HIGHWAY RESEARCH RECORD

Number 126

Land Use Forecasting Concepts

7 Reports

Subject Classification
83 Urban Land Use

HIGHWAY RESEARCH BOARD
DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING
Washington, D. C., 1966
Publication 1369

Department of Urban Transportation Planning

Pyke Johnson, Chairman
Washington, D.C.

HIGHWAY RESEARCH BOARD STAFF

James A. Scott, Urban Transportation Planning Specialist

COMMITTEE ON LAND USE EVALUATION

(As of December 31, 1965)

F. Stuart Chapin, Jr., Chairman
Research Director, Center for Urban and Regional Studies
University of North Carolina
Chapel Hill

Joseph R. Stowers, Secretary
Civil Engineering Department, Northwestern University
The Technological Institute
Evanston, Illinois

Frederick T. Aschman, Executive Vice President, Barton-Aschman Associates, Incorporated, Chicago, Illinois
Robert E. Barraclough, Chief, Land Development Section, Tri-State Transportation Committee, New York, New York
Salvatore V. Ferrera, Chief Economist, Chicago Area Transportation Study, Chicago, Illinois
William L. Garrison, Director, The Transportation Center, Northwestern University, Evanston, Illinois
David A. Grossman, Acting Assistant Director for Program Review, Community Action Program, Washington, D. C.
John R. Hamburg, Associate Director, Upstate New York Transportation Studies, New York State Department of Public Works, Albany
Calvin S. Hamilton, Director of Planning, City of Los Angeles, Department of City Planning, Los Angeles, California
Walter G. Hansen, Alan M. Voorhees and Associates, Inc., Washington, D. C.
Britton Harris, Associate Professor of City Planning, Institute for Urban Studies, University of Pennsylvania, Philadelphia
C. David Loeks, Director, Twin Cities Metropolitan Planning Commission, St. Paul, Minnesota
David D. Longmaid, Director, Penn-Jersey Transportation Study, Philadelphia, Pennsylvania
Ira S. Lowry, Logistics Department, The RAND Corporation, Santa Monica, California
John K. Mladinov, Director, Puget Sound Regional Transportation Study, Seattle, Washington
Kenneth J. Schlager, Chief Systems Engineer, Southeastern Wisconsin Regional Planning Commission, Waukesha
Lowdon Wingo, Jr., Resources for the Future, Inc., Washington, D. C.

Foreword

The seven papers appearing in this issue of the RECORD will be of special interest to those concerned with growth and structure of urban areas and with the development of land use models for transportation planning. As a group, these papers constitute a sampling of work in progress relating to land use concepts and models.

The first three papers deal with several related considerations of theory and its application in modeling land use processes. Harris' paper keynotes for those in the urban research community some of the fundamental concerns that all researchers must eventually cope with in using the scientific method. He restates the elements of this method and in the process sets forth a set of basic guidelines in theory-building, indicating how theory and application are intimately tied together. He advocates an emphasis on scientific rather than mission-oriented research, an emphasis that stresses real world systems, behavior of decision units, and the interaction between systems, subsystems and decision units, and an approach that replaces the segmented view of urban phenomena with a more unified effort of disciplines and researchers, and one with a clear and frank recognition of the importance and expectation of the necessity of experimentation in this research effort. Lowry's commentary on the Harris approach helps to clarify some of the issues of the day and at the same time serves to highlight the timeliness and key importance of this paper.

Garrison's paper follows naturally from Harris' discussion and poses some of the difficult decisions that model-makers face in moving from a conceptual approach to model-building. One group of decisions covered relate to behavioral entities appropriate to land use models, the recognition of planning behavior in forecast models and problems of levels of aggregation. Another group of decisions discussed concerns the nature of processes represented in the model and the structure of the model and such problems as cross-section analysis, measurement, and the specification of variables.

Schlager's paper, the third in the group of three theory-related paper, at once offers a context for the review of some of the issues raised by Harris and Garrison, and at the same time presents in succinct and convincing form an optimizing type of land use model. The model makes use of a recursive programming framework and outlines an approach for simulating the behavior of land developers, households and government in the residential development process.

The Swerdloff and Stowers paper is concerned with testing various models using a common set of data, and the Boyce and

Cote paper focuses on the need for confidence statements of forecast variables to permit more valid comparisons to be made among models. The Hill and Brand paper provides illustrative examples of model methodology and discusses the proper use of linear regression analysis in land use modeling. The final paper by Hadfield and Orzeske deals with a land use inventory technique, the use of a sampling approach to updating land use from aerial photography.

