THE IMPLEMENTATION OF A JOINT DISAGGREGATE DEMAND MODEL IN AN URBAN SIMULATION.

by

VINCENT R. WORMS
Ingenieur Diplome de l'Ecole Polytechnique .
(1974)

Submitted in partial fulfillment
of the requirements for the degree of Master of Science.

at the
Massachusetts Institute of Technology
(June 1976)

Signature of Author .... Signature redacted ....
Department of Civil Engineering,May 21,1976

Certified by .. Signature redacted ....
Thesis Supervisor

Accepted by ...... Signature redacted ....
Chairman, Departmental Committee on Graduate Students
of the Department of Civil Engineering

ABSTRACT

THE IMPLEMENTATION OF A JOINT DISAGGREGATE DEMAND MODEL IN AN URBAN SIMULATION.

by
VINCENT R. WORMS

Submitted to the Department of Civil Engineering on May, 1976
in partial fulfillment of the requirements for the degree of
Master of Science

This work is an attempt to construct a new urban simulation model through a synthesis of two previously developed approaches:

1) An existing Land Use Model, the "NBER urban simulation model", characterized by a high level of disaggregation in its descriptions, the existence of an explicit demand/supply equilibrium mechanism and of a time-dynamic adjustment process.

2) A disaggregate joint choice demand model which achieves a behavioral representation of households' location, automobile ownership and mode-to-work preferences.

While synthesizing the two approaches, a number of conceptually appealing features have been designed.

1) A fair degree of behaviorality is achieved by the use of the joint disaggregate description of the consumers' choices.

2) A good deal of realism is gained from a "new" market-oriented process of matching the demand and the supply, which primarily relies on a price-adjustment mechanism.

3) Numerous urban phenomena are explicitly taken into account. This gives the model added potential as an evaluation tool for a fair number of public policies. Special emphasis was placed on accounting for the impact of the transportation policies; the way is open to a further integration of the simulation into a comprehensive transportation and Land Use Model.

A case study has been implemented. It consists of a four years simulation of the development of Washington, D.C. urban area. The purpose was not as much to develop forecasts per se, as to develop an evaluation of the feasibility of the model.

The quantitative results are not fully satisfactory because of some inaccurate initial data, and a probable misspecification of the demand model, leading to a lack of sensitivity to price-adjustments.

However, it has been proven that the model satisfactorily reflects the behaviors of the suppliers, the preferences of the consumers and the trends of the market.

Furthermore, the computational requirements are reasonable. Improvements have been suggested in order to assess the model's reliability, and to improve it to the level of an actual public policy-evaluation tool.

Thesis Supervisor: Steven Lerman

Title: Assistant Professor of Civil Engineering

ACKNOWLEDGEMENTS

I would like to extend my gratitude to the people who have assisted me in the preparation of my thesis:

I am especially grateful to my advisor, Steven Lerman: his knowledge of the subject and the support he gave me on many occasions have been invaluable.

William Apgar, from the National Bureau of Economic Research, has provided me with useful information: I did appreciate his assistance.

Many thanks to my friend Jessie Jacobson for his repeated contributions.

My gratitude to Johannah who provided precious help in the editing.

Thanks to Sally who typed this paper.

Thanks to all my friends who have helped me to keep my optimism on some occasions.

Table of Contents.

Table of Contents (cont.)

List of Tables and Figures

Chapter I - Introduction

The title of this thesis highlights the fact that it is one step in the widely studied field of urban simulation. Within that framework, the contribution has been determined by the desire to focus on the way such models can be developed for the analysis of urban transportation policy questions: the impact-orientation as well as the modeling tools that have been introduced reflect the emphasis which has been put on a "transportation type approach".

However the study remains basically an urban simulation. Therefore, it is desirable to spend some time in the introductory chapter summarizing what the purposes are of such simulations, and which basic issues are to be addressed in their implementations. The remainder of the chapter consists of an overview of the study and a brief summary of the following chapters.

1.1. Purpose of an urban simulation

The concept of "urban simulation" is identified with the idea of utilizing the processing capacity of a computer to represent various and interconnected urban phenomena: change and growth of the housing stock, modification of the structure of employment, moving and migrations, modifications of the transportation network, etc...

The complexity of these interconnections explains the contributions of the computer: three objectives are stated in order of increasing ambition.

1) to synthesize the understanding of the urban development. A large number of studies have focussed on the behavior of the various actors and submarkets that constitute a city. These analyses are useful to the city-planner only to the extent that the interrelationships between them are also represented. This effort to synthesize is mandatory: an example among others is the impact of a new transit line on the housing stock of a zone serviced by the line. A primary impact would be the choice of that residential zone by a non-auto-owner, but a counter-effect could be the migration out from the zone by inhabitants discouraged by the increased density and congestion. Both effects should be represented. Because of the multiplicity of such interconnections, it is impossible to have a complete analytical knowledge of the phenomena. The only alternative is to "teach" the computer the "rules of the game" played by the various actors, and to have it simulate their behaviors in a given environment.

2) to forecast the pattern of the city development, given all necessary exogenous data. This is achieved by the use of an iterative process. The outcome of each period is the initialization framework for the next iteration. It should be noted that the word "exogenous" has no meaning per se, but rather depends upon the level of sophistication of the model. For example, the growth of population-serving employment can be given exogenously, or internally generated

in connection with the increase of population.

3) the last purpose deals with policy evaluation. Since the model can forecast the impact of a number of policies (housing, prices, transportation, etc...), a comparative analysis of these policies can be performed. In order to be given a value, the criteria of evaluation must be defined as functions of explicit attributes. Hence, as far as some quantitative tools can be defined to evaluate a program, the urban simulation potentially provides a powerful tool for the city-planner.

There is no need to underline the harm of a model which misleads the decision-maker because of its poor validity. Two conditions of validity deserve special interest:

- the representation should be reproductible, in location,to take advantage of other studies, but essentially in time so that the model keeps its validity throughout the iterative process. This argument advocates strongly in favor of a behavioral representation as opposed to a non-causal model structure.

- the model should not forego or understate any important relationship. In particular, one should avoid any approximation that would make inaccurate the description of a phenomenon. This has been a constant concern in the present study. For example, it has been attempted to describe faithfully the transportation impacts: the demand model does take them into account in the process of generating the consumer's choices.

These requirements necessitate the addressing of some important issues in the designing of an urban simulation.

1.2. Three basic issues in the designing of an urban simulation:

1) A crucial point is to appreciate whether a model relies on a correlative approach (such as the consideration of trends), or reflects the very nature of the observed phenomena by using a behavioral description of them. As discussed above, this issue is at the root of the reliability of the model. If the model comprises a demand sector, a supply and an equilibrium routine, this issue of behaviorality should be present in all three of them:

- "Supply" in a city is a generic term which covers as well the supply of total employment, the existence of population-serving employment, the network and means of transportation, the supply of housing stock, etc... The modeling of employment in a behavioral way is difficult because the decisions depend on a lot of exogenous factors which are not yet satisfactorily known.

The supply of housing is probably easier to represent. It requires the understanding of the economic behavior of landlords and developers. It is also the most important component in describing the physical evolution of the city.

• The demand-modeling involves a lot of issues: it must consider a vast number of types of decision-makers, a great number of choices (such as residential location, auto ownership, means of transportation...) with several alternatives for each choice. Moreover, those choices are fundamentally interconnected, so that each of them cannot be considered individually without losing some accuracy. They should rather be generated jointly.

• Finally, the "market clearing" is the equilibration of the supplies and demands, basically through the determination of equilibrium prices (salaries for the market of employment, rentals for the housing-market...)

Very different approaches have been followed by previous researchers in this field. One important issue is the choice between an optimization technique (i.e. the optimization of an aggregated utility) and a dynamic adjustment mechanism. Again, the essential criterion to choose among these approaches is the expected reliability of the model. In that respect, chapter three explains the choices that have been made in this study.

2) The preceding remark apply in fact to all demand/supply models; an urban simulation deals with specific phenomena, among which are the impacts of the transportation network. Everybody recognizes its close interconnection with city-planning. In M.L. Manheim's terminology,

this involves "type III" but primarily "type II" relationships. Type III relationships describe the modifications of the network decided by the policy-makers to improve the level of services. Type II applies to the impact of the transportation system on the structure of the city: for example, it explains comparative growth of various zones having different levels of service. Therefore, it is essential that this latter type of relationship can be built in a model of urban simulation. This raises a difficulty that has been mentioned about the second condition of validity of the model: if Type II relationships are most of the time present in the existing models, their impacts are often underrepresented in the formulation.

3) The third and last issue that will be pointed out here is the problem of computational capacity. Despite the huge improvements which have been achieved with respect to the size and efficiency of the big computers, an urban simulation remains a very large problem to handle: this is because of the great number of parameters involved. Some basic dimensions are the geographical zoning, the interzonal transportation data, the stock of housing and corresponding prices, the types of households and the characteristics of their choices, etc... The search for greater accuracy induces one to use a highly disaggregate description, which means a multiplication of the dimensions: storage and computation problems get out of hand very fast. Therefore,

the modellers who meet this problem have to find methods to keep the model under control. Since such methods always include some sort of approximation, attention must be paid not to lose here the accuracy that has been gained elsewhere.

1.3. Overview of the study and remaining chapters

All the preceding remarks are central issues in the present study: this study is an attempt to build a new urban simulation characterized by what was thought to be the best trade-off between exhaustiveness, behavioral validity, and feasibility.

Two basic components of the model have not been created, but are borrowed from two sources:

- the supply sector and several other linked formulations are derived from a simulation currently implemented by the National Bureau of Economic Research (NBER). This simulation is discussed in chapter two.

- the demand sector relies on a study by S. Lerman, "A disaggregate behavioral Model of Urban Mobility Decisions", 1975.

On the basis of those two works, the purpose of the study has been to prove the feasibility of a model benefiting from the two contributions. This has caused some problems, because of the large differences between the two approaches: the demand model was based on an entirely disaggregate approach and the consideration of joint choices, while developers of the NBER simulation adopted a more global point of view (e.g. the market clearing process is the maximization of an aggregate utility), and sequential choices.

Thus, the study has basically consisted of putting together these two approaches:

. first, conceptually, in addressing the important issues which have been discussed in section 1.2. (these developments are the content of chapter three)

. second, practically, by formulating the model itself.

Since the purpose of the study was to undertake a feasibility study, it was necessary to implement the model on an example problem. The goal however is not to develop the model at a level of detail suitable for application in a specific city, but rather to appreciate whether or not the objectives have been met (computational feasibility, strong behavioral orientation, sensitivity to transportation impacts, convergence of the process, meaningfulness of the results). Therefore the case study is a computationally simplified form of the model, implemented for the city of Washington D.C. , Chapter four describes the characteristics of this case study.

Finally, Chapter five draws conclusions about the model, its strengths and weaknesses. It addresses the improvements which should be achieved and indicates the author's point of view about the potential of this new urban model.

Chapter 2: Relationship with other existing land use models

The art of urban modeling is still in an early enough stage so that no single approach has prevailed. There is a plurality of models which differ more or less widely, and when designing a new formulation, it is useful to compare it to the alternative formulations. It clarifies the fundamental choices that have been made.

Because of the relationship with the NBER model, the following review of existing land use models addresses two issues: How does the NBER formulation compare to other models? What is the relationship between the NBER and the new model? Therefore, this review does not pretend to be comprehensive: more complete ones can be found in several papers which are referenced at the end of section 2.1.

2.1. The National Bureau of Economic Research Urban Simulation and the other land use models

Several issues differentiate the existing models: the degree of behaviorality, the level of disaggregation, and several other basic characteristics.

2.1.1. Causal versus correlative models:

The concept of behavioral modeling means obtaining a reliable description of the spatial allocation by capturing the residential location behavior in quantitative terms.

In opposition to that, the statistical models rely only on a number of correlations between **several** urban attributes: the rationale is that these relations are stable enough to allow their extrapolation and hence, to generate predictions.

The most noticeable in that category is the "Empiric" model, which assumes an additive linear construct. Several characteristics make its usefulness questionable: the households make their decisions according to a general accessibility criterion, instead of a known workplace, the prices of housing do not influence their choices significantly. Several other weaknesses are connected with a lack of behavioral content: they finally encompass the whole applicability of the model.

In general, it seems that future development of the art will favor the behaviorally oriented approach.

2.1.2. The level of aggregation

Another trend has appeared in favor of capturing the individual behavior and then aggregating the forecasts to obtain macro-descriptions: this is to be opposed to the classical aggregate prediction-models.

This latter category includes the pioneering "Lowry model" which has been at the root of a number of important developments. The Lowry location of residences is a function of the location of employment and the transportation behavior of workers. This allocation relies on a non-origin/destination-oriented measure: the accessibility of each residential zone to its surrounding employment. The expected residential density in the zone is derived from it. Implemented in 1964, the Lowry model has been followed by a sequence of so-called *"Lowry-derivatives";* among them is the "Time Oriented Metropolitan Model" (TOMM). As opposed to the Lowry model, this one achieves some disaggregation of the locating population: first, a total number of households is assigned to each zone, then this number is distributed among household types.

The NBER model has performed a much higher level of disaggregation. It has a number of other characteristics which are viewed as improvements in the state of the art, and are now discussed.

2.1.3. Other characteristics

Three other issues provide the opportunity to compare the NBER formulation with some of its predecessors: the time-orientation, the consideration of supply, and the so-called "Monocentric assumption".

The time-orientation

Another factor limits the usefulness of the Lowry Model for the purpose of forecasting: the model generates an instant city. It does not consider the path of adjustment between the current status and the forecast equilibrium. In terms of urban phenomena, this means that it does not take into account the existing housing stocks, and does not describe how those stocks will be modified. Because of the importance of that foregone issue, the models which generate only a long-run equilibrium cannot actually be relied upon as forecasting tools and guides to public policy.

This problem is solved by using a behavioral description under an incremental form as opposed to a static perspective: the modifications are described period by period. To reflect a realistic adjustment, only a portion of the households or employment are moving in each period: the description of the housing stock and residential distribution is progressively evolving. In fact, the city is viewed in a constant dynamic disequilibrium which is modified by the current period decisions of demanders and suppliers of urban goods.

The TOMM model, which has already been referenced, has that time-orientation as opposed to the Lowry Model from which it is derived.

The NBER formulation does consider the issue of time-adjustment. It incorporates an explicit description of the current housing-stock. A special "filtering" submodel describes its evolution over time: a decline in quality in the case of under maintenance; an improvement of quality when the unit is rehabilitated. The supply sector itself generates the modifications of the structures and the new constructions, according to "market signals" computed in each period.

Quoting the authors of the models, this improvement of a time-dynamic description obviates that embarrassing issue of "destroying cities every night and rebuilding them the next morning"!

Consideration of the supply:

A considerable improvement has taken place with another Lowry-derivative, the "Bay Area Simulation Study" (BASS). Until then, the models were basically demand-oriented; supply considerations were often limited to the available stock of developable land. In reality, the supply activity has a significant impact on the shape of the urban development. The choices of landlords and developers influence the demands, and hence the physical changes in the city. A typical example is the discouraging of an over-demand in a zone through a process of price increase.

Some models, like the Herbert and Steven's "Penn-Jersey Transportation Study", include the housing supply but only as one factor among others orienting the demand of housing. The BASS model was the first one to estimate the supply activity internally, and match it with the demand to predict the locational decisions.

As stated before, the NBER modellers have dealt explicitly with that issue. There is an endogenous supply activity which is described in section 2.2.

The monocentric assumption

The concept of a "center of the city" has always had a role in the attempts to describe the urban spatial structure. An underlying factor is the higher price of land for zones easily accessible to the center: there is a trade-off between a larger amount of residential land and reduced commuting times for people working in the center.

This is at the root of the monocentric urban land use models: they have provided valid descriptions of such phenomena as the central-city declines in population.

The most famous is the **Alonso**'s formulation: he applies a consumer-theory to the housing-demander. Each household maximizes its utility by choosing among three goods:

- employment accessibility in terms of distance from a theoretical central workplace so-called "Central Business District" (CBD).

- the amount of residential land

- a composite of all other commodities consumed by the household.

The trade-off between housing costs and transportation costs depends on the residential space consumed, and characteristics of the household (essentially the income).

Further developments of Alonso's work have proved to be useful tools to analyze a number of urban phenomena. The "bid rent" curves Alonso derives show the greatest rent a household wishes to pay for land, as a function of the distance of that lot of land from the CBD, for a fixed value of the utility. Extensions of that approach permit such factors as congestion, pollution, tax-impacts, etc., to be analyzed. They allow one to relax the monocentric assumption by the consideration of non-CBD workplaces. However, such changes typically lead to the loss of a direct analytical solution for the land market equilibrium in the city.

Even so, there is a serious loss of realism in viewing the employment as concentrated in a very small portion of the urban space. The more promising land use models have totally departed from this assumption; they represent the employment as spread over the total surface; then, work-trip-based models locate the families in relationship with their workplace: work-trip data, such as travel-time and cost, are associated with the choice of a residence.

