MEMORANDUM
RM-2878-FF
NOVEMBER 1961

A FIRST APPROXIMATION TO A RAND MODEL FOR STUDY OF URBAN TRANSPORTATION

J. F. KAIN AND J. R. MEYER

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM

RM-2878-FF

November 1961

A FIRST APPROXIMATION TO A RAND MODEL FOR STUDY OF URBAN TRANSPORTATION

J. F. Kain and J. R. Meyer

This is a working paper. It may be expanded, modified, or withdrawn at any time.

The RAND Corporation

1700 MAIN ST. • SANTA MONICA • CALIFORNIA

PREFACE

In August, 1960 the Ford Foundation made a grant to The RAND Corporation to undertake an exploratory study of urban transportation. One objective of this study is to develop a generalized model of an urban complex, for studying the intricate interrelationships between transportation and the spatial organization of economic activities. This Memorandum is a first step in the model's design.

Model building being the evolutionary process it is, this document -- written in the summer of 1961 -- is already out of date in a real sense. A number of modifications of this suggested framework are currently being debated at RAND, and there is little doubt that the detailed specifications of the final version of the RAND model will differ substantially from those presented here. Nonetheless, since this document was the basis for initiating the first stage of the data collection and analysis, the general characteristics of the final model are likely to be influenced significantly by this first approximation. Our objective in circulating this working paper to urban researchers is to obtain discussion, evaluation, and criticism of our model-building efforts early enough in the study for useful suggestions to be incorporated in the final design.

CONTENTS

PREFACE iii

Section
I. INTRODUCTION 1

II. DEFINITIONS AND NOTATION 7
Endogeneity and Exogeneity 7
Status Variables 8
Input Variables 9
Notation 9

III. WORKPLACE LOCATION AND NONRESIDENTIAL LAND-USE MODELS 11
Industry 11
Industrial Workplaces 14
Commercial and Other Residentially Related Establishments . . 17
Central Office and Related Functions 21
Location of Government Employment and Land Use 23

IV. INDUSTRIAL LAND-USE CONSTRAINTS 25

V. EMPLOYMENT DISTRIBUTIONS 29

VI. LAND VALUE SURFACES 31
Commercial Land Value Surface 33
Residential Land Value Surfaces 33
Industrial Land Value Surface 34

VII. RESIDENTIAL DISTRIBUTIONS 35

VIII. RESIDENTIAL LAND-USE CONSTRAINTS 39
Totality Checks 39
Residential Density Check 39
Color Check 40

IX. STATUS VARIABLE MODIFICATIONS 43
Physical Structure 43
Land Vacancy and Availability 45
Purchasing Power Distributions 46

X. TRANSPORTATION MODELS 47

XI. CONCLUSION 51

I. INTRODUCTION

It hardly can be overemphasized that this is a first approximation to a model for RAND's study of urban transportation. A number of the suggested variables are based on the most intuitive and informal of theory, and are included mainly because of a conscious effort to make the list of suggested hypotheses as inclusive as possible. Despite this effort, however, it also seems highly likely that some important and useful variables are still omitted. In short, the model's ultimate form may bear little empirical or conceptual resemblance to the one advanced here.

Even as it stands, however, the model does serve some useful purposes. It provides at least a few informal guide lines on how the study should proceed, filling an important void in such a field as urban studies, where there is so little formal theory, especially economic or sociological. While the behavioral relationships suggested in this first approximation fall short of constituting a full "structural model" of urban economic and social processes, they do provide a systematic framework for collecting some of the information needed to construct more complete and basic models.

Furthermore, even if we did have available a well-developed body of theory, previous experience with statistical measurement in the social sciences strongly suggests that the models would need extensive adaptation before they could be evaluated or tested against available historical data. Nature seldom performs all the needed experiments, and data collectors almost invariably are serving other masters than the researcher. This process of adapting models and data to one another has been at least partially anticipated here; specifically, the design of this first approximation was

heavily conditioned by knowledge about the availability of data.

In fact, one of the more striking traits of this model is its being based almost exclusively upon relationships to be estimated from cross-section data. This differentiates it sharply from most similarly large models, particularly econometric ones, that have preceded it. Two types of cross-sections figure prominently in the projected estimation procedures: those based primarily on one city and therefore mainly reflecting within-city variations in the relevant variables; and those based on observations over many cities and, accordingly, mainly reflecting between-city variations. Each source will have its advantages and disadvantages and appropriate areas of application. For example, the intra-city data, obtained mainly from origination-and-destination and related studies, will provide greater detail on the geographic configuration of different activities and transportation choices but will have limited generality because origination-and-destination surveys of the needed quality are available only for a few cities, such as Detroit, Chicago, Pittsburgh, and Philadelphia. Inter-city cross-sections, on the other hand, will have greater generality but much less geographic detail; they also possess the large advantage of providing a sufficient number of observations to estimate statistical relationships for activities that conventionally appear only at one, two, or a few locations in any one city. At this stage, therefore, it seems highly probable that the inter-city cross-sections will be most useful for analyzing the locations of industrial, central office, and specialized commercial activities, while the intra-city data will be the main source on the locations of residences and of the more common commercial and manufacturing enterprises.

The two different types of cross-sections also provide a possible means of incorporating a wider range of experience into the studies. For example,

the inter-city data should incorporate a considerably greater range of historical stages of urban development and, correspondingly, of different technological effects and possibilities. This feature also raises the possibility of determining different time-rates of response from the cross-section data without supplementation from time-series sources. Care must be exercised, however, to insure that the different sources of information are not used in such a way that internal inconsistencies will afflict the final model. This care must be taken both in the theoretical specification of the model and in adapting the data to a form suitable for statistical testing.

Illustrative of these consistency problems is that of specifying and estimating "adaptation periods." For example, one projected use for the RAND urban transportation model is to simulate urban growth or adaptation processes in six-month increments, six months being an inherently short-run (less-than-complete) adjustment period in the life of a city. Since cross-section estimates are usually long-run by comparison, they might be considered incompatible with the proposed simulation. In truth, however, cross-section estimates may or may not reflect long-run adjustments, depending on the formulation of the statistical model. Specifically, parameters estimated for a simple cross-section model with only a few basic explanatory variables, such as price or quantity, are likely to reflect a fairly full adjustment; as more and more variables are introduced into the typical cross-section model, however, these full-adjustment coefficients will be partitioned or spread over the additional variables. If these extra variables represent elements that require longer adjustment periods, the coefficients for the original or prime variables increasingly tend toward representation of a short-run adjustment. Put another way, the additional variables are "shift" variables that move a

