Difficult Decisions in Land Use Model Construction

W. L. GARRISON, Northwestern University

THE ART and the science of formulating and using models in studies of urban land uses require a large number of decisions—decisions with respect to data to be used, scale and detail, methods of estimation and projection, and other activities in model construction and utilization. The purpose of this discussion is simply to identify and discuss certain of these decisions that appear to be especially difficult. In this paper, decisions are separated into two classifications: (a) those having to do with the inclusion of "notions about behavior" in models, and (b) decisions bearing on the formal structures of the models. The discussion of behavior emphasizes problems of possible actions by planners within models, problems of identification of entities to which behavior is attributed, the related problems of levels of aggregation, and problems of estimation occasioned by spatial correlation of behavior. In the discussion of the formal structures of models these topics are again emphasized. This emphasis, however, will focus on how decisions about behavior have counterparts in decisions about formal structures of the models.

The two classes of decisions—decisions about behavior and decisions about the formal structures of models—are by no means completely inclusive of all decisions that must be made in model construction. No attempt is made here to catalog all of the decisions that the model builder must make. The decisions emphasized in this paper are thought of as difficult ones mainly because of the interdependence between what is assumed about behavior and the subsequent problems that appear when the formal structure of the model is determined. Other considerations bearing on the judgment that these are difficult decisions include lack of guide lines in the literature concerning how these decisions ought to be made, and lack of adequate attention given to these problems in many current studies. In this author's judgment, at least, there is lack of adequate attention to these problems.

It should also be noted that this paper is included in a panel of papers dealing with the science and the art of land use model construction. In a broader context, this paper is set within a large, rich, and rapidly developing literature. What is presented here is a fragment that is of high priority to this writer and that has meaning in reference to the literature. The reader should remember that much has been accomplished in the last decade in this area of work. Reference to Lowry (1) and the papers in two recent publications (2, 3) will provide an overview of the status of this work.

INCLUDING NOTIONS OF BEHAVIOR IN MODELS

An example may help define what is meant by including behavioral notions in models. Assume that a model generates as one set of its outputs the industrial development of sites within and surrounding a city. Then, how firms select industrial sites would be a notion of behavior that might be replicated in a model. More generally, behavioral notions are concepts related to how decisions are made by the entities within the model. Behavioral notions are included in models when the actions of entities within the model are related to behavioral ideas.

Some Brief Notes

At first glance, the selection and inclusion of relevant behavioral considerations in a model present no special problems: simply examine the desired output of the model

Paper sponsored by Committee on Land Use Evaluation and presented at the 45th Annual Meeting.

and in check-list fashion assure that appropriate behavioral considerations have been replicated within the model. This check-list type evaluation would be a most interesting exercise, and it might well point up some gaps in thinking with respect to a specific model. The evaluation might bog down in the semantics of what we take to be knowledge of behavioral relationships. Since in this world everything seems to be related to everything else, decisions with respect to redundancy and pertinence would soon become quite critical. As Berelson and Steiner (4) show, large numbers of behavioral findings are available to the developer of a model.

The relative importance of the problems which a model attempts to solve, along with the sensitivity of solutions to the outputs of the model and to the weights of goals, may provide guides to the kinds of behavior to receive priority in inclusion in models. A presumption is that when given the criteria by which a system should be evaluated, entities can be identified and assigned a specific priority. Relationships that associate entities, together with those that determine the states reached by entities, can be identified on behavioral grounds. Consequently, weighting of goals provides a guide to the behavioral content of the model.

This seems simple, but in dealing with urban areas there is no consensus on how we should react to problems, much less to the priorities that ought to be given to treatment of problems. Besides, there is a lack of consensus on the willingness to trade off solutions of one class of problem for solutions of others. Some of the features of this situation are manageable; others are not.

The difficulty of identifying and attaching priorities to solutions of urban transportation problems illustrates this situation. Some persons identify the urban transportation problem with congestion during rush hours. Others maintain that the problem is the low profitability of mass transportation. Still others claim that it is the very large investment requirements over the next decade or so that pose the problem.