Contents

The Uses of Theory in the Simulation of Urban Phenomena

BRITTON HARRIS*, Professor of City and Regional Planning, University of Pennsylvania

A brief examination is made of the part played by theory in human development. Theory and practice are contrasted and a parallel contrast is drawn between science and engineering. An effort is made to demonstrate that in the modern world it is becoming increasingly difficult to draw a sharp line between these various opposed pairs of activities. It is also concluded that theory and practice have a well-known reciprocal relationship related to scientific induction and deduction. An examination is made of the creative jump from inductive generalization to the generation of new theories or new ways of looking at the world which will sustain extensive deduction and generate new hypotheses which are testable. Three criteria for useful theories are discussed: capacity for manipulation, fruitfulness, and economy. Other criteria such as realism, comprehensibility, and comprehensiveness are examined and found to be only partly applicable. The discussion is illustrated throughout with examples from land use and transportation simulation. Conclusions are drawn regarding future directions of research and some of the desirable characteristics of research establishments.

•MOST OF us who are engaged in one form or another of transportation and land use research have focused a very large proportion of our efforts on simulation. This means that we have devoted our efforts to reproducing in recognizable form certain aspects of human behavior and the performance of mechanical systems or a combination of the two. We have done this generally in order to make predictions, and we have been interested in the accuracy of predictions in order to assist our agencies or other policy-makers in making decisions. It is the aim of this paper to provide a brief review of some of the ways in which theory can be of assistance in improving the similitude of simulations, and consequently the accuracy of predictions and the wiseness of decisions.

Since there is a good deal of popular jargon which tends to imply that practical activities are useful while activities dealing with theory tend to be nonproductive, I intend to devote a part of this discussion to what might be called paradoxically a down-to-earth defense of these impractical activities—and to some extent I shall oppose what I would call crackpot realism with what might be termed realistic idealism. As Bertrand Russell has said, "Nothing is as practical as a good theory."

In very simple terms, theory is a general statement about the real world. In these simple terms, for example, the Pythagorean theorem is one of many consequences of the theory of Euclidean geometry. As such it makes its own general statement about the properties of right-angle triangles on plane surfaces, and has had tremendous

---

*Formerly Research Coordinator, Penn-Jersey Transportation Study.

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

practical influence in surveying and engineering. This theorem provides the basis for all the well-worn formulas of elementary trigonometry. There are two ways in which, however, we need to qualify this simple-minded definition of theory, and it is these qualifications which may tend to give the notion of theory some of its other-worldly character. First, when we say that theory concerns the real world, we have to include in that real world the minds and ideas of men. Thus, theory may deal to some extent with technology and concrete things on the one hand, and on the other hand with mental constructs which are seldom or never encountered in the physical world outside of men's minds until they have been written down. The real world of mental constructs is a very important one, and in the end has many practical applications. The extension of the trigonometry of measurement into trigonometric functions, for example, is the basis for other large parts of engineering. The second qualification is that a theoretic statement about the real world may not be, to the layman at least, a recognizable mapping of the real world, and the nature of the correspondence between the theory and the world and the consequences of the theory may not be readily expressible in everyday language. This sometimes makes it difficult for the layman to conclude at first glance that the theory is in any sense realistic or has any practical consequences.

There is of course an intimate relationship between theory and science or between the verification of theory and the scientific method. Since theories consist of statements about the real world, their degree of correspondence with this reality can be tested. Where the real world in question is one of mental constructs, as in logic and mathematics, the testing may be of a special and somewhat different nature, based on internal relations between constructs. It is not in general a requirement of the development of conceptual systems and their theory that any direct correspondence with material phenomena should be established, but it has frequently proved to be the case that after short or long periods of development, such concepts have found important and unforeseen applications to theories of phenomena. This course of events is analogous to, but not the same as, the laboratory development of methods and devices which for a long while remain mere curiosities, but which ultimately become technologically important.

We live in an age of rapidly expanding scientific endeavor. Science and the scientific method are being newly applied to old systems of human thought such as ethics and philosophy, and in these areas, the boundaries of untestable contention are constantly being narrowed. We now know that because of the atomic nature of matter only a finite number of angels can dance on the point of a pin; we also feel greater confidence in the rigor and cogency of philosophy. At the same time, new groups of phenomena are becoming the subject of science. Testable rather than speculative hypotheses are developed about these phenomena, and these hypotheses are organized in increasingly unified systems, frequently of a quantitative nature. Some of this movement towards new applications of the scientific method is occurring in the social sciences, and in this field the two tendencies to reduce the area of philosophical and ethical speculation and to systematize our understanding of objective phenomena go hand in hand.

It is hardly necessary to review the practical ways in which the advances of science during the last two centuries have greatly increased man's control over his natural environment through the application of science to technology. It is more useful to point out, first, that not only is science successful in an objective sense, but also that it is widely accepted publicly and politically, as may be judged by the governmental and private resources which are devoted to it, and second, that the growth and prestige of science have not simplified but have complicated the distinction between practical and theoretical endeavors. The customers of the scientific establishment are basically interested in results and frequently have shorter time horizons than the scientists. This dichotomy expresses itself in the distinction between science and technology as disciplines, and organizationally in the distinction between mission-oriented research and theoretically oriented research. As Alvin Weinberg (9) has recently emphasized, the objectives of mission-oriented research are externally imposed upon the scientific community by the real or imagined needs of society and by society's control over expenditures, while the objectives of theoretical research are largely generated within the scientific community.