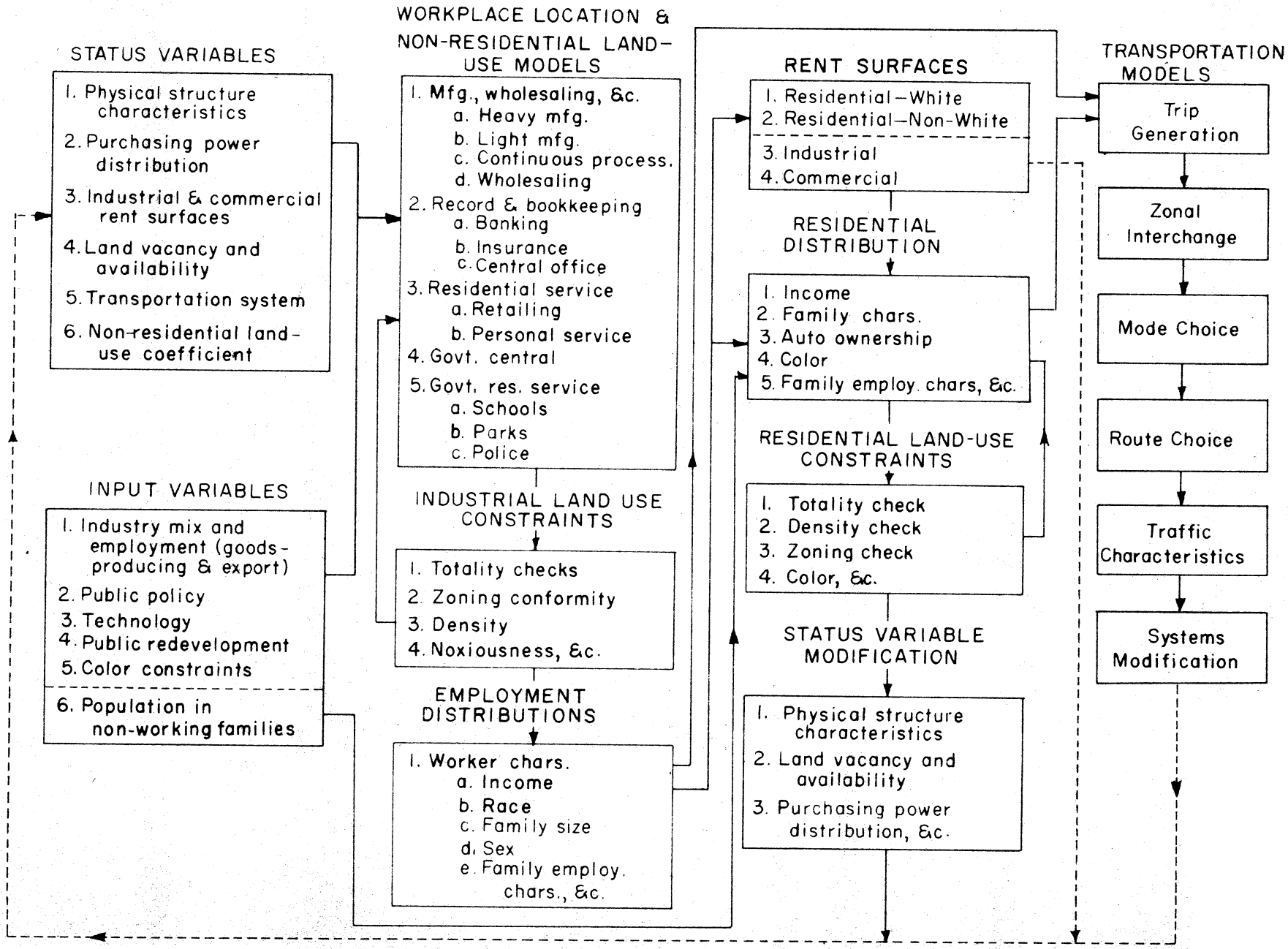
short-run adjustment function along a long-run path defined by the relationship between the dependent variable and the prime causal variable or variables. A good example of such a shift variable is the capital stock in a cost function. With capital stock excluded, a regression of cost against quantity of output using cross-section data would yield a fair approximation to a long-run cost function; while with capital stock included, the cost-quantity relationship would approximate a conventional short-run cost function, i.e., the adjustment of costs to quantity with capital stock held constant.

Furthermore, even if only long-run relationships could be estimated from available data, this is not at all necessarily undesirable. It may merely mean that the objectives of the simulation should be changed. Instead of emphasizing the analysis of small incremental changes, attention might be directed to studying the long-run effects on urban structure of altering underlying technological or choice assumptions. For example, the ultimate effect on a city of eliminating all public transit is in many ways as interesting as determining the effects over time of introducing an increasingly better public system. The point is, simply, that applications must be brought into adjustment with what is feasible, given available data. It is not easy, of course, to determine exactly what is feasible before considerable experience with the data has accumulated. Moreover, an exact specification of the model and its possible analytical uses, while it would be helpful, is not a vital requisite to initiating the empirical studies outlined here. Answers to some questions can be, and probably should be, delayed until a later stage in the research procedure, especially those that make little difference in the early design of the study. The initial stages of the required empirical research will be virtually the same under almost any assumptions about eventual objectives. In short, the model described here

is almost necessarily both a first and a rough approximation.

Any description of the model at this point will of course reflect this tentativeness. Two different descriptive approaches are presented in subsequent pages, one referring mainly to the design of immediately needed empirical studies and one to the eventual synthesis of these empirical results into an analytical framework. This synthesis or overall design is represented by a flow diagram (Fig. 1), which is a quick survey of the model's general configuration. As can be seen, the model is essentially recursive. In its causal ordering, it begins with an exogenously specified or predetermined industry configuration, including employment levels, and proceeds to a distribution of this employment over the urban surface; the next step is to determine the effect of this new employment distribution on rents; the new rent surface and the employment surface, along with several exogenous influences, then jointly determine a residential distribution; from the residential distribution, the model branches into a calculation of predetermined variables needed for the next iteration, on the one hand, and into a conventional transportation planning sequence on the other.

A second description of the model is provided by the appended list of equations, variables, and their descriptions, which comprise a far more detailed specification of the model's characteristics than does the flow chart. In fact, the essentials of the proposed research program are embodied in the hypotheses incorporated into these functions. The empirical results obtained from the investigations necessary to define these functions more precisely should constitute independent contributions, justifiable in and of themselves, regardless of whether they can eventually be fitted into a serviceable over-all model.



Model for study of urban transportation

Fig. 1

II. DEFINITIONS AND NOTATION

While this Memorandum makes a conscious effort to limit the use of specialized jargon, a few special terms are unavoidable. Some of those in Fig. 1, for example, may be intelligible only to persons familiar with "machine simulation," which employs electronic computers in model building. The next few paragraphs therefore attempt to explain some of the more uncommon notations and usages.

ENDOGENEITY AND EXOGENEITY

Status variables are lagged variables specified for the first iteration and subsequently modified endogenously, -- that is, within the model -- for each subsequent iteration. In the flow chart these operations are indicated by or occur in the status-variable modification subroutine, the transportation-systems modification subroutine, and the rent routine. Input variables are variables exogenously specified for each iteration period. These can be manned controls which permit tests of various policies, experimentation with different growth patterns, etc.

The distinction is not always clear-cut between exogeneity and endogeneity and, therefore, between status and input variables. Changes in the transportation system, for example, will probably include changes in both the status and input variables. Changes in travel times for the network used by the various location models and rent surface models will probably be made endogenously. Structural changes in the network will probably be made exogenously. Similarly, the proposed model makes certain modifications of the color constraints endogenously. Other modifications of these constraints are conceived as being introduced exogenously. For example, the proposed model permits endogenous peripheral expansion of nonwhite areas by response

to demand pressures, while non-peripheral expansion of the nonwhite market is exogenous.

STATUS VARIABLES

In Fig. 1, "physical structure characteristics" refers primarily to the qualitative characteristics of the housing supply. If specifications of the quality characteristics of nonresidential structures are required, they could be introduced here as well. The purchasing power and population distributions in the period t-1 are required to locate residentially based retailing, personal services, and governmental activities. Industrial and commercial rent surfaces for the previous period are used in the workplace-location models. In the model, rents are defined as that portion of total rental outlays or value paid for land, i.e., total rents or values minus rent or value attributed to buildings. Only the industrial and commercial land-value surfaces for the previous period are introduced into the model as status variables, since the residential land-value surfaces are generated within the period that they are used. This implies a causal structure in which residential land-values are assumed to increase or decrease in anticipation of changes in demand actually being physically recorded.

Land vacancy and availability as a status variable represents a specification of vacant land zoned for various uses as of the end of the previous period. This variable enters the various workplace-location and land-use models and into determination of the rent surfaces.

The transportation system is described by different times required to travel from i to j on the network using alternative modes. The nonresidential land-use coefficients are required in the model to convert spatial employment distributions to nonresidential land-use distributions. It is highly probable

that these will be modified by differences and changes in rents, the distribution of vacant land, etc.