An illustrative example is another Lowry derivative: the "Projective land use model" (PLUM). In this model, the probabilities of a trip to each destination, starting from one given zone, are calculated by the consideration of the radial length of the trip, and comparative attractiveness of zones associated to the same length. That interzonal probability matrix is applied to the distribution of employment to forecast the residential location.

The NBER approach, too, relies on a spatial distribution of households with workplace in a given zone. This distribution is generated by application of the "consumer theory", under the form of the minimization of a "gross-price", as explained below.

This brief overview of some existing land use models did not pretend at all to be comprehensive: the few mentioned models have been chosen to highlight the affiliations and origins of the NBER simulation. A more complete review can be found in different papers such as:

- "Seven models of urban development" by I.S. Lowry
- "Quantitative models of Urban development" by B. Harris.
- "Urban land use and Transportation Models: a state-of-the-art summary" by S.T. Putnam.
- Chapter 3 of S. Lerman's PhD. thesis.

Documentation about the NBER model, for the preceding and following issues, is available in:

- "The Detroit Prototype of the NBER Urban Simulation Model" by G. Ingram, J. Kain and R. Ginn.
- Progress reports are also published by the Bureau to acknowledge the changes occurring.
- Recent issues are discussed in a presentation by J. Kain and W. Apgar to the "Committee on Urban Economics Conference", Santa-Fe 1976.

2.2. The NBER simulation and the "new model"

The NBER urban simulation program has been initiated in 1968 and the first calibration was based on a study of the city of Detroit. Then, it appeared that more encouraging results would come out with the representation of Pittsburgh: "Pittsburgh I" was implemented in 1971. The current version, "Pittsburgh II" is far more sophisticated. In 1974, the U.S. Department of Housing and Urban Development had provided funding for the following research: to explore the impact on housing prices and quality of a direct cash assistance program for low income families.

The present section provides a rather brief overview of the NBER specifications which have been maintained in developing the new model, and a more detailed one of those which have been significantly changed.

Finally, an evaluation of the NBER strengths and weaknesses is made, and some future developments of the model are discussed. In the next three sections, the "new model" is referenced, but the descriptions apply to the NBER simulation primarily.

2.2.1. Specifications which are maintained

The following components are these portions of the NBER model for which no structural changes are implemented. However, two kinds of differences remain:

- the capacity constraint: as far as available data and computational means were concerned, the "new model" could not compete with the NBER implementation. For example, a seven times smaller number of housing types and a six times smaller number of household types has been used.

- the emphasis is not put on the same issues. The NBER, mostly because of its being currently funded by the "US Department of Housing and Urban Development" (HUD), is concentrating on a housing program evaluation: in particular, this justifies the consideration of many housing types. The new model, as stated in the introduction, is more transportation-oriented and focuses on level-of-service variables. Three submodels of the new formulation have that characteristic of not differing significantly from the NBER approach:

The Employment Location and Population submodels:

These provide exogenous data about

- the change of employment, for each of nine basic industries, and per zone. (the distinction of various industries is dropped in the new model).
- the consecutive changes of population, per workplace, income and education class of the primary worker. (the classification of households by education, as well as age of the head, is not maintained). The passage between the two steps relies on an empirically estimated matrix of job turnover and retirement.

The Mover and vacancy submodel:

It generates the households' moving decisions in the period, and finds the total number of housing demanders and vacant units. Those data are input into the demand allocation routine which in turn provides information to the supply sectors. For reasons largely developed in the next chapter, this sequential order as well as the demand allocation itself are entirely modified in the new version: the content of the mover and vacancy submodel however, is unchanged.

Filtering and Supply submodels:

These two processes represent the renovation and new construction activities of the suppliers, who are assumed to follow a profit-maximization behavior. For that purpose the income expected from any housing-stock change is compared to the costs or/and loss of future revenues attached. to the change. A profit maximizing integer-programming technique could have been used; for reasons of simplicity, a ranking and enumeration procedure has been preferred: each operation is assigned a "profit-rate", the form of which assumes that the risk involved and capital committed are proportional to the total cost of the activity.

The supply is subject to various constraints which have been initially or progressively established: they are zoning constraints, activity constraints, and one constraint of not exceeding the expected needs for the period: in the NBER model, these needs are estimated per housing-market, as opposed to the new model where there is a total substitutability between the types, and one unique demand constraint.

Another piece of information needed by the supply sector is the expected rentals. These are determined by the market clearing and price-adjustment process of the last period (this procedure has been changed in the new model; this issue is discussed in the next section).

In the early version of the NBER, these prices were indicated to both suppliers and demanders of the next period; in the latter version, namely "Pittsburgh II", the price-prediction is more sophisticated: the supply sector takes into account expected changes in neighborhood quality. This factor is thought to be of a considerable importance in the determination of prices in the medium-long run.

2.2.2. Demand Allocation and Market-Clearing submodels

The essential modifications brought to the NBER perspective amount to respecifying these two submodels. As explained in the next chapter, these two issues underlie the overall philosophy of the model.

The Demand Allocation

One role of the simulation is to assign the housing-demanders to the various markets: this requires generating the demands, then matching them with the availabilities per market. In the NBER simulation, since such a "market" is defined by one of the 44 residence-zones and 27 housing-bundles (50 in Pittsburgh II), the demand allocation would be a formidable task if a demand estimation procedure was to be calibrated for each "market".

Alternatively, the demand is first allocated to the various housing-types; then it is directly assigned to the different zones through a market-clearing process performed housing-type per housing type. (This order has been chosen because the type of housing is a more easily discretized dimension than is the location).

The assignment of households to housing-types is probabilistic: probabilities are generated, and interpreted as the proportions assigned to each type. The criterion which determines the choice of a type is called the "gross-price": it is a measure, aggregated all over the zones, of the housing and transportation costs, for the considered housing bundle. (The aggregation procedure is at the root of a major criticism encountered by the NBER formulation: this issue is largely discussed in the next chapter.)

On the basis of the gross-prices, the allocation is performed through a procedure which has been modified over time:

- initially, in the Detroit Prototype of the NBER, linear demand functions were used.

- In Pittsburgh II, the choice-theory is applied with a Logit formulation. This happens to be also the tool used in the new model to specify the demand (but in addition, the sequential process is dropped in favor of a joint-choice model).

The Market-Clearing and price adjustment

The market clearing is an assignment to residential zones of households having done the preceding choice of a bundle. An optimizing technique is used, and consists of a linear programming. The objective function is a total transportation cost to be minimized, subject to the availability-constraint in each zone.

This procedure assumes an individual behavior that attempts to optimize a collective aggregate: this assumption is grounded only in the case of a perfect market (this has rooted the choice of an alternative technique in the new model: a household makes its decision according to its individual preferences and the market data, primarily the prices).

At the same time that demand is assigned to zones, information is generated to modify the rentals which will be in effect in the next period. The dual variables of the program are the market-signals that cause a price to increase or not. The adjustment procedure itself is sophisticated:

. each zone is assigned a locational rent for the considered submarket, by a manipulation of its shadow-price.

. the "one-period price" is defined in the marginal zones, that is, a zone where an oversupply is observed.

. the "one-period price" in a non-marginal zone is obtained as the sum of the value in the marginal zone plus the locational rent of the zone.

. the "expected prices" are the result of an exponential smoothing between the one-period price and its previous value, so that only persistent trends can have an impact on the final prices.

Before undertaking the description of the new model in the next chapter, the following section summarizes the characteristics of the NBER simulation, and gives an insight into the future of its implementation.

2.2.3. Characteristics and implementation of the NBER simulation

The first part of this chapter has highlighted the way the NBER model has benefited from most of the recent improvements in the art of urban simulation:

- it is basically a work-trip oriented model: the workplaces are determined prior to the locational decision: the multiplicity of the employment-locations is recognized.

- it is a dynamic-adjustment model, which explicitly takes into account the existing housing-stocks and the process of their evolution.

- it involves a supply sector and matches the demand with it to assign the population to residential zones, and to modify the current prices.

- it performs a high level of disaggregation in the classification of the housing demanders, as well as in the screening of their alternative choices.

Besides, it is thought that the behaviorality of the model could be improved in two respects:

. another specification of the demand allocation, grounded on more attributes than the sole gross-price : many other factors play a role in the consumer-preferences, (the convenience of shopping in the zone, the proportion of non-white residents, etc....). Above all, a joint choice of zone and housing-type is recommended.

. Another assignment process and price adjustment might be more realistic in reflecting an imperfect market. This modifies profoundly the meaning itself of the market-clearing. It is thought to gain more behaviorality thanks to a more market-oriented description.

The following potential modifications are considered by the designers of the model:

- the consideration of a non-unique price of land, to take into account the imperfect substitutability of the construction activity all over the surface of a zone.

- making endogenous the employment location process, essentially the population - serving employment.

- some sophistication of the profitability forecast: the expected rental could be respecified as a function decreasing with the number of supplied units (instead of being held constant in each period, like it currently is).

- a modification of the demand allocation: people would choose first a kind of "macrolocation", that is one out of a small number of parts of the city; then, they would choose jointly a housing-type and their precise residential location.

Chapter 3: Theoretical Issues in the Adopted Formulation

The introductory chapter has pointed out the existence of several factors which do characterize a model because they relate to important choices of formulation. The current chapter addresses four of these basic issues:

- the capacity problem
- the joint disaggregate demand model
- the supply sector
- the market clearing process

3.1. The capacity problem

3.1.1. Introduction

Once the framework of a model is defined,there is the problem of the computational capacity needed to implement it. Whatever are the performances of the most recent computers, an urban simulation is an extremely complex process, which raises the problem of capacity.

This study does not include an implementation of the designed model at a full scale. It should rather be understood as a "feasibility study", which involves a simplified number of alternatives as well as decision-makers. However, it is desirable to address here the problem, for two reasons:

- Since the resources available for the study were at the same reduced-scale as the complexity of the description, the issue of capacity was still meaningful.

- For the potential users of such techniques - with the resources of a large city-administration - the costs involved by the model are more acceptable, but there is still the issue of not exceeding the capacity of even larger available computers.

This is why the adopted formulation tries to solve the difficulty at the reduced scale. The purpose is that the model could be implemented at full scale with no structural modification.

At the implemented level the problem is posed in the following terms. As it will be explained in section 3.2., the demand sector comprises

* 288 types of users
  - . 12 types of households
  - . 24 possible work-places

* up to 480 alternatives
  - . 4 types of housing
  - . 24 possible residential locations
  - . 2 modal choices
  - . 3 levels of auto-ownership (exclusive of the zero auto alternative if "car-to-work" is the modal choice.)

Therefore, if one would keep track of the distribution of all users among all available alternatives, this would require:

288 x 480 = 138,240 pieces of information

Several of these arrays (distribution of all inhabitants, of movers, of current house-seekers...) would be necessary to implement the various submodels.

That size was clearly out of reach at the level of the study; the same problem would be worse at the full scale. For a comparison, the NBER Detroit Prototype involves

72 household classes, instead of 12

27 housing types, instead of 4

which, by itself, multiplies by 40 the number of pieces of information. No further evidence is needed to justify the use of approximation methods, both in the feasibility study and the potential level of utilization. The method described now suits both cases. It aims at reducing the amount of storage needed.

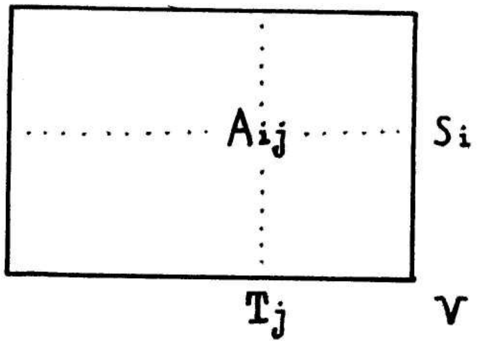
3.1.2. Reducing the amount of stored information

The method utilized in several parts of the model is based on the following idea: to decrease the size of the matrix to be stored, one should try to decrease the number of dimensions handled at the same time, and store a larger number of smaller matrices.

This can be clarified with the case of a two-dimensional matrix: all other manipulations are only a recurrent process of this first one, as it will be shown.

Therefore, considering a two-dimensions matrix (A_{ij}) where :

p is the number of values related to the first dimension q is the number of values related to the second dimension. The purpose is to decrease the p * q numbers to be stored: instead, (p + q) numbers will be stored, <u>the p horizontal sums Si, and q vertical sums Tj</u>.



The problem is: having stored the Si and Tj, how to create an "acceptable" value of Aij? All that can be done is clearly an approximation, since (p * q - (p + q)) pieces of information have been lost in the process.

But this was the whole point!

In mathematical terms, a system of (p + q) equations with p * q unknowns is to be solved, which means (p * q) - (p + q) degrees of freedom. The best that can be done is to point out one acceptable particular solution.

The system is:

$$\text{p equations}\left[\begin{array}{l} A_{11} +\dots+A_{1i} +\dots+A_{1q} = S_1 \\ \\ A_{p1} +\dots+A_{pi} +\dots+A_{pq} = S_p \end{array}\right.$$

$$\text{q equations}\left[\begin{array}{l} A_{11} +\dots+A_{j1} +\dots+A_{p1} = T_1 \\ A_{1q} +\dots+A_{jq} +\dots+A_{pq} = T_q \end{array}\right.$$

in reality, there are only $p + q - 1$ equations and $(p * q) - (p + q) - 1$ degrees of freedom, since there is the compulsary relationship:

$S_1 + \ldots + S_p = T_1 + \ldots + T_q$ = sum V of all elements in the matrix.

Fortunately, there is a simple form of a particular solution, which is:

$$\boxed{A'_{ij} = \frac{Si * Tj}{V}} \tag{1}$$

It is straight forward to check that this solution satisfies

$$\sum_i A'_{ij} = \frac{T_j}{V} \sum_i S_i = T_j$$

and

$$\sum_j A'_{ij} = \frac{Si}{V} \sum_j Tj = Si$$

Though the value defined by (1) cannot be taken as the actual value of A_{ij}, one can have an estimation of the validity of the approximation. If the A_{ij} in each given line (respectively column) are constant, the formula (1) gives the actual value A_{ij}. As a matter of fact;

if $A_{ij} = \frac{Si}{q}$ for each i and j, then the formula (1) gives

$$A'_{ij} = \frac{Si * Tj}{V} = A_{ij} \quad \begin{array}{l} (Si = q * A_{ij}, \\ V = T_1 + \cdots + T_q = q * Tj, \\ \text{since all } Tj \text{ are equal.}) \end{array}$$

In the general case, the validity of the formula (1) is related to the variability of the element a_{ij} in the line i and column j: this variability can be reflected by the ratio of the variance σ_i in line i (or σ_j in column j) to

the mean value $\frac{Si}{q}$ of that line (or $\frac{Tj}{p}$ of that column).
The validity of the formula (1) can therefore be qualitatively appreciated by the criterion:

$$\frac{\sigma_i}{Si/q} * \frac{\sigma_j}{Tj/p} = pq \frac{\sigma_i \; \sigma_j}{S_i \; T_j}$$

A typical satisfactory value would be .20

This method is not applied in the case of a two-dimension matrix (which is usually easily stored), but in the case of a multidimensional matrix. An example is borrowed for the model: it is the matrix RES (j,H,K,i) of the number of households

- with workplace in Zone j
- of household type H
- living in a housing of type K
- with residential Zone i

The full information would involve

24* 12* 4* 24 = 27,648 cells, which is too large.

The alternative has been to store the matrix RES(H,K,i) (1152 cells) and the matrix RES (j,i) (576 cells), and to apply the formula (1) by stating

$$\text{RES }(j, H, K, i) = \frac{\text{RES }(H,K,i) * \text{RES }(j,i)}{\text{RES }(i)}$$

This saves 27,648 - 1,152 - 576 = 25,920 cells!
In practical terms, the approximation consists in considering that the proportion of inhabitants of zone i who work in zone j is not dependent on their household characteristics,

nor their type of housing. The choice of the overlapped subscript i is extremely important: it must reflect the simplification which makes the most sense. The fact of overlapping decreases the bias introduced by the approximation. In each case, the approximations have been designed for the best trade-off between storage-saving and accuracy.

3.2. The joint disaggregate demand model:

In the introduction chapter, it has been pointed out how critical is the attempt to build a true behavioral model to insure the validity of the simulation. In the light of the other approaches reviewed briefly in the preceding chapter, the reader will be in position to appreciate the attempt to construct the model on as causal as possible description of the urban phenomena.

- This is true for the supply sector. The adopted NBER approach, as it has been seen, has a strong economic orientation: this reflects the profit maximization behavior of landlords and developers.

- This effort has to be still more emphasized on the demand side: the demand involves a large number of choices and decision-makers. The final choices interact not only with the characteristics of the households, but also between themselves in a complex process of allocation of time and money resources. Therefore, it does require a certain level of sophistication for a demand model to reflect the preferences of the consumers.

