difference between time and distance as the basis for determining accessibilities. The criteria used in determining the independent variables were the percentage variance explained by the set of variables (i.e., R^2), and the ratio (t) of the regression coefficient to its standard deviation for each variable. Results are summarized in Table 1 (40). Apparently considerable hand coding had to be done on such a fine grid; a computer was used to calculate accessibilities and the regressions.

Tests were run on another urban area of comparable size and on two smaller towns, with similar results. Apparently no tests were made to determine the degree of fit attained when the parameters (B-values) for one city are used on the other city, or for one time period on the other time period. (X-values for two different years are used in the final equations.) Thus, no indication is given of the stability of the parameters over time. Some questions may also be raised concerning the predictability of X_2, X_5 and X_6 in the future year. According to illustrations presented (40), there appears to be some tendency to underestimate Y_1 and Y_2 in the central areas and to overestimate them in outer areas. The use of this model in prediction may, therefore, be question-

TABLE 1

RELATIVE INFLUENCE OF MIXES OF 14 AND 8 VARIABLES IN EXPLAINING TOTAL LAND IN URBAN USE AND DWELLING DENSITY, GREENSBORO, 1960[a]

Independent Variable[b]	Total Land in Urban Use				Dwelling Density			
	t-Value for Mix		Rank for Mix		t-Value for Mix		Rank for Mix	
	14 Var.	8 Var.	14 Var.	8 Var.	14 Var.	8 Var.	14 Var.	8 Var.
Marginal land not in urban use ('48 and '60)	-15.30	-16.66	2	1	-6.35	-7.39	5	3
Travel distance to nearest major street ('60)	-10.06	-11.61	4	4	-2.08	-5.15	11	6
Availability of sewerage ('48)	5.09	8.65	9	7	2.81	7.24	8	5
Distance to nearest available elementary school ('60)	-8.40	-10.44	6	6	-5.73	-7.29	6	4
Zoning protection ('48)	7.02	3.08	8	8	1.07	-0.12	NS	NS
Assess value ('48)	9.47	13.54	5	3	14.44	18.37	2	1
Accessibility to work areas ('60)	1.56[c]	15.50	NS[c]	2	-0.39[c]	9.49	NS[c]	2
Proximity to nonwhite areas ('48)	2.06	–	12	–	9.58	–	3	–
Proximity to blighted areas ('48)	0.30	–	NS	–	7.44	–	4	–
Total travel distance to high value corner ('60)	2.84	–	11	–	-2.81	–	8	–
Proximity to mixed uses ('60)	23.49	–	1	–	16.08	–	1	–
Distance to nearest playground or recreation area ('48)	-7.07	–	7	–	2.34	–	10	–
Distance to nearest convenience shopping area ('48)	4.40	–	10	–	-5.00	–	7	–
Residential amenity ('60)	11.67	10.93	3	5	-1.94	-4.97	NS	7
Multiple regression coefficient (R)	0.817	0.778			0.727	0.669		
Multiple determination coefficient (R^2)	0.667	0.605			0.529	0.448		

[a]NS = not significant (t-value less than 2).
[b]In order of input.
[c]Significance blocked by strong correlation with several independent variables, particularly those deleted in 8-variable test.

able; the authors point out that it is useful mainly in showing the significant independent variables for their proposed model.

The model (40) is to be a procedure, recursive in time, which distributes an increment of residential development during the first time period to zones of an urban area according to their attractiveness and capacity, reevaluates the attractiveness at the end of the first time period according to the new distribution and the exogenously specified changes, distributes the second increment for the second time period, reevaluates...etc.

The attractiveness of a zone is to be determined by certain priming factors, which are, according to indications of the regression analysis described (excluding those felt not to be predictable), the following: (a) accessibility to work areas; (b) travel distance to nearest major street; (c) distance to nearest available elementary schools; and (d) availability of sewerage. The effect of these priming factors also depends on a certain time lag element.

The development is to be distributed by a Monte Carlo process and conditioned on capacities determined by existing development and by zoning and density restrictions The output is the expected pattern of the distribution and intensity of development in the study area.

The method by which the priming factors are to be translated into attractiveness and thence to probabilities useful in the Monte Carlo simulation is not described in the report, but would appear to involve a technique more elaborate than linear multiple regression. Since the model has not been formulated there are no indications yet of its predictive accuracy or its computational feasibility. The accuracy is, of course, conditioned by the possiblity of predicting future work areas exogenously since these are parts of the priming factors.

GROUP 2—RESEARCH-ORIENTED TECHNIQUES

Rand Corporation

The Rand model (19-28) is not intended as a forecasting tool for predicting land use in particular zones of a metropolitan area or for direct solutions to policy problems, but rather as a framework for research into the relationships between transportation and land use (Fig. 5). As described in general terms (19), the model is recursive in time. It makes a prediction of land use and transportation variables for one time period (as short as 6 months) and uses the results of this prediction (endogenous variables), along with externally specified (exogenous variables), to predict the values for the next time period. The solid lines in the flow chart trace the sequence of steps for one time period and the dashed line represents the changes made in the status variables to be used for the next time period. The prediction starts with an exogenous industry configuration comprising employment levels, distributes this employment over the metropolitan area, and determines land rents over the area. It then calculates worker characteristics in each employment area and, using the rent surface and other variables, finds the residential locations of workers in each work place. Adjustments are then made in the status variables, including the transportation system, before the next prediction.

The actual mathematical functions used for the various calculations are not stated; rather the relationships are stated as $Y = f(x_1, x_2, \ldots, \text{etc.})$, where y and x are dependent and independent variables, respectively. In some cases alternate formulations of the relationships are presented for use where certain classes of one type of variable or the type of data available for deriving parameters lend themselves to an alternate approach. For example, percentage of an industry's workers employed in a zone is predicted for industries of high employee density but percentage of land used for that purpose is predicted for low employment density. Once the input variable (Fig. 5) and the initial values of the status variables have been specified, the model uses the following relationships (in this order) in each zone.

1. Percent of workers employed in a given industry depends on industrial land value, distance from CBD, presence of transportation facilities (for both people and goods

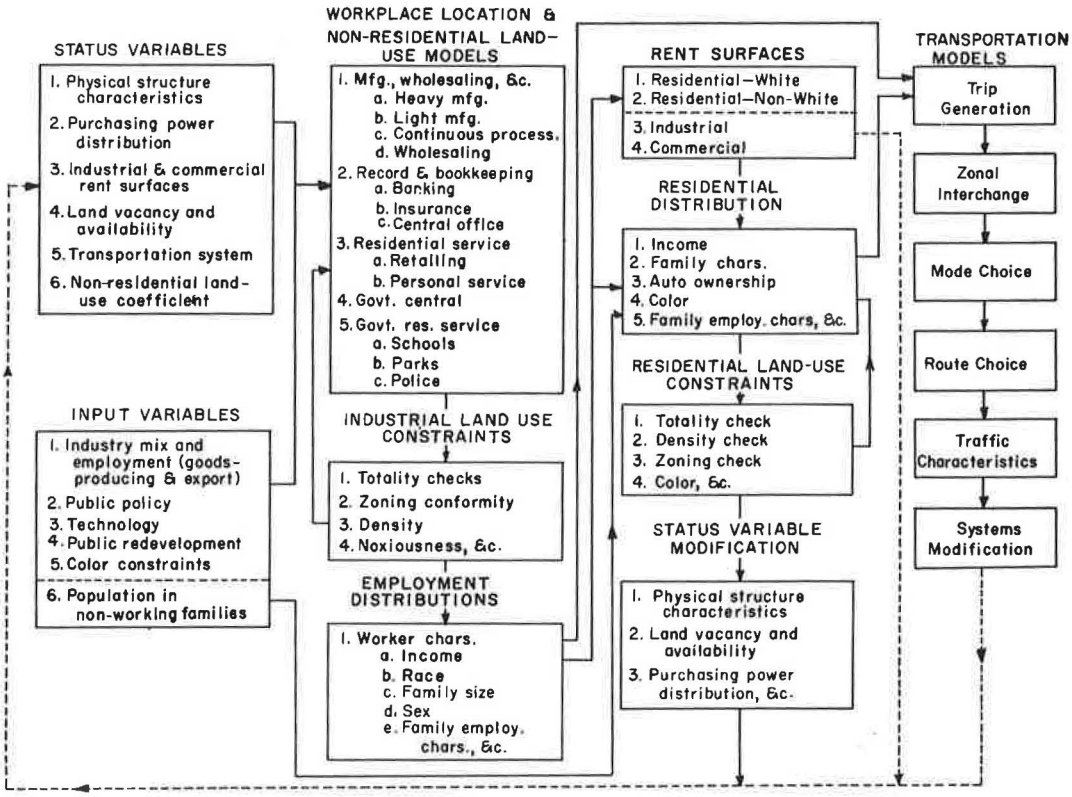

Figure 5. Model for study of urban transportation.

movement), travel time to the urban boundary, vacant industrial land, zoned industrial and commercial land, and workers' supply potential.

2. Number of workers employed (or floor space) in a given commercial or other residentially related activity depends on accessibility to purchasing power, competing commercial floor space in adjacent zones, commercial land value, and zoned industrial and commercial land. A special function which takes into account city population growth is used for CBD retailing.

3. Percent of workers in central office or related functions depends on city female employment, city commercial and manufacturing employment, travel times to city boundary, to major airport, and to best residential district in the city.