INPUT VARIABLES

Aside from retailing and service activities, the industry and employment mix in the model is specified exogenously in each iteration period. Employment in retailing, government, personal service, and other service industries is determined endogenously.

The public-policy variables included as input variables in this model are primarily zoning, density, and land-use controls represented in the flow chart as industrial and residential land-use constraints. This should enable some simulation of the effects of public policies on urban development. Experiments can be performed with urban growth under very rigid land-use controls, very loose land-use controls, or with completely unconstrained land-use patterns.

Technology as an input variable refers primarily to changes in transportation technology. Other kinds might also be considered, however, such as changes in firm production functions leading to different employment mixes. Public redevelopment refers primarily to exogenous spot changes in physical structure characteristics. Location of the population in nonworking families is treated in the model analogously to the industry-location models; the number of nonworking families is specified exogenously and then distributed to residential areas through a special non-working-population location model.

NOTATION

At first glance, the model's notation appears formidable, but actually it has a moderately simple pattern. The model has three basic land-occupancy

sectors: industrial, commercial, and residential. Land use, vacancy, zoning, rent, and other variables pertaining to these three sectors are consistently designated by the super-scripts I, C, and R, respectively. Similarly, the general subscripts i, j, k, l, m, and p always have the same meaning: i designates residential locations; j, workplace locations; k, industries; l, characteristics (i.e., stratifications) that pertain to residential site choices; m, characteristics that pertain to trip generation; and p, different cities (so that p appears only in inter-city cross sections). Summation or aggregation over these groups is noted, moreover, by the simple convention of dropping the subscript for the group over which the summation is performed. For example, W_{ijk} designates the workforce in industry k working at site j and resident in area i; W_{ij} indicates the total work force (in all industries and for all work places) residing at i.

Greek letters generally are used to indicate factors or proportionality and other parameters, either arbitrarily specified or empirically estimated. For example, α's designate land-use-to-employment proportions, while β usually indicates regression coefficients.

III. WORKPLACE-LOCATION AND NONRESIDENTIAL-LAND-USE MODELS

The Workplace-Location and Nonresidential-Land-Use studies, vital to the proposed model, are designed to be independent substudies. They will emphasize the relative requirements by various kinds of work activities for access to freight and passenger transportation facilities, and access to the resident population taken in relation to their requirements for land. Firms in both the same and different industries use a variety of production processes which employ these inputs in varying proportions. It is reasonable to expect many of their location decisions to be strongly affected by the cost and availability of these critical inputs. A high priority will be given to an analysis of the locational implications of fairly recent developments which for many industries may be making the price and availability of land more important in the choice of locations. These include changes in material-handling methods, the trend toward "continuous" manufacturing processes, electronic data processing, and the increasing requirements for employee parking. These studies will employ a mix of inter-and intra-city cross-section analyses. The selection of method for each group will depend on whether there is a greater need for fine detail about structure within cities or for less fine detail about differences in structure that are observed for cities in different stages of development and technology.

INDUSTRY

As the industry-location models are presently structured, the objective is to estimate the numbers of workers employed, and the acres of land used by each industry in each employment sub-area. Because of the distributional properties of the data, the proportion of each industry's total employment

or land use is used as the dependent variable, rather than the numbers of workers employed or acres of land used in each sub-area. Using the proportion of each industry's employment in each sub-area as the dependent variable also permits direct comparisons of cross-sections fit for cities with different-sized labor forces and different industry mixes. To insure internal consistency of the model, a totality check is made of the estimated proportions of each industry's employment or land-use assigned to each employment sub-area. If the sum of estimated employment or land-use in an industry is less than 100 per cent, proportional additions are made to each estimated employment sub-area. If the sum exceeds 100 per cent, proportional reductions are made. The routine for performing this totality check is shown in Fig. 1 as part of the industry land-use constraint and consistency checks, but it should be understood that this check, and any required distribution, must be performed before the other constraint checks. By the use of empirically estimated coefficients of land use per worker, employment estimates are converted to estimates of land use by industry in employment areas. This procedure is designated as Variation (a) below. For some industries this procedure will be reversed, with the proportion of total acres used by each industry estimated as the dependent variable, designated as Variation (b) below.

The choice between land used and the proportion of workers as the dependent variable will be determined by data availability and by industry characteristics. It is likely that the spatial distributions for industries with high worker-densities per acre can best be estimated by using the proportion of the industry's total employment employed in each sub-area, while for industries with low workers-per-acre ratios, the more accurate

estimates can probably be obtained by recourse to the land-use variables. For example, central office employment probably can best be estimated by using the proportion of each industry's workers employed in each sub-area as the dependent variable. If the land-use Variation (b) were used, very small errors in the estimate of land use per worker, or of the worker-per-acre coefficient, would cause very large errors of estimation for the employment distribution. By way of contrast, if the area in parks were estimated from distributions of park employees, very large errors in estimating the land-use surface might result from small errors of estimation for the coefficients of the employment surface. With Variation (b), the estimated land-use distributions are converted to spatial-employment distributions, using coefficients of workers per acre for each industry and work sector.

Variation (c) represents a reformulation of the industry-location problem for estimation from inter-city cross-section data. As noted, these data should be mainly useful in industries where there are too few observations in individual cities to obtain reliable estimates using an intra-city cross-section.

Variation (d) represents a very crude attempt to use time-series information embedded in a simple stochastic model as a means of forecasting workplace locations where suitable cross-section data are not available. It is to be hoped that such a recourse is not necessary.

INDUSTRIAL WORKPLACES

Variation (a)

$$\frac{W_{jk}}{W_k} = f\left[R^I_{j(t-1)},\ D_{j-c},\ RS_j,\ HA_j,\ T_{j-B(t-1)},\ T_{j-RM},\ Z^I_{j(t-1)}/Z^I_{(t-1)},\right.$$
$$\left. Z^c_{j(t-1)}/Z^c_{(t-1)},\ PT_j,\ \frac{V^I_j}{Z^I_j},\ PW_{j(t-1)},\ \text{etc.}\right]$$

$$L_{jk} = \alpha_{jk} W_k \left(\frac{W_{jk}}{W_k}\right)$$

where $R^I_{j(t-1)}$ = industrial land value in zone j at the end of premised time period;

D_{j-c} = distance of center of zone j from central business district;

RS_j = presence of railroad siding (dummy, 0, 1);

HA_j = presence of highway arterial (dummy, 0, 1);

PF_j = presence of port facilities in zone j (dummy, 0, 1);

$T_{j-B(t-1)}$ = time to urban boundary or cordon;

T_{j-RM} = time to railroad marshalling yard;

$\frac{Z^I_j}{Z^I_{(t-1)}}$ = percentage of all-city total of acres zoned for industry;

$\frac{Z^c_{j(t-1)}}{Z^c_{(t-1)}}$ = percentage of all-city total of acres zoned commercial;

PRT_j = presence of public rapid transit (dummy, 0, 1);

PT_j = public transit other than rapid (dummy, 0, 1);

V^I_j = vacant industrial land in zone j;

$PW_{j(t-1)}$ = worker supply potential at point j at the end of preceding period;

L_{jk} = land used by industry k in zone j;

$\frac{W_{jk}}{W_k}$ = percentage of industry k work force employed in zone j;

W_k = total work force in industry k;

α_{jk} = land used per worker by industry k in zone j.