It appears that these distinctions, while valid and useful for analyzing and discussing the problems of science and technology, can be unduly overemphasized. Many factors tend to blur the differences. Technology and engineering themselves are becoming more scientific in their basic methods, and consequently engineers are becoming scientists. Mobility between the professions tends to inject mission-orientation and social responsibility into the scientific community, which was in any case never detached from these values. The tremendously accelerated pace of science tends to shorten the scientist's time horizon and bring it more into accord with that of the decision-maker. Finally, the complexities of real life which face decision-makers are driving them away from simplistic common-sense judgments and in the direction of a more comprehensive and quantitative approach to the problems which they face.

It is in fact the magnitude of the problems of societal control in a period of rapid change and development which is providing the impetus for the truly scientific development of the social sciences. Problems such as maintaining peace, feeding billions of people, reducing racial discrimination, and organizing great cities require powerful instruments of control over men and machines. These problems of control can no longer be resolved by an engineering approach which is overwhelmingly oriented towards physical, inanimate, machine systems. Engineers working in transportation planning must pay increasing attention to problems of human behavior, and it is rapidly becoming evident that the relevant behaviors are not only in the fields of transportation demand and transportation system utilization, but also in the field of land use development and locational choice. In a sense, therefore, the planning-engineering professions find themselves working on a frontier of science. This is the area of social behavior and social control, in which the application of the scientific method has been unduly retarded. In order to explore what implications this situation will have for their work, we must therefore take a closer look at some of the elements of this method.

We are used to the idea that man and the other higher primates are endowed with an innate curiosity which leads them to explore their total environment in an apparently insatiable but not entirely purposive way. There is usually no a priori identifiable useful payoff in some of the exploratory activities of monkeys and children, and one is tempted to make an analogy with the data-collection propensities of social science research and transportation studies. It is also perfectly clear, however, that in man at least, curiosity extends beyond the accumulation of data about the environment. First, even the childlike exercise of curiosity involves the exploration of cause and effect. The experimenter will employ some of the simpler ploys of the scientific method to discover what worked when and where. And second, there is frequently an effort to generalize; there seems to be a tendency to seek out analogies and similar situations in which earlier findings and elementary theories can be tested. Thus we have in a primitive form the four main steps in some classic descriptions of scientific method:

1. Induction: the collection of information and its organization into patterns;
2. Generalization: a restatement of the cause and effect relations behind the patterns, or a redefinition of the patterns themselves in a more abstract form which includes the observations as a special case;
3. Deduction: the search for new special cases previously unstudied, as suggested by the more general statement, or theory; and
4. Testing: a check to see whether the new cases perform as predicted—if not, the theory must be revised.

This schema, while useful for analytic purposes, does not correspond in its rigid division of steps with the way in which scientific investigations actually proceed. We will use these categories as a basis for discussion, specifically maintaining, however, that the classification is artificial and if pressed too far, actually harmful.

The testing of theories about the world of phenomena raises special problems in the social sciences which should be generally understood before we take up other aspects of the scientific method. In this discussion I use the term testing in preference to the more usual verification because in principle no theory can be established, but only disestablished. A theory does not have verity, but verisimilitude. There are of course many theories which are outstandingly successful and for which there have been an almost unlimited number of successful tests, while the unsuccessful tests are nonexistent or occur only under well-defined special conditions. An especially significant case in which a theory may be regarded as firmly established because no counterexample exists is the correspondence between the counting numbers of everyday experience and the invented set of positive integers as defined in modern algebra by the use of the Peano postulates or otherwise. The theory states that these two systems, one from real life and one from mathematics, have the same form, and no exception to this theoretical statement is known or is likely to be discovered. A second example is the Newtonian statement based on the laws of motion, the law of universal gravitation, and Euclidean geometry. This theory relates the mechanics of the real world to a mathematical system of differential equations which Newton, in fact, was forced to invent. Up to the point where by relativistic considerations the Euclidean geometry no longer obtains in real space, there are no exceptions to this theoretic statement, and it too may be regarded as firmly established. In this latter case, it is important to note that the correctness of the theory has only been established by eliminating or controlling extraneous factors such as air resistance which affect the theoretically defined unimpeded motion of observed bodies.

The special problems of the social sciences arise, as is well known, out of the difficulties of pursuing this experimental method in which most variables are held constant (the ceteris paribus assumption of economics), while a limited subset of possible variables is manipulated. Where a large number of variables is involved, this difficulty can sometimes be overcome if a large number of diverse observations is available to the researcher, but this is unfortunately not the case with regard to the study of the development and the manipulation of large urban areas. Here the case material is limited in extent, and experimental manipulation is both extremely slow and vastly expensive with regard to the aggregate phenomena. Experimental cum statistical methods are only possible with respect to smaller elements of the total system. In these regards, science as related to total development of the function of the urban system is in most respects analogous with astronomy, which has a few cases of major interest, subject matter which is inaccessible to experimental manipulation, and the capacity for studying in the physics laboratory elements which do not aggregate by simple addition into the whole. The conclusions of the science of astronomy are not yet, however, directly useful in the guidance of societal action on a large scale—even though expenditures on the space program exceed expenditures on urban development.