This is the crucial point where the model departs from the NBER perspective, and borrows the approach developed by S. Lerman in "A Disaggregate Behaviorial Model of Urban Mobility Decisions." His thesis includes an exhaustive description of the characteristics of the model,and the reader is referred to it for a complete understanding of the approach. In the present paper, only the most basic points will be discussed. These can be grouped according to three issues:

- the choice model: a multinomial Logit
- A disaggregate model
- A joint-choice model

3.2.1. The choice model: A multinomial Logit

The socio-economic behavior of the consumers of urban goods and services is captured by a utility-oriented approach. The consumer is assumed to be rational and to maximize his utility. On that basis, two objectives must be achieved:

-a) the utility must be specified to take into account all relevant factors in the preferences of the consumers, including their socio-economic characteristics.

-b) Assuming that the utilities are correctly specified, the entire process of forecasting the aggregated choices must be defined. The inherent uncertainties in the description of the utilities have an important result, namely, the probabilistic nature of the choice predictions.

a) Specification of the utilities:

The next paragraph explains the use of a "choice theory" according to which the consumer chooses among a "set of alternatives", each alternative being characterized by a list of attributes. On that basis, "specifying a utility" means to determine the list of relevant attributes, while keeping two important remarks in mind.

1) According to the mathematical form described in the next paragraph, an attribute having the same value for all alternatives must be made "alternative specific": its specification or its parameter must be differentiated as to which alternative is chosen. Otherwise it has no influence at all on the choice prediction. (The utility is specified under an additive form but is not an absolute value index: such an attribute would alter all utilities by the same amount, but the differences would remain unchanged).

2) A choice has been made between two possibilities:

. calibrate one utility (i.e. find out the parameters of its attributes) for each individual household type.

. calibrate only one utility, but take into account the characteristics of the consumer by some so-called "socio-economic variables" such as income, family size, etc..

the second option has been selected and has the advantage of being more simple. The other alternative is describing more accurately the differences between the various consumers.

With the chosen option, the same remark as the above applies: the socio-economic variables should not appear as such, but its influence must be differentiated as to which alternative is chosen. This should be achieved in a way that reflects the behavior of the choice-maker. For example, it is assumed that the income of the household is influencing the choice of housing, auto-ownership, and mode of transportation, by the intermediary of the "remaining income" after taxes, transportation and housing expenses. This "remaining income" represents the money resources which are available for general living expenses, leisure, luxury items... Hence, the choice of an alternative as opposed to another is modifying the impact of the socio-economic income-variable in the utility.

This chapter includes a brief review of the demand attributes: since they should be considered in connection with the set of alternatives, i.e. the "mobility bundle", their description belongs to section 3.2.3.

b) Forecasting the aggregate choices:

Beside the specification of the utilities, the demand modeler is facing several other choices of formulation. In the present case, three decisions are to be described to fully characterize the demand model:

- the use of the "choice theory"
- the use of a random utility model
- the use of a Logit formulation

1) <u>"Choice theory" versus "consumer theory"</u>:

The latter is at the heart of the classical "microeconomic theory": the consumer is choosing among a variety of goods and services, and is maximizing his utility by deciding the optimal amounts of each consumption, under an income constraint. The basic characteristic of that approach is the assumed continuity of choices. It is appropriate to describe the consumption of infinitely divisible goods. When the choices are inherently discrete, the concern of behaviorality leads to a preference of a discontinuous approach which reflects more faithfully the nature of these individual choices. These are represented by a set of discrete alternatives. An example where the consumer is faced with discritized alternatives, would be the following choice between:

- living in a walk-up apartment near workplace and owning no car, or

- living in a suburban area, in a single family dwelling and owning a car used for work-trips. In this case, a household is clearly confronted with discontinuous choices which cannot be described by the "consumer theory".

Hence, the definition of the choices is twofold:

- the dimensions of the alternatives are to be described,
- the alternatives are discretized to form the "choice set" available to each consumer.

The first issue is addressed in section 3.2.3. where the "mobility bundle" is defined; the second step refers to the issue of aggregation, developed in section 3.2.2.

2) The use of a random utility model:

For practical reasons, the number of attributes representing an alternative is limited; hence, this cannot be expected to reflect exhaustively the rationale of the consumers. Confronted with the same alternative , two different households may choose differently, for reasons not taken into account by the utility function. In order to maintain validity, the forecasts made on the basis of the utilities cannot be deterministic: only the probability of choosing an alternative can be inferred. That is the root of using what is called a "random utility model".

Then, to generate aggregate forecasts, the probability of a given alternative is interpreted as the expected share of the total group which will choose that alternative.

Practically the randomness is introduced by adding a "disturbance term" to the deterministic utility. The sum of the two is the "random utility" on the basis of which the choice is made deterministically - in that respect the choice is the one which maximizes the random utility.

This approach allows one to take into account factors that cannot be captured explicitly.

3) The Logit formulation

The use of a random utility model necessitates one more step: Some assumptions about the form of the disturbance term. The abundant literature on that topic is not discussed here; only the basic components of the adopted formulation are referenced. A first assumption is the independence of the disturbance across individuals; this excludes imitation or leadership effects.

Another essential assumption is the independence of the disturbances attached to different alternatives. This simplifies a lot the utilization of the model and makes it feasible to use the method for practical purposes. The assumption is grounded if much attention is paid to specifying properly the choice set, while avoiding any pair of correlated alternatives. Some attention must be paid in specifying the utilities: their deterministic part should capture as much as possible the similarities between two alternatives.

The last step is to assume an identical distribution of the random terms, according to a "Weibull distribution". This justifies the use of the logit formulation, according to which the probability of choosing the alternative i in the choice-set S is determined by the following fundamental formula:

$$P(i : S) = \frac{e^{Vi}}{\sum_{j} e^{Vj}}$$

where V is the deterministic part of the utility function. This formulation has been adopted, and is quite easy to manipulate. It has a further important characteristic known as "independence of irrelevant alternatives": the ratio of the probabilities of two alternatives taken in a subset is the same as when they are taken in the total set.

This provides a lot of flexibility. The model can be calibrated on a subset. Then, without changing its specification, it can be applied to the entire set of alternatives.

The logit formulation has been used successfully for an increasing number of studies: a proportion of them deal with transportation, like the work by T. Adler and M. Ben-Akiva, "A Joint Frequency, Destination and Mode to Work Model for Shopping Trips".

3.2.2. The issue of aggregation

The concept of aggregation applies to both consumers and alternatives.

Level of aggregation of the choice-makers

The model is a disaggregate choice model, which means that the aforementioned choice theory applies to the individual decision-maker; aggregated forecasts are then obtained by summation.

The use of disaggregate models has increased in the recent years due to the qualities of that approach. In the present case, these qualities are basically a higher statistical validity, but it is not true that the disaggregation improves the behavioriality of the model. The improved validity comes from the fact that no aggregation-bias is introduced prior to the application of the choice theory. However, this perspective is only theoretical, because it assumes that the model is applied separately to each individual characterized by his own set of socio-economic variables. That would be obviously unmanageable in most situations. Therefore, there is a need for some aggregation, which necessitates the definition of the following two:

- the appropriate decision-level
- the suitable grouping of these decisional units.

The decision level that has been chosen as the most realistic is the household defined as "a group of persons living in the same dwelling and sharing the same economic decisions".

The suitable grouping of households is an important issue, and must be such to avoid any important "within-variances" inside the groups. This leads one to define a greater number of household types, but it heavily increases the capacity requirements: the chosen number should be consistent with the overall size of the model.

In the present model, it was chosen as 12 instead of the 72 types of the NBER formulation. However, one dimension has been added: it is the race of the people in the household. The figure of "12" is the result of the cross-classification of:

* 3 income groups: below $8000 of annual income, between $8000 and $14,000, and above $14,000.

* 2 ranges of household sizes: two or less persons in the household, or more than two.

* white or non-white household.

In fact, the disaggregation of the decision-units goes much further than that. Each household type is also classified according to "its" workplace: this will be discussed in the description of the "mobility bundle" of section 3.2.3. What must be said here is that the model generates the choice of households satisfying two conditions:

.1) the household has one breadwinner: this excludes the case of the non-working, but not the multiworker families: their choices are made considering the "primary worker".

-2) The breadwinner currently has a job in a given workplace. The choice of his residence, described in the model, is heavily dependent on the consideration of time and cost of his trips to work.

Since the workplace is chosen <u>prior</u> to the model, it has to be an input for the mobility decision-making. As a consequence, the decisional unit is defined by

- one of the 12 household types.

- one of the 24 workplaces. (the city is divided in that number of zones, for both purposes of workplace and residential zoning).

This provides quite a disaggregate description of the population, but it implies an increased computational cost. It is thought that an accurate representation of the behaviors is achieved that way.

Furthermore, the model escapes a serious source of error which often deteriorates disaggregate predictions: once the probabilities of choice for each decision-maker are forecast, the distribution of those has to be known in order to reach aggregate results; this distribution is often not available, and some bias takes place.

On the contrary, this model keeps track of that distribution throughout, and the problem of aggregation is solved by a simple addition: no error is introduced, there is only an initial aggregation-bias coming from the consideration of segments of population instead of individuals. Because a fair level of disaggregation is achieved, this bias is presumably quite small: this is documented in F. Koppelman's paper "Travel predictions with Disaggregate Choice Models".

Aggregation of the alternatives

The multidimensional nature of the alternatives is addressed in the next section. For the same reasons as before, each dimension cannot be described by a huge number of values of its attributes. For example, the available choice of residential location and housing cannot take into account every vacant unit in the city. Instead, each dimension is discretized.

For the preceding example, the discretization amounts to a choice between:

- 24 residential zones (the same as before)

- 4 housing types: single-family dwellings (with the option to own or rent), garden-style units, or walk-up apartments.

This poses a problem of consistency. Since the demand model forecasts the probabilities of choosing each alternative, the grouping of these alternatives should be taken into account: in other words, if it is decided to group two zones into one, the probability attached to this new zone must be the sum of the two previous probabilities. This requirement is easily satisfied with the Logit formulation (a detailed development of that topic is in S. Lerman's thesis, in chapter five): a "size-variable" is added to the utility of the alternative.

- The parameter of that variable has to be "1"

- Its definition necessitates the knowledge of what is perceived as the "size of the alternative". The area of the zone is in a sense a size-variable, but it is not a good measure of the attractiveness of the zone for people willing to live in a given housing type. It makes more sense to consider some volume of the stock of that housing type in the zone:

*if one uses the number of vacant units, it does reflect the size of the available alternatives. However because of the mathematical form of the model, this would implicitly amount to subjecting the demand to supply limitations. In particular, no demand could be generated in a submarket where the supply itself is equal to zero. That would destroy the very meaning of the overall formulation: the demand submodel should generate demands ex-ante, that is reflecting the actual preferences of consumers. Then, it is the role of the market-clearing to match that demand with the supply: this is done by an explicit procedure which essentially causes an adjustment of prices.

* Therefore, the adopted size-variable is the sum of the currently occupied plus the available stock of the given housing-type in the zone. Involving the total stock obviates the problem of the preceding page. The introduction of the available units - including new constructions - allows the demand to be generated in a zone where that housing-type was not existing previously.

3.2.3. The joint choice model

The last important characteristic of the demand is to be represented by a joint choice model. The alternatives are multidimensional and the decision is assumed to be made jointly: this can imply making some concessions like "living in a single-family dwelling, but far away from the workplace".

Choice between a joint and a "sequential" formulation:

These are constant options of choice-modelling. (This issue must not be confused with the possibility of using or not using a sequential estimation of a jointly structured model!) The sequential form assumes that decisions at one step are contingent upon the decisions of a previous step; for instance, the comprehensive set of decisions made by a household is an example of sequential choices. As a matter of fact:

- it has been discussed that the initial decision relates to the employment location. This takes place for the long run, like several years.

- Conditional on it, the "mobility bundle" is a medium-term consideration: its description is at the root of the studied model.

- Depending on that step, the so-called "travel choices" are made by the households: they deal with frequency, destination, mode, time of day and route of non-work trips. They are clearly sequential to the preceding step: as a matter of fact, the selection of a mode to go shopping is heavily dependent on the number of autos available. The auto ownership decision is made earlier in connection with "strategical" choices such as "allocating money to buy a car versus living in a more expensive house".

For some other cases, the true behavior is a simultaneous choice of several items which interact in a complex way, and cannot be considered as hierarchically ranked. Here, the use of sequential formulation is a misspecification; it cannot reflect the true behaviorality. It is thought to be the case of the mobility bundle, the description of which points out now the interconnection of its various components.

"The mobility bundle"

The contents of this section have been referenced several times before; it describes the components which underlie the joint choice demand model. Since this model is utilized as it has been calibrated in S. Lerman's thesis,the reader should refer to that study to get a more detailed description.

The mobility bundle comprises four elements which are all medium-term decisions, and closely interact:

- The residential location has been described as a choice between 24 zones.

- Related to that choice is the one of the housing-type. For instance, living in the center is associated with a smaller probability of staying in a single-family structure. Furthermore, the same housing-type corresponds to different lot-sizes in the different zones. The set of a housing type and a residential zone will be referenced as a "market".

- The decision of auto-ownership is linked to the previous choices: primarily, there is less rationale (because of a better transit network), and more drawbacks (because of higher costs and congestion) to own cars in the central zones. As opposed to a number of other attributes of that dimension (like quality, price...), only the number of cars determines different alternatives:zero, one, or more than two cars are the levels used. The cost is the same for all kind of cars, but is taken into account in the demand process through a fixed annual charge per car.

- Finally, the "mode-to-work" depends heavily on the preceding choice; it belongs to the mobility bundle as opposed to other purpose-trips which belong to the subordinate level of "travel choices". The rationale is that it influences the most other decisions such as the location of residence. An illustrative example is the zones not serviced by transit, where only car-owners can decide to live. Only two alternatives, car-to-work and transit, were considered, since they cover a very large proportion of the cases.

Sensitivity to Transportation impacts:

This issue has been at the root of the essential modification brought to the NBER formulation, that is adopting the joint choice demand model. This point illustrates the advantage of jointly representing decisions when they are thought to be simultaneous. For the purpose of comparison, a brief summary of the NBER perspective is useful:

The theory is that consumers make a sequential choice of the housing type, then of the residential zone.

The first step is performed essentially by comparison of the "gross-prices" of the various types. These include the rental and a comprehensive transportation cost. For one household and its given workplace, one gross price is computed for each housing type K, and each residential zone; then, one unique gross-price is associated to each type K: it underlies the probability of choosing that type. This unicity necessitates some method to aggregate over the various zones. In the Detroit Prototype, that was achieved by a weighted averaging of gross prices. In the latter version, the measure is the gross price of the unit at the lowest fifth percentile.

Whatever is the method, it decreases the sensitivity of the model to transportation impacts, because local phenomena are not well represented. Since the documentation of this fact involves the entire model, it is easier to clarify it through an example:

Assume that the authorities want to subsidize a high-rise housing program for low income families, with workplaces in the central zones. Transit Line, connected to the center, is extended to a close suburb: it is expected to implement there the construction program. Most of the families in the segment have no car; they are Transit captives, and the Transit line offers to them a new possibility of residential location:

- In the NBER perspective, the attraction effect of that market in that zone will be diluted by the aggregation of the gross prices. Hence, the high-rise structure market as a whole will capture only a fraction of the segment which it would actually capture. As a matter of fact, more people in the reality would choose this housing-type because they would have simultaneously selected the new zone.

- In the new model, there is no a priori subdivision into housing markets: the housing-type and location are decided jointly. In the example, the true fraction of the segment will choose the new opportunity. It will generate the supply through a price effect (compensated by the subsidies): finally, the physical prediction of the model will differ from the NBER.

Therefore, the model offers more possibilities to appreciate the effect of various transportation policies. Because of that, some emphasis has been put on the Transportation-related data:

- In the inputs, several types of information are given for all origin/destination pairs: the distance, in-vehicle and total travel times, and the costs for car and transit; all of them play a role in the demand process; they can easily be modified to reflect any transportation policy.

- Several outputs are related to transportation issues:

*the auto ownership per zone and household type or housing type.

* the share of transit ridership per residential zone.

* the share of car-to-work trips per zone and household type or housing type.

* for auto-trips, the total Vehicle-Mile-Travelled for work-trips per day (V.M.T.) is a relevant measure: it is used to evaluate several policies such as energy-saving or anti-pollution policies.

* one measure of satisfaction of the people is tested: the "generalized speed", reflecting the velocity with which the population of one zone is, on the average, transported to work. The measure seems more useful than the existing "isochronic lines" (a set of points from which the center of the city can be reached within a certain time): these are too center-oriented, instead of considering the actual work-destinations, which can be circumferential. The "generalized speed" is the average over all destinations and the two modes of all speeds (defined as the ratio of physical distance over actual total travel time). The averaging is weighted by the corresponding flows. The criterion can easily be modified to take into account differences of perception of out-of-vehicle and in-vehicle travel times. The present definition was chosen to reflect true figures of speed.

It is expected that the speed is higher for long distance trips (the influence of access-time being proportionally lesser):the comparison should be made between zones of the same category, like central, close-suburbs, far-suburbs.