Variation (b)

$$\frac{L_{jk}}{L_k} = f\left[R_{j(t-1)},\ D^c_j,\ RS_j,\ HA_j,\ T^B_{j(t-1)},\ T^{RM}_{j(t-1)},\ (Z^I_j/Z^I)_{t-1},\ (Z^c_j/Z^c)_{t-1},\ PRT_j,\ PT_j,\ PW_{j(t-1)}\right]$$

$$W_{jk} = \alpha'_{jk}\left(\frac{L_{jk}}{L_k}\right) L_K$$

$$\alpha'_{jk} = \frac{1}{\alpha_{jk}} \text{ (but estimated directly; see comments above).}$$

Variation (c)

$$\frac{W_{jkp}}{W_{kp}} = f\left[\frac{R^I_{jp}}{R^I_p},\ R^I_{p(min)},\ \frac{W^I_p}{W_p},\ RS_{jp},\ HA_{jp}\ PF_{jp},\ W_p,\ T_{p(C-B)},\ PRT_{jp},\ PT_{jp},\ \frac{C_p}{W_p},\ PCY_p,\ \frac{W_{p(60)}}{W_{p(50)}},\ \frac{W_{p(50)}}{W_{p(40)}},\ \frac{V^I_{jp}}{Z^I_{jp}},\ \frac{Z^I_{jp}}{Z^I_p},\ \text{etc.}\right]$$

where all major variables are defined as before except that a jp subscript indicates the j^{th} zone in the p^{th} city, and j can now be, like k, a variable of stratification; i.e., j does <u>not</u> necessarily have to vary from observation to observation within a sample -- only p. For example, industry k and zone j could define a group that would contain observations on p cities; alternatively, k alone could be used as a classification variable with p x j observations in the sample;

and W_p^I = industrial employment in city p;

W_p = total employment in city p;

$T_{p(C\text{-}B)}$ = time from central business district to cordon boundary of city p;

C_p = cars in city p;

PCY_p = per capita income in city p;

$W_{p(60)}$ = total employment in city p in 1960;

$W_{p(50)}$ = total employment in city p in 1950;

$W_{p(40)}$ = total employment in city p in 1940.

Variation (d)

$$W_{jk} = \left(\frac{\overline{W_{jk}}}{W_k} \right) W_k$$

$$\frac{W_{jk}}{W_k} - {}_1jk = E_\beta \left(\frac{W_{jk}}{W_k} \right) (t\text{-}1) + E_\partial \, \pi_{jk} = \left(\frac{\overline{W_{jk}}}{W_k} \right)$$

$$E_\beta + E_\partial = 1.0$$

$$L_{jk} = \alpha_{jk} W_{jk}$$

where ljk = is a random component;

π_{jk} = percentage of total land zoned for industry k purposes in zone j;

$\left(\overline{\frac{W_{jk}}{W_k}}\right)$ = estimated ratio of industry k work force in zone j; and

E_β, E = empirically estimated parameters (assuming needed time series are available).

COMMERCIAL AND OTHER RESIDENTIALLY RELATED ESTABLISHMENTS

The first three of these models, Variations (a) through (c), differ fundamentally from those for industry location in not assuming the level of the particular activities as given. That is, the models attempt to explain the level of activity as well as its distribution. The fundamental hypothesis is that retailing and certain other private and public activities are brought into being at specified residentially oriented locations as population and incomes in the neighboring residential areas reach certain levels. As a corollary, the location decision cannot be separated from the decision to create or invest in such activities.

Variations (a) and (b) are simply models of purchasing-power potential, spliced to fairly conventional rent-zoning location models. In (a) and (b), purchasing-power potential is also corrected for the provision of similar services in nearby areas. Accordingly, a better nomenclature might be "unrequited purchasing-power potential models." The main difference between (a) and (b) is that (a) assumes a specific weighting structure for aggregating income (purchasing-power potential) over an area to a specific point (saying that it is inverse to the travel time from the point), while (b) is an attempt to estimate these weights empirically.

Variation (c) is simply a discrete change version of the same models. It is an effort to integrate into the model the common observation that new retailing activities tend now to appear in shopping-center developments of two distinct sizes. The smaller are centered as a rule around a super food market, and include a few other specialty shops and perhaps a variety store. The larger are complete shopping centers with a department store as the main activity.

Variation (d) differs in using inter-city as well as intra-city data and in not integrating the estimation of the level of activity directly into the location model. The hypothesis would be that the division of retailing between downtown and suburban areas is mainly a function of a city's age, growth, and other economic and social characteristics.

In short, in variation (d) differences between the central business district (CBD) and the outlying area are hypothesized to be the result of differences in the historical development of urban areas and gross differences in their structure, land-use patterns, and transportation systems. The proposed independent variables would measure characteristics of transportation systems, spatial arrangement of land uses, the density of residential development, the spatial distribution of employment, and the timing of historical development in the different cities. In variation (d), moreover, total demand for the activity under analysis would have to be either exogenously specified, as in the case of manufacturing and export industries, or estimated from an additional analysis of inter-city data; and the distribution of activity outside the CBD would require a location model like those specified earlier for industry.

Variation (a)

$$W_{ik} \text{ or } F_{ik} = \mu + \beta_1 Y_{i(t-1)} + \beta_2 \sum_j \left(\frac{Y_j}{T_{ij}} \right) - \beta_3 {}_j \frac{F_{jk(t-1)}}{T_{ij}} + \beta_4 \left(\frac{Z_i^I}{Z^I} \right)_{t-1}$$

$$+ \beta_5 \left(\frac{Z_i^c}{Z^c} \right)_{t-1} + \beta_6 R_i^c \quad (i \neq j)$$

$$Y_i = \sum_i W_{il} \bar{Y}_{il}$$

$$L_{ik} = A_{ik} F_{ik} \quad \text{(if } F_{ik} \text{ is the dependent variable)}$$

$$W_{ik} = \frac{1}{\alpha_{ik}} L_{ik}$$

where F_{ik} = floor space used by commercial activity k in zone i;

Y_i = weighted aggregate income of zone i;

$\bar{Y}_{il}$ = average (or central) income of income class l in zone i;

τ, Ψ = arbitrary coefficients;

A_{ik} = ratio of land use to floor space in activity k in zone i;

and R_i^c = commercial land values in zone i.

Variation (b)

$$F_{ik} = \mu_1 + \beta_1 Y_i + \beta_2 Y_{i+1} + \beta_3 Y_{i+2}, \ldots, + \beta_{1+j} Y_{i+j} - \beta_{2+j} F_{i+1,k}$$

$$+ \beta_{3+j} F_{i+3k}, \ldots, + \beta_{1+2j} F_{i+jk} + \beta_1' \frac{Z_i^I}{Z^I} + \beta_2' \frac{Z_i^c}{Z^c} + \beta_3' R_i^c + e_{ik}$$

(and previous identities),

where Y_{i+1} = aggregate income in first time zone beyond i;

Y_{i+2} = aggregate income in second time zone beyond i, etc.;

and $F_{i+1,k}$ = aggregate floor space in activity k in first time zone beyond i, etc.

Variation (c)

F_{ik} is to be estimated by models like (a) and (b) above. Then test for

$$(1)\quad F_{ik,t} - F_{ik,t-1} \gtrless \overline{F^s};\ \text{and}$$

$$(2)\quad F_{ik,t} - F_{ik,T} \gtrless \overline{F^m}\ .$$

If (1) is satisfied, $F_{ik,t} - F_{ik,t-1} = \overline{F^s}$; otherwise zero.

If (2) is satisfied, $F_{ik,t} - F_{ik,t-1} = \overline{F^m}$; otherwise zero.

By definition:

$\overline{F^s}$ = the floor space in an average new, small, food-market shopping center;

$\overline{F^m}$ = the floor space in an average new, large, department-store shopping center.