There is nothing contradictory in the identification of multiple facets of a problem. Sets of inputs are used to produce sets of outputs, and different persons at different times single out different inputs and/or outputs for comment. Difficulties for the analyst occur when an attempt is made to value these inputs and outputs at different places, at different times, and with reference to different people. It is one thing to observe that inputs and outputs can be stated; it is quite another to assign specific values to these inputs and outputs. This latter complexity makes it difficult to identify behavioral relations to be given priority for inclusion in models.

A critical point for consideration is the manner in which the activity of planning is included in the model. In many models planning plays a somewhat passive role. In many transportation planning studies, for instance, models are used to forecast urban land use changes and other expansions. In these studies planning activity is undertaken in order to supply the transportation facilities implied by the forecast. This procedure is followed in spite of the ample evidence in the transportation literature that the leverage on urban growth and development available from transportation is very high. Planners and the planning activity should play an active role in spelling out the alternative paths of development that may be achieved through transportation investment.

Another place where planning should enter the model is in instances where the sets of outcomes from the model ought to be predetermined by planning, rather than left to an evolution of present patterns.

Planning is ordinarily undertaken either to provide goods and services to the public sector or to provide special protective activities. Protective activities may apply to the elimination of obnoxious activities from quiet, residential streets. Such activities may refer to the managing of bundles of activities with high levels of interdependencies, such as those combinations of activities which can produce a desirable residential neighborhood. If planning makes any sense at all, then it affects urban growth patterns from time to time, and planning activities are essential elements in forecasting models.

This is, I believe, an important thought. A model may include a variety of entities. Some might behave in ways replicated by stochastic processes of a purely random sort. Behavior of other entities might be replicated by diffusion mechanisms or market mechanisms. In addition, planning behavior, itself, is a kind of behavior that should be replicated in the model.

One feature of land use models is that from situation to situation problems are attacked at different levels of generality and of aggregation. This is true both within models and between models. For instance, industries are sometimes classified as either basic or nonbasic. Then inferences are made from relationships between city sizes and percentages of all workers employed in basic activities. At this high level of aggregation, basic activities are found to decrease in percentage of importance when city size increases. In other situations, researchers might choose to attack a problem through reference to specific industry sectors. In still other situations, they might work with specific industries and with specific plants and firms. At each of these levels of aggregation, the investigator's statement of what is taken to be the pertinent behavioral relations would change. For a specific firm, in contrast to the statement made earlier about basic industry, one might be interested in the perception of the market in the eyes of the entrepreneurs, resources available to the firm, the payoffs the entrepreneurs are seeking, and the combination of all these factors into the behavior of the firm. One consequence of different levels of discussion and modeling is redundancy—redundancies occur because problems are identified at different levels of detail. To continue the example, at one place the researcher may use a notion relative to the behavior of the firm. At another place a behavioral notion may relate to entire industries. The notion relating to the firm is, in a sense, redundant to the notion about the industry; and the reverse is true.

Later in this discussion reference will be made to aggregations by geographic areas. Here again redundancies occur, for a notion relating to large areas may be redundant to notions about small areas; and the reverse may be true. An example here might be notions about travel generations by traffic zones, census tracts, households, and individuals.

As an examination of Berelson and Steiner (4) will show, very little is known about relationships between behavioral notions identified at different scales, different levels of generality, or different levels of aggregation.

Remarks up to now have been skimpy and preliminary in character. However, the following are some decisions that emerge:

1. Decisions must be made with respect to the behavioral ideas to be incorporated and the relative emphasis that they should be given. These decisions are difficult because it is difficult to attach values to the outcomes generated by the model and trace these back to the levels of emphasis that should be given to alternative behavioral ideas.

2. Decisions are required with respect to the levels of aggregation at which topics in the model will be treated. One choice is to speak of the behavior of aggregates of entities; another choice is to make statements about entities themselves. Aggregation decisions affect the kinds of and levels of generality of behavioral notions that may be treated in a model.

3. Decisions must be made with respect to the place of planning-behavior in a model. A model may be a forecasting device providing estimates of growth which are then used as a basis for planning; and/or a forecasting model may include within it those planning activities which are required to give proper direction to the outcomes and reliability to the forecast.