Such criteria may be useful to document a transportation planning: an illustrated view of the problem is provided by drawing maps which reflect the information (some examples are included in chapter four).

3.3. The supply sector:

This section is briefer since,on the supply side, there are not so many differences with the NBER formulation. It addresses the method of generation of the demand (as opposed to its allocation precedingly described), the supply process itself, and some issues of price and profitability.

3.3.1. Generation of the demand:

The supplies of housing are assumed to anticipate what will be the demand for housing in the next period. This is generated in the "mover submodel", then input in the supply routine itself. It comprises four sources;

- demographic changes result in the creation of new households in search of a dwelling. They also cause the dissolution of households because of the mortality or other reasons. The latter process generates available units, but no addition to the demand.

- The intrametropolitan moving (for which statistics are available by economic group , essentially the age of the head). It increases both the demand and the number of vacant units.

- The in-and-out-migration to and from the city, the balance of the two, together with the balance of the demographic changes, explains the population growth of the city.

All the preceding data are available in the form of rates. Applied to the population figures per household type and workplace, they give the mobility forecasts for the next period.

- They are added to the residues of the last period to form the total demand: these residues and the "non located people", that is, the people who have not found the housing they wished to. At the same time, there are some "vacant units" which have met no demand in the last period. The existence of these two categories was the chosen formulation, as opposed to a "forced allocation" of non-located people to some vacant units. The latter is what happens in the real world, followed by a new attempt to move in the next period: the formulation has the same effect in all respects, except that theoretical view of provisional "hotel residents": this criterion is one way to evaluate the degree of convergence of demand and supply, as is discussed in chapter five.

3.3.2. The supply process

A percentage is added to those forecasts to take into account the imperfection of the market: it can be termed the "dynamic vacancy rate" since it reflects the needed oversupply to locate all the housing-demanders. The total indicates to the suppliers a maximum, but not a minimum, number of units to provide. It is called "total demand".

The behavior of the supply sector is now explained:

The supply is a medium-term issue, as opposed to the demand. Building a new housing unit takes more time than searching for a suitable dwelling. This is basically why the supply process takes place only once per period, while the demand and market clearing are iterated to reflect a short-term equilibrium. (A "period" is any medium-term lapse of time, which is chosen in relationship with the horizon of the simulation. The implemented case is a four years simulation, and the period is equal to one year).

The description of the supply relies on the assumption of profit-maximization; this determines:

-1)the maximum number of operations, which is the previously introduced "total demand". This departs from the NBER formulation, since there is one overall constraint instead of one per housing-type: therefore, it assumes a substitutability of housing-types in the preferences of the consumers. This assumption is linked to the jointness of the choice process: because the type and the location are two dimensions at the same level of the mobility bundle, an alternative can be substituted for another, even if they differ in the two characteristics. In that light, the suppliers do not segment their predictions of demand; they only rely on the information of the last period, communicated by the prices.

This issue is discussed further in chapter five.

-2) the order in which they are undertaken, which is the order of their profitability: the latter is reflected by the ratio of the future inflows (that is the stream of rentals) over the current and future outflows;

* the current outflows are the construction or modification costs.

* the future outflows are the foregone rentals if a unit is transformed into another; if the construction is made from scratch, they are the price of land, made equal to the present value of theoretical "land rentals" for computation convenience. That issue involves a consideration of prices and capitalization factor which is addressed in the next section.

Finally, the supply activity is subject to various constraints which reflect the real-world situation: they have been elaborated to take care of some unrealistic outcomes, in particular by the designers of the NBER simulation.

There are three of them:

-1) One input constraint; in the case of a transformation, it is obvious that the number of new units is limited by the number of vacant input units. It is not a one-to-one correspondance if the lot-sizes are different for the two structures.

If the activity is construction from land, a constraint is imposed to reflect the reasonable amount of activity tolerable within a period. Only 10% of the remaining developable land can be built in each year.

-2) Another limitation reflects the same concern of a maximum construction activity. Since one type of structure can be built from several inputs, it is also imposed that the total stock of each type in each zone must not increase by more than a given "market absorption factor" of 10%.

-3) Finally, there is a constraint to perform a realistic distribution of the activity all over the city. No more than 20% of the total output of one type can be built in one given zone. The results show that all those constraints are operating. For example, the first limits the amount of new construction in the center; the third regulates the building of single-family dwellings in the suburbs.

3.3.3. Some issues of price:

The rationale of landlords and developers is based on the prices they expect for the rentals, or selling prices expressed as the present value of future rentals. This point is the basic link between the market-process and the physical development of the city. The prices are the vectors of information between the short-run variations of the demand and the medium to long-run modifications of the city. Several difficulties are raised from the attempt to use prices which at the same time:

- are those vectors of information

- are realistic representations of the true market-prices. Some of these difficulties are explained below.

<u>Rental prices</u>:

The rentals are the outcome of the market clearing, as explained in the next section. Basically, the prices decrease or increase accordingly to an over- or undersupply; this latter case can be generated by two situations:

- there is no supply in the considered market due to a lack of profitability. The demand will put a pressure to increase the price until the operation becomes profitable. The market-clearing is designed to reflect that increase realistically in timing and magnitude.

- there is a positive but insufficient supply, because of one of the constraints; here the real-world consequence is not the same price increase as before: this would not increase the supply, but only discourage the demand until the equilibrium is reached at some unrealistic high price. In reality, the price would level somewhere, and the market would provide some other type of information than the price to discourage the demand.

Several remedies have been thought of:

* A control can be placed on the prices to avoid sharp increases. For instance, it can be smoothed with the preceding values. However, that would be at the expense of the market-equilibrium, which is at the core of the model.

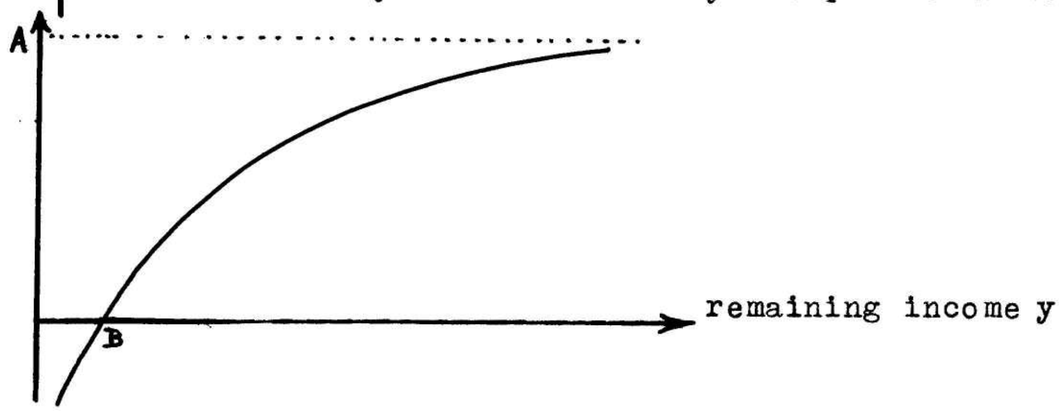
* In opposition to the aforementioned choice of taking a "size-variable" equal to a total number of units, this variable can be taken as the number of available units: this would carry the information and discourage the demand without increasing the price to an exaggerated degree. However, for the reason stated in section 3.2.2., it destroys the meaning of the market-clearing in the general case. That is why that alternative has not been adopted. Therefore, the determination of the price has not been changed: the success of the convergence process has been preferred to the realism of price. Therefore, in some cases, the price may be over-estimated because of the very nature of the model (however, it does not affect the validity of the prediction for the physical stocks):

furthermore, the more sensitive the demand is to a price-increase, the less disturbing it is.

In that respect, the current formulation of the choice model can probably be improved: all costs are influencing the demand through one variable "remaining income". It is the after-tax revenue of the household less the transportation and housing costs. The variable itself is the logarithm of the remaining income: this probably causes a too small sensitivity to prices. The elasticity should be bigger, especially for small values of the remaining income. An alternative is to use an harmonic function of the remaining income y:

$$A \left(1 - \frac{B(H)}{y} \right)$$

where B(H) is a "base remaining income" for the households of type H.
A is the utility of an "infinitely cheap alternative."



In the Logit formulation, this would be the same as using the simple variable B(H)/remaining income, because the additive constant A has no impact on the probabilities of choice.

Price of Land:

Another difficulty arose with the determination of the price of land. The theory is as follows:

Two units of the same housing-type but in two different zones have different values (practically, their rental prices are not the same).The difference is a locational rent which can be described as the variation of the price of land under that housing type (this defines the surface of land price within an additive constant which is the land price at the fringe of the city: the latter can be taken equal to the agricultural land price).

The price of land in a zone is defined as the weighted average of the price of land under the various housing-types.

The difficulty is that it generates unrealistic high values. Two explanations are available:

- the difference of rental is also due to a differential of quality. By attributing it solely to locational rent, one overestimates the price of land.

- it may be relevant to consider the various prices of land under one housing-type rather than the averaged value. One reason is that all lots are not usable for all housing-types, because of possible zoning constraints. The unique price was assuming a total substitut ability of land,which is not realistic.

The initial formulation has been empirically manipulated to reduce the land prices to reasonable values: an exhaustive study of the problem seems necessary to come out with a more adequate representation.

The capitalization factor:

This is needed to transform a capital expense into a stream of future values, or vice-versa. It takes place when computing the profitability ratios of the suppliers of housing. It has been seen that the denominator may include a capital expense such as a purchase of land, a construction or transformation cost. This expense is made homogeneous to an annual rental by applying a rate which is the time-preference rate of landlords and suppliers.

This rate has been estimated as follows:

- 4% of pure profit rate before income tax (3% after)
- 2% as a risk-premium for the non-liquidity of the investment
- 4% of depreciation cost, assuming that the maintenance is supported by the landlord.
- 1.5% for the charges to be paid by him (heating, electricity...)
- 1.5% for property taxes

This amounts to a 13% rate, or a 7.7 capitalization factor.

For owners, there is a saving of 1% income tax, so that the rate is 12% and the factor 8.3. For the land, there is no depreciation nor charge: the rate is 7.5% and the factor 13.3.

It seems a posteriori that those rates may have been overestimated by something like 1% to 2%.

3.4. The Market-Clearing:

This part again departs entirely from the NBER approach. The theoretical basis is different, a consideration of convergence is added and the equilibrium prices have another meaning than in the NBER model.

3.4.1. The theoretical basis

The fundamental choice is to use a dynamic price adjustment mechanism as opposed to an optimizing technique, like the NBER linear programming approach. In the theoretical case of a perfect market, the two methods are equivalent. Two reasons have been thought of to prefer a dynamic adjustment:

- If the optimization technique had been chosen, the NBER minimization of an aggregate gross price would not have been chosen. The demand choices involve many other considerations that the prices (as many as there are components in the utility function). Therefore, an aggregate utility would have been maximized.

This would have posed a major problem: the available utilities are only the deterministic parts of the true "random utilities"; only these latter represent the actual preferences of the consumers (the first ones are intermediaries to compute the probabilities of the different choices). Unfortunately, they are unknown, which is the least one can expect from disturbance terms!

- In reality, the market is not perfect. Basically, there are time-lags and viscosities. It is then more realistic to model the true interactive assignment, and to represent those imperfections, than to rely on an optimizing technique.

The key-factor in the chosen approach is the confrontation of demand and supply on each "market", that is in each zone and for each housing type. This process is a short-term one. Furthermore, there is a continuous pressure on prices reflecting the every-moment preferences of housing-seekers. A model cannot capture that continuity. It was thought that a satisfactory description is to <u>implement several market-clearings in one period in order to reach the equilibrium.</u> In each step, the prices are modified according to the excess supply or excess demand; then, the demand allocation model is iterated and the demand progressively adjusts to the supply situation. It is important to underline that this does <u>not</u> involve the supply activity which is implemented only once per period.

These iterations have a high computational cost. In fact they comprise the major cost of the simulation. However, it is desirable to maintain that repetition for two reasons:

-1) If no iteration takes place and the modification of prices is merely input in the next period, the model fails its major objective; there will be no equilibration at all in the short-run, and a weak trend of convergence in the long-run. That is an unacceptable departure from the reality where there is an equilibrium, within a small range of market-friction.

-2) If the demand is iterated only once to meet the supply, the convergence will be improved, but still be quite imperfect. Furthermore, to have some efficiency, the model would require an abrupt manipulation of the prices; this can lead to excessive housing values. A bigger number of smaller iterations enjoys more realism.

The adopted formulation is to iterate until either of the two occurs:

- the convergence-test, explained in the next section, is met, or

- the number of three iterations is reached, which means four implementations of the demand allocation.

In each iteration, the prices react to the under- or oversupply, expressed as a percentage p. Furthermore:

. the market reacts only beyond a disequilibrium point of 5%.

. After 5%, the adjustment takes place with a decreasing elasticity when the disequilibrium increases. For extreme values, there is a maximum change E depending on the sense of the adjustment.

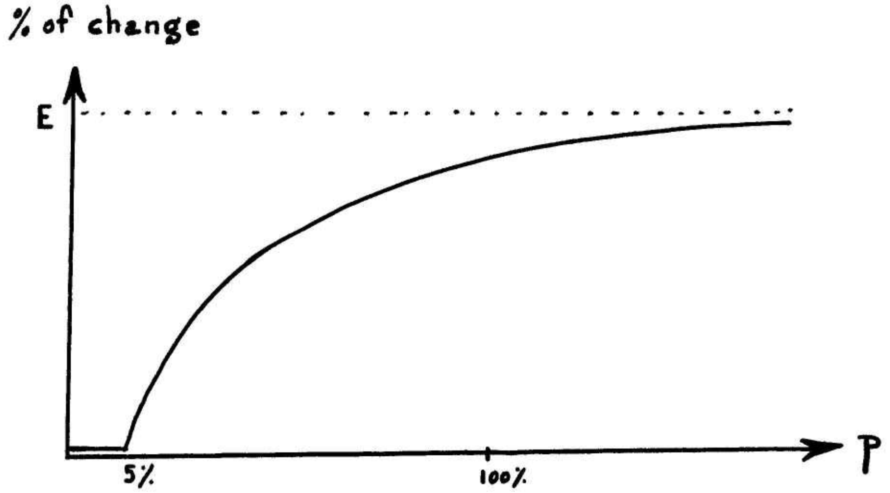
- it is equal to 20% per iteration when the price increases. This means a maximum of 73% increase within a period, if the disequilibrium pertains despite the market mechanism.

- In the case of an oversupply, it is thought that the prices do not decrease with the same elasticity. Landlords and developers "protect" somehow the market against noticeable collapse. Per iteration, the maximum is 10%, or 33% for three consecutive iterations.

The response curve has the following form.that relies on a summary empirical calibration.

$$\% \text{ change} = E * (1 - .17/(.12 + p))$$

where
E is the maximum percentage of response.
p is the percentage of disequilibrium.



3.4.2. The convergence test:

Whatever is the efficiency of the market mechanism, it cannot be expected to realize a formal equilibrium. This would require an infinite number of iterations; besides, there is in the real world some friction which prevents the market from being in a perfect state of equilibrium. Consequently, there must be a range in which the convergence is considered as realized. When this test is satisfied, no further market-clearing iteration is required (otherwise, they take place up to the maximum number of three). The test is defined as follows:

First is computed an "average rate of undersupply" for the entire city: it has been seen that the supply activity is limited by the number of "total demanders", but is not required to reach that number. Therefore, the ratio of final availabilities to "total demanders" is less or equal to one: its complement to one is the average rate noted "AVED".

In that light, the equilibrium means that the pressure exerted by the excess demand is equally spread over all the "markets"; it should not be a zone i and housing type k for which the demand is satisfied while a large undersupply takes place elsewhere. Each market should absorb a portion of the total undersupply, proportional to its size, that is the size of its supply.

In mathematical terms:

DEMD (k,i)	are the	demand on the "market" (k,i)
AVAL (k,i)		supply

TDEM, TAVL are the average values.

At the equilibrium:

$$DEMD(k,i) - AVAL(k,i) = \frac{AVAL(k,i)(TDEM - TAVL)}{TAVL}$$

or :

$$\frac{DEMD(k,i) - AVAL(k,i)}{AVAL(k,i)} = \frac{TDEM - TAVL}{TAVL}$$

The left hand side is the local rate of disequilibrium p(k,i)
The right hand side is the average rate AVED

That is for the perfect equilibrium. The "convergence test" is met when the two figures do not differ by more than 3%.

$$\left| p(k,i) - AVED \right| \leq .03$$

That is not quite yet the formulation of the convergence test; instead of implementing it for each "market" (k,i), some substitution is allowed: the test has to be met only for each housing market k. The four figures which are submitted to the test are averaged over the zones. The weights that have been chosen are the size of the supply in each zone: it makes the test easier to meet since the "hard" zones are usually undersupplied and their rate is affected by a small weight.

A more demanding test would be to weight according to the demand in each zone. As it is, the test is:

$$\sum_{i=1}^{24} \frac{AVAL(k,i)}{TAVL} \left| p(k,i) - AVED \right| \leqq .03$$

for each housing-type k.