Variation (d)

1. To estimate retailing or similar activity in central business district use:

$$\frac{W_{ckp}}{W_{kp}} \text{ or } \frac{W_{ckp}}{F_{kp}} = f\left[Y_p,\ P_p,\ \frac{P_{p(60)}}{P_{p(50)}},\ PRT_p,\ PT_p,\ \frac{W^I_p}{W_p},\ \text{etc.}\right],$$

Where Y_p = aggregate disposable income in city p;

P_p = population in city p;

C = designator for CBD.

2. Estimate total activity either exogenously or by similar inter-city cross-section study, and distribute non-CBD activity with intra-city models like industry location models (a) and (b), including a consistency check.

CENTRAL OFFICE AND RELATED FUNCTIONS

These locators are similar in many ways to those discussed in the previous section. Historically, a large proportion of their employment was located in the central business districts of American cities. As with retailing, changes in commuting costs, and changes in the numbers and kinds of workers required, growing out of advances in data processing techniques, communications, and transportation, have either changed or have threatened to change the historical locational equilibrium of these functions.

The distribution of central office and related activities probably could be investigated most profitably as a function of inter- rather than intra-city differences. The model proposed here reflects this belief. The rapidity of technological change in passenger transport and communications also suggests that this is an area in which outside, special studies of technological possibilities merit priority. Some of the more patently obvious of these possibilities, such as increased use of inter-city air travel and long-distance telephoning, are at least partially incorporated into the model.

$$\frac{S_{jkp}}{S_{kp}} \text{ or } \frac{E_{jkp}}{E_{kp}} \text{ or } \frac{O_{jkp}}{O_{kp}} = f\left[W_p,\ T^{PT}_{j\text{-}B,p},\ T^{A}_{j\text{-}B,p},\ D_{jB,p},\ \frac{W^{C}_{p}}{W_p},\ \frac{W^{M}_{p}}{W_p},\ PCY_p,\ \frac{W^{F}_{p}}{W_p},\ \frac{W_{p(60)}}{W_{p(50)}},\ \frac{W_{p(50)}}{W_{p(40)}},\ T_{j\text{-}ap},\ T_{j\text{-}BRD},\ \text{etc.}\right]$$

$$W^{O}_{jk} = \delta_{jk} \left(\frac{S^{j}_{pk}}{S_{pk}} \right) S_{pk}$$

$$L^{O}_{jk} = \alpha^{O}_{jk} W^{O}_{jk} ,$$

where:

S_{kp} = sales of firms in the k^{th} industry with a central office in the p^{th} city;

E_{kp} = employment of firms in the k^{th} industry with a central office in the p^{th} city;

O_{kp} = number of central offices of the k^{th} industry in city p;

S_{jkp} = sales of firms in the k^{th} industry with a central office in zone j of city p;

E_{jkp} = employment of firms in the k^{th} industry with a central office in zone j of city p;

O_{jkp} = number of the k^{th} industry's central offices in zone j of city p;

j = city zones or rings ranging from the central business district to exurbia;

W^{F}_{p} = female employment in city p;

W^{C}_{p} = commercial employment in city p;

W^{M}_{p} = manufacturing employment in city p;

$T^{PT}_{j\text{-}B,p}$ = average time from j^{th} zone to boundary of city p by most rapid available public transit;

$T^{A}_{j\text{-}B,p}$ = average time from j^{th} zone to boundary of city p by automobile;

$D_{jB,p}$ = distance from zone j to cordon in city p;

$T_{j\text{-}ap,p}$ = time from zone j to nearest major airport in city p;

$T_{j\text{-}BRD,p}$ = time from zone j to best residential district in city p;

W^{O}_{jk} = number of workers in central office and related functions (e.g., legal, etc.) in zone j for industry k;

δ_{jk} = factor of proportionality, W^{O}_{jk}/S^{j}_{k} ;

L^{O}_{jk} = land area used by central office workers in industry k of zoning;

α^{O}_{jk} = factor of proportionality, L^{O}_{jk}/W^{O}_{jk} .

LOCATION OF GOVERNMENT EMPLOYMENT AND LAND USE

Government employment and land use are heterogeneous and important urban activities. In the Chicago metropolitan area, for example, government consumed a third of the total acres in urban use in 1956, excluding the area used for streets and alleys. If streets and alleys are included, the figure rises to one-half. As for heterogeneity, the government sector includes both large public office buildings which occupy fairly small pieces of land but employ thousands of workers, and public open spaces which cover thousands of acres but have very low worker densities. Furthermore, the locational choices by governments may result from a calculus very different from that employed by private decision-makers. Political power relationships, a consensus of what constitutes an ideal city by its residents or planners, and other variables not usually incorporated in economic analysis, may have a significant impact on the location of government employment and land uses. The extreme heterogeneity of this sector probably

means that the models for locating government workplaces and land uses will be more variegated than those of any of the previously described sectors. In fact, they will probably employ nearly the full array of models presented for the other three sectors: industry, commercial and other residentially related establishments, and the central office and related functions. For this reason these models are not repeated at this point. Some of the appropriate applications should be fairly obvious. The location of large federal, state, and local office buildings is probably best analyzed with one of the central office models. Parks, schools, etc., probably will use a variant or variants of one or more of the population-potential models.

IV. INDUSTRIAL LAND-USE CONSTRAINTS

The precise role to be played in the model by zoning and other government and private land-use constraints is far from clear at this time. We expect to lean heavily on the proposed study of government actions affecting land use to provide effective guidelines for the appropriate use of these constraints. In any event, it seems desirable to conduct many totally unconstrained runs of the model. If on the other hand the proposed study of zoning and land-use controls provides effective guidelines as to the current and future practices of land-use control, and suggests the likelihood of stringent enforcement and maintenance of these standards despite severe market pressures, a number of runs of the model should conform to the best and most widely accepted practices. Entry point for these zoning constraints to the model would be as input or exogenous variables in each period.

The proposed study of land-use constraints includes an investigation of both historical and current zoning and planning practices, with particular emphasis on enforcement and on the responsiveness of zoning and other land-use regulations to market processes. Designing objective analytical appraisals of these matters will not be easy, but their importance to a meaningful model of the type proposed, as well as to an understanding of the interaction between market and non-market forces affecting land use, justifies the effort.

The mechanics of the zoning check are fairly simple. A certain number of acres of land zoned as industrial or commercial are available for use in each period by these industries. If the estimated number of acres required for all commercial or industrial users exceeds the amount zoned for them,

the excess is assigned to the nearest employment sub-area having available acreage. In addition to the general zoning restrictions, certain noxious users may be excluded from many areas zoned industrial.

The worker-density check establishes outer limits for the employment distributions. Their purpose is to prevent minor errors of estimation of the employment surface from cumulating within the model. If worker densities per acre within the model exceed a specified maximum, they are to be truncated, and the excesses distributed to the nearest employment zone where worker densities are below a prescribed level. Hopefully, the model will seldom use this routine.

(a) Totality Check

Test for $\sum_j \frac{W_{jk}}{W_k} = 1.0$.

If < or > 1.0, distribute excess or deficiency proportionally over all j.

(b) Zoning Check

Test for $L_j^I = \sum_k L_{jk} + V_j^I \leq Z_j^I$ where k is taken over all I's

and for $L_j^C = \sum_k L_{jk} \leq Z_j^C$ where k are all C's.

If not satisfied, distribute excess to nearest zones as possible.

(c) Worker-Density Check

Test for $\sum_k W_{jk} \leq \bar{W}_j$ where $\bar{W}_j$ is a specified maximum and k is all employment activities, i.e., I, C, etc.

If not satisfied, distribute excess to nearest zones as possible.

(d) Noxiousness

Special zoning check where $Z_j^I = 0$ for I.