It may however be argued that the disadvantages of the social sciences in establishing an experimental method have been greatly exaggerated. This argument is advanced on the grounds that the most important tests of theories concern their ability to predict new phenomena or phenomena not previously studied in detail, and to extrapolate the effects of causes beyond the ranges in which the causes were originally observed. It is curious to note that the literature of engineering and the social sciences abounds with warnings as to the dangers of extrapolation. If we wish to use, as we are almost forced to do, the power of extrapolation of a theory as a test of its credibility, then this cautionary advice is a frank confession that the relationships being extrapolated have no theoretical basis whatever. From the point of view of testing theories, the social scientist should welcome rather than shun opportunities for extrapolation, since this will provide his main basis for justifying a theory or for designing improvements in it.

Any acceptance of this criterion for testing theory tends to indicate the ultimate futility of a complete reliance on induction for generating theories. Even in the physical sciences, a fairly thorough knowledge of a particular range of joint variation of phenomena does not guarantee any adequate knowledge of cause and effect or even any complete description of relationships outside the range of observation, and this is even more true of the social sciences. Most of us are thoroughly familar with the situation which arises when we get a good fit of a polynomial to a set of observed data, only to find that it behaves extremely erratically outside the range of observation. Poincaré (6) has pointed out that if we had a complete knowledge of a portion, however small, of a continuous function, we would have a knowledge of the behavior of the function over its entire range. We could attempt to reach this happy state by developing the function as an infinite series and fitting all its coefficients statistically. Unfortunately, this procedure requires an infinity of observations (and a larger infinity than N_0). Even more important, our observations must be free of errors of measurement, and our function must include all relevant variables, each of which must be measured. It is quite clear that even the process of induction from observed phenomena must be guided by theoretical concepts based on previous experience which will suggest the ranges and objects of observation and the character of the functions to be fitted. A simple engineering example might be found in the difference between the parabolic curve and the catenary. These curves arise under different circumstances in the construction of suspension bridges, and lie extremely close together over a certain range of the variables. To distinguish between them by induction would be a hopeless task, especially since the formula for a catenary would not intuitively occur to a statistician, yet in extreme cases the distinction is important for engineering design. The differences are well defined a priori on the basis of a theory which may have been suggested by observation, but which does not spring automatically from it. The difficulties of induction are conclusively delineated, in fact, by the difficulties which arise in social science research in selecting functions for curve fitting and interpreting the results. Linear models are most frequently used because of their simplicity, perhaps with the justification that the linear approximation to some unknown function is not unreasonable over the range of the observations. The function being unknown means that theory is out the window. Perhaps a polynomial is used as some sort of an approximation to a Taylor expansion of a function. In this case, the catenary is defined as a parabola. Where the choice of function is deduced from a priori considerations and not merely to satisfy goodness of fit, we are suddenly in the realm of deduction rather than induction.

Deductive thinking is of very high value in science. Examined closely, the antinomy between induction and deduction is somewhat artificial; on the one hand, induction is almost never initiated without some kind of prior theory, however naive, which suggests areas of investigation, relevant variables, and the form which functions might take—while on the other hand, if deduction is unsuccessful and does not result in a confirming instance of the theory on which it is based, then the contradictory evidence may be the basis for a new round of induction. But the importance of deduction as a part of the scientific method remains in spite of the partial artificiality of its separation from induction. At the start of the process of deduction, the investigator is forced to make a statement of a general nature about the real world; in other words, he must formulate a theory. The motive for this formulation frequently comes from psychological forces very closely related to induction and to the search for generality. To follow the processes of deduction suggested by the theory, the investigator must search out new areas or new modes of application of the theory. It is useful to him in defining supposed cause and effect or functional relationships and variables to be investigated. In considering, therefore, the nature and power of the deductive process, we are led naturally to the final and perhaps the essential part of our discussion of theory construction, that of generalization, or the actual formulation of the theory. We must consider this in the light of all of the processes and problems which have been discussed previously.

Generalization is the bridge by which the scientist or theoretician crosses over from induction, or the observation of reality, to deduction, or the testing of theories and their application to new phenomena. For this reason, I rather like the name transduction, which is sometimes applied to it. No matter how often this bridge is crossed in the course of a scientific investigation, the act of transduction always involves some invention on the part of the investigator. The psychology of invention in this field is intricate and fascinating, but a discussion of it is out of place here. The sources of this invention may, however, be better understood through a consideration of its inherent nature.