3.4.3. Characteristics of the equilibrium

The equilibrium is now explicitly defined. Within the test of convergence, the allocation of demanders to each "market" is known; two issues have not been addressed; at what prices is the equilibrium met? How are the various types of housing-demanders distributed? The choice of a market-clearing process is not neutral in those respects:

Existence of equilibrium prices

Since the mechanism relies on a price adjustment, it is important to know whether they are one, or more than one, vector of prices that allow supply and demand to balance. This necessitates some mathematical development about the impact of rentals upon the probabilities of choice:
A "market" (k,i) is represented by a unique subscript l.

A decisional unit (a household of type H with workplace in j) is represented by m.

S_l : total supply of l (it exists L "markets" in total)

D^m : total demand of m

D^m_l : demand of the market l by m

S : total supply

D : total demand

The equilibrium defined in the preceding section is written:

$$\frac{D_1}{S_1} = \frac{D_2}{S_2} = \ldots\ldots\ldots = \frac{D_l}{S_l} = \frac{D}{S} \qquad (1)$$

since that is equivalent to the equalization of the pressures on the different markets, which is written:

$$\frac{D_1-S_1}{S_1} = \frac{D_2-S_2}{S_2} = \ldots = \frac{D_l-S_l}{S_l} = \text{AVED} \left(= \frac{D-S}{S}\right)$$

The logit formulation implies:

$$\frac{D^m_l}{D^m} = \frac{e^{U^m_l}}{\sum_{l'} e^{U^m_{l'}}} \qquad (2)$$

where U^m_l is the deterministic utility of the alternative l: the rental R_l associated with it plays a role through the sole variable "logarithm of remaining income", with a parameter a; it is written:

$$a \, Ln(Y^m_l - R_l).$$

Y_l^m includes all other revenues and expenses, among which the Transportation costs (the Transportation subalternatives are not considered in (2); there is no loss of generality for the current development; it suffices to consider that the suscript m designates a decisional unit and one Transportation alternative. This simplifies the following calculations).

If V_l^m is the exponential of the sum of all other variables, then:

$$e^{U_l^m} = V_l^m (Y_l^m - R_l)^a \qquad (3)$$

and (2) becomes:

$$D_l^m = D^m \left[\frac{V_l^m (Y_l^m - R_l)^a}{\sum_{l'} V_{l'}^m (Y_{l'}^m - R_{l'})^a} \right] \qquad (4)$$

Since $D_l = \sum_m D_l^m$, the condition of equilibrium (1) becomes:

$$\boxed{\sum_m D^m \left[\frac{V_l^m (Y_l^m - R_l)^a}{\sum_{l'} V_{l'}^m (Y_{l'}^m - R_{l'})^a} \right] = S_l * \frac{D}{S}, \text{ for each } l} \qquad (5)$$

For the purpose of analyzing the prices R_l, (5) is to be viewed as a system of L equations to the L unknowns R_l.

In reality, the system comprises only (L - 1) independent equations: the L equations are linked by a trivial equality obtained by summations of left and right hand sides:

$$\sum_m Dm * 1 = \sum_l S_l * \frac{D}{S}$$

which is nothing else than:

$$D = \frac{D}{S} * S$$

Therefore, there is one degree of freedom in the system, and the vector of prices is known only when one of them is fixed.

In other words, when the market clearing iterations lead to equilibrium, this can be performed by a non-unique modification of rentals. It is of importance to analyze what reality reflects that non-unicity of a solution. For that purpose, some knowledge is useful about the following issue: how two vectors of rentals compare, which realize the same equilibrium?

It is provable that all prices move in the same sense. If $\bar{R}$ and $\bar{P}$ are two solution - vectors:

$R_l < P_l$ for all l, or $R_l > P_l$ for all l.

demonstration:

The situation $R_l > P_l$ and $R_k < P_k$ is now proved impossible.

for all m: $$\frac{D^m_l\,(R)}{D^m_k\,(R)} = \frac{(Y^m_l - R_l)^2}{(Y^m_l - R_k)^2} < \frac{(Y^m_l - P_l)^2}{(Y^m_l - P_k)^2} = \frac{D^m_l\,(P)}{D^m_k\,(P)}$$

by summation, that would imply:

$$\frac{\sum_m D^m_l\,(R)}{\sum_m D^m_k\,(R)} < \frac{\sum_m D^m_l\,(P)}{\sum_m D^m_k\,(P)}$$

which is illogical since, according to (1), both sides are equal to D_i/D_K

Consequently, one solution-vector is derived from any other by an increase or a decrease of all rentals in the vector (it can also be shown that the higher rentals in the vector increase - or decrease - less than the lower: the demonstration involves a differentiation of the system (5) and is somehow long to report here).

Interpretation of the non-unicity of prices:

A market-clearing interation is a price adjustment to respond to a lack of balance between demand and supply: since the process itself leads to a new equilibrium, it also determines which solution-vector will be reached. It is essentially in the response-curve of section 3.4.3. that that information is buried.

What reality should reflect this formulation? The market has some viscosity: the house-seekers do not have a perfect information: it takes time - or forever - for them to discover all the alternatives; the equilibration between under- and over-supply does not function perfectly: this viscosity puts more or less pressure on the markets, and causes the prices to increase accordingly. Consequently, the market will reach an equilibrium at one level of prices, or at a higher one, according to its level of imperfection.

That reality is to be captured adequately by the market clearing process. This is important because not only the level of prices will depend on the mechanism of price adjustment: the allocation of "markets" to housing-demanders will also be modified, and hence the physical evolution of the city.

As a matter of fact, if the general level of price is modified upwards, the weight of the rental will be heavier for the lower-income family who will shift progressively to cheaper housing-markets: in other words, the income-segregation is reinforced.

In conclusion, the designing of the market-clearing process is of much importance to capture the true behavior of the market. The exposed formulation is thought to be qualitatively sound: it is, in no way, supposed that the figures used are accurate: there is still a serious need for a statistical work to improve the representation of the real-world phenomena. This is true for the price-determination and several other "innovative" issues that have been exposed in that chapter. It is recalled that this work is a feasibility study rather than a true-prediction effort: the next chapter documents that issue of feasibility with the implementation of a case-study.

Chapter 4 - The implementation of a case-study

Thus far, the theoretical grounds for a new formulation of urban simulation have been exposed. This chapter consists of a feasibility study and describes the results that have been obtained.

4.1. The Case - Study:

This section includes the description of the modeled city and of the data needed to initiate the simulation.

4.1.1. The modeled city

The city of Washington has been chosen: in fact, the entire "Statistical Metropolitan Standard Area" (SMSA) is considered as the urban field of experimentation.

This choice has been made on several grounds:

- It is a large city where all relevant urban factors are represented. In particular, there is an important transit network which services more than half of the residential zones. The remainder of the zones are available alternatives only for car owners.

- A fair amount of data are available. The two most useful sources have been various Census tables, and the 1968 "Washington home interview survey". This latter provides a lot of information about the housing and transportation decisions of the households. Since a large portion of the initial data comes from that source, the implemented simulation is starting in 1968.

- The most important reason by far is the availability of the demand-model discussed in the preceding chapter: the logit formulation requires a large calibration effort which would have been out of the reach of this thesis. Fortunately, S. Lerman, who has built this demand specification, has implemented it on the same example as Washington, D.C., 1968. The behavioral structure of this model allows some transference, but rather in time than in space. Furthermore, certain variables apply exclusively to Washington: an example is provided by the per-pupil school-expenses variable, which take positive values only outside the District of Columbia. Hence, it was recommended to start from the same case study, in order to use the calibrated model as it was.

The issue of aggregation of locational alternatives has been addressed in chapter three; twenty-four zones have been used. They respect the existence of circumferential rings: three of these rings are located inside the square of the District of Columbia, two of them are outside the D.C. area and delineate a close and a far suburban area.

Each ring is divided in several zones which have been designed in a trade-off between two objectives:

- the zones should be of a comparable size of population.
- they should be as compact as possible to insure the meaningfulness of the transportation data aggregated in each zone.

The map on the next page illustrates the selected zones.

4.1.2. The initial data:

It is recalled that the model represents a time-recurrent process; starting with the data of the last period, the model performs the following operations:

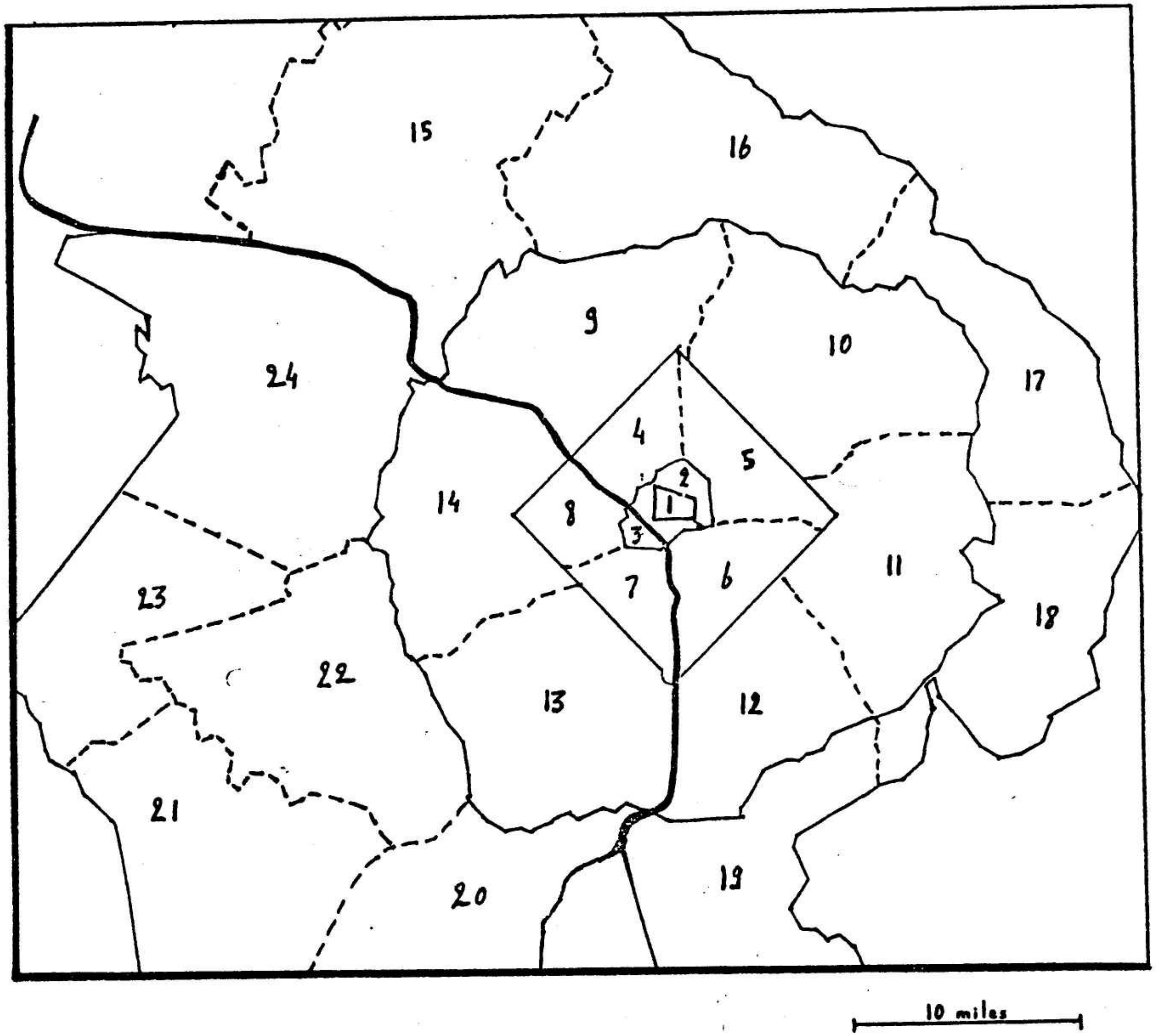
-it generates the moving households and stock of vacated units

-it describes the supply activity on the basis of the last period prices

-it predicts the demands and matches it with the supply in a market clearing iterative process, until a convergence test is met (or a maximum number of iterations performed).

-it finally assigns the households to vacant units, and operates a book-keeping function: all the classifications of the residents are updated, vacant units and "non-located people" figures are ready for the next period iteration, together with prices.

This clearly indicates that the amount of information desired at the time-horizon requires an equivalent amount of initial data; this has implied a substantial data- col-



(figure 4.])

lection effort.

The simulation requires another kind of information to perform the time-recurrence: this consists of demographic and employment-modification data per period.

The implemented simulation has a four year time horizon: this was thought to be adequate to observe some of the relevant transportation impacts.

4.2. Results of the implementation

Before describing the results, the form of the outputs is briefly discussed.

4.2.1. Form of the outputs:

They are of two types:

- The model keeps track of the distribution of the households among a large number of categories: housing type, workplace, residence zone, level of auto ownerhip, mode-to-work, and a number of cross classifications of the preceding. The storage requirement is multiplied by three, since these classifications apply at three stages in each period: the initial number of families, the number of movers and the number of relocating people in the category. This amounts to a huge qunatity of information: it would not be appropriate to print all those results.

- A selection of the forecasts is effectively printed; it has been chosen to allow a simple interpretation. They include some information provided in each period, and time-horizon predictions.

* In each period, a table reflects the results of the consecutive submodels: these are a distribution of housing demanders by household type and workplace, a count of the constructions and transformations per zone and types, a matrix of available units per zone and type, the results of the convergence-test achieved in the market-clearing iterations, the final distribution of the demand per zone and type, the classification of "non-located people" per workplace and household-type: and the important evolution of rentals and price of land per zone. Some of these tables are enclosed to illustrate the results interpreted in the next section.

* At the end of the simulation, the important variables which describe the urban areas' evolution are represented by a series of zonal distributions. These include:

. population data, like the density in the total population, the proportion of black households

. rentals and the price of land data

. transportation data: modal split, auto-ownership, and the criterion introduced of "generalized speed". This information is provided by both tables of values and an illustrative format of output: a mapping of the city where a lighter or darker shading represents higher or lower values of the measure.

Some mapping for the four year time horizon are included to allow the reader to make his own opinion about the informational content of that format.(see Figs. 4.2-4.6)

4.2.2. Interpretation of the results:

It must be clarified that the purpose of this section is not at all to draw city-planning oriented conclusions about the city of Washington. An actual utilization of the model for its final forecasting purpose would require anyway the previous definition of an explicit policy-testing program.

Alternatively, the objective of this section is to describe the time-recurrent implementation and to analyze the behavior of the various submodels.