V. EMPLOYMENT DISTRIBUTIONS

Identities are needed to convert the spatial distribution of employment by industry to distributions of worker or household characteristics. The equation below gives both this procedure and the aggregation of worker characteristics over industries for each employment sub-area. The employment surface provides, for each workplace, frequency distributions of relevant household or worker locational characteristics such as sex, race, family size, incomes, etc., for use by the residential location model. These identities should be obtainable from the origin-and-destination studies being used to obtain estimates for the residential location models, or from analysis of published statistics giving labor-force characteristics by industry. The first procedure is the simpler of the two, but has less generality than the second.

WORKER CHARACTERISTICS IDENTITIES

$$W_{jl} = \alpha_{kl} W_{jk}$$

where l subscripts represent worker characteristics relevant to residential location decisions.

VI. LAND-VALUE SURFACES

Economic rents for location have played a critical role in most theories relating transportation and location. Despite this fact, no significant effort has been made to measure empirically differences in land values within metropolitan areas systematically or at any scale, and to relate these differences to characteristics of the transportation system, differences in land utilization, and other explanatory variables.

Land values or location rents are included as explanatory variables in all the location models outlined below. It is expected these will be most relevant for residential location and for certain kinds of manufacturing, wholesaling, and other land-extensive industries. It is probable that they will be less useful as explanatory variables for central office functions and other space-intensive uses. At any rate, a statistical land-value or location-rent study is obviously an important part of the research program required to complete the model proposed here.

Separate land-value surfaces are included for commercial, residential, and industrial uses, each land use being treated as a separate identifiable sub-market within the urban area. The presence of zoning controls supports such a procedure. At the same time, it is well known that zoning tends to break down under market pressures and that there are speculative and other kinds of spillovers between these markets. Insufficient evidence is available to decide the extent of these spillover effects at this time, but experience obtained from the statistical cross-section land-value studies outlined here should provide some guides on how to decide on the best empirical and theoretical means of dealing with this estimation problem.

Furthermore, there are still other reasons for suggesting that the industrial, commercial, non-white residential, and white residential land-value surfaces should be estimated separately. Commercial and industrial users occupy only a small proportion of the land in urban areas. In urban Chicago, commercial land amounts to less than four per cent of the total developed land; land devoted to manufacturing occupies only slightly more than four per cent.

It is probable, however, that their usage of certain types of land is very high. They occupy a large proportion of the acreage along railheads, major through-highways, major arterials, deep-water facilities, and at the focal points of public transit systems. These locations are differentiated for their use by the existence of these and other facilities. For industrial and commercial users there exists a positive utility for these locations. Land values in these limited, spatially differentiated areas would thus be higher even without the absence of institutional controls limiting the number of such locations. A statistical analysis possessing a fine enough grain might be capable of describing and explaining the spatial differentiation of these limited, but critical high-value areas preferred by and occupied by nonresidential users; but the level of aggregation forced by the characteristics and availability of other kinds of empirical information required for the model, as well as the capacity and cost limitations of the computer, prevent the use of enough spatial detail to explain differences in land values in these very small areas using a single land-value surface. Splitting the land-value surface into three or more subsurfaces would be one way to incorporate greater spatial detail without having to disaggregate to impossibly small units of land area.

Through stratification by use or zoning, enough spatial detail can be included. Land values can be separately estimated for each stratum. The model uses a bookkeeping system for each sub-area which keeps account of the number of acres of each type of differentiated land area within each sub-area, and the price of land per square foot for each type. The proposed study of government actions affecting land use should also assist in determining methods of treating these problems.

The objective of the rent study will be to provide estimates of land values for each sub-area in the model in terms of variables generated within the model, or input variables supplied exogenously for each period. In keeping with the general survey design, multivariate analyses will be used to try to explain those differences, using as explanatory variables the distribution of employment and distributions of other nonresidential activities, characteristics of the transportation system, and characteristics of the stock of physical structures.

COMMERCIAL LAND-VALUE SURFACE

$$R_i^C = f\left[D_i^C,\ HA_j,\ \frac{Z_i^I}{Z^I},\ Y_i,\ {}_j\frac{Y_i}{T_{ij}^{\Delta}},\ \frac{V_i^C}{Z_i^C},\ \frac{Z_i^R}{A_i},\ PRT_i,\ PT_i,\ \frac{Z_i^C}{A_i},\ \text{etc.}\right].$$

RESIDENTIAL LAND-VALUE SURFACES

$$R_i^{WR},\ R_i^{NWR} = f\left[P_{edi},\ A_i,\ \frac{V_i^R}{Z_i^R},\ H_i,\ D_i^C,\ \frac{Z_i^I}{A_i},\ \frac{Z_i^C}{A_i},\ PRT_i,\ PT_i,\ T_{i\text{-}C},\ HA_i,\ T_{i\text{-}R},\ \text{etc.}\right],$$

where P_{edi} = excess demand potential for zone i;

A_i = total acreage in zone i;

H_i = housing stock characteristics;

V_i^R = vacant residential land in zone i;

T_{i-C} = time from zone i to central business district;

T_{i-R} = time from zone i to major shopping center;

R_i^{WR} = land rents or values for white residential areas in zone i;

R_i^{NWR} = land rents or values for non-white residential areas in zone i;

and

Z_i^R = zoned residential land in zone i.

INDUSTRIAL LAND-VALUE SURFACE

$$R_j^I = f\left[D_j^C,\ RS_j,\ HA_j,\ T_{j-B},\ T_{j-RM},\ \frac{Z_j^I}{Z^I},\ \frac{Z_i^I}{A_i},\ \frac{Z_i^C}{A_i},\ PRT_j,\ PT_j,\ \frac{V_j^I}{Z_j},\ PW_j,\ \text{etc.}\right].$$

VII. RESIDENTIAL DISTRIBUTIONS

For transportation planning, it is tremendously important to comprehend the locational choices of households and the spatial distribution of residences. In Chicago, for example, residences take up more than 40 per cent of all developed land excluding streets and alleys. Residences comprise 55 per cent of all trip destinations. The objective of the residential location model is to determine the relationship between the location of a person's workplace, the location of shopping centers and other nonresidential centers to which he values access, his income, family characteristics and race, land values, the relative prices and availability of alternative transportation media, and his choice of residential location. As presently conceived, these analyses will be cross-sectional regression studies in which the dependent variable will be the proportion of workers of a particular class employed at a given workplace and residing in each residential sub-area. Conceptually, the locational choices of a particular class of workers are seen to depend on the location of their workplaces, the time required to reach alternative residential areas from the workplace by various modes, the time required to reach various nonresidential centers from each of the residential areas, zoning and use characteristics of the residential zone, and location rents or land values in each residential area. Likely variables for stratifying workers employed at each center to obtain the worker classes are income, family size, sex, number in household employed, etc. Theoretically, we should expect a number of complex tradeoffs of travel times, rental expenditures, dollar expenditures for transportation, space consumption or density, trip-making to nonresidential centers, consumption of differing quality levels of housing, etc., which should determine

varying locational patterns for different classes of workers. For example, if two workers had similar incomes, similar tastes, and valued their time similarly, but if one preferred to live in an apartment and one in a ranch-style home, the worker preferring the lower-density area would be willing to incur larger transportation costs if residences are permitted in areas with lower land values. A worker living in an apartment has less incentive to reside in low-land-value areas. This is because greater dollar savings are available to the low-density dweller from residing in the low-land-value areas. Similarly, households which make large numbers of trips to large shopping centers or cultural and recreational centers have a positive incentive to live as closely to these centers as is consistent with satisfying their other objectives.

Worktrip origins and destinations and zonal interchanges are obtained directly from the residential location model. For the transportation models, it is necessary to estimate zonal interchanges for non-work trips as well. The usual way to do this is to use residential densities and nonresidential land usage to estimate the numbers of origins, destinations, and zonal interchanges. Some variant of this technique will probably be used in the model for generating non-work travel. It will probably often be necessary to convert the distribution of worker classes used in the location model to distributions of trip-generation classes for generating volumes of non-work trips. The equation below describes both this procedure and the summation of workplaces and worker classes for each residential sub-area.