4. Government employment functional dependence is not specified but would probably employ similar models.

5. The commercial and industrial distributions are checked for overflows in the constraints and these are assigned to nearest zones with available space.

6. Worker characteristics relevant to location and trip-making decisions are calculated for each employment class and summed over all classes.

7. Commercial, residential and industrial land values are estimated separately and depend in general on the distribution of employment, purchasing power (commercial only), the transportation system and physical characteristics of structures.

8. The residential distribution of workers employed in each zone depends on rents, travel times, zoning restraints and racial prejudice, and the previously calculated worker characteristics. Nonworking households are distributed separately.

9. The residential land-use constraints work much the same way as the industrial land-use constraints.

Total population for each time period is not specified explicitly in the model as an exogenous variable but is implicit in the specification of total employment and of household characteristics of workers (i.e., family size).

Most of the parameters of the model are to be derived by cross-sectional data, that is, data at one point in time for the same metropolitan area. Other relationships will be derived from data in more than one city. The model was intended more as a research tool than as a predictor of land-use variables in particular zones of an area at a particular point in time. As formulated it has a certain drawback for use in prediction: there will probably be a discontinuity between the initial and the first predicted distributions even though the changes during a short time period (6 months) are really slight. This would result because the model predicts the absolute value of the land-use parameters instead of increments of change; in the former case the error is some percentage of the initial value, whereas in the latter the error is a percentage (though perhaps somewhat larger) of only a small increment.

As indicated, the general model (19) does not yet comprise specific forecasting equations. No calibration or testing has, therefore, been possible, nor can it be specified at this time what computational equipment would be required to apply the model. Data requirements may be inferred from the list of general relationships on the previous pages.

Augmenting this general approach, more specific studies (20-28) have been reported by Rand personnel, based on transportation and urban development data from a number of cities. One such report (20) deals with the locational choice of a household and how it varies with location of employment, income, and family size, and how the amount of its residential space consumption is determined. Empirical evidence based on data from 40,000 Detroit households is presented and logically explained on the basis of a simple theoretical model.

A locational rent function, best interpreted as a price per square foot of residential space of a stated quality and amenity, is postulated as decreasing with distance from the center and leveling off in outer zones. The negative of the slope of the locational rent function times the household's residential space can be interpreted as the marginal savings in rent per square foot as residential distance from the center and from the workplace increases. Higher space consumption will give higher marginal rent curves ($q_3 > q_2 > q$, in Fig. 6). A marginal transportation cost function can also be drawn (t(d)). A household will locate away from its workplace and from the center so that its marginal rent savings for a given space consumption will be equal to marginal transport costs and its total budget allocated for transportation plus rent is exhausted. The higher the budget, the greater the space consumed, and the farther away from the center a person employed in an inner zone will reside.

The Detroit area was divided into six concentric zones by distance from the center. Data on work trips and worker and household characteristics were available by origin-destination pair. Consistent patterns were found, in general harmony with the theory. For example, residential distributions tended to be farther out as income increased. When the data were examined by zone of employment it was found that this relationship no longer was as strong for outer employment rings. The relationship was determined by ranking the data for each employment zone by income class (occupation), listing for each residence zone the percentage of workers in that class residing in that zone, and examining the correlation between income and percentage by the Spearman coefficient of rank correlation. An example is shown in Table 2 (20). When the coefficient goes from largely negative to largely positive with increasing distance of the residence ring (as for employment ring 2), it indicates that higher income workers locate further out. The relation-

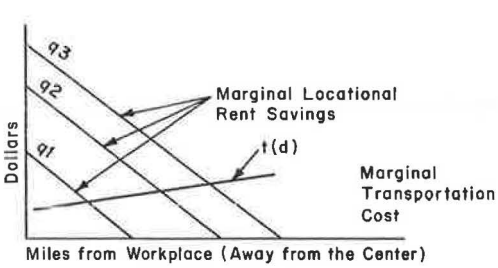

Figure 6.

TABLE 2

RANK ORDER COEFFICIENTS BETWEEN OCCUPATIONS RANKED BY INCOME AND BY RATE OF RESIDENTIAL SELECTION FOR ALL EMPLOYMENT AND RESIDENCE RINGS[a]

Employment Ring	Residence Ring					
	1	2	3	4	5	6
			(a) All Workers			
CBD		-0.86[b]	-0.74[b]	0.69	0.83[b]	0.98[b]
2		(-0.81)[b]	-0.67	0.83[b]	0.89[b]	0.71[b]
3		-0.50	(-0.64)	0.76[b]	0.88[b]	0.76[b]
4		-0.57	-0.71[b]	(-0.48)	0.88[b]	0.89[b]
5		-0.47	0.04	0.29	(-0.33)	0.19
6		-0.33	0.21	0.42	0.90[b]	(-0.61)
			(b) Male Workers Only			
CBD	(-0.69)	-0.81[b]	-0.74	0.57	0.95[b]	0.95[b]
2		(-0.50)	-0.55	0.88[b]	0.88[b]	0.90[b]
3		-0.76[b]	(-0.95)[b]	0.95[b]	0.86[b]	0.83[b]
4		-0.88[b]	-0.57	(-0.17)	0.48	0.19
5		-0.43	0.29	0.29	(-0.38)	0.38
6		-0.40	0.36	0.36	0.81	(-0.40)

[a]Figures are in parenthesis where residence and employment rings are the same.
[b]Differs significantly from zero at the 0.05 level.

ship levels down with outer employment rings since the rent curve is no longer as steep and provides less reason for different location.

Similar empirical evidence and logical explanations too lengthy to explain in this review are offered for locational phenomena associated with family size, space consumption, sex, and race. The model is not formulated in terms of statistically estimated parameters but is used rather as an explanation of the major relationships in the data. Thus, the results are not directly applicable to a land-use forecasting model except as an insight into the behavioral process of choosing a residence, given a workplace.

In two other papers (21, 25), Kain describes the derivation of regression equations to express endogenous variables, relating to workers' choice of residence, travel mode, and travel time, in terms of certain exogenous variables describing the worker characteristics in each employment zone. The exogenous variables are mean income, a proxy variable for locational rent, percentage of males, number of workers in a single family, level of transit service, and family size. Endogenous variables are percentage of workers residing in different types of housing, car ownerships, mean travel time, and percentage driving, using transit, and other modes to get to work. The equations are recursive; i.e., certain endogenous variables are used exogenously in estimating other endogenous variables. The recursive order is type of residence, car ownership, modal split, travel time. Finally, the equations are aggregated and all endogenous variables are expressed as linear functions of the original exogenous variables.

In a further paper (27), Kain reports on the testing of a model similar to one discussed previously (20). The test is carried out by determining whether certain statistics can be explained as the logical outcome of the model. The model explicitly considers several kinds of cost trade-offs available to urban households in maximizing their real income. The first kind is between housing costs and transportation costs; the second is between time and money costs in the journey to work. The model also deals explicitly with the effect of racial discrimination on the operation of the housing market. Thus, further insight is provided into behavioral characteristics in residential location.

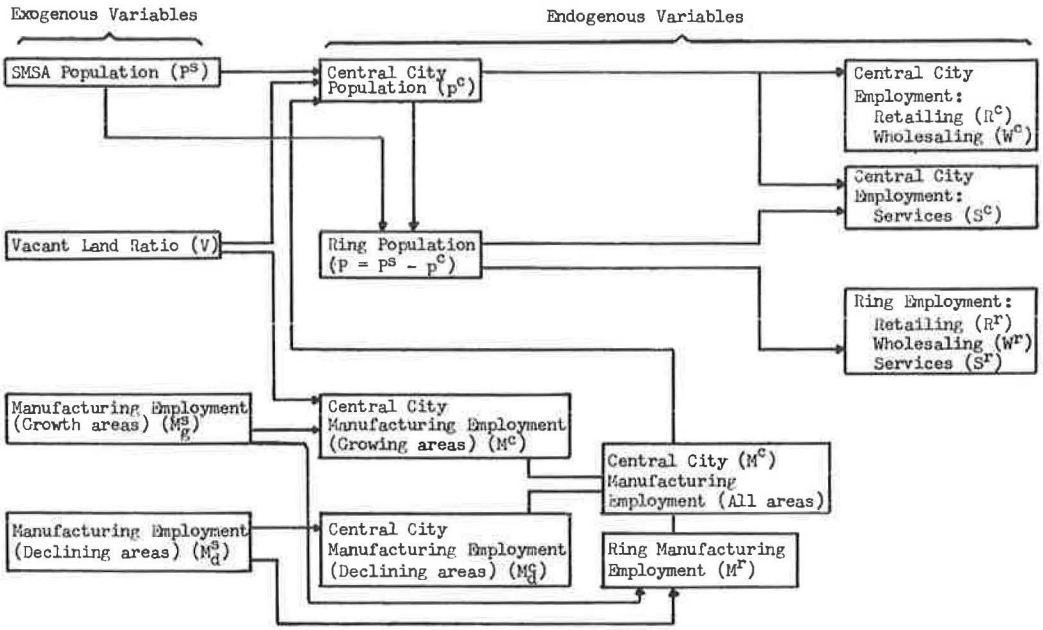

Figure 7. Model of population and employment changes.

In a comparative urban development study (26), Niedercorn and Kain analyze the 39 largest metropolitan areas in the United States with regard to changes in population and types of employment in central cities and ring areas. A recursive model is employed, whereby certain exogenous variables are used to predict some endogenous variables which, in turn, predict other endogenous variables.