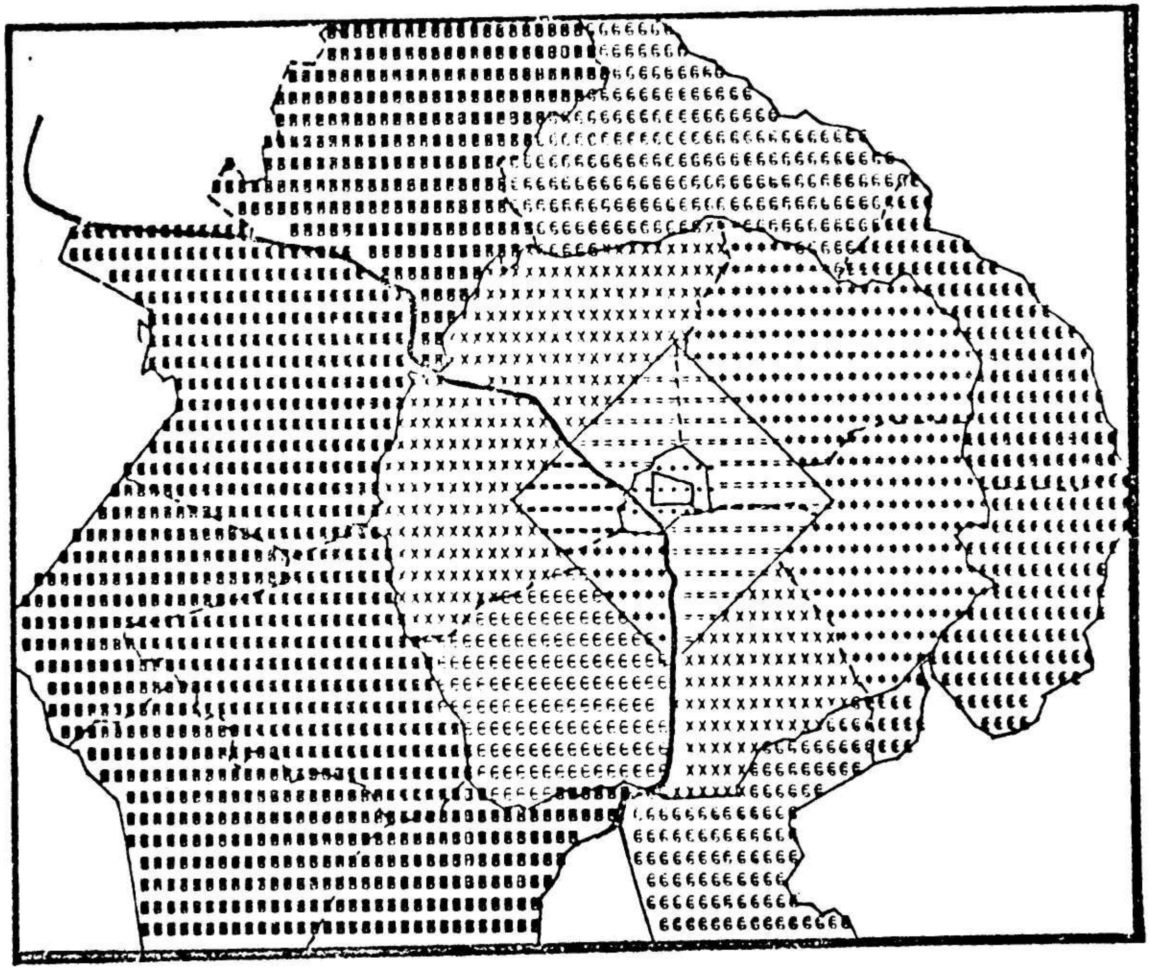
In the first period, the mover submodel generates the number of movers which amounts to a proportion of 23% of the total number of households. This high value illustrates the U.S. high rate of mobility; since this is the source of the modeled urban evolution, it can be expected that within five years, the model will be significantly modifying the shape of the city; however, this is in accordance with the reality, and still avoids the criticism of an instantaneous relocation of all inhabitants. (see Tab.4.7)

The supply activity takes two forms;

- the construction of new units from land;owner oc-

```
'GENERALIZED SPEED' :

12.34
12.11
14.20
14.89
15.42
15.02
20.55
18.51
22.41
21.62
19.75
23.18
24.94
24.06
35.56
28.74
20.29
20.87
28.47
36.25
34.26
32.16
36.04
33.20
```

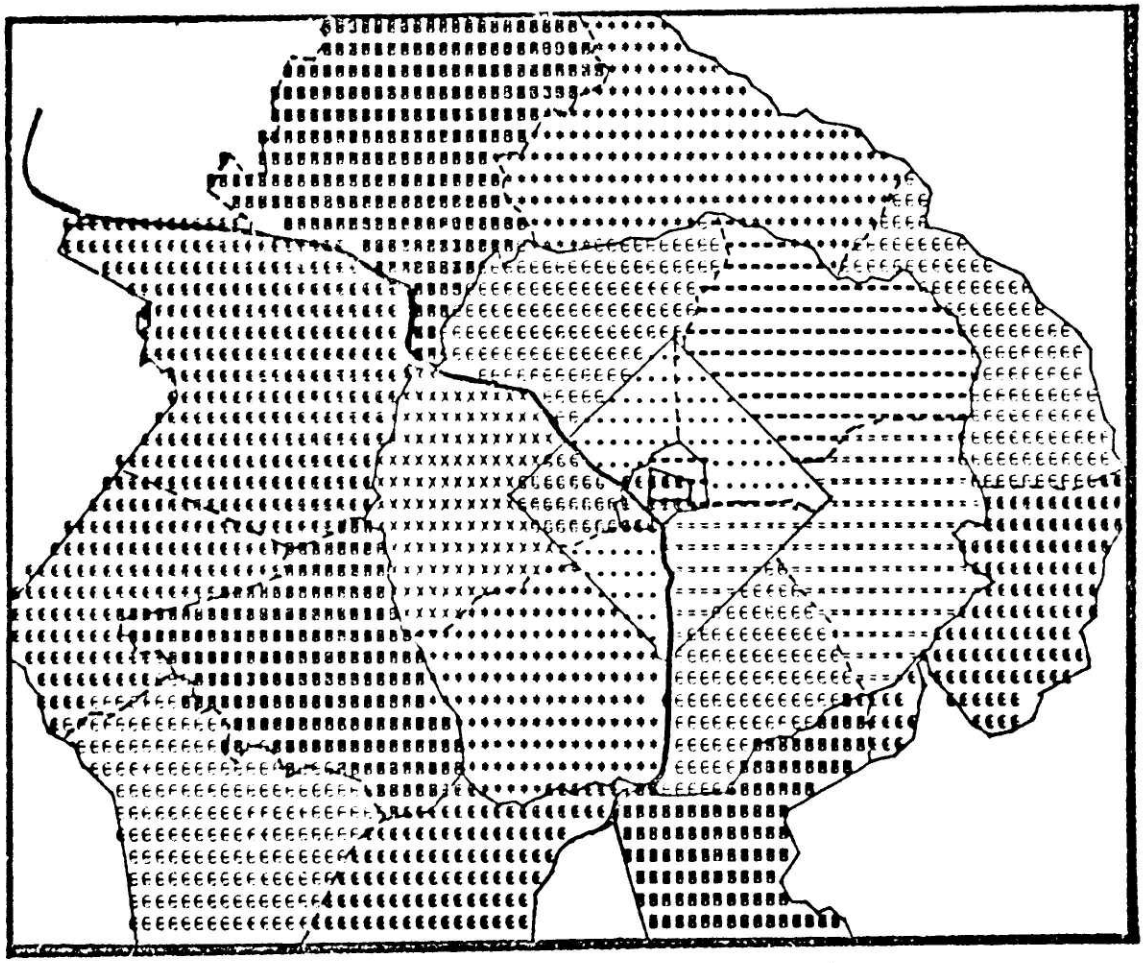


```
THE 'GENERALIZED SPEED' IS: 21.8

THE TOTAL V.M.T. PER DAY IS  18412496
```

(Figure 4.2)

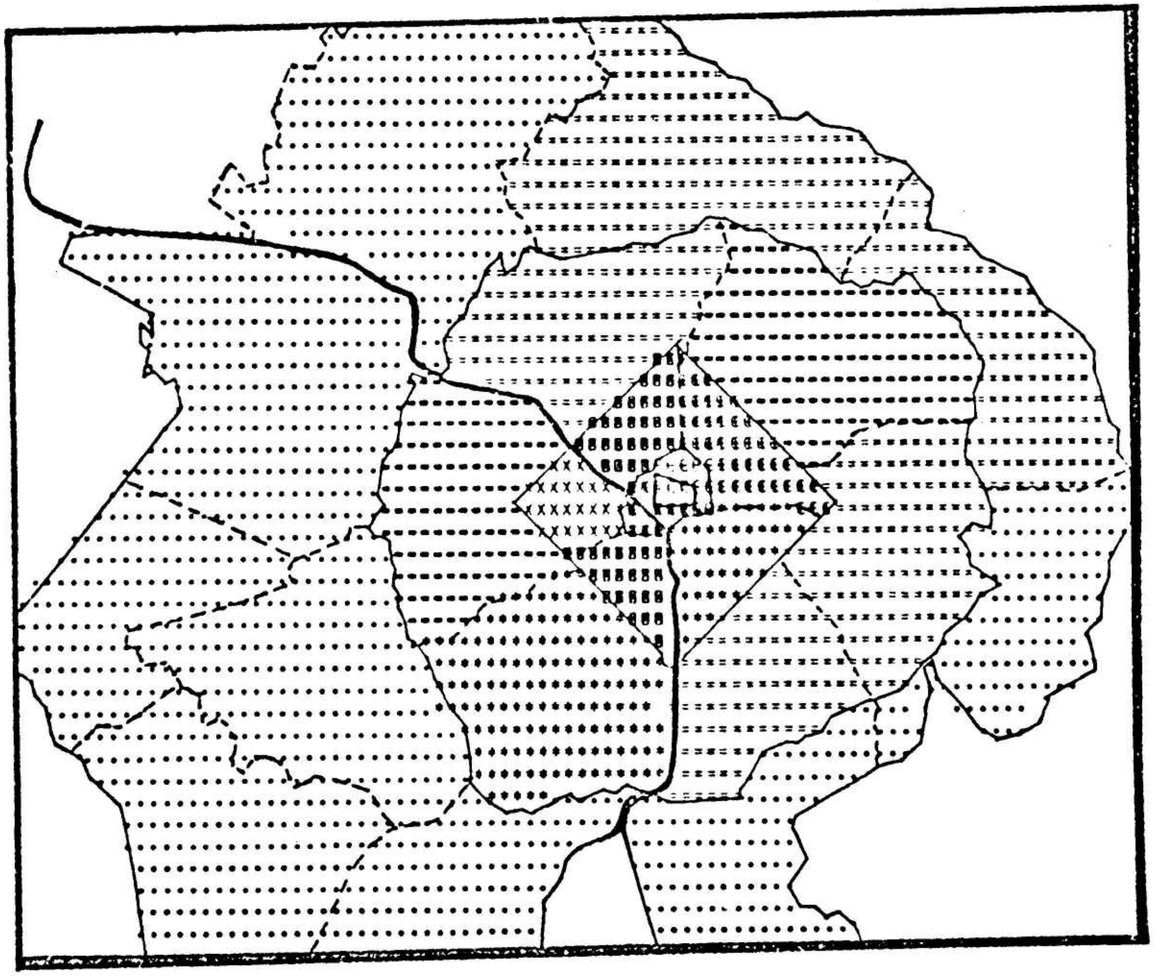
```
PRICE OF RENT :
6914.96
3045.90
6246.51
1C60.57
1166.C9
2030.04
1C32.83
5343.21
5132.17
2780.C0
2261.33
4808.56
3108.81
4361.34
7636.12
3755.C7
5096.75
6213.05
7258.14
6381.68
5054.54
7920.C0
6224.50
6429.19
```



```
THE AVERAGE PRICE OF RENT IS :+3830
```

(Figure 4.3)

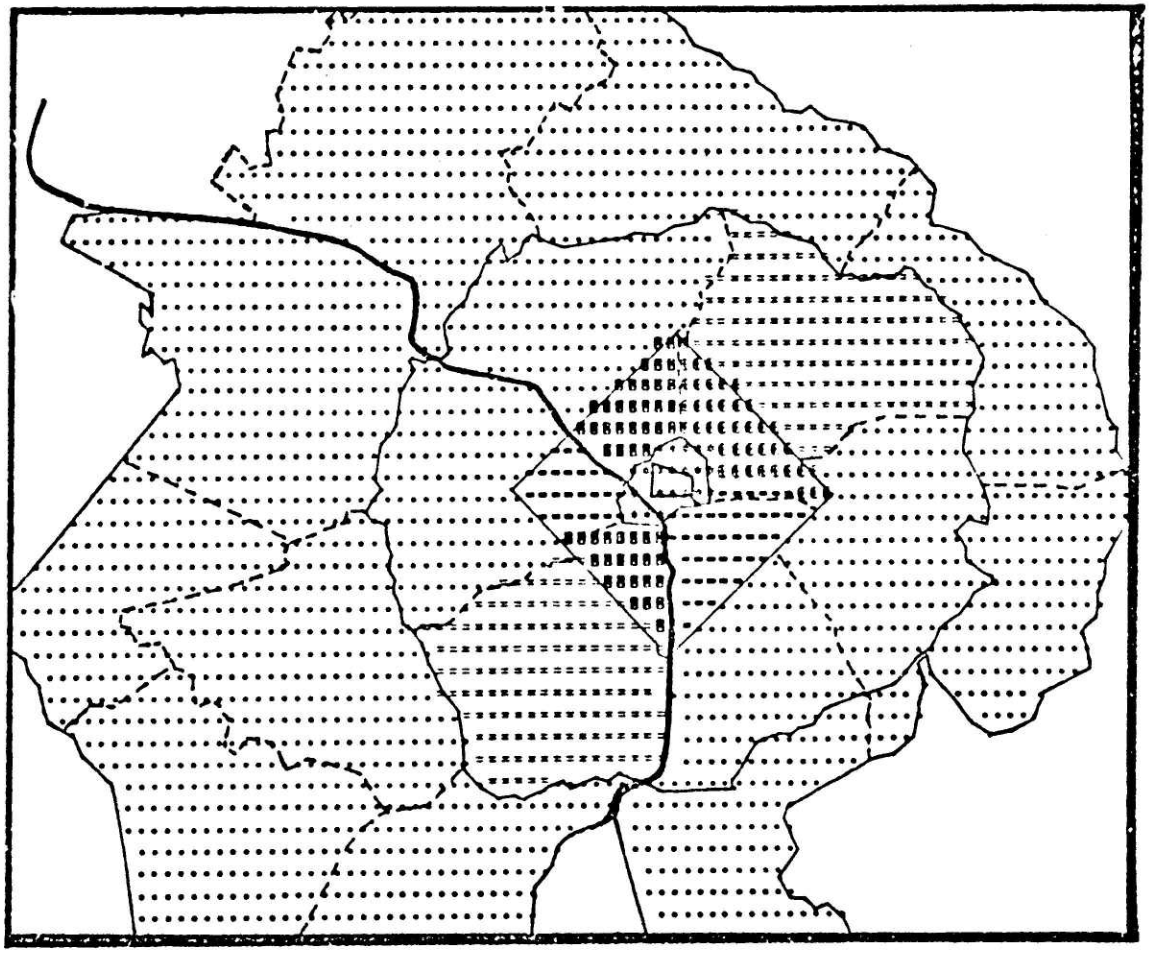
PRICE OF LAND :



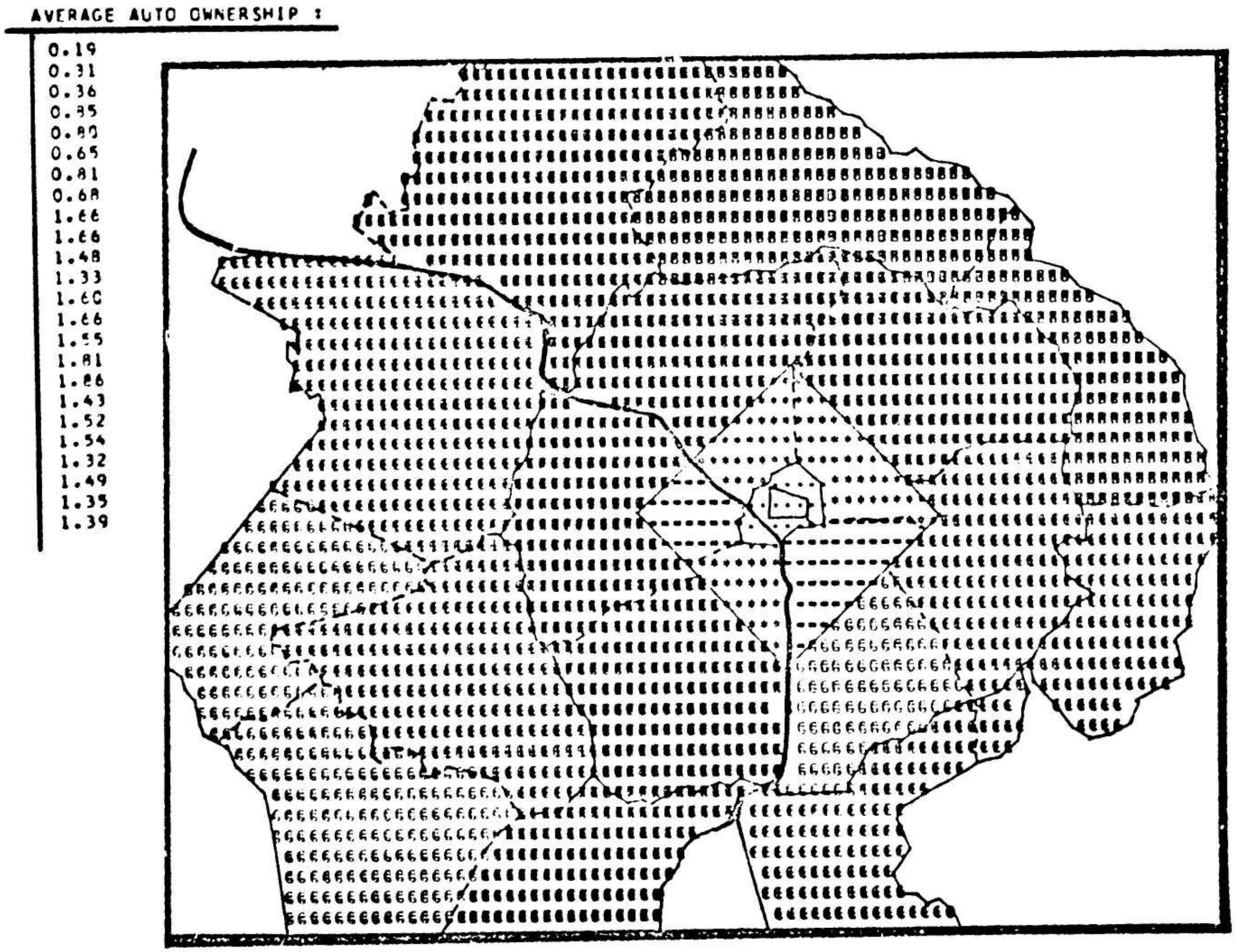
(figure 4.4)

RESIDENTIAL DENSITY :

21.38
21.46
22.50
54.78
46.05
16.91
58.81
15.37
6.21
8.46
7.48
6.67
10.87
7.16
2.53
3.54
4.46
2.53
2.53
2.97
2.53
2.74
2.52
2.76



(figure 4.5)



THE AVERAGE AUTO OWNERSHIP IS : 1.3

(figure 4.6)

cupied single family dwellings are built in the suburban zones. The construction of garden-style housing and high-rise structures takes place both inside the District of Columbia (the small lot-size per unit compensates for the high price of land and makes the operation profitable), and also in some suburbs.