The construction or invention of a theory involves in essence a precise statement regarding formal relationships, usually including relationships of cause and effect. There is an infinity of possible formal statements of relationships which may be made in their most abstract form in the language of mathematics or logic. Such statements regarding relationships are purely formal and have no reference to the real world. Within the sciences dealing with concepts, they may in fact be developed quite independently of the real world. The problem of theory construction or invention is, then, to make the correct identification between a real phenomenon and a mathematical or logical statement regarding relationships. There are three possible ways in which this may be done, two of which are merely suggestive and one of which tends to satisfy rigorous scientific requirements.

1. An analogy may be recognized at the level of phenomena. Thus, for example, a city may be compared with an organism—say a jellyfish. This analogy is scientifically useless unless two conditions are met: (a) the comparative object (the jellyfish) must have a form which has been clearly and logically defined; and (b) the object compared (the city) must be unequivocally said to be theoretically identical. In this case, we have identified a correspondence of the third type below, but otherwise we have merely made a statement which is useful for heuristic purposes.
2. An analogy may be recognized as between a phenomenon and a mathematical or logical construct, but may indeed be very loosely defined. Thus, for example, the gravity formula recognizes an analogy between the decay of trip frequency with distance and the power function X^{-a}. This analogy is extremely crude, seizing as it does on the most obvious and easily manipulated of a host of monotonically decreasing non-negative functions. No statement of the gravity model, to my knowledge, states any causal relationships which would generate this particular function in preference to others. I think that we may designate an analogy between phenomena and a logical function as a homomorphism, meaning a similarity of form.
3. An important qualitative change is introduced if a scientist identifies a particular phenomenon as having a clearly defined logical form. The definition of form may have already been made either in the development of logic and mathematics and unrelated to phenomena, or in connection with the development of theory dealing with some other and perhaps completely unrelated phenomena. The use of formal statements pertaining to other phenomena is indeed often suggested by analogies between the phenomena themselves. On occasion, the study of phenomena and the formulation of ideas about cause and effect necessitates the invention of a new relational calculus. This has been the case in Newtonian mechanics and quantum mechanics. In any event, the essence of a theoretical statement is to identify an isomorphism (identity of form) between a set of phenomena and a logical or mathematical relational system. Thus, the Schneider (7) model for trip distribution, in contradistinction to the gravity model, makes a rigorous statement that the decay of trip frequency is isomorphic to the negative exponential function and consequently also to the radioactive decay of fissionable elements; and Schneider identifies the precise cause and effect relationships on which the isomorphism is based.

It should of course be clear that the borderline between homomorphisms and isomorphisms is blurred, partly because it refers to the motivations and psychology of the scientists. A theory which is in fact generated as an analogy must be presented as an isomorphism, and a badly conceived isomorphism may turn out to be only an analogy. The appropriate distinction can be made only upon close examination of the theory and of its results.

In formulating a theory to serve as a bridge between induction and deduction, the analyst has a number of guides as to desirable features of his formulation. Some of the most significant criteria tend to conflict with one another while others reinforce each other, depending on circumstances. All arise out of the general characteristics of the scientific process as we have outlined it.

The outstanding criterion, of course, is that the theory should be correct, that is, that the theory if testable should pass its tests successfully. This is the basis for the essentially practical nature of science, that is, that it says true things about the real world. We have seen that such truth is impermanent, always awaiting contradictions. If these arise, it frequently is retained by circumscribing the generality of the theory, limiting the circumstances in which it applies, and creating new and more general theories to apply to other combinations of circumstances. This criterion of correctness is in general overriding. However, practical considerations frequently lead to the generally indefensible practice of applying theories which have been inadequately tested or which have known errors.

Thus a theory, if testable, must pass its tests. In fact, the untestable theory is a nontheory, and testability is therefore an important criterion in theory construction since, like Milton, science has "no use for a fugitive and cloistered virtue. . . . " There are of course examples of very important scientific theories, such as Einstein's theory of relativity, which when published appeared very difficult if not impossible to test. These difficulties in relation to a reputable and exciting proposal often serve as a spur to experimental work. In the field of social sciences, however, this particular type of nontheory has two other manifestations. One of these is a normative monitory description of how things should be done; the other is a literary or pseudostatistical description of the real world. The fact that these things are merely masquerading as theory can easily be exposed by searching for critical tests which could deny their validity. If such tests do not exist, the so-called theory is in fact a nontheory. Occasionally tests acceptable to the authors of the theory will be so circumscribed with restrictive conditions and assumptions as to expose the fact that the theory is of extremely limited application and has dubious stature.