The prediction sequence is illustrated in the flow chart in Figure 7 (26). Growing areas are distinguished from declining areas, and the parameters are estimated separately for each equation by regression. The equations specify annual changes both as input and as output variables, so that predictions are performed by predicting the exogenous variables for each year (i.e., changes in manufacturing employment and population for the metropolitan area and ratios of central city vacant land to total central city land area) and integrating the equations over time. Unfortunately, the subdivision is only into two zones, central city and ring, but the model might be used as an order of magnitude check for a prediction model.

Niedercorn and Kain (23) have also used data from 37 of the largest metropolitan areas in the United States to derive a simple model explaining changes in employment in central cities and concentric ring zones of general merchandizing stores (department stores with large numbers of employees per establishment) and food stores (small numbers of employees per establishment). The equations were of the form:

$$E_{kj} = \alpha_{kj} P_j Q_{kj} + \alpha_{ki} P_i Q_{kj} + \alpha_{k_4} \cdot T \qquad (6)$$

where i and j are only two zones, the center city and the ring, and

E_{kj} = employment in industry k in zone j,
P_j, P_i = population in i and j,
Q_{kj} = ratio of employment to retail sales in industry k in zone j,
α_{kj}, α_{ki}, α_{k_4} = parameters, and
T = time trend variable.

The parameters were estimated by least squares fit of the data on change in the variables during two different time periods, 1948 to 1954 and 1954 to 1958, to the first difference of the equations stated previously:

$$E_{kj} = \alpha_{kj} P_j \Delta Q_{kj} + \alpha_{kj} Q_{kj} \Delta P_j + \alpha_{ki} P_i \Delta Q_{kj} + \alpha_{ki} Q_{kj} \Delta P_i + \alpha 4 \qquad (7)$$

where $\alpha 4$ represents the effect of a unit time change. On the whole, the results appear to conform to a priori expectations on employment locations.

Study by Lowry, Rand Corporation

Lowry, in work carried out on behalf of the Pittsburgh Regional Planning Association, proposes a model (46) which is intended for predictions at 5- or 10-year intervals with a maximum (horizon) of 25 years. The proposed model is not recursive and involves nine types of simultaneous equations solved by iteration. Certain activities and the amount of land they use will be distributed exogenously by hand. Those activities treated endogenously must have the following characteristics: (a) demand for their output mostly within the region, (b) high orientation to the regional market for location, (c) market composed of a large number of independent units, (d) historical statistics separable from those of related exogenously located industries. Based on these criteria, exogenously located industries were taken to include all manufacturing, wholesaling, public utilities, research facilities, central administrative offices, government, hospitals, most outdoor recreation, agriculture, and extractive enterprises. This is well over half of the employment of the Pittsburgh region. The model predicts employment in 14 retail and service categories and residential population, and distributes these activities to zones of 1 sq mi. The output for each zone is summarized as follows:

1. Employment in m lines of retail and service trade;
2. Quantity of land in use in m lines of retail and service trade;
3. Number of resident households; and
4. Quantity of land in use by resident households.

The following notation is used in describing the model:

A = area of land (1,000 sq ft);
E = employment (number of persons);
N = population (number of households);
T = index of trip distribution;
Z = constraints;
U = unusable (land);
B = exogenous sector;
R = retail and service sector;
H = household sector;
k = category of establishments within a sector;
m = number of such categories;
i,j = individual tracts of land, defined by the grid system;
n = number of such tracts; and
a, b, c, d, e, f, g = unspecified functions or parameters.

The equations comprising the model are:

$$A_j = A_j^U + A_j^B + A_j^R + A_j^H \qquad (8)$$

$$E^{Rk} = a^k(N^H) \qquad (9)$$

$$E_j^{Rk} = b^k \sum_{i=1}^{n} \left[\frac{c^k N_i^H + d^k E_i}{T_{ij}} \right] \qquad (10)$$

$$E^{Rk} = \sum_{j=1}^{n} b^k \sum_{i=1}^{n} \left[\frac{c^k N_i^H + d^k E_i}{T_{ij}} \right] \tag{11}$$

$$E_j = E_j^B + \sum_{k=1}^{m} E_j^{Rk} \tag{12}$$

$$A_j^R = \sum_{k=1}^{m} e^k E_j^{Rk} \tag{13}$$

$$N^H = f\left(\sum_{j=1}^{n} E_j \right) \tag{14}$$

$$N_j^H = A_j^H . g . \sum_{i=1}^{n} \left(\frac{E_i}{T_{ij}} \right) \tag{15}$$

$$N^H = \sum_{j=1}^{n} N_j^H \tag{16}$$

Subject to the following constraints:

$$\begin{matrix} b^k = 0 \\ b^k > 0 \end{matrix} \text{ for } \sum_{i=1}^{n} \left[\frac{c^k N_i^H + d^k E_i}{T_{ij}} \right] \begin{matrix} < \\ \geq \end{matrix} \begin{matrix} Z^k \\ Z^k \end{matrix} \text{ and}$$

$$g = \frac{Z_j^H}{\sum_{i=1}^{n} \frac{E_i}{T_{ij}}}$$

These nine equations express the following relationships and identities:

1. Total available land in each zone equals the sum of land in each use (Eq. 8);
2. Total regional retail and service employment by category is a function of the number of households in the region (Eq. 9);
3. Zonal retail and service employment by category is proportional to an accessibility measure weighting households and employment access (Eq. 10);
4. Total retail and service regional employment by category equals the sum of zonal employment in the category (Eq. 11);
5. Total employment in each zone equals the sum of exogenously located employment plus employment in the endogenous categories (Eq. 12);
6. Land used by each employment category is proportional to the employment in that category, and the sum of land in each use in each category in a zone equals the total land in use in that zone (Eq. 13);
7. Regional household population is a function of total employment (Eq. 14);
8. Number of households in each zone equals the total area available for residential use times accessiblity to employment times a scale factor g (Eq. 15); and
9. Number of households in the region equals the sum of the number of households in each zone (Eq. 16).

There are also two constraints: (a) a minimum size of establishment in each category such that if calculated employment in a zone does not exceed this minimum, the employment is assigned to other zones; and (b) a maximum residential density for each zone. There are altogether $4n + mn + 2m + 2$ unknowns and the same number of equations, where n = number of zones and m = number of employment categories.

Lowry proposes an iterative technique for solution of the equations but does not present a theoretical or empirical analysis of the conditions of convergence. He expresses some doubts about possible instabilities caused by the constraints. There is no constraint on maximum density of retail establishments, and, theoretically, retail employment can be assigned to a zone without available land. Lowry interprets this as possibly meaning mixed use of multistory buildings.

Considerable attention is given to the discussion of parameters required by the model. These will not be determined simultaneously but by statistical analysis of the individual equations taken separately. All necessary data are obtainable through the census and through surveys carried out by the Pittsburgh Area Transportation Study. These parameters include minimum size of establishment, maximum residential density (from zoning), retail employment as a function of population (from national data), relative weight of residential accessibility and employment accessibility for retail employment (regression analysis), employees per acre for different activities, and ratio of total employment to total number of households. The accessibility measures will use the same parameters as those derived for interaction models describing distribution of the appropriate type of trips for the given activity.

There is probably no unique solution to the land-use pattern as formulated. This depends on the actual sequence of the iterative steps. Since this is not a recursive model which distributes increments of change, there appears to be the usual danger of inconsistency of the initial period prediction with the existing pattern of development. However, Lowry asserts that this type of model is superior for long-run predictions. There is also the common danger that if the constraints are operative during predictive phases, the statistically fitted parameters may begin to lose their meaning by distortion.

Although no testing of the model has been described in the literature, the author has indicated that the model has been calibrated based on 1958 Pittsburgh data. It was then applied using 1958 regionwide control totals as input, and the calculated values of subregional land use were compared with observed 1958 values. No formal statistical evaluation of the deviations has yet been made, but the results of the test are generally encouraging. A description of this test was to be published in the fall of 1963. There are presently no definite plans for applying the model operationally, but the possibility exists that it will be utilized as a planning tool for the Pittsburgh region.

Study by Hansen for the Boston Regional Planning Project

Metropolitan residential extension is defined (51) as "the new occupancy of intrametropolitan open sites by urban housing during specified time periods." Equations are tested to examine independent variables, using data obtained from the U. S. Census and local agencies for the Philadelphia metropolitan region during the periods 1940 to 1950 and 1950 to 1956.

Four independent variables were used to describe conditions in each analysis subregion:

1. Residential settlement (B)—population at the beginning of the period divided by total area;
2. Centrality (C)—distance from the intrametropolitan center;
3. Residence accessibility (M); and
4. Employment accessibility (E).

Three different measures were used to describe the dependent variable, residential extension:

1. Dwelling construction volume (Q)—total number of new units on open land;
2. Dwelling capacity utilization ratio (V)—Q divided by the product of open site area and economical dwelling unit density; and

3. Incremental population density (Z)—population change divided by total area.

Three types of functions were tested:

1. Linear

$$Y = A_0 + A_1 X_1 + A_2 X_2 \ldots + A_m X_m \tag{17}$$

2. Exponential

$$Y = A_0 A_1^{X_1} A_2^{X_2} \ldots A_m^{X_m} \tag{18}$$

3. Power

$$Y = A_0 X_1^{A_1} X_2^{A_2} \ldots X_m^{A_m} \tag{19}$$

where A_i ($i = 0, \ldots, m$) are parameters.