Since households having no wage earner have no specified workplace, a special model is required for locating them. For example, previous workplace, if available, might be a possible substitute for workplace as a

variable determining residential choices of retired persons. It should be pointed out, however, that there is much less of an incentive to tie together the workplace and residence of nonworking households. A major reason for wanting to do so for working households, in addition to the strong theoretical and empirical justifications for such a procedure, is the importance of the journey-to-work in metropolitan transportation problems. The journey-to-work is by far the most frequent kind of trip made in urban areas. It is also extremely concentrated in time, leading to very peaked demands for transportation facilities.

RESIDENTIAL DISTRIBUTIONS

$$\frac{W_{ijl}}{W_{jl}} = f\left[R_i^{WR}, R_i^{NWR}, T_{ij}^{A}, T_{ij}^{P}, H_i, D_i^{R}, D_i^{C}, \frac{Z_i^{R}}{A_i}, \frac{Z_i^{I}}{A_i}, \frac{Z_i^{NWR}}{Z_i^{R}}, A_i, \frac{V_i^{R}}{Z_i^{R}}, Z_i^{R} - V_i^{R}, \text{etc.}\right]$$

$$V_{im} = \lambda_{il} W_{il} ,$$

where T_{ij}^{A} = travel time from place of work j to residence site i by auto;

T_{ij}^{D} = travel time from place of work j to residence site i by public transit;

Z_i^{NWR} = implicit zoning constraint on non-whites;

W_{ijl} = workers of residential characteristic l (a stratification variable) working at j and residing at i;

W_{jl} = total workers of residency class l working at j.

W_{im} = total workers with trip-generation characteristic m residing at i; and

λ_{il} = ratio of $\frac{W_{im}}{W_{il}}$, empirically determined.

NOTE: One j will represent all non-working households and probably will require somewhat different explanatory variables.

VIII. RESIDENTIAL LAND-USE CONSTRAINTS

TOTALITY CHECKS

The residential-location totality check is very similar to the industrial-location totality check. The sum of the estimated proportions of workers by each workplace and class is computed over all residential sub-areas. If the sum for any workplace and worker exceeds unity, the proportion assigned to each residential area is reduced proportionally. If it is less than unity the proportion assigned to each area is increased.

(a) Totality checks:

$$\text{Test for} \sum_i \frac{W_{ijl}}{W_{jl}} = 1.0.$$

If $<$ or $>$ 1.0, distribute excess or deficiency proportionally over all i.

RESIDENTIAL DENSITY CHECK

The residential-density check serves much the same function as the worker-density check. It gives an outer bound within which it is hoped the quality of estimation procedures will be high enough to permit the model to operate unconstrained. If the quality is not uniformly high enough, this outer bound will prevent spurious results or extreme values in only one place or a few places in the model from too badly biasing the remaining portions of the model.

(b) Residential-density checks:

$$\frac{W_i}{Z_i^R \quad V_i^R} \blacktriangleleft \bar{W}_i =$$

a specified maximum residential density (in household units per acre).

$$W_i = \sum_{j,l} W_{ijl}$$

COLOR CHECK

Even if it is not intellectually pleasing to most of us, it is a fact that non-whites in American society are badly restricted in their consumer choices. This restriction is most evident in the housing market. In nearly all United States urban areas, segregation of the non-white and white markets for housing is almost absolute, with important implications for the travel patterns in these urban areas. Much of the journey-to-work cross-hauling observed in American cities is to be explained by the journeys of non-whites from their workplaces to their residences, generally located in the older close-in parts of urban areas.

Both the systematic and less-systematic studies of the dynamics of the non-white and white markets generally affirm that the non-white market typically expands peripherally, as demand pressures rise because of rises in the non-white population.

Less frequent is leapfrog growth, with the non-white market jumping all-white areas and establishing nuclei elsewhere in the urban area. Undoubtedly, integration of the two markets is the least common expansion pattern of all.

The model proposes to simulate the processes of housing-market segregation and peripheral growth. The residential density checks play a special role in this mechanism. Informal color-zoning constraints are used in the same manner as zoning constraints elsewhere in the model. Nonconforming uses--i.e., non-whites in areas informally zoned against non-white residence--are relocated to the nearest sub-area permitting non-white occupancy.

Following the zoning check for non-white conformity, a residential density check is made of sub-areas within the non-white market. If these permissible densities are exceeded by any sub-area within the non-white market, the excess population is allocated to the closest sub-area or sub-areas within the non-white market. When residential densities exceed the maximum throughout the entire non-white market, this is taken to signal the presence of intolerable demand pressures within the non-white market and to indicate the probability of non-white peripheral expansion. When this occurs the model adds additional area to the non-white market. If residential densities are exceeded in the newly added portion of the market, whites are moved to nearby residential areas.

Forms of nonperipheral growth could also be handled within the model. New non-white markets probably have a fairly high probability of forming near new employment concentrations located at substantial distances from the non-white market, especially if the industries have very large non-white labor forces. It would be possible to locate, within the model, satellite non-white housing markets near such industries if their non-white labor force reached a certain level, or preferably if the total travel miles by non-whites employed there exceeded a certain number. This second index has the advantage of including a measure of both the number of demanders and the extent of their inconvenience.

The model as presented here eschews these more elaborate solutions and introduces nonperipheral growth exogenously. Later research may make it possible to remedy this situation.

(c) Color check:

$$\text{Test for } \sum_{j,l} W_{ijl} \leq \overline{W}_j \, Z_i^{NW} \qquad \text{where } l = \text{all NW (non-white) classes.}$$

If not satisfied, distribute to nearest zones as possible, i.e., to zones where inequality is satisfied.

However, if $\sum_{i,j,l} W_{ijl} > \sum_{i} \overline{W}_i \, Z_i^{NW}$, then implicit non-white zones must be altered by letting non-whites into nearest white areas. If alteration is required, check densities again for both non-whites and whites and move whites as required to nearest available zone.

IX. STATUS VARIABLE MODIFICATIONS

PHYSICAL STRUCTURE

The theory implicit in this formulation of the model postulates that a household, in selecting a residential location, maximizes the satisfaction obtainable from an extremely complex bundle of residential services. It follows, therefore, that the household makes a series of complicated tradeoffs between alternative goods in this bundle in order to achieve this objective. The variables included in the residential location model represents a first effort to list, on the basis of this theory and on the basis of previous empirical work, a few of the more strategic variables.

The list includes a number of variables which describe the physical characteristics of the structures located in each residential sub-area in the previous period. This distribution is introduced into the model as a lagged endogenous variable. The purpose of the physical structure modification function is to modify a structure-quality index (or indices) for the succeeding iteration period in response to changes introduced into the model.

It is postulated that the quality of the structures within a sub-area depends on the characteristics of the residents, the net residential density, the incidence of certain nonresidential uses, etc. Underlying these postulates is a theory of structure deterioration. It is argued that most construction in American cities can, if desired, be maintained at high levels through adequate maintenance expenditures and periodic renovation. Thus it follows that structures drop in quality because owners let them.

To some extent, at least, it is believed that there is an unequal spatial distribution of demand by various income classes in urban areas, which leads to an unequal distribution of demand for structures of various quality. This view holds that the level of housing quality is a function of the expenditures for housing quality, the level depending on family incomes and other household characteristics.

For at least two reasons, nearby nonresidential activities are also likely to encourage permissive deterioration. First, certain types of nonresidential activity (heavy industry, for example) undoubtedly lower the quality of a residential site by their very proximity. Second, owners of property near other nonresidential activities are likely to spend little money on maintenance if they expect that their property will eventually be absorbed as those activities expand.