The testability of a theory is a special case of a more general property of useful theories—productiveness or fruitfulness. Important theories in the development of science not only answer the problems posed in the initial stages of induction, but are pregnant with consequences which are only dimly seen by their inventors and which lay the basis for a wide variety and a great number of deductive experiments. Axiomatic systems in geometry, algebra, and logic exhibit this property. The tremendous accomplishments of modern mathematics follow (although not effortlessly) from a very limited set of carefully considered initial assumptions. Similar examples exist outside of conceptual systems. The quantum theory, which was invented to explain anomalies in black-body radiation, has found innumerable applications to phenomena as diverse as photoelectricity and solar spectrography, and is in fact a key element in all modern physics. The social sciences and the planning-engineering professions are somewhat lacking in such key theories, but some nominations could be made. These might include, for example, marginal substitution and general equilibrium concepts from economics, and applications of general systems theory. In any event, while it is somewhat difficult to define the process by which a theorist comes upon a fruitful theory with many applications while attempting to solve a more limited and more particular problem, it is apparent that solutions of this type are unusually desirable. At the least, theory builders should have this objective in view, especially since this state of mind leads to stripping any problem to its most essential elements, and thus may simplify as well as lead to greater generalities.

Simplicity is in fact an ancient and honored criterion for choosing between otherwise equipotent theories. Occam's Razor, named after a fourteenth century English philosopher, dictates that theories should contain the minimum possible number of hypotheses, and many of the more durable theories elegantly exhibit this characteristic. Because of the large number of conditions, relations, and variables which occur in social science research, this condition is difficult to meet here and frequently conflicts with requirements of realism and testability. It is nevertheless a desirable characteristic, not only for reasons of elegance, economy, and generality, but also for practical reasons which will be discussed later. Here there is a special pitfall which social science researchers can dig for themselves by the use of modern computational techniques. It has been suggested that, had computers been available at the time of Copernicus, the ease of computation of epicycles might have removed the practical difficulties which led to the construction of the elegant and economical heliocentric theory and the Newtonian theory of celestial mechanics. By the use of computers in the descriptive system of Ptolemy, navigational tables could have been constructed to any required degree of accuracy, and the practical impetus for the Copernican and subsequent revolutions would have been removed. I feel that we fall into the same trap when, as with the introduction of K-factors into the gravity model, we constantly patch up a nonexistent or inadequate theory with computational amendments.

A requirement which follows from testability and which is necessary to it is the requirement of manipulability. The experimental method in the social sciences is, as we have said, forced to rely on paper experiments, and for these our professions commonly talk about the use of models. To quote Harmer Davis, "A model is a smaller copy of the real thing, as the woman said about a model husband." This pointed definition does not permit us, however, to distinguish between a mathematical model and a simulation model on the one hand, nor between a simulation model and a theory on the other hand.

As we have emphasized, there are in principle distinct sources of an understanding of cause and effect in the real world and of formal representations of relationships in the world of mathematics and logic. Science is in many respects an effort to establish isomorphisms between these distinct realms. If we refer to a linear programming model of warehouse location, we are referring to just such an assertion about an isomorphism. We might then be correct in speaking of a mathematical model of warehouse location. Frequently, however, people speak of the linear programming model, and more generally of mathematical models in the abstract without relation to any particular real world phenomena. I would submit that this application of the word model is incorrect, though lamentably ineradicable, because the mathematical linear programming model is not a model or a smaller copy of anything.

The distinction between a theory and a simulation model is somewhat more subtle and difficult. A theory in fact could also be said to be a logical or mathematicl model of the phenomena to which it refers. It is smaller, it is a copy, and it is of the real thing. Yet this identification of a theory with a model somehow goes against the grain. On the basis of very serious consideration, I have redefined models as they are used in the simulation of social and economic events in a way which tends to provoke outraged reactions, but which I believe withstands serious examination and criticism: a model is an experimental design based on a theory. Let us examine the implications of this definition somewhat more carefully.

As is well known to workers in our fields, there are many theories which are testable in the sense that a critical experiment can be designed—but which remain untestable in the sense that the data requirements are for practical purposes excessive and involve presently unobservable variables, or perhaps most important, they cannot be cast in a form which will fit into a computer and run economically. These practical considerations do indeed provide a spur to all kinds of experimental ingenuity, and they should by no means dominate the process of theory construction.

In the process of developing a theory, there are many applications of experimental design in which the theorist must invoke models. First, in exploratory or inductive investigations, he is quite apt to use a severely truncated or patently inadequate experimental design such as a multiple regression model to explore relationships and to provide information as to the direction of his next transductive steps.

1. In a more developed form he will use a model more closely corresponding to theory inductively to establish the parameters of relationships.
2. He will use a model for testing in the deductive sense in order to determine the applicability of his theory under a wider range of conditions.
3. Used scientifically in a context of projections, the model will provide experimental evidence as to the consistency of the theory and possible inductive evidence as to the sensitivity of the real world to changes in conditions.

It may be a matter of scientific but not practical indifference to the scientist that the projective use of models also is important to decision-makers.

One may choose to make a distinction between the value of theory building and the value of experimental work with models, imputing a higher value to the first of these activities. However, in the tradition of British and American experimental science, the theorist usually has some responsibility for making feasible the tests of his ideas, and it is only the boldest and most brilliant innovator in pure theory who can expect others to accept a division of labor in which they will devise feasible tests for his impractical fomulations. It is this experimental difficulty which often leads to emphasis on the false dichotomy between theory and practice, which can only be overcome by a long-run view of the value of theory and by a nice sense of the potential contributions of new theories whose testing and application may appear outrageously difficult.