The efficiency of the relationships was investigated according to four criteria:

1. Explanatory precision—ratio of corrected residual error to the mean of the dependent variable distribution for the linear functions, and the corrected residual error for the other functions;
2. Explanatory completeness—corrected coefficients of bivariate (r^2) and multivariate ($\bar{R}^2$) determination;
3. Parametric reliability (C)—ratio of the regression coefficient to its standard error, and
4. Parametric plausibility—conformance of regression coefficient signs to accepted theory.

Several hundred bivariate and multivariate regression equations were tested and their parameters were estimated by least squares. The efficiency of the equations varied greatly. The explanatory precision for equations employing Q tended to be somewhat higher than that of equations employing V or Z; however, Q tended to be poorer than V or Z in explanatory completeness. The use of multivariate regression equations instead of bivariate equations produced varying effects on efficiency but, generally speaking, moderate gains were realized in the explanatory indices at the expense of substantial reduction in parametric plausibility and reliability. The efficiency of the power equations tended to be moderately to substantially higher than that of the linear and exponential forms, which were similar to each other in overall performance. Exponential equations tended to be better than linear ones for all of the independent variables except residential settlement. Some of the results are shown in Table 3.

The subregions tested were in the Philadelphia Standard Metropolitan Statistical Area. The number of subregions in the basic set was 44 with an average size of 78.1 sq mi. A more disaggregated set of 71 subregions was tested and it was found that the explanatory precision and the explanatory completeness decreased moderately and the parametric reliability increased slightly. The question of the best exponents to use in the indices of accessibility was investigated. Accessibility (to employment, for example) was expressed as

$$E_i = \sum_{j=1}^{n} T_j \, (2.5 + D_{ij})^{\theta} \tag{20}$$

where

T_j = the number of jobs in region j;

2.5 = terminal time constant; and

D_{ij} = airline distance to region j from i.

TABLE 3

LEVELS OF r^2 AND R^2 FROM REGRESSION EQUATIONS YIELDING DEVELOPMENT RATE V AS ALTERNATIVE FUNCTIONS OF FOUR DEVELOPMENT FACTORS, 1950-56[a]

Factors	Linear	Exponential	Power
(b) Level of R^2			
Residential settlement (B)	0.7042	0.3401	0.8003
Centrality (C)	0.3728	0.6715	0.6415
Residence access (M)	0.7229	0.7929	0.8183
Employment access (E)	0.6090	0.8183	0.7845
(b) Level of $\bar{R}^2$			
B, C, M, E	0.8023	0.7909	0.8752

[a]N = 44 zones.

It was found that $\theta < 0$ gave superior explanatory efficiency to $\theta < 0$ and that the best values were from 1.5 to 2.0 for the linear and power functions, but from 0.5 to 1.0 for exponential functions.

Residential extension accounted for 88.7 percent of additions to the stock of dwelling units in the Philadelphia metropolitan region between 1939 and 1957. The character of residential extension is homogeneous relative to other measures of population settlement so that regression equations on this variable are likely to be fairly efficient, particularly since the study excluded all portions of the area predominantly in urban use at the beginning of the period and those including mass housing projects. Thus, the author has developed meaningful relationships accounting for a large percentage of residential land-use change but with some loss in generality from not accounting for the rest of the change. For use in prediction, it would be desirable to adapt the equations to conform to a control total for the area and possibly to capacity limitations in each subregion.

A predictive technique is also proposed (51), based directly on the derived regression equations. This technique, for which the use of electronic computing equipment would be desirable, is not operational, as it has not been subjected to historical testing. However, the technique is operable in that it could be applied to the Philadelphia region based on the derived relationships.

Study by Bogue, University of Michigan

Bogue (53) explores some aspects of the hypothesis that great cities or metropolises dominate the social and economic organization of technologically advanced societies. No mathematical models are presented for the prediction or description of metropolitan land use, but there is considerable discussion of geographical distribution of population and economic activity and its relationship to interdependence between metropolitan areas and their hinterlands.

Sixty-seven cities in the United States were selected which had populations over 100,000 in 1940 and were not parts of larger metropolitan areas. The entire United States was then subdivided in such a way that each area was assigned as the hinterland of one of the cities. The areas were subdivided according to distance from the central city and according to 30° sectors. Each sector was put into one of three classes:

1. Intermetropolitan if an intermetropolitan thoroughfare went through it to the central city;
2. Subdominant if not intermetropolitan but if it contained a major hinterland city; and
3. Local if neither of the above.

Patterns of dominance were studied by seeking consistent nonrandom differences in dependent variables for the following four-fold classification of independent variables: distance; sector type; size of metropolitan community; and size of hinterland city. The dependent variables were population density and the following indicators of specialization: per capita retail sales, per capita receipts from services, per capita wholesale sales, and per capita value added in manufacturing, as well as other measures of economic activity. Census data from 1939 and 1940 were used as a basis for the statistical analyses.

One of the major conclusions reached was that central cities tend to control the economic conditions of life in communities surrounding them by a higher degree of specialization in such functions as services and wholesaling and by their ability to foster industrial development in their immediate vicinity by providing favorable combinations of factors or production.

Population density is concentrated above the national average up to points 65 mi from the center in all areas, and up to 165 mi for areas of central city population over 500,000. It was found that central cities specialized in retail, services, and wholesale trade and that larger central cities specialize less in manufacturing than the inner zones of the hinterland and vice-versa for smaller cities. Locations outside the central city but within distances of 35 mi had a low retail sales index, though at longer distances sales kept up with population requirements. With large central cities specialization in services declined steadily with distance; with smaller central cities this decline was more abrupt but rose at a distance of 45 mi. Wholesale trade specialization was extremely concentrated in the central city. Manufacturing specialization is not characteristic of only the central city but extends through the metropolis and inner zones of the hinterland to within 45 mi of the center. In short, every zone in the hinterland was dependent on the metropolis for wholesale trade and services and the outer zones were dependent on the metropolis and inner zones for manufacturing. The central city was more specialized in services and wholesale trade than the principal hinterland cities but not in retailing and manufacturing. With increasing distance from the center, specialization of the principal hinterland cities in retail, services, and wholesaling as well as the hinterland city's trade area tends to increase.

This study suggests some measures which might operationally define the importance of different parts of a region and their relationship to the metropolitan center. In addition, it raises questions about the reasons for the concentration of certain types of activities in certain areas of a region and the implications of changing technology with regard to the continuance of this pattern. It must be kept in mind that the study was done on data of about 23 years ago. The author gave little statistical attention to time trends or to data on interzonal communication and trade, which, though harder to obtain, are the real measures of metropolitan dependence.

Study by Pendleton, University of Pittsburgh

The purpose of this study (56) was to estimate the value that residents of the Washington, D. C., metropolitan area place on highway accessibility to the CBD. The approach was to analyze through multiple regression a cross-section of sales prices of sampled residential properties sold during the first nine months of 1961 with financial assistance from the Federal Housing Administration.

Pendleton presents some information on accessibility which may be useful in land-use forecasting. From data of the Washington metropolitan area it was found that three variables—a job accessibility index (similar to that developed by Voorhees), the 1955 driving time, and the 1959 driving time—were correlated to miles from the CBD or log miles from the CBD, plus a constant term, with a coefficient of determination (R^2) between 0.84 and 0.94. From this it was concluded that accessibility may be measured in minutes, index points, or miles, and that the house value estimates should be roughly similar whatever measure is chosen.

It was found that subdividing the region by sectors and introducing dummy sector variables did not add significant explanatory power to the equations. It was also found that CBD job orientation (percentage residents working in the CBD) was significantly related to distance when the data were examined by sector. The subdivision was made to eliminate the effect of one sector's being more strongly CBD-oriented than another.

Finally it was found, by multiple regression equations relating house price to accessibility and certain house-quality variables, that accessibility, time, and distance from CBD can make a significant difference in selling price, i.e., about $444 more for a house 3 mi out than a house 4 mi out, and $206 more for a house 7 mi out than a house 8 mi out. These relationships exhibited coefficients of determination (R^2) of 0.86 for all three accessibility measures.

Pendleton draws some approximate conclusions based on these relationships concerning the value of job accessibility ($2.33 per hundred index points) and the monetary value of driving time ($0.0126/min). He points out that the latter estimate is considerably lower than generally accepted values and finds it difficult to account for this discrepancy.

The accessibility relationships derived show the importance of the CBD as a focal point for urban development. However, reported sector differences indicate the danger of assuming circular symmetry as a basis for urban models.

GROUP 3—CONCEPTUAL TECHNIQUES

Study by Garrison, Northwestern University

Garrison does not present a regional land-use model in the usual sense but limits himself to a general description of urban simulation (43) and a theoretical discussion of development at a freeway interchange or at several freeway interchanges in a region (44). Much of the latter paper (actually a set of three papers) is devoted to possible types of approaches to the problem and their advantages and disadvantages. The two heuristic models which he discusses in some detail are summarized in the following paragraphs.