- the transformation of single family dwellings for the purpose of changing their tenure: a proportion of the tenant-occupied units become owner-occupied. The magnitude of that transformation is surprising: it is thought that this prediction is unreliable and caused by inaccurate initial data (the selling price of the owner-occupied houses being overestimated).

After the supply, the total number of available units is slightly smaller than the number of housing-demanders: there will be an inherent undersupply during the period. This latter is larger if one considers the buffer stock necessary to locate all the housing-demanders in the context of a non-perfect market (this is the meaning of the aforementioned "dynamic vacancy rate"). In that perspective, the "percentage of supply" rate indicates that only 27.3% of the total needs before construction are met by the supply sector in period one.

Then, the demand allocation and market clearing process are iterated. The "rates of disequilibrium" of the

```
NEW CONSTRUCTIONS IN PERIOD   1

     ZONE      TYPE 1     TYPE 2      TYPE 3      TYPE 4      LAND

       1        167        108          0           0         18
       2        725       -724          0           0          0
       3        174       -173          0         300          2
       4        310       -309          0           0          0
       5        718       -717          0           0          0
       6       2662      -2661        319           0          6
       7        105       -104          0           0          0
       8       2296       -929          0           0         95
       9       2662      -2661          0           0          0
      10       2662      -2661          0           0          0
      11       1224      -1223          0           0          0
      12       1140      -1119          0           0          0
      13       1787      -1855        347           0          0
      14       2065      -2064        347        1113         35
      15        632          0          8           0        252
      16        436       -435         73           0          4
      17         53        -52          0           0          0
      18        235         -1          2           0         93
      19        309         -2          3           3        122
      20        688       -101        166          20        243
      21        270         -2          0           5        106
      22        961        -88          0           0        349
      23        658         -3          6           0        261
      24         20        -19         69           0          4

 TOTAL NEW CONSTRUCTIONS =     7905

 PERCENTAGE OF SUPPLY =  0.273
```

```
RENTALS AND PRICE OF LAND IN PERIOD   1

     ZONE      TYPE 1     TYPE 2      TYPE 3      TYPE 4      LAND

       1        2506       3527        2467        1693      12744
       2        2228       3263        2607        1171      15321
       3        2192       2859        2678        1098      17641
       4        2094       2485        2570         905      38329
       5        1866       2167        2113         855      29490
       6        2329       2464        1555         903      11591
       7        2471       2202        2150         803      35579
       8        2565       2278        1547        1085      12963
       9        3000       2887        2425         915       7029
      10        2683       2868        2463         939       8140
      11        2092       2278        1809         857       6573
      12        2461       2276        1543        1033       5746
      13        2406       1837        3022         997      10858
      14        2471       2278        2750        1137       7858
      15        3818       2663        1250        1260       1655
      16        3824       2653        2605         868       4797
      17        4485       2485        2254        1163       6082
      18        3315       2245        1220        1260       2402
      19        3986       3295        1196        1025       2215
      20        3064       2070        2911        1012       2528
      21        2431       2428        1569        1014       1950
      22        3890         07        2455        1320       2490
      23        2641       2064        1413        1260       2137
      24        4079       2070        2455         885       3397
```

```
TOTAL DEMANDERS = 192246

TOTAL NEEDS = 211470

TOTAL AVAILABILITIES = 190430

NON LOCATED PEOPLE = 46930
```

(TABLE 4.7)

```
MARKET-CLEARING ITERATIONS IN PERIOD   1                  ITERATION   2
                                                              DESEQUILIBRIUM ON HOUSING-MARKET   1:     0.164
 THE AVERAGE PERCENTAGE OF UNDERSUPPLY IS :0.01               DESEQUILIBRIUM ON HOUSING-MARKET   2:     0.174
                                                              DESEQUILIBRIUM ON HOUSING-MARKET   3:     0.127
                                                              DESEQUILIBRIUM ON HOUSING-MARKET   4:     0.199


 ITERATION   1                                            ITERATION   3
    DESEQUILIBRIUM ON HOUSING-MARKET   1:     0.179           DESEQUILIBRIUM ON HOUSING-MARKET   1:     0.145
    DESEQUILIBRIUM ON HOUSING-MARKET   2:     0.189           DESEQUILIBRIUM ON HOUSING-MARKET   2:     0.153
    DESEQUILIBRIUM ON HOUSING-MARKET   3:     0.150           DESEQUILIBRIUM ON HOUSING-MARKET   3:     0.104
    DESEQUILIBRIUM ON HOUSING-MARKET   4:     0.227           DESEQUILIBRIUM ON HOUSING-MARKET   4:     0.169
```

four housing-markets are as high as 20%, which indicates that demands and vacancies do not balance very well. This is particularly true for the high rise structure market.

Again, this surprising disequilibrium of the market is due to inaccurate initial data which biases the distribution of forecast demands. The comparison of the two tables "Total Availabilities" and "Demand in period 1" allows to find out for which housing-types and zones there is the least balance between supply and demand.

As a consequence, there is a large number of "non-located people": it has been explained before that if people who cannot locate in the zone and type of housing they wished, become "non-located", they will try to relocate in the next period. (The rationale of that formulation is largely discussed in chapter five). Because of the disequilibrium, the "convergence-test" developed in section 3.4.2. is not met: the demand and market-clearing iterations are performed up to the maximum number of three. In that process, significant price adjustments take place. The demand reacts and tends to match better and better the availabilities: this is clearly documented by the fact that the "rates of disequilibrium" are decreasing iteration after iteration.

Since the supply generation is not iterated within a period, only half of the distance to be covered to reach the equilibrium is actually covered. Therefore it is not surprising that the rates after the third iteration are still high, in the order of 10% to 16%.

The undersupply in period 1 results in a general increase of the rentals for period 2: more supply activities are profitable. In particular, there are much more new land constructions; the change of tenure for single-family dwellings has practically ceased: the operations of the last period have led to an oversupply of owner-occupied units; their selling-price in several zones has decreased and most of the transformation from tenant-occupied to owner-occupied are no more profitable in period 2. In that respect, it is interesting to note here what happens in period 3: some transformations back from owner-occupied to tenant-occupied units are observed. It means that the adjustment of prices between the two types has gone a little too far in some zones. This documents the necessity to iterate the market-clearing process, as discussed in section 3.4.1.: the lower the number of iterations there are, the higher the probability of overshooting the equilibrium.

The results of the Demand/Market Clearing iterations in period two are interesting: they prove that the whole model is reflecting the trend of the market to be in equilibrium. Two criteria show that the necessary adjustments have taken place between period 1 and period 2:

- the market clearing rates of disequilibrium are significantly lower, in the range of 10% to 16% for the first iteration, 6% to 12% for the third. This still exceeds the 3% convergence-test, but the model is clearly self correcting the initial misspecification of data.

- the number of non-located people has decreased by 25%

However, such adjustments have necessitated some modifications of prices which deserve attention: the single-family dwellings in zones 15, 16, and 22 for example, or the garden-style apartments in zones 8, 13 and 14 have been in a continuous state of undersupply: the prices have increased as much as they were free to do, which amounts to a tripling of prices within two years! This is unrealistic, and due to a misspecification of the demand, as explained in the next chapter. (see table 4.8)

In period three, the supply is still more active: the bulk of new construction takes place for the garden-style apartments, in the District of Columbia and in the first outside ring. The total number of availabilities now exceeds the number of movers, but not the "total needs" which include a buffer stock. The rates of disequilibrium are again significantly decreased: 4% to 12% in the first iteration, 2% to 10% in the third. Two of the four markets already meet the "test of convergence". The number of "non-located people" is divided by two. As in the previous periods, there is no subsidy, which means that all the households have been able to find at least one housing such that their "remaining income" after taxes, housing and transport-

```
NEW CONSTRUCTIONS IN PERIOD  2

  ZONE     TYPE 1    TYPE 2     TYPE 3     TYPE 4      LAND

    1        140        39        268        132        16
    2        -93        39        250          0         1
    3        -38        39        124          0         2
    4        -71        39        116          0         0
    5        -39        39        484          0         9
    6       3222        39        362          0       234
    7          0         0        267          0         5
    8       2362        39        139          0       183
    9       3717        39        226          0       759
   10        -38        39        484          0        19
   11        -38        39         94          0         3
   12          0         0        127          0         5
   13       1752     -1751        484          0        19
   14          0         0        484          0        19
   15        712        39          0          0       299
   16       2045      -397         85          0       663
   17        523       -37         15          0       193
   18        262         0          0          0       104
   19        348        25          0          0       148
   20        763       -75        190          0       295
   21        297        21          0          0       126
   22       1082        39        131          0       454
   23        741         0          0          0       296
   24        740        -8         80          0       296

 TOTAL NEW CONSTRUCTIONS =    21366

 PERCENTAGE OF SUPPLY =  0.454
```

```
 TOTAL DEMANDERS = 241957

 TOTAL NEEDS = 266152

 TOTAL AVILABILITIES = 240464

 NON LOCATED PEOPLE = 34324
```

```
RENTALS AND PRICE OF LAND IN PERIOD   2

  ZONE     TYPE 1     TYPE 2     TYPE 3     TYPE 4      LAND

    1       3013       5841       4113       1499      14790
    2       1599       3048       2912       1479      16240
    3       3764       4882       4510       1850      19676
    4       1846       2913       3858        738      36509
    5       1530       2357       2369        690      20862
    6       1840       2650       2368        751      11853
    7       1916       2829       3116        646      34178
    8       3304       3796       3257       1085      13157
    9       3718       4606       4009       1195       7042
   10       2222       3542       3707        750       8185
   11       1960       3159       2358        702       6441
   12       3878       3801       3261       1547       5673
   13       2145       3141       4052        836      11343
   14       3476       3704       4555        964       8338
   15       5858       4542       2151       1260       1737
   16       3845       4650       4084        815       4676
   17       5704       4250       3657       1845       6784
   18       4945       3829       1300       1260       2367
   19       6055       5048       1617        756       2771
   20       4177       3574       4286        780       2778
   21       3165       2567       1560        784       1810
   22       5892       5237       4049       1320       2666
   23       4033       3559       2148       1260       2067
   24       4878       3572       3554        712       3438
```

(TABLE 4.8)

```
MARKET-CLEARING ITERATIONS IN PERIOD   2

 THE AVERAGE PERCENTAGE OF UNDERSUPPLY IS :0.01



 ITERATION   1
    DESEQUILIBRIUM ON HOUSING-MARKET   1:     0.167
    DESEQUILIBRIUM ON HOUSING-MARKET   2:     0.103
    DESEQUILIBRIUM ON HOUSING-MARKET   3:     0.095
    DESEQUILIBRIUM ON HOUSING-MARKET   4:     0.165

 ITERATION   2
    DESEQUILIBRIUM ON HOUSING-MARKET   1:     0.179
    DESEQUILIBRIUM ON HOUSING-MARKET   2:     0.076
    DESEQUILIBRIUM ON HOUSING-MARKET   3:     0.071
    DESEQUILIBRIUM ON HOUSING-MARKET   4:     0.137

 ITERATION   3
    DESEQUILIBRIUM ON HOUSING-MARKET   1:     0.092
    DESEQUILIBRIUM ON HOUSING-MARKET   2:     0.055
    DESEQUILIBRIUM ON HOUSING-MARKET   3:     0.058
    DESEQUILIBRIUM ON HOUSING-MARKET   4:     0.119
```

ation costs, is positive. (See table 4.9)

The situation in period 4 is not very different from period 3. The rates of disequilibrium are smaller and, except for the high-rise apartment market, very close to meeting the convergence-test. However, it should be noted that, because the simulation is incremental instead of creating an instant-city, the model takes some time to absorb any disequilibrium. This is a quite realistic characteristic of the market. (see table 4.10)

Since period 4 is the last in the implemented simulation, the corresponding maps and tables of data are printed (see figures 4.2 to 4.6). The map of "generalized speed" is particularly explicit: it shows that the transportation level of service, by car, transit or both, is better in the west belt of the urban area than in the eastern one. The North-East corridor has particularly low figures of "generalized speed". It is interesting to note that this map has some resemblance to the map of rentals: high rents seem to have some correlation with good transportation levels of service.

The map of rentals and the map of prices of land are very different-looking: though the rentals are higher in the suburban areas, the price of land is still higher in the central zones, because of a more dense utilization of the land. It is the map of residential density which shows resemblance to the map of prices of land.

```
NEW CONSTRUCTIONS IN PERIOD  3

   ZONE      TYPE 1     TYPE 2     TYPE 3     TYPE 4      LAND

     1         -60        206        300         53        15
     2        -272        209        276         53         0
     3         197        208        141         53        29
     4        -183        147        130          0         0
     5        -209        208       1620          0        32
     6        -209        208        397          0         7
     7        -208        208        297          0         5
     8        2835        208        160         53       215
     9        3162        208        255          0       683
    10        -207        208       1653          0        66
    11        -207        208        105          0         4
    12        1294        208        146         53       303
    13        -207        208       1784          0        71
    14        2641        208        771          0       599
    15         787        208         10          0       397
    16        2255        208         97          0       989
    17         579        128         18         53       285
    18         292         10          0          0       119
    19         386         28          3          0       165
    20         852        163        213          0       417
    21         335         23          0          0       142
    22        1195        155        149          0       546
    23         821         32          7          0       340
    24         817         66         91          0       357

TOTAL NEW CONSTRUCTIONS =   29480

PERCENTAGE OF SUPPLY =  0.608
```

```
RENTALS AND PRICE OF LAND IN PERIOD   3

   ZONE      TYPE 1     TYPE 2     TYPE 3     TYPE 4      LAND

     1        5203       6779       6608       2506     17976
     2        2766       3019       3057       1411     16897
     3        3872       7369       7478       2962     23333
     4        2402       2650       5020        617     35140
     5        1498       2257       2214        555     27717
     6        1570       2603       2257        631     11513
     7        1941       2688       3378        579     12630
     8        3590       5573       5517       2025     13692
     9        4418       6666       6479       1387      7049
    10        2270       3842       4555        608      8036
    11        2428       3043       2431        590      6126
    12        4020       6032       5366       1973      6138
    13        2692       3429       5415        797     11169
    14        3429       5511       7280       1039      8374
    15        7071       7255       3522       1260      1335
    16        4132       6043       5982        559      4460
    17        6277       6755       5581       2480      6489
    18        6076       3501       2078       1260      2331
    19        7164       6491       2158        617      2313
    20        5443       5596       7217        607      3121
    21        4533       3229       1560        604      1729
    22        7074       8069       6761       1320      2910
    23        5708       5181       3045       1260      2044
    24        5760       5635       6430        984      3469
```

```
TOTAL DEMANDERS =236010

TOTAL NEEDS = 259610

TOTAL AVAILABILITIES = 240578

NON LOCATED PEOPLE = 18056
```