Economic Behavior

In the sense that a model deals with economic things and/or includes economical (maximizing or minimizing) calculations, all models are to a large extent economic. Use of economic ideas in model construction to organize treatment of these economic things and calculations is immensely appealing. This greatly simplifies statements of expectations concerning the behavior of individual entities identified in the model, and statements concerning the properties attached to the sum of their behavior. Constructing a model within the framework of economic behavior means that the well-developed apparatus of economic analysis can provide a basis for evaluation of alternatives, estimation of the model, and understanding of possible advantages and pitfalls of social action based on the model.

Two types of activities about which economic notions have been replicated within land use (and transportation) models may be noted: (a) the consumption of transportation and (b) the choice of the points in space from which the transportation system is utilized. The latter activity is a central one in land use models. In the instance of entities that are households, choices affecting the places from which transportation is consumed include those of residential site selection, selection of place of employment, and selection of places for shopping, education, and recreation. Problems in the demand for transportation include those of the selection of transportation equipment, time of travel, routes employed, and frequency of travel. In the instance of the industrial firm, relevant decisions of the first type are those of location of the firm with reference to locations of suppliers and markets, employees and service personnel, and other firms. Transportation utilization decisions include type of vehicle to be used, and routes. It is clear, of course, that selection of the point from which the transport system will be used affects transportation demand, both for the household and for other kinds of entities.

Consider the transportation demand question first. There is every reason to lean very heavily upon economic concepts within a rather constrained range. Transportation activities are carried out quite frequently so that the user of the transportation system has a good knowledge of the alternatives available. It seems reasonable to assume that within the range of alternatives available, choices are economic. This is a simplifying assumption of great interest, for it simplifies the assignment of traffic to networks and, at least in theory, permits the assignment of values at the margin to increments of transportation system capacity. Nonetheless, our ability to make this assumption does not carry us very far with transportation and land use problems. A number of supply alternatives are available; and, for each alternative, decisions must be made about the facilities provided, their prices, and their locations. The economizing of travel behavior on the demand side is not very helpful when these supply questions are discussed. Answers to supply questions involve considerations of income transfers, new developmental alternatives, and evaluation of new technological alternatives which are outside the experiences of current users. The latter cannot be weighted by observing the economic behavior based on the choice of current transportation systems.

Adopting economic behavior as a framework for reproduction of the site selection process would appear to be very logical. Here it is assumed that sites are differentiated in value because of their locations. It is also assumed that the competitive struggle for strategic sites maintains these values and assures the assignment of uses to optimal sites—optimal within the range of choices available at a given time. A problem here is that this mechanism, in some sense, may not work very well. The site selection process is carried out relatively infrequently in the life cycle of the firm or of a family. There is some reason to believe that the range of alternatives considered by firms and households is very limited and that location-derived land values do not play a large role in the decisions being made. One point bearing on this is the fact that site rents tend to be a rather minor part of firm or household budgets. In a decision situation where information is poor, these minor variations of budgets may tend to go unnoticed in a decision, in comparison with real or imaginary non-monetary advantages and disadvantages of sites.

Inferences

With respect to utilization of the transportation system, the number of system configurations and technological alternatives from within which choices can be made is very great. Consequently, it is not very meaningful to consider transportation usage patterns except in contexts of other considerations, such as constraints on investment and capacity, staging of investment, nature of pay-offs from the investment, and so on. A critical item enters here: the behavior of the planner in such a decision situation. This returns us to one of the themes emphasized earlier in this discussion: decisions about the representation of planning activities, as such, within the models are critical. In this instance, they provide the context within which economic behavior takes place.

It is also critical to make proper decisions with respect to the context within which the site selection process takes place. From one urban site to another, differences in travel costs may be relatively small compared to differences in the costs of sites occasioned by other considerations. Examples in industrial location are provided by presence or absence of adequate sewage, influence of zoning on possible land use, availability of large parcels, and prestige. A variety of neighborhood amenities plays a large role in residential site selection, as do prestige, discrimination, and other factors. Factors of this type may over-ride location-travel questions.

This paper is not intended to be a review—it is intended simply to stress some of the difficult decisions that must be made in land use model construction. Nonetheless, it might be useful to mention that reference to existing land use models will point up a rather chaotic situation resulting from the way behavioral questions are included in models. Some studies represent the site selection process strictly in reference to economic choice; other studies have represented the diffusion or spreading out of the structure of land uses from the existing city. Some studies permit different projections in response to the choices available to the planner. In other studies, plans accommodate the current trend.