The first model deals with industrial and residential land development around a single newly constructed interchange between a freeway and an arterial. The first type of industry (or residential subdivision) is selected by a Monte Carlo process; its lot size is similarly selected. Part of it is then located on vacant land in such a way as to minimize airline distance to the interchange; the rest of the lot area is located contiguously by a systematic process. The Monte Carlo game is repeated until all the land within a square surrounding the interchange is used. Different runs (using different random numbers) are tried and the resulting patterns of development are examined for common characteristics. Garrison has tried this out with a desk calculator using artificial data for the probability distributions used in the Monte Carlo games, but has not presented the results because of the arbitrary nature of the probability distribution.

The second model is an economic one, in which firms are located near different highway interchanges and workers for these firms are drawn from various residential areas according to a linear program which minimizes production cost plus workers' transportation cost subject to four types of constraints: (a) that enough workers are employed in each industry to meet the total demand for its products; (b) that the number of workers from each residential area does not exceed its labor force; (c) that the capacity of the connecting roads between each residence and work area is not exceeded; and (d) that the total amount of land used at each interchange does not exceed what is available there. Certain extensions and generalizations of the program are also discussed.

The notation for use in the linear program is as follows:

a_k = land used per employee in industry k;

i, j = residential and workplace locations (i.e., expressway interchange), respectively;

s_i = total land available at j;

x_{ijk} = number of workers from the ith residential area to the jth working place, in industry k;

k_{ij} = capacity of the route between i and j for the use of the ij movement only;

y_{jk} = output of the kth industry at j;

d_k = demand for the output of the kth industry;

b_k = number of employees required per unit of k output;

e_i = number of persons in the labor force in the ith residential area;

c_k = unit cost of production in the kth industry; and

t_{ij} = the cost per worker of transportation from i to j.

Thus the total cost of production and transport per worker going from i to j employed in industry k is

$$f_{ijk} = t_{ij} - \frac{c_k}{b_k} \tag{21}$$

We can also set

$$Y_{jk} = \frac{1}{b_k} \sum_i x_{ijk} \tag{22}$$

The linear program then becomes

$$\text{Minimize} \sum_i \sum_j \sum_k f_{ijk} x_{ijk} \tag{23}$$

subject to

$$\sum_i \sum_j x_{ijk} \geq b_k d_k;$$

$$-\sum_j \sum_k x_{ijk} \geq -e_i;$$

$$-\sum_k x_{ijk} \geq -k_{ij};$$

$$-\sum_k \sum_i a_k x_{ijk} \geq -s_j; \text{ and}$$

$$\text{All } x_{ijk} \geq 0$$

The model is a static one, designed to predict a settlement pattern at a point in time. The projection of the composition of a city over a long period of time would be broken down into a series of short projections. The model would be used for each short projection to allocate developments required by demands during that period.

Garrison proceeds to interpret the dual of the linear program as a maximization of the value of labor minus certain costs associated with its use, and explains that such a pattern would not emerge in practice because of imperfections in the market. These are mainly minimum wage controls and also the inability to charge tolls on highways. Apparently a more serious objection to such a solution is that it implies that each residential area is connected to each work area by a road of given capacity, thus not accounting for the joint use of facilities by different trips. Also, product distribution costs and transportation costs of raw materials are not considered in the program even though firms are likely to be sensitive to these as well.

Study by Wingo, Resources for the Future, Inc.

Wingo (47, 48) attempts to organize related subjects such as demand for urban space, population distribution, land values, and transportation costs into a theoretical framework which can simulate various phenomena associated with urban growth and spatial characteristics.

Wingo's model (47) consists of the following equations which can be used to calculate the spatial pattern of population and rents if certain simplifying assumptions are made.

1. A space demand curve relating the amount of space per household to the rent of land per square foot:

$$q - \left(\frac{r}{\lambda}\right)^{\frac{1}{\eta}} \tag{24}$$

where

q = quantity of land per household ($1/q$ = density);
r = rent per square foot; and
λ, η = parameters.

2. A relationship which gives cost per trip to the center of the region as a function of distance to the center and total population. This is combined with an identity stating that rent per household at point i equals the difference between annual trip cost at the farthest point of settlement in the region and cost at point i, yielding the equation:

$$R = \varphi (S, N) \tag{25}$$

where

R = rent per household,
S = distance from center, and
N = total population.

3. An integral giving total population enclosed by an area as a function of population densities and the position of its boundaries m:

$$N = \int_0^m \frac{\sigma'}{q} ds \tag{26}$$

where σ' is a parameter depending on the shape of the region.

4. The definition

$$q = \frac{R}{r} \tag{27}$$

which when combined with Eq. 24 yields

$$\frac{1}{q} = \left(\frac{R}{\lambda}\right)^{\frac{1}{\eta - 1}} \tag{28}$$

Thus, the relationships can be summarized as

$$N = \int_0^m \sigma' \left[\frac{\varphi (S, N)}{\lambda}\right]^{\frac{1}{\eta - 1}} ds \tag{29}$$

The assumptions are that all households work at the center, have the same space demand curve, and try to minimize their own transportation cost.

The author elaborates on the transportation cost function and introduces the concept of ingression loss to describe delays incurred when many people converge on one point at the same time in a system of limited capacity. He discusses the economics of the journey to work, the value of time, and the determinants of the space demand function. There is further discussion of the model and some applications when some of the simplifying assumptions are relaxed, but it is evident that gains in generality are offset by losses in descriptive simplicity. Consequently, it would appear that the model cannot now be applied to land-use forecasting except in a fairly gross sense.

Study by Alonso, Harvard University

Alonso (49, 50) describes a simple static model of the market for urban land developed along the lines of the traditional rent theory of von Thuenen which states that in a simple agricultural market where all of his product is sold at one point, the profit realized by a farmer per acre of crop equals the selling price for one acre's crop minus the production cost, minus the transport cost to the market. This profit can be capi-

talized by the landowner in the form of rent. The farther from the market, the more costly the transport, thus the less the profit, or rent. The curve for rent as a function of distance from the market, therefore, has a negative slope and a maximum value at zero distance from the market. When two or more crops must be grown, each has its own bid rent curve. Those crops with the steepest bid rent curves (largest transport cost per unit distance per acre of crop) tend to occupy the more central locations.

This analysis is extended by Alonso (50) to the location of business firms where not only the transport costs and other operating costs but also the volume of business changes with distance from the center. Residential households behave in a somewhat similar manner as they maximize their satisfaction in different locations.

If a curve of actual land prices, the equilibrium land rents, is given, an individual chooses his location as follows. He has a set of his own bid rent curves which differ by constants in the Y direction (parallel curves). Each curve represents a different level of accomplishment, profit for the businessman and satisfaction for the household. The lower curve is a higher level since he has to pay less for rent. The point of location and the level of accomplishment are determined by the point where this family of curves is tangent to the land value curves (Fig. 8). The point of intersection will depend roughly on the steepness of the bid rent curves. Wealthy people have less steep bid rent curves than poor people since they can afford more land and, therefore, the transportation cost per mile per acre of land (which determines steepness) is less. Thus, the wealthier people will tend to locate farther out.

Here the author's argument appears to be only one possible explanation. Others consider the dynamics of the market. For example, the poor will live close to the center since they occupy the deteriorated structures where redevelopment is too expensive; open land farther out is the only place where it is economical to build houses for the newly wealthy or moderately wealthy.

Alonso's model is purely theoretical and is developed for the simple case of a one center city with transportation in all directions. He hopes it will be useful in providing a logical structure for econometric models used for prediction.

Study by deCani, University of Pennsylvania

Three stochastic models are presented (54): a pure migration model; a birth, death and migration model; and a predator-prey model. All are formulated using differential equations and associated techniques for their solution.

In the first model an area whose total population is constant consists of two zones, A and B. Population migrates from one zone to the other at different rates proportional to the population of the other zone. Expressions are derived for the mean and variance of each zone's population and it is shown how both the mean and variance approach finite limits as time goes on.

In the second model, each zonal population not only migrates to the other in a fashion similar to the first model but also reproduces and dies off at different rates proportional to population in the same zone. In this case, populations grow exponentially and the initial population distribution has less and less of a percentage influence on the mean value as time goes on.

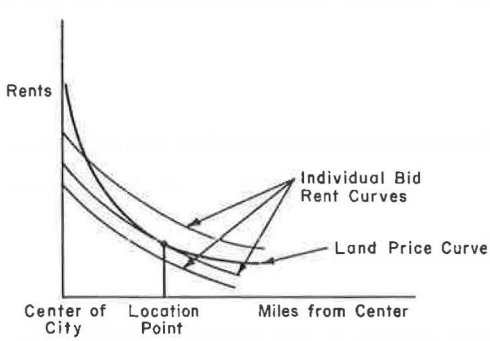

Figure 8.

In the third model, a region has two populations, each of which has different birth and death rates as in the second model, but one of which tends to drive the other population from the scene. It is shown how the mean value of the second population tends to zero in a finite time while the first population continues to grow.

The author indicates how such models can be solved explicitly. However, it is evident that the explicit solutions become

increasingly difficult as complications are introduced into the assumptions about the way in which the populations interact.

Study by Horwood, University of Washington

Horwood's work (55) represents an application of a highly idealized analytic formulation to describe a smoothly varying population density surface for urban areas. An equation is presented which describes population density as a function of radius and sector with respect to the city center. Population density is assumed to vary as the inverse square of distance from the center and as a periodic function of an angle measured with the center as vertex and a major arterial as the axis. The differential equation is of the form:

$$\frac{dP}{dA} = \rho = \text{population density} \tag{30}$$

and

$$\rho = \frac{K}{r^2} = \frac{1}{2}\left[\frac{K'}{r^2} + \frac{K}{r^2}\right] + \left[\frac{K'}{r^2} + \frac{K}{r^2}\right] \cos 4\theta \tag{31}$$

where

r = distance from the center of the city;

$\frac{K}{r^2}$ = characteristic of the maximum density in a circumferential direction; and

$\frac{K'}{r^2}$ = characteristic of the minimum density in a circumferential direction.