(TABLE 4.9)

```
MARKET-CLEARING ITERATIONS IN PERIOD   3

THE AVERAGE PERCENTAGE OF UNDERSUPPLY IS :-.02



ITERATION   1
   DESEQUILIBRIUM ON HOUSING-MARKET   1:      0.073
   DESEQUILIBRIUM ON HOUSING-MARKET   2:      0.036
   DESEQUILIBRIUM ON HOUSING-MARKET   3:      0.049
   DESEQUILIBRIUM ON HOUSING-MARKET   4:      0.123

ITERATION   2
   DESEQUILIBRIUM ON HOUSING-MARKET   1:      0.054
   DESEQUILIBRIUM ON HOUSING-MARKET   2:      0.028
   DESEQUILIBRIUM ON HOUSING-MARKET   3:      0.040
   DESEQUILIBRIUM ON HOUSING-MARKET   4:      0.116

ITERATION   3
   DESEQUILIBRIUM ON HOUSING-MARKET   1:      0.046
   DESEQUILIBRIUM ON HOUSING-MARKET   2:      0.022
   DESEQUILIBRIUM ON HOUSING-MARKET   3:      0.032
   DESEQUILIBRIUM ON HOUSING-MARKET   4:      0.108
```

```
NEW CONSTRUCTIONS IN PERIOD   4

  ZONE      TYPE 1     TYPE 2      TYPE 3      TYPE 4       LAND

    1         -48        169         333         89          13
    2        -522        478         306         89           0
    3         219        409         157         89          46
    4        -162        163         145          0           2
    5        -872        873        1772          0          35
    6       -1087       1088         440          0           8
    7        -419        419         330          0           6
    8        2779        615         181         89         241
    9        2089        940         285         59         615
   10       -1087       1088        1825          0          72
   11        -816        817         117          0           4
   12        1419        855         166         89         460
   13        -996        997        1969          0          78
   14        2911        769         854          0         769
   15         868        245          13          0         444
   16        2487        291         108          0        1116
   17         638        143          21         89         316
   18         322         12           2          0         132
   19         426         31           4          0         182
   20         910        182         236          0         457
   21         372         26           0          0         158
   22        1317        172         167          0         604
   23         906         36           8          0         376
   24         902         75         102          0         395

 TOTAL NEW CONSTRUCTIONS =    33588

 PERCENTAGE OF SUPPLY =  0.764
```

```
 TOTAL DEMANDERS = 227044

 TOTAL NEEDS = 249748

 TOTAL AVAILABILITIES = 239398

 NON LOCATED PEOPLE = 15073
```

```
RENTALS AND PRICE OF LAND IN PERIOD    4

  ZONE      TYPE 1      TYPE 2      TYPE 3      TYPE 4       LAND

    1        8973        7445        7108        2343       21303
    2        4750        3056        3201        1107       17725
    3        4777        7013       11118        3933       28204
    4        2685        2852        4550         561       33455
    5        2261        2143        2210         448       26518
    6        2099        2541        2411         541       11255
    7        3269        2815        3442         457       30556
    8        5011        6847        8935        2130       14674
    9        5616        6677        8746        1636        7344
   10        3043        3967        4817         509        8595
   11        3335        2841        2550         555        5853
   12        4760        6473        8268        2176        6746
   13        3130        3380        5573         977       11459
   14        3719        5587        8512        1435        8855
   15        7291        8573        5563        1260        1555
   16        4619        6043        6349        1268        4342
   17        6759        8738        8509        2019        6799
   18        6385        3596        1548        1260        2273
   19        7375        7032        2158         479        2374
   20        5894        7023        8347         492        3515
   21        5189 [illegible]6       1560         446        1656
   22        7468        9117       10236        1320        3269
   23        6216        6181        6448        1260        2019
   24        6218        7156        8519        1573        3603
```

(TABLE 4.10)

```
MARKET-CLEARING ITERATIONS IN PERIOD   4

 THE AVERAGE PERCENTAGE OF UNDERSUPPLY IS :-.05



 ITERATION    1
    DESEQUILIBRIUM ON HOUSING-MARKET   1:      0.092
    DESEQUILIBRIUM ON HOUSING-MARKET   2:      0.021
    DESEQUILIBRIUM ON HOUSING-MARKET   3:      0.022
    DESEQUILIBRIUM ON HOUSING-MARKET   4:      0.108

 ITERATION    2
    DESEQUILIBRIUM ON HOUSING-MARKET   1:      0.072
    DESEQUILIBRIUM ON HOUSING-MARKET   2:      0.012
    DESEQUILIBRIUM ON HOUSING-MARKET   3:      0.019
    DESEQUILIBRIUM ON HOUSING-MARKET   4:      0.103

 ITERATION    3
    DESEQUILIBRIUM ON HOUSING-MARKET   1:      0.057
    DESEQUILIBRIUM ON HOUSING-MARKET   2:      0.015
    DESEQUILIBRIUM ON HOUSING-MARKET   3:      0.018
    DESEQUILIBRIUM ON HOUSING-MARKET   4:      0.090
```

Finally, the map of auto-ownership shows the strong disincentive which exists for owning cars inside the District of Columbia.

The implementation of the case-study has pointed out several conclusions:

. two kinds of improvements are to be made. The collection of more accurate initial data is rather simple as opposed to the respecification of the demand model, which is a considerable task.

. The imperfection of the quantitative results does not encumber the validity of the model. It has been proven that this model very satisfactorily represents the behavior of the suppliers, the preferences of the consumers, and the trends of the market. The sample of results that has been discussed shows the richness of the interpretation which is made feasible by the available forecasts. Finally, the computational requirement for the implementation is reasonable: 11.6 minutes of execution on IBM 260-270 , for a cost of $77. Furthermore, for the purpose of larger implementation, the author is convinced that the computational efficiency of the simulation can be improved.

Chapter 5 - Conclusions and Recommendations

The first conclusion is clear to the author of this paper: constructing a new type of urban simulation is a large task. It was beyond the scope of a master's thesis to design a formulation which would have all the qualities required for an operational policy-evaluation tool. The author is perfectly aware that there remains a lot of work to do in order to improve the reliability of the predictions. However, this study has proven that the simulation is technically working, and that all the market reactions are represented qualitatively, if not also quantitatively. Furthermore, it is thought that the approach is very promising for its final purpose, that is, to be a forecast and evaluation tool for urban and transportation planners. Its strong market-orientation and the comprehensiveness of its description should insure both its reliability and usefulness.

At the present point of its construction, it is thought that this chapter of conclusions should be restricted to two kinds of issues: what conclusions can be drawn and what recommendations can be made concerning the implementation of the method? In the long run, what is its potential usefulness?

5.1. Conclusions concerning the model-implementation

For the purpose of future improvements of the model, it is useful to highlight the main difficulties that have been met, and to make some recommendations about possible improvements.

5.1.1. The "internal coherence-test":

This issue is at the same time a cause of difficulty and one of the biggest assets of the model: it deals with its market-orientation, as is explained below:

The model relies entirely on the representation of a market mechanism, as opposed to what can be called a "forced allocation". This latter approach characterizes most of the operating models, including the NBER in some respects:

- 1) A fair proportion of the existing models are purely demand-oriented: they forecast the preferences of the consumers, but do not match them with some supply data. Therefore, they escape the issue of assigning the households to a vacant stock which a priori does not meet perfectly the demands. The accuracy of their predictions is still to be proven.

- 2) The NBER simulation does include a supply sector, but still does not have what is referenced below as an "internal coherence-test": in practical terms, the formulation does not allow the results to show large inconsistencies, even if these inconsistencies were to disclose some sort of misspecification. Three remarks support this statement:

-a). There is indeed a demand submodel which indicates the choices of the consumers among the different housing types. However, the total demand for housing of a given type cannot be smaller than the corresponding supply: this supply is generated subject to the constraint of not exceeding the expected demand. This is one of the constraints of the supply-submodel, as it was mentionned in chapter two.

This "safety" seems to be somehow at the expense of realism. The supply is clearly a longer term activity than the demand, and it is not likely that the suppliers could wait for the current demand to orient their construction and transformation activities. At best, they obtain an aggregate forecast of the needs and an approximate prediction of the shares of the various housing-types. Their main indicator remains the market prices: these both determine the profitability of their undertaking and reflect the preferences of the consumers. For the concern of behaviorality the prices should be the essential determinants of the content and magnitude of the supply activity. In the NBER simulation, the worst that can happen is an undersupply for certain submarkets: this results in an increase of price which lasts until the profitability is sufficient to generate a supply. But there is no possibility to observe a shift of the demand that would result in a large oversupply in one market and undersupply in another. However, it is thought that the aforementioned constraint has been modified in the current NBER version, in the sense recommended above.

-b). Once the households are assigned to housing-types, there is no matching of their locational preferences with the stock of availàbilities per zone: in the NBER model, these preferences are just not disclosed! As explained in section 2.2., the market-clearing operates what can be called a "forced assignment" to existing availabilities. That is quite different fromgenerating the demands ex-ante, then matching them with the number of vacant units. This is clearly documented by the fact that the linear-programming optimization is performed subject to the constraints that the allocation to each zone does not exceed the corresponding supply. This approach fails to faithfully reflect the reality, for the following reason:

the optimization does capture the pressures that the market is exerting. The shadow prices generate a modification of the rentals accordingly. However, this adjustment is well represented only if the market-pressures are so, which in turn assumes that the optimization algorithm captures the true behaviors of people. This assumption is, at least, not obvious because the market is not perfect,and it is likely that the optimization will assign some households to some locations which differ from their optimal individual choice. In the new model, such households which cannot satisfy this choice become "non located" (this amounts to reflecting the likelihood that they will try to relocate in the coming periods). In the NBER approach, they are assigned to a place; the only way people can be "non located" is because of the overall undersupply, but not because of the zonal undersupply in the location of their choice. What is somehow misleading is not only that the behavior of such unsatisfied households is foregone: there is also no criterion for warning the user of that issue. On the contrary, the new model provides an estimate of the convergence: if the total number of "non-located people" increases significantly over the total undersupply, the convergence is not achieved. The market-clearing "rates of disequilibrium" provide a cross-checking.

-c). In the NBER approach, since the balance of supply and demand is insured by the preceding mechanism, it is feasible to manipulate the prices in order to keep them in realistic ranges. It is the case of the rentals which are exponentially smoothed with the previous values, in order to avoid abrupt changes.

3) The new model does not have those safeguards: since the prices are the unique factors of equilibrium, they must be free to move as much as it is necessary to achieve the equilibrium. If their variations were exogenously constrained, the model would not make sense anymore: therefore, the convergence-ability of the model is preferred to the realism of prices. However, the two are achieved at the same time if the specification is correct. If it is not the case, the prices will presumably depart from reasonable ranges to attempt to realize the equili-

brium: this is therefore a way to check the validity of the model, referenced as an "internal coherence test". In the implementation, described in chapter 4, this criterion has indicated that there is some misspecification, presumably in the demand sector. Section 5.1.2. addresses the possible improvements to be made. Once the coherence test gives satisfactory results, the model gains a fair amount of reliability: there is then a fair probability that the market mechanism is well reflected. Indeed, it does not mean that no misspecification remains, but a first type of serious misprediction is avoided.

5.1.2. Recommendations

In this section, conclusions are drawn by analyzing what is unsatisfactory in the result of chapter 4, in the light of the preceding considerations.

The first observation to be made is a low convergence in the initial period: this can be checked with the too high rates of disequilibrium of the market-clearing, or the excessive number of "non-located people".

It is thought that this is due to the initial data: if some of these depart from their actual values, it is expected that the same will happen with the demand or the supply, and no convergence can be achieved in the initial periods. After a few years, the equilibrium process will have brought back the system to a new equilibrium: this can be biased by the initial misspecification.

Some data have been obtained through an averaging over large zones of wide-spread values: those figures are inherently unreliable estimates at the disaggregate level. In the model developed, this is particularly the case for the initial values of the owner-occupied single family dwellings.

The second observation involves the coherence test-issue: the aforementioned lack of convergence is progressively suppressed because of the equalizing trend of the model. However, it was observed that some prices have to be raised unreasonably in order for the equilibrium to be achieved. As it was largely developed in the preceding section, this discloses a misspecification: it is thought that this deals with the <u>insufficient sensitivity to prices.</u> The unique variable taking prices into account is the logarithm of remaining income: it has been seen in chapter three that this formulation presumably is inadequate. One alternative was proposed in that chapter to improve the sensitivity to prices.

Another possibility would be to introduce a new variable reflecting how a rental compares with what the household is willing to pay in average for housing expenses. It can take the form of a ratio such as the actual rental over that average housing expense, specified per household type.

However, that new formulation departs from the underlying idea that all the "primary expenses" (taxes, housing, transportation) are substitutable, and that only the remaining income is a relevant factor.

Finally, it is recommended that the formation process for the price of land be revised: it has been seen that the locational rent concept alone leads to excessively high values. In the implementation, the values were exogenously maintained within reasonable bounds. It is desirable to design a reliable internal price-generation mechanism: the issue is important because the price of land is a crucial factor of the supply, and hence influences the whole process of equilibration.

5.2. Potential of the model

These long-term considerations rely upon the hope that the model can be improved soon so as to reach the twofold objective of realism and convergence. Thereafter, its reliability as a forecasting tool would be proven. In that expectation, two comments can be made about its usefulness:

-1) The first comment does not address the general question of appreciating the usefulness of urban land use models: the model shares with all its predecessors the big hopes and major criticisms which they alternatively generate. The comment deals with its original features.

One characteristic of the model is the comprehensiveness of its description. Because of the several elements of the mobility/bundle, the model generates predictions about a number of dimensions;

For example, the knowledge of auto ownership and mode-to-work makes possilbe a number of analyses:

- one has been discussed in chapter 4: it is the number of vehicle-miles travelled per day for work trips (VMT). This criterion is linked with a number of phenomena that the policy-maker is interested in analyzing. The pollution level is one example, the consumption of gasoline is another. Therefore, the model is useful to evaluate the impact of public policies upon those two important issues. In a multidimensional analysis, it can be sought to optimize an aspect of the performance of the transportation system, subject to the constraint that the total VMT remains within a given range.

- some other analyses have not been explicitly formulated, but can very easily be on the basis of the information generated by the model.

Since the magnitude and price of several consumptions of urban goods and services is described, one can appreciate various policies by comparing the increased utility of the consumers, and the public cost involved.

This can apply to housing programs and the impact of public subsidies on the demand; it also allows various transportation policies to be evaluated: the interzonal work-trip matrix amd model splits being available, it is possible to evaluate the impacts of

. a change in transit fares.

. an increase of auto-costs (for example through an increased gas tax).

. the improvement of parts of the transportation network.

-2) As stated in the introduction, it is clear that most of the potential of the model is transportation-oriented: the <u>second comment</u> about its usefulness also deals with transportation. More precisely, it addresses the potential integration of urban transportation network models with urban simulations. As referenced in Putnam's paper, this has been already attempted, with the following components:

- the urban simulation was the Lowry derivative "PLUM" which has been described in chapter two.

- The chosen transportation package was developed by the Planning Sciences Group of the University of Pennsylvania, 1973. Several results have been obtained and would not have come out from the implementation of one of the two models alone. This is particularly true in subareas experiencing rapid development.

Because of its strong transportation orientation, the new model is suitable to take place in an integrated package.

- The network equilibrium model would generate the flows and levels of service in each period: this would reflect the dynamic congestion effects which are currently foregone by the urban simulation.

- The latter would forecast the distribution of activity which results from the status of a transportation system. These relationships are typically ignored by the transportation flows - equilibration models.

- A last step would be necessary to insure the completion of the simulation: the status of the transportation network itself is subject to changes. These so-called "type III" relationships are rooted by public policy-makers, in consideration of the following two:

  . the current level of service

  . the predicted distribution of activity

The usefulness of the model is to forecast the effect of alternative policies for the purpose of comparing them. Therefore, the very exogenous factor is the description of the policies to be tested.

However, for several-years simulation, the initial plan can be followed by a number of linked policies which depend on the observed evolution: these cannot be initial inputs and could perhaps be internally generated by the integrated model. For instance, the model could automatically react to the level of service signals.

The following illustration deals with the implementation of a new subway line: the initial decision would be exogenously input. The extension of the line to a fast developing subarea could be internally generated according to the growth of the population in the area.

This field of research will presumably raise a lot of difficulties but is thought to be very promising: the simultaneous use of models addressing complementary issues of the urban development is certainly the appropriate way to gain a robust knowledge of the whole process.

The author expresses his hope that the "new model" will receive further consideration, and maybe take place one day in the ambitious construction of an integrated Land Use Model.

BIBLIOGRAPHY

*Adler, T. and Ben-Akiva, M. "A Joint Frequency, Destination and Mode to Work Model for Shopping Trips". Transportation Research Record, 1975.

*Alonso, W. "Location and Land Use", Harvard University Press, Cambridge, 1964.

*"Bay Area Transportation Study Commission". Locational Model System, 1968.

*Harris, B. "Quantitative Models of Urban Development" in "Issues in Urban Economics". Johns Hopkins Press,1968.

*Herbert, J. and Stevens, B. "A Model for the Distribution of Residential Activity in Urban Areas". Journal of Regional Science 2. No. 2. 21-36.

* Ingram, G. Kain, J., Ginn, R. "The Detroit Prototype of the NBER Urban Simulation Model"; Columbia University Press, 1972.

*Koppelman, F. "Travel Prediction with Disaggregate Choice Models", PhD dissertation, Department of Civil Engineering, MIT. June, 1975.

*Lerman, Steven. "A Disaggregate Behavioral Model of Urban Mobility Decisions", PhD dissertation, Department of Civil Engineering, MIT. June, 1975.

*Lowry,I. "A Model of Metropolis" RM - 4035RC, Rand Corporation, Santa Monica, 1964.

*Lowry, I. "Seven Models of Urban Development:A Structural Comparison; Santa Monica, Cal. Rand, September, 1967.

*Manheim, M. "Fundamentals of Transportation System Analysis", Preliminary Edition, Third Revision, August,1975.

*NBER Publication. "Progress Report on the development of the NBER Urban Simulation Model and Interim Analysis of Housing Allowance Programs". J. Kain, Editor, December 1974.

*NBER Publication. "First Interim Report on Contract to Improve the NBER Urban Simulation Model"; J. Kain, W. Apgar, R. Ginn, G. Ingram.

*Putnam,S. "Urban land Use and Transportation Models: A state of-the-Art Summary", Transp. Re. Vol. 9. Pergammon Press 1975.

*Traffic Research Corporation. "EMPIRIC Land Use Forecasting Model Final Report", February, 1967.