There are other criteria which may or may not be useful for the construction and selection of theories but which are frequently in the minds both of scientists and practical people. We have mentioned that for a variety of reasons a theory may not correspond directly with intuitive and popular ideas about the nature of reality. In this case, the theorist or scientist may be accused of being unrealistic and may feel a social obligation to change or tone down his theory in the direction of greater realism. Such a compulsion grossly distorts the role of the scientist, which is to identify a genuine isomorphism between the behavior of the real world and a set of mental constructs. Frequently he has to invent the mental constructs in order to disclose the isomorphism. Many of the most pregnant ideas of the physical and biological sciences, such as the quantum theory, the theory of relativity, or the independence of heredity from environment, run counter to widely held and deeply rooted popular ideas. Yet, without the discovery of these theories and their application to everyday life, the world would have given up a great deal of progress. A search for naive realism is counterproductive in science.

Frequently even though a naive demand for realism may be abandoned, the critics of science will continue to take refuge in an unthinking insistence on comprehensibility. In the field of social sciences, this insistence is based on two circumstances. First, every critic is a member of society, a user of cities, and a participant in the political process. Hence he feels intuitively that by virtue of this special status he and most other informed people ought to be able to understand directly all of the theories which purport to define the operations of society, of cities, and of politics. In my view, it would be equally ridiculous to say that because we are all made of protein, we should understand at a glance the theories of molecular biology. A second circumstance resides in the fact that a great deal of social science research is conducted in such a way that the scientists are close to the administrators, the administrators are close to the decision-makers, and the decision-makers are close to the voters, with no clear separation of function. Because of the personal and normative nature of the communication between these groups, each link in the chain feels that he ought to know all about what the adjacent link is doing. We may contrast the somewhat more impersonal relationships which govern research and development in industry. The laboratory scientist may understand solid-state physics in detail. The corporation executive will understand the main directions of this research and its potentialities. The sales department understands the capability of the resultant product, and the customer chooses in the market place between the products of competing technologies and competing companies. The man in the street could not care less about the crucial role of, say, quantum mechanics in the production of his transistor radio. Probably when social science theories produce as effective results as quantum mechanics, the administrators, policy makers, and voters will be less inclined to ask questions and more inclined to judge by results.

A possible requirement for theory which requires brief mention is more likely to be generated by the scientist than by the layman. As a result of the complexity of social phenomena which requires holding other things constant, and as a result of the drive for generalization which is inherent in theory building, there is a considerable drive to create theories which are comprehensive. This drive encounters resistance on two fronts. A comprehensive theory may in certain cases become so general as to say nothing about everything. Even if this is not the case, the more comprehensive theories may be the most difficult to manipulate for purposes of testing and application. An important part of theory building is therefore a nice sense of discrimination as to when comprehensive theories are necessary and when they may be appropriately avoided by discretion in the subdivision of the problem into manageable parts. In policy-related sciences improper subdivision of the policy-making problems may result in suboptimization, but a subdivision of the problems of the real world and its functioning for purposes of study need not entail this danger.

In the preceding sections of this discussion, I have developed my ideas with regard to the scientific construction of theory, mainly with respect to the problems of simulating events in the real world of mass behavior, in the use of transportation facilities, and in the choice of locations, even though this concern has been in the main implicit rather than explicit. There are two other areas related to public decision-making in which theories of a different kind will have to be developed. Transportation and planning literature already recognizes the need for the development of more general theories of decision-making. In crude terms, the questions to be answered by such theories are: what are we planning for; what trade-offs are involved in the public decision process; and what values will our plans satisfy? In more sophisticated terms we are dealing with difficult problems of public discount rates, collective consumption, spillovers and externalities, the aggregation of utilities, and the reconciliation of conflicting interests. It is hoped that the improvement of theories and models in this general area may be expected to reflect backward into the planmaking process so that sketch planning procedures are replaced by optimizing procedures, and optimizing is not limited to narrow engineering criteria but is extended to the most general of social objectives. I think it is also predictable that as we explore the problems of decision-making, planning, and optimizing more thoroughly, we will discover that there are ferocious computational problems which arise in the design process as a result of the huge combinatorial variety which exists in the possible combinations of policies and future conditions. Our fraternal theorists in the field of mathematical programming may be able to make contributions of a theoretical nature with practical applications which are related to the needs of decision-making. It is also probable that a clearer formulation of these needs will influence this development of what are essentially design models.

We have now reached the vantage point of a somewhat shaky and perhaps imperfect understanding of some of the processes of science, from which we may view the needs and accomplishments of experimental simulation of transportation systems and land use systems and the behavior of their users. I will not here belabor the point which is now becoming widely accepted in principle—that in many policy-making contexts we are dealing with these systems not independently, but as a part of the larger urban metropolitan system or regional system. I will emphasize the fact that most theories of locational behavior contain ideas about transportation costs and convenience, and consequently that locational models must contain as submodels some replication of the transportation aspects of the system. It will also prove useful in the discussion which follows to consider the salient features of all these problems together from the point of view of theory construction, drawing freely upon examples from any field wherever they may be appropriate.