There are four circumferential peaks each at a major arterial. An integral is stated (but not evaluated) giving the total population between two angles θ_1 and θ_2 and two arcs of radii R_1 and R_2. The total population within the ring bounded by $r = R_1$ and $r = R_2$ is given by:

$$P = \pi\left[\frac{1}{R_1} - \frac{1}{R_2}\right]\left[K' - K\right] \tag{32}$$

No empirical or theoretical basis is presented for the equation, which appears to serve only as an abstract description of urban space.

COMMENTS AND CONCLUSIONS

The various land-use models reviewed may be classified according to the following functional characteristics:

1. Whether the output variables characterize the state of the system in each zone (e.g., population, employment, and total land in each use) or changes in the state of the system over a certain period of time (by addition the state of the system can then be obtained);
2. The presence or lack of control totals (such as total regional population in future years) which are predicted exogenously;
3. The presence or lack of zonal capacities, determined exogenously, which must not be exceeded;
4. The dependence of one set of predictions on another previously determined set for the same time period (e.g., the prediction of population change which is then used to predict change in retail employment);
5. Time recursiveness, as opposed to a system which produces a direct forecast for a target year 20 or 30 years in the future with no intervening stepwise forecasts;
6. The degree to which the model is made to conform to economic rent theory in the determination of a locational pattern;

7. The degree of stratification or disaggregation which is employed in the solution (e.g., the prediction of the location of different types of households on the grounds that they either have different trip characteristics or different locational characteristics);
8. The extent to which parameters of the model are determined simultaneously in one statistical fit, as compared to their individual determination by a number of individual statistical analyses; and
9. The extent to which accessibility measures play a part in the model.

Some general observations may be made on the merits and disadvantages of these characteristics of the reviewed models.

1. The difficulty with predicting the state of the system as opposed to increments of change is that short-term projections could be in great error. For example, if we wanted to forecast the system a year from now we would do well to utilize information on the present state of the system and make small changes; a model which works from first principles would tend to introduce its statistical errors into the projections and probably produce results considerably different from those of today. However, as longer term projections are made, the present state of the system is less of an indication of the future state, and such projection may have some merit, particularly in a very gross sense for large areas.
2. Most of the models rely on an exogenous prediction of total manufacturing employment in the region. Some models also use this as a predictor of total population, which then serves as another control total. Retail employment usually does not have a control total and depends on population of individual zones. There is usually no mechanical problem with control totals since it is only a question of adjusting in one step the individual zonal figures by a factor to make their total equal the exogenous control total.
3. Most of the models make some adjustments for zonal capacities of population and employment. This is usually in the form of past observed values for population per residential acre times the number of residential acres, and similar figures for employment. The population density figures and residential acres are often set by existing zoning and subdivision control regulations. The problem here is that zoning laws are subject to change, especially under demand pressures, and to set an absolute inviolate capacity limitation requires some assurance of its applying in the future. Application of zonal capacity restraints usually involves some type of iterative procedure unless the predictive technique makes use of a mathematical function which asymptotically approaches a maximum value. A function of this nature, such as the hyperbolic tangent function proposed by Donald Hill of TRC for this purpose, may be used to simulate development of one or more land-use categories subject to zonal capacities, without the requirement of an iterative procedure. Zonal capacities are essential to the forecasting process, but they have to be recognized as assumptions about governmental regulations which may change in the future.
4. There is a tendency in the various models to predict manufacturing employment and land use first and to use this information as input to the population distribution phase which, in turn, is used as input to the retail employment distribution phase. There appears to be a definite logic in distributing retail employment after population; indeed, some empirical evidence is available on the lag of retailing establishment decentralization after the decentralization of population in the RAND study of food and general merchandizing employment (23). The arguments for predicting manufacturing before population are not so convincing. It is argued that there are only a limited number of locations suitable for certain types of industries and that these are relatively independent of population. This is held to follow from the fact that travel is becoming faster and that industry will not have much trouble recruiting workers wherever it locates, as long as it can pay a good wage. However, Penn-Jersey proposes first to predict residential land use and then to use the residential rent surface as an input into industrial land-use distribution.
5. It seems to be generally recognized that direct data flow from the end of one forecasting period to the beginning of the next is necessary, if not within the model, then at least by some type of manual updating. It is necessary because the inertia of

an urban system is such that any forecast which disregards the immediately previous state of the system may predict sudden and perhaps oscillatory changes in land use which are not reflected in real life.

6. The equation systems involved in economic rent theory become rather complex if they are not idealized to the point of having only one or two centers of employment and one or two homogeneous population groups. Even the Penn-Jersey linear program version of rent theory utilizes a set of simplifying assumptions (e.g., amenity of different land parcels for different household types) whose statistical foundations could be rather dubious; in addition, the Penn-Jersey linear program approach has been adapted to predict increments of growth instead of the total settlement pattern, as the formal rent theory model would have it do. The linear programming approach is, in general, questionable as a simulating technique because of its property of maximizing or minimizing some aggregate function (such as total rent-paying ability of all residential locators). Economic theory holds that each individual is trying to maximize his economic position, so that a linear program cannot be said to simulate locational behavior unless it can be shown that the sum of all actions to maximize individual economic position is synonymous with a maximization of the aggregate economic position. Such a relationship does not seem to be readily demonstrable for residential or industrial locators; it is difficult to evaluate the possible effects of this on the predictive realism of a linear programming model. Land-use models based on linear programs may also be questioned because of their reliance on economic motivations and their consequent exclusion of all input variables which cannot be meaningfully translated into economic terms. This limitation may exclude some variables (for example, aesthetic considerations) which have a significant effect on urban development.

7. The desirable degree of stratification of input variables depends to some extent on the type of model being used. Generally, the more behavioral the model (i.e., the more it attempts to simulate actual locational decision-making processes), the more stratification is required, since different classes of a certain variable (e.g., residential population) exhibit quite different locational tendencies. A limit on useful stratification is, of course, reached when stratified groups become so small as to be statistically unstable. A possible means of overcoming this problem is to regroup the stratified classes, by techniques such as latent class analysis or factor analysis, into fewer classes which behave alike functionally. Data availability may also be a limit on the degree of stratification possible. Similar considerations affect the degree of output variable stratification, although the overriding consideration here is the use for which model output is required.

8. None of the 14 models described in this paper comprises a fully integrated formulation which allows the simultaneous forecasting of all urban variables pertinent to regional planning studies. That is, each of the models is either one submodel, dealing with one set of variables such as residential, industrial, or commercial activities, or a number of such submodels applied serially to obtain the desired output values. An attempt to apply any set of submodels for comprehensive urban forecasts suffers from two weaknesses: (a) assumptions must be made about which submodels should be run first, i.e., which variables are primary and which are secondary in locational characteristics; and (b) relationships and parameters must be determined separately for each submodel, leading to questions (which are difficult to answer) concerning their reliability when applied in concert. Therefore, there appears to be some advantage in a model which handles all variables simultaneously and thus allows the derivation of a self-consistent set of parameters. Such a model may, however, suffer other difficulties such as parametric instability and difficulty of interpretation. Although it is a laudable goal, the fully integrated model may prove, therefore, to be difficult to achieve as an operational technique.

9. All of the models reviewed show some dependence on accessibilities, whether explicitly in terms of travel times to various types of activities in other zones, or implicitly in terms of distance to the center of the urban region. The use of distance as a measure of accessibility precludes model sensitivity to changes or proposed changes in transportation systems, which generally have little effect on travel distances but may have profound effects on travel times, costs and convenience levels. A general

disadvantage of concentric ring approaches is that they usually require ad hoc adjustments to account for manifest circular asymmetries in most urban regions.

A number of relationships derived with respect to the reviewed techniques show strong evidence of general applicability. These include the dependence of rents and transportation costs on residential locational decisions, the tendency of retail growth to follow fairly closely behind residential growth, and in general, the effects on various land-use categories of accessibilities to other land-use categories.

Based on the foregoing considerations, it is felt that a number of the concepts noted in the reviewed techniques should be studied for possible inclusion in the model developed for BRPP planning analyses of the Boston region. These include a model which forecasts changes rather than absolute values; the use, wherever possible, of exogenously determined regionwide control totals; the inclusion of a mechanism for simulating the effect of zonal holding capacities; a model which can deal with all variables simultaneously, based on relationships which have been calibrated simultaneously rather than separately; a time-recursive model using fairly short (of the order of 2 to 10 years) forecasting periods; a model containing as many behavioral relationships as may feasibly be used without an inordinate amount of mathematical complexity and/or data stratification; a fairly high degree of data stratification in initial analyses followed, wherever possible, by regrouping of data classes into fewer functionally similar groups; and a model which is sensitive to accessibilities as measured by travel times by all major travel modes, and possibly by travel costs.

Although none of the reviewed techniques provides a comprehensive model framework having all the desired properties mentioned, they represent a fund of experience and insight to draw on during the current Boston land-use model development project.