The final variable included in the structural modification function is a net residential density variable. As a part of permissive deterioration, landlords often follow a policy of density-increasing conversions which tends to speed the rate of deterioration. The process of slum development generally includes low-maintenance policies, occupancy by lower and lower socio-economic groups, and an increasing density of occupancy.

A very interesting exercise for the model will be to begin in the first period with a uniform quality level for each residential sub-area, and determine if the model generates spatial distributions of housing quality similar to those observed in United States urban areas.

(a) Physical structure

$$H_i = f(W_{il}\text{'s}, \frac{W_i}{Z_i^R}, \frac{L_i^I}{A_i}, \frac{L_i^C}{A_i}, \text{etc.})$$

LAND VACANCY AND AVAILABILITY

The withholding of land or the unwillingness to develop land depends on or results from some complex dynamics of the urban land market. Expectations of future speculative returns as a reason for withholding land for development, the influence of various tax policies, etc., are extremely difficult to build into a model of this sort. Despite this, they significantly affect the urban use pattern.

Similarly, a number of effects of extreme durability and the interdependence of adjacent residential sites affect the use or re-use patterns in urban areas. It is commonplace for high-priced vacant land located in the middle of a slum to remain undeveloped for many years. The explanation is complex; the eventual re-use value of the site is high and it thus has a high reservation price. But its value for current development is low because of the run-down condition and generally low quality of adjoining properties. At the same time, the acquisition costs of adjoining properties, because of the high gross income resulting from high and possibly illegal densities, prevent immediate re-use.

A good statistical study of vacant land distributions for a number of cities using as independent variables indices of assembly costs, quality characteristics of the area, proximity of higher uses leading to high reservation prices, land values, zoning regulations, proximity to goods-moving transportation facilities, etc., will provide insights both into

the dynamics of the urban market, and the role of speculation both judgmentally correct and incorrect, etc.

The eventual objective is to obtain a description of something similar to a supply function of urban space, with emphasis on the relationships between the long-run and short-run functions. Our approach will be to attempt to estimate the time-path adjustment from cross-section data.

(b) Land vacancy or availability

$$V_i^R = f(R_i^R,\ P_{EDi},\ \frac{L_i^C}{A},\ \frac{L_i^I}{A},\ D_i^C,\ HA_i\ PRT_i,\ Z_i^R,\ A_i,\ \text{etc.}).$$

$$V_i^C = f(\text{ditto but } \frac{L_i^R}{A_i} \text{ for } \frac{L_i^C}{A_i})$$

$$V_i^I = f(\text{ditto but } \frac{L_i^R}{A} \text{ for } \frac{L_i^I}{A} \text{ plus } \frac{L_i^I}{I}\ HA_i,\ PRT_i,\ Pt_i,\ D_i^B,\ \text{etc.}).$$

PURCHASING POWER DISTRIBUTIONS

A number of location models use lagged distributions obtained from the residential distributions in the previous period. For example, the retailing and personal services location models require a population distribution weighted by income or other family characteristics as an explanatory variable.

To minimize storage requirements and facilitate efficient computation of the model, a number of these status variables obtained from the residential distributions will probably be computed at this point.

(c) Purchasing power distributions, etc.:

To minimize storage requirements it is probably convenient to compute $Y_i = \sum W_{il}\ \bar{Y}_l$ and similar variables at this point.

X. TRANSPORTATION MODELS

The intellectual orientation of transportation planners as well as the specific nature of their assignment has led to a heavy investment of resources in the development of traffic models. Trip generation, zonal interchange, and assignment methods have been carefully worked out and represent a high level of technical development. The more sophisticated of these models employ complex networks, minimum time-path-assignment techniques, diversion curves, etc., to simulate route choice and the effect of congestion and other variables on route choice. Considering the nature of their problem--providing a transportation plan for a specific metropolitan area--these planners have done a creditable job of developing and testing these models. At this point, both the state of the art in this area and the rate at which improvements are being effected appear satisfactory. In this study, therefore, the particular focus is on developing a better balance between forecasting models and structural models, and between general models, appropriate for evaluating general transportation policies, and specific models, appropriate for studying the transportation problems and for evaluating alternative policies for a unique urban area.

For these reasons little development work of a technical kind is projected as a part of our model-building process. Instead the intention is to borrow heavily from the technical work carried out by the Chicago Area Transportation Study, the Penn-Jersey Transportation Study, etc. New consideration will be given mainly to possible modifications of these existing transportation models to increase their theoretical and behavioral content and to making them more consistent with the RAND model's emphasis on the interdependence of land use and transportation.

For example, some further analysis might be given to the modal choice problem, to try to develop a better understanding of the substitution relationships between the private automobile, public transit, and public rapid transit. An important potential use of the RAND model is the study of the effect of alternative transportation systems and different modal mixes on the urban form. The advocates of one or another transportation media have blatantly advertised the ability of their wares to affect the land-use patterns and the spatial organization of the city. To some extent the residential and nonresidential location models ought to provide some useful insights concerning the modal choice problem.

Even though little development work is anticipated in the area of transportation planning models, the RAND research program does include a substantial engineering effort which deals with the important systems modification routine listed in Fig. 1 under "Transportation Models."

RAND engineers have for some time been systematically collecting and organizing information describing the characteristics of transportation vehicles as well as of entire urban transportation systems. This engineering effort will aim at describing and measuring the components of transportation systems in actual use and in various stages of acceptance, and serious proposals for the future. This evaluation will cover the range of transportation systems including pedestrians, automobiles, buses, vertical-takeoff-and-landing aircraft, rail cars, and other proposed vehicles. This exercise should thus provide a basis for a system-analysis approach to evaluating competitive transportation systems, finding desirable mixes, and pointing out the influence of design parameters upon the over-all system costs. Unproven, novel alternative configurations and designs will also be gauged by a cost comparison with existing systems.

This systems analysis will provide both input and status variables for the model. The description of currently used vehicles and components will provide a basis for experimentation in the model with transportation systems having varying characteristics. Combination of expected technologies, obtained from the second part of the engineering study, and the model, should enable us to glean some insights into the nature of changes in the spatial organization of today's urban communities which are to be expected from changes in the costs and characteristics of transportation.

XI. CONCLUSION

The model just outlined is obviously a melange of many techniques of estimation and analysis. In large part, it is an econometric model, unusual, however, in being based almost exclusively on relationships estimated from cross-section data. There are parts of the model, moreover, where constraints and other analytical complexities are involved and the model at these points is more akin to so-called computer simulations.

Perhaps the most important aspect of the model, however, is its almost exclusive preoccupation with structural estimation and diagnosis as an objective. Specifically, the model is not intended as a forecasting tool in the narrow sense of foretelling what future values will be for specific variables in specific localities. Such an objective is worthy enough but its realization would dictate a fundamentally different design. For example, in a pure forecasting model the strong autocorrelation probably present in many of the model's variables could and should be utilized as a simple means of improving forecast accuracy.

Nor is the model intended to provide direct and quick solutions to specific policy or planning problems. Hopefully, it will provide some helpful insights into such problems, but the model is so geographically general or unspecific that its translation to the planning processes of a particular urban area may be difficult.

What the model is intended to do is to provide a systematic framework or survey design for studying the relationships between transportation and the spatial organization of the urban area. Hopefully, this will provide some consistency in the efforts of a number of independent researchers, so that the results from these diverse empirical studies can

be synthesized into a more interdependent analytical entity. An objective of the model at this juncture is to insure that the resources employed on the study are fully exploited and that the end product adds up to something more than the sum of its component parts. Furthermore, experiments with the final model should provide not only some helpful insights into policy formulation but some very direct evidence about where the next round of research effort will be most productive. In short, it is hoped that the research programs represented by the model will provide a foundation for further systematic study of the empirical relationships that determine urban structure.