STRUCTURE OF THE MODEL

A model is simply the device within which notions about relationships are made explicit, in terms of available data and in terms of the questions to be answered. The model's structure is the manner in which relationships are identified, expressed, and related one to another. A model of land utilization should be responsive to the relationships of a behavioral type emphasized earlier in this paper. Other classes or types of relationships might also be included. Most models are to be used for studies of development as well as of growth, and relationships indicating technologically feasible alternatives would be included. Necessity for computation of the cost of subsystems and the cost of systemwide alternatives, and the presence of budgetary constraints may require sets of fiscal relationships. Elasticities of demand might require representation; and systems effects may require including appropriate circuit and mode relationships.

Additional topics for consideration in model construction are those of the appropriate entities which are to be measured and entered into the model, the methods by which the parameters of the model are to be estimated, the degrees to which the model is to be self-adaptive and/or recursive, etc. Each of these topics is a complex subject, and no one of these can be covered adequately in a short paper. Rather, the present discussion will continue the difficult decisions theme by picking out certain aspects of the model building problem which seem to pose especially difficult problems.

Many Relationships Required

A special feature of land use models is their strong cross-section orientation. The models depend upon cross-section data for the variability from which parameters of relationships may be estimated, and the developments projected by land-use models clearly have a two-dimensional spatial content. These aspects of the analysis pose difficult problems.

It might be constructive to consider a specific simple case, say a model of the form $y_{ij} = a + bx_{ij} + e_{ij}$, where the i subscript refers to the i th time period and the j subscript referred to the jth entity observed. In the instance of a time series, j is constant and i takes on values for the appropriate time periods. In other words, the entity on which the observations are made is unchanged, and variability in y is strictly that occasioned by different levels of x (and of e_{ij}, j constant). However, in the instance of cross-section analysis, the subscript i becomes constant, and observations are made on different entities. Here variability may be occasioned by entity to entity differences, as well as by the variability occasioned in y by different levels of x (and of e_{ij}, i constant).

This discussion is somewhat oversimplified. Nevertheless, it points up that because of entity to entity differences, cross-section analysis may require measurement

of many properties of entities in order to measure the resultant and related variability in the outcome.

Entities may tend to change in time, too. Generally, however, entity to entity differences are greater in cross-section work. This point has been stressed by Klein (5).

If a large number of properties of entities are measured, it follows that a large number of relationships must be treated. Measurements cannot be included willy nilly in a model unless relationships are specified. In turn, specification of relationships may require reference to behavioral notions. This is a point where technical aspects of the analysis—the cross-section structure of the model—have impact upon the kind and quantity of behavioral relationships entering the model.

Aggregation

It was previously stressed that difficult decisions must be made relevant to levels of aggregation used in the model. The problem of aggregation is further complicated by problems of working with cross-section data. Cross-section data deal with entity to entity contrasts; and if entities are aggregates made up artificially, say political areas, then contrasts of interest may be obscured by variability introduced by the artificial method of aggregation.

Total variability within a set of data may be broken down into the variability within aggregates and variability among aggregates. However, when only the variability among aggregates is available for study, there may be systematic biases in comparison with findings that might be obtained from the total set of data. Aggregates may be formed in many ways—for a given set of data, findings might differ from time to time when different methods of aggregation are used. One kind of aggregate used frequently in urban work is the census track. Other arbitrary geographic areas are also commonly used. In many instances these aggregates are very poorly suited to accommodate the kinds of comparisons required in model estimation.

Aggregation has been discussed from several standpoints. It was stressed that aggregation should be responsive to the purposes of the model and the entities which the model was presumed to replicate. Another set of problems in relation to aggregation has also been mentioned. Here is still another instance where decisions are difficult. Decisions made on behavioral grounds about aggregation may affect ease of determination of parameters of the model; and in turn, decisions about aggregates made to enable parameter fitting may affect interpretation of behavioral relationships.