REFERENCES

1. Stevens, B. H., and Coughlin, R. E. A Note on Inter-Areal Linear Programming for a Metropolitan Region. Jour. of Regional Sci. Assoc., Spring 1959.
2. Herbert, J. D., and Stevens, B. H. A Model for the Distribution of Residential Activity in Urban Areas. Jour. of Regional Sci. Assoc., Fall 1960.
3. Herbert, J. D., and Stevens, B. H. A Model for the Distribution of Residential Activity in Urban Areas. Penn-Jersey Transp. Study, PJ Paper No. 2, 1960. 57 pp.
4. Harris, Britton. Some Problems in the Theory of Intra-Urban Location. Penn-Jersey Transp. Study, PJ Paper No. 3, Apr. 30, 1961. 30 pp.
5. Almendinger, V. V. Topics in the Regional Growth Model: 1. Penn-Jersey Transp. Study, PJ Paper No. 4, Apr. 1961. 22 pp.
6. Harris, Britton. Regional Growth Model Activity Distribution Sub-Model. Penn-Jersey Transp. Study, PJ Paper No. 7, June 6, 1961. 12 pp.
7. Harris, Britton. Computation of Accessibility Measures. Penn-Jersey Transp. Study, PJ Paper No. 9, Aug. 14, 1961. 2 pp.
8. Harris, Britton. PJ Area Systems. Penn-Jersey Transp. Study, PJ Paper No. 14, Feb. 8, 1962. 10 pp.
9. Regional Data. Penn-Jersey Transp. Study, PJ Paper No. 18, July 26, 1962. 8 pp.
10. Harris, Britton. Linear Programming and the Projection of Land Uses. Penn-Jersey Transp. Study, PJ Paper No. 20, Nov. 1962. 30 pp.
11. Harris, Britton. Experiments in Projection of Transportation and Land Use. Penn-Jersey Transp. Study, Traffic Quarterly, pp. 305-319, Apr. 1962.
12. Harris, Britton. The General Scheme of Penn-Jersey Work. Penn-Jersey Transp. Study, P-J Program Rev. Memo No. 1, Feb. 20, 1963. 8 pp.
13. Harris, Britton. The Penn-Jersey Regional Growth Model. Penn-Jersey Transp. Study, P-J Program Rev. Memo No. 2, Feb. 26, 1963. 10 pp.
14. Bruck, H. W. The Simplified Distribution Model. Penn-Jersey Transp. Study, P-J Program Rev. Memo No. 5, Mar. 20, 1963. 10 pp.
15. Exploding Land Development Measured by PJ. Penn-Jersey Transp. Study, PJ News, Feb.-Mar. 1963. 4 pp.

16. A List of Standard Area Data Files. Penn-Jersey Transp. Study, Spec. and Ref. Manual, Pt. 1, Apr. 1963. 6 pp.
17. Variables Lists. PJ Star Files Ref. Manual, Pt. 2. 72 pp.
18. Loewenstein, L. K. Community and the Cost of Housing in Philadelphia. Penn-Jersey Transp. Study, Traffic Quarterly, pp. 302-319, Apr. 1963.
19. Kain, J. F., and Meyer, J. R. A First Approximation to a Rand Model for Study of Urban Transportation. Rand Corp., Nov. 1961. 42 pp.
20. Kain, J. F. The Journey-to-Work as a Determinant of Residential Location. Univ. of California, Dec. 1961. 40 pp.
21. Kain, J. F. A Multiple Equation Model of Household Locational and Tripmaking Behavior. Rand Corp., April 1962. 67 pp.
22. Kain, J. F. A Report on an Urban Transportation Model, Some Progress and Some Problems. Rand Corp., June 1962. 26 pp.
23. Niedercorn, John H., and Kain, J. F. Changes in the Location of Food and General Merchandise Store Employment Within Metropolitan Areas, 1948-1958. Rand Corp., Aug. 1962. 23 pp.
24. Zwick, C. J. Models of Urban Change: Their Role in Urban Transportation Research. Rand Corp., Oct. 1962. 17 pp.
25. Kain, J. F. A Contribution to the Urban Transportation Debate: An Econometric Model of Urban Residential and Travel Behavior. Rand Corp., Nov. 1962. 37 pp.
26. Niedercorn, John H., and Kain, J. F. Suburbanization of Employment and Population, 1948-1975. Rand Corp., U. S. Air Force Acad., Jan. 1963. 38 pp.
27. Kain, J. F. Commuting and the Residential Decisions of Chicago and Detroit Central Business District Workers. Rand Corp., Apr. 1963. 43 pp.
28. Niedercorn, John H., and Hearle, Edward F. R. Recent Land-Use Trends in Forty-Eight Large American Cities. Rand Corp., June 1963. 37 pp.
29. Creighton, R. L., Carroll, J. D., and Finney, Graham S. Data Processing for City Planning. Spec. Issue of AIP Jour., pp. 96-103, May 1959.
30. Hamburg, J. R., and Creighton, R. L. Predicting Chicago's Land-Use Pattern. Spec. Issue of AIP Jour., pp. 67-72, May 1959.
31. Chicago Area Transportation Study—Volume II—Data Projections. July 1960. 133 pp.
32. Hamburg, J. R., and Sharkey, Robert H. Land Use Forecast. Chicago Area Transp. Study Paper No. 3.2.6.10, Aug. 1, 1961.
33. Voorhees, A. M. The Nature and Uses of Models in City Planning. Spec. Issue of AIP Jour., pp. 57-60, May 1959.
34. Hansen, W. G. How Accessibility Shapes Land Use. Spec. Issue of AIP Jour., pp. 73-76, May 1959.
35. Barnes, Charles F., Jr. Integrating Land Use and Traffic Forecasting. Highway Research Board Bull. 297, pp. 1-13, 1961.
36. Hartford Area Traffic Study Report, Vol. 1, Connecticut Highway Dept., July 1961. 65 pp.
37. Voorhees, A. M. Urban Growth Characteristics. Urban Land News and Trends. Vol. 20, Dec. 1961. 8 pp.
38. Voorhees, A. M., Barnes, Charles F., Jr., and Coleman, Frances E. Traffic Patterns and Land-Use Alternatives. Highway Research Board Bull. 347, pp. 1-9, 1962.
39. Appendix to November 1, 1962 Report to the President—Volume III—Traffic Forecasting. National Capital Transp. Agcy., Nov. 1, 1962.
40. Chapin, F. Stuart, Jr., and Weiss, Shirley F. Factors Influencing Land Development. Univ. of North Carolina, Aug. 1962. 101 pp.
41. Chapin, F. Stuart, Jr., and Weiss, Shirley F. Urban Growth Dynamics. Univ. of North Carolina, 1962. 475 pp.
42. Chapin, F. Stuart, Jr. Taking Stock of Techniques for Shaping Urban Growth. Univ. of North Carolina, AIP Jour., pp. 76-86, May 1963.
43. Garrison, W. L. Toward Simulation Models of Urban Growth and Development. Proc. of IGU Symp. in Urban Geography, Lund, pp. 91-108, 1960.

44. Garrison, W. L. Land Uses in the Vicinity of Freeway Interchanges—Models of Land Use Developments and Related Traffic Flows. Northwestern Univ., May 1961. 74 pp.
45. Lowry, Ira S. Residential Location in Urban Areas. Ph. D. thesis at Univ. of California, Berkeley, 1959.
46. Lowry, Ira S. Design for an Intra-Regional Locational Model. Pittsburgh Regional Planning Assoc. Economic Study of the Pittsburgh Region, Working Paper No. 6, Sept. 1960.
47. Wingo, Lowdon, Jr. Transportation and Urban Land. Resources for the Future, Inc., 1961. 132 pp.
48. Wingo, Lowdon, Jr. An Economic Model of the Utilization of Urban Land. Papers of Regional Sci. Assoc., Vol. 7, 1961. 16 pp.
49. Alonso, William. A Model for the Urban Land Market: Location and Densities of Dwellings and Businesses. Unpublished Ph.D. dissert. at Univ. of Pennsylvania, 1960.
50. Alonso, William. A Theory of the Urban Land Market. Papers of Regional Sci. Assoc., Vol. 6, 1960. 10 pp.
51. Hansen, Willard B. Residential Extension in a Metropolitan Region: A Regression Analysis of Subregional Development Rates in the Philadelphia Area During the 1940-50 and 1950-55/56 Periods. Ph.D. dissert. at Univ. of Pennsylvania, 1961.
52. Hansen, Willard B. An Approach to the Analysis of Metropolitan Residential Extension. Jour. of Regional Sci., Vol. 3, 1961. 19 pp.
53. Bogue, Donald J. The Structure of the Metropolitan Community. Ph.D. thesis at Univ. of Michigan, 1948.
54. de Cani, John S. On the Construction of Stochastic Models of Population Growth and Migration. Jour. of Regional Sci. Assoc., Winter 1961.
55. Horwood, Edgar M. A Three-Dimensional Calculus Model of Urban Settlement. Highway Research Board Bull. 347, pp. 143-146, 1962.
56. Pendleton, William C. Relation of Highway Accessibility to Urban Real Estate Values. Highway Research Record No. 16, pp. 14-23, 1963.
57. Selected Bibliography on Land Use and Traffic Models. Spec. Issue of AIP Jour., p. 104, May 1959.
58. Berman, B. R., Chinitz, Benjamin, and Hoover, Edgar M. Projection of a Metropolis, Technical Supplement to the New York Metropolitan Region Study. Harvard Univ., 1960. 119 pp.
59. Blumenfeld, Hans. Tidal Wave of Urban Expansion. AIP Jour., Winter 1954.
60. Blumenfeld, Hans. Are Land Use Patterns Predictable? Spec. Issue of AIP Jour., pp. 61-66, May 1959.
61. Cohen, M. H. The Relative Distribution of Households and Places of Work: A Discussion of the Paper by J. G. Wardrop. In Theory of Traffic Flow, pp. 79-84. Univ. of Chicago, 1960.
62. Colwell, Robert C. Interactions between Transportation and Urban Renewal Growth. Housing and Home Finance Agency, Jan. 1963. 17 pp.
63. Clark, C. Urban Population Densities. Jour. of Royal Statistical Soc., Series A, Vol. 114, Pt. 4, 1951. 6 pp.
64. Curry, L. The Geography of Service Centers with Towns: The Elements of an Operational Approach. Proc. of IGU Symp. on Urban Geography, Lund, pp. 31-54, 1960.
65. Dacey, M. F. Analysis of Central Place and Point Patterns by a Nearest Neighbor Method. Proc. of IGU Symp. on Urban Geography, Lund, pp. 55-76, 1960.
66. Davis, H. S. The Economic Impact of the Proposed Locations for Interstate 40 in the Holbrook Area. Presented at 11th Ann. Conf. on Roads and Streets, Arizona Univ., pp. 80-85, Nov. 1961.
67. Davis, J. Tait. Parkways, Values and Development in the Washington Metropolitan Region. George Washington Univ., 1962. 19 pp.
68. The Future Metropolis. Essays originally published in Daedalus, Jour. of Amer. Acad. of Arts and Sci., 1961. 253 pp.
69. Fagin, Henry. Transportation Systems Planning as an Influence on Urban Land Uses. In The Dynamics of Urban Transportation. 12 pp.

70. Foote, Nelson, Abu-Lughod, Janet, Foley, Mary M., and Winnick, Louis. Housing Choices and Constraints. McGraw-Hill Action Series, 1960. 450 pp.
71. Fuchs, Victor R. The Determinants of the Redistribution of Manufacturing in the United States Since 1929. New York Univ., Rev. of Economics and Statistics, pp. 167-178, May 1962.
72. Futterman, Robert A. The Future of Our Cities. 1961. 348 pp.
73. Gibbs, J. P. Urban Research Methods. 1961. 576 pp.
74. Gruen, Victor. Cities to Doughnuts. Victor Gruen Assoc., Nat. Civic Rev., Mar. 1961. 3 pp.
75. Hadfield, S. M. An Evaluation of Land-Use and Dwelling-Unit Data Derived from Aerial Photography. Highway Research Board Bull. 347, pp. 121-132, 1962.
76. Hancock, Macklin L. Transportation and Organic Urban Design. Project Planning Associates Ltd., Traffic Quarterly, pp. 5-23, Jan. 1963.
77. Hooper, J. W., and Zellner, Arnold. The Error of Forecast for Multivariate Regression Models. Econometrica, pp. 544-555, Oct. 1961.
78. Hoover, Edgar M. Location Theory and the Shoe and Leather Industries. Cambridge, Mass., Harvard Univ. Press, 1937.
79. Hoover, Edgar M. Location of Economic Activity. New York, McGraw-Hill, 1948.
80. Hoover, Edgar M., and Vernon, Raymond. Anatomy of a Metropolis. 1959. 345 pp.
81. Howlett, B. E. Determining Urban Growth and Change from Aerial Photograph Comparisons. Puget Sound Regional Transp. Study. Jan. 1963. 11 pp.
82. Hoyt, H. The Effect of the Automobile on Patterns of Urban Growth. Homer Hoyt Assoc., Traffic Quarterly, pp. 293-301, April 1963.
83. Huff, David R. A Topographical Model of Consumer Space Preferences. Papers of Regional Sci. Assoc., Vol. 6, 1960. 16 pp.
84. Isard, Walter. Location and Space Economy. M. I. T. Press, 1956. 350 pp.
85. Isard, Walter. Methods of Regional Analysis. New York, John Wiley and Sons, 1960. 784 pp.
86. Isard, Walter, and Reiner, Thomas A. Regional Science and Planning. Regional Sci. Assoc. Papers, Vol. 8, pp. 1-36, 1962.
87. Levin, M. R., and Grossman, David A. Industrial Land Needs Through 1980. Land-Use Report 2, Planning Services Group, May 1962. 87 pp.
88. Levin, M. R., and Abend, N. A. The Boston Regional Survey. Traffic Quarterly, pp. 177-192, April 1963.
89. Losch, August. The Economics of Location. Yale, 1954. 508 pp.
90. Mayfield, Robert C. Conformations of Service and Retail Activities. An Example in Lower Orders of an Urban Hierarchy in a Lesser Developed Area. Proc. of IGU Symp. on Urban Geography, Lund, pp. 77-90, 1960.
91. Mitchell, Robert B., and Rapkin, Chester. Urban Traffic, A Function of Land Use. Columbia Univ., 1954.
92. Mitchell, Robert B. The New Frontier in Metropolitan Planning. AIP Jour., No. 3, Aug. 1961. 10 pp.
93. Morrill, R. L. Simulation of Central Place Patterns over Time. Proc. of IGU Symp. on Urban Geography, Lund, pp. 109-120, 1960.
94. Mohring, H. D., and Harwitz, Mitchell. Highway Benefits—An Analytical Framework. Univ. of Minnesota and Northwestern Univ., 1962. 203 pp.
95. Mumford, Lewis. The Highway and the City. 1953. 246 pp.
96. Muth, Richard F. The Spatial Structure of the Housing Market. Papers of Regional Sci. Assoc., Vol. 7, 1961. 13 pp.
97. Nelson, L., Ramsey, C. E., and Verner, C. Community Structure and Change. Univ. of Minnesota, Cornell Univ., and Florida State Univ., 1960. 464 pp.
98. Urban Systems and Economic Development. Oregon Univ., June 1962. 126 pp.
99. Philbrick, Allen K. Analyses of the Geographical Patterns of Gross Land Uses and Changes in Numbers of Structures in Relation to Major Highways in the Lower Half of the Lower Peninsula of Michigan. Michigan State Univ. and Michigan State Highway Dept., 1961. 81 pp.
100. Forecasts and Plans. Vol. 2, Pittsburgh Area Transp. Study, Feb. 1963. 177 pp.

101. People, Jobs and Land 1955-1975 in the New Jersey—New York-Connecticut Metropolitan Region. Regional Plan Assoc. Bull. 87, June 1957. 70 pp.
102. Rouse, J. W. Transportation and the Future of Our Cities. In The Dynamics of Urban Transportation. 7 pp.
103. Row, Arthur, and Jurkat, Ernest. The Economic Forces Shaping Land Use Patterns. Spec. Issue of AIP Jour., pp. 71-81, May 1959.
104. Schnore, Leo F. The Growth of Metropolitan Suburbs. American Sociological Rev., Vol. 22, Apr. 1957. 8 pp.
105. Reasons for Relocation, New Jersey Manufacturing Firms. Seton Hall Univ., July 1955. 21 pp.
106. Siegelman, L. R. A Technical Note on Housing Census Comparability, 1950-1960. AIP Jour., pp. 29-39, Feb. 1963.
107. Silver, J., and Stowers, J. Population, Economic, and Land Use Studies in Urban Transportation Planning—An Evaluation of Current Practices and Recommendations for a Future Program. U. S. Bur. of Public Roads, May 3, 1963.
108. Smith, Larry. Impact of Urban Transportation on Commercial and Industrial Site Choices. In The Dynamics of Urban Transportation. 10 pp.
109. Sweetser, Frank L. The Population of Greater Boston Projects to 1970. Boston Univ., Economic Base Rept. 2, June 1959. 16 pp.
110. Sweetser, Frank L. Projections of Greater Boston's Population to 1970 and 1980. Boston Univ., Economic Base Rept. 9, July 1962. 35 pp.
111. Thomas, E. N. The Stability of Distance-Population-Size Relationships for Iowa Towns from 1900-1950. Proc. of IGU Symp. on Urban Geography, Lund, pp. 13-30, 1960.
112. Trevithick, Morris. Land Use Analysis in Metropolitan Transportation Studies—A Summary of Methods. ITTE Intradept. Working Paper No. 5, Dec. 1962.
113. Ullman, E. L., and Dacey, M. F. The Minimum Requirement Approach to the Urban Economic Base. Proc. of IGU Symp. on Urban Geography, Lund, pp. 121-145, 1960.
114. Vernon, Raymond. Metropolis—1958. Interpretation of findings of N. Y. Metropolitan Regional Study, 1960. 228 pp.
115. Warner, Sam B., Jr. Streetcar Suburbs—The Process of Growth in Boston, 1870-1900. Joint Center for Urban Studies, 1962. 208 pp.
116. Webber, Melvin M. The Engineer's Responsibility for the Form of Cities. Traffic Eng., Oct. 1959.
117. Webber, Melvin M. Transportation Planning Models. Traffic Quarterly, July 1961.
118. Zettel, Richard M., and Carll, Richard R. Summary Review of Major Metropolitan Area Transportation Studies in the United States. Univ. of California, 1962. 65 pp.