Type of Relationships

A return to the comparison between cross-section and time-series analyses will provide orientation to another difficult problem, and related requirements for decisions in behavioral content and in model statement. Model estimation is simplified if the e_{ij}'s can be assumed to be normally and independently distributed, with mean of zero and constant variance. (Specifically, $E(ee^t) = I\sigma^2$, where e is a column vector; e = either $\{e_i\}$ for time series or $\{e_j\}$ for cross-section series. E refers to expected value. I is the identity matrix and σ^2 is the variance of the process.) If this is not the case [that is, $E(ee^t) = D \neq I\sigma^2$], then the process is not stationary and purely random. It is known that if a process is not stationary and purely random, then ordinary estimation procedures will lead to estimates of coefficients with biased parameters. If the process is not stationary and purely random but can be specified (i.e., in terms of matrix D), then a method of estimation is available. There has been considerable experience in time-series work with autocorrelation processes and estimates involving consideration of this type.

In the instance where the residual e_{ij}'s are from cross-section analysis (i = constant), the situation is somewhat more complex than in the instance of time-series analysis. In time-series analysis processes may be presumed to be one-directional and one-dimensional. The processes move forward along a time dimension. In cross-section analysis, on the other hand, the processes are in two-dimensional space and they may work in any direction. Because of this increased complexity, it is somewhat more difficult to work with cross-section processes than with time-series processes.

Work by Geary (6) on the analysis of residuals from cross-section analyses and spectral analysis of series (7, 8) illustrates difficulties and also suggests approaches.

It was emphasized that the site selection process is influenced by such matters as prestige, level of neighborhood amenities, zoning, and other phenomena. A difficulty is that these phenomena are linked in complicated ways in two-dimension space. They are specifically the kinds of phenomena which result from processes that are not stationary and purely random with reference to the cross-sectional frame.

Residuals, e_{ij}'s, from an analysis would not meet the specifications of being stationary and purely random if some of the processes that had spatial character pertinent to the analysis were omitted from the analysis. This points up another way in which the decisions about behavioral content of the model carry over into problems about the actual structure of the model. With respect to the structure of the model, one must make decisions about assuming or not assuming certain properties of the residuals from parameter estimation. These decisions cannot be separated from decisions that must be made about the behavioral content of the model.

SUMMARY

Somewhere in the process of model development the planner must make certain decisions about the kinds of behavior that will be replicated in the model. Decisions concerning where planning fits into urban growth, what entities are assumed to take on behavior, and what kinds of behavior are to be included, were emphasized because they interrelate with decisions about the structure of the model. Decisions with respect to entities that take on behavior and levels of aggregation influence the methods that must be used in estimating and the number of variables to be included. Decisions related to the nature of processes, and especially to spatially autocorrelated processes, influence errors of estimation, particularly estimates of the properties of the coefficients of the model.

The decisions and problems discussed are very limited ones when compared to the totality of our knowledge and experience with model construction and use. Nonetheless, these decisions reappear in land use model construction tasks, and they require more systematic thought. What is needed is more understanding of the nature of the processes involved in urban growth, development of ways of associating these processes with entities in the urban growth process, and continued improvement of ways of estimating models. This is a field where wisdom has increased greatly with experience and where experience is increasing by leaps and bounds. The prognosis is good for improved handling of these difficult decisions.

REFERENCES

1. Lowry, I. S. A Short Course in Model Design. Jour. Amer. Inst. of Planners, Vol. 31, pp. 158-66, 1965.
2. Urban Development Models: New Tools for Planning. Britton Harris, ed., Jour. Amer. Inst. of Planners, Vol. 31, No. 2, 1965.
3. Urban Transportation Planning Techniques and Concepts. Highway Research Record 102, 1965.
4. Berelson, B., and Steiner, G. A. Human Behavior: An Inventory of Scientific Findings. Harcourt Brace and World, 1964.
5. Klein, L. R. A Textbook of Economics. Row, Peterson and Co., 1953.
6. Geary, R. C. The Contiguity Ratio and Statistical Mapping. The Incorporated Statistician, Vol. 5, pp. 115-142, 1954.
7. Ocean Wave Spectra. Proceedings of a Conference. Prentice-Hall, 1963.
8. Priestly, M. B. The Analysis of Two Dimensional Stationary Processes with Discontinuous Spectra. Biometrika, Vol. 51, pp. 195-18, 1964.