The range of our interest in these phenomena covers a wide span from very large and complex total systems through subsystems which may be defined in engineering terms, in social and economic terms, or in spatial terms, down to the smallest elements of the system. These last may be mechanical components, but the greater interest attaches to decision units—a man driving a car, a family looking for a home, or a corporation deciding to build a new establishment. At each of these levels, different problems arise regarding the appropriate form and content of research.

The broadest view, of the overall system as a whole, is probably not in itself highly productive, but it is a starting point for certain applications of general systems theory which later affect our view of the components and the elements. General systems theory with respect to the total urban, metropolitan, or regional system will ultimately play a direct role in decision models. Meanwhile, it can be particularly useful in defining the appropriate limits of a system and in guiding the structuring of the problem in such a way that its decomposition into subproblems dealing with subsystems will entail a minimum of distortion. Up to now in both transportation and land use analysis these two problems have been approached largely by intuition and induction. I do not feel that the results have been seriously wrong, but a systematic and better informed approach might provide some surprises and prove a useful guide to research design.

With respect to subsystems properly defined and considered as systems in their own right, general systems theory may very well contribute powerful methods for dealing with system stability as a planning objective and with homeostatic or equilibrating tendencies within systems as handles for both planning and analysis. My own intuitive feeling is that concepts of equilibrium animate a great deal of research and theory in land use and transportation analysis, but that these concepts are inadequately explored and not sufficiently explicit. For example, transportation analysis and the assignment of traffic to networks with capacity restraints imply a whole pattern of equilibrating behavior on the part of individuals which may or may not lead to system equilibrium and may or may not be related to various forms of optimization. These questions have been very lightly explored by brute force iterative methods in modeling experiments, and their full implications remain to be seriously examined. In land use growth model simulations based on trend data, there is also a set of unexplored assumptions about tendencies to equilibrium. Whether such an equilibrium exists or ought to exist has in fact been very slightly examined in theory. Needless to say, one-shot sketch planning or design models and "instant cities" such as Ira S. Lowry's Model of a Metropolis (4) are constrained to use either simultaneous determination or optimizing, and it seems likely that the former method contains some optimizing assumptions in its behavioral parameters. More generally, I feel that land use behavior as well as land use system performance can hardly be explained without a consideration of land marked equilibirum and simultaneous determination—all of which pose major problems for system theory.

There are a number of interesting problems which arise out of the communication between subsystems and between elements and subsystems and out of the mechanisms by which equilibrating, disequilibrating, and determining forces are transmitted to and from decision units. The organs of the body communicate information leading to action by nerve impulses and those maintaining homeostasis by chemical messengers; what are the messengers in a large city or region? Many of these questions will arise again in the discussion of the behavior of decision units below, but there is some advantage in taking an overview in the context of systems. It is quite apparent that the generic name for these messengers will be information, and it seems quite likely that some gains for theoretical clarity will be achieved if a systematic application of communications theory can be made about the diffusion of information through and about the systems under study. The applicability of this concept is already apparent in the most elementary consideration of the stability of traffic flow systems, and these ideas can probably be extended to land use systems and larger transportation systems. Considered in the communications context, there is some merit in merging the study of decision units with a priori considerations from different disciplines as to what information is likely to be important and available. At one extreme this type of merger leads to a consideration of the individual's reaction to the visual environment as developed in studies by Lynch (5) and others. At a different extreme, economics suggests that prices are the messengers by which important economic information regarding, say, the housing market is transmitted. Between these extremes lie many combinations of phenomena which are observable, influential in behavior, and to some extent predictable as consequences of other developments.

The importance of prices as a messenger and of the allocation of money to different purposes (i. e., of economic behavior) in private decision-making is so great that it deserves special attention. It is a curious fact that in spite of this a priori importance of monetary phenomena, they have really received relatively little emphasis in transportation and land use planning and analysis. For somewhat understandable reasons, transportation planners have been reluctant to explore the importance of pricing policies in alternative transportation systems. Surely, however, this reluctance should not extend, as it frequently does, to the omission of cost factors and the exclusive emphasis on time-distance which is frequently found in network analysis, trip distribution, and even modal split. Fortunately, this default is not universal. In land use analysis, the problem is perhaps even more severe. Housing rents and values are the medium through which consumers communicate with each other their willingness or unwillingness to compete for space, and more commonly land prices are the medium of communication in the competition of residential, industrial, and public uses for land. Yet in the research field, housing value and land prices very seldom appear as variables. So pervasive is this omission that expensive and otherwise useful surveys of locational, social, economic, and housing variables by the Penn-Jersey Transportation Study and the Tri-State Transportation Committee are partly vitiated by the failure to inquire as to housing value or rent. It must be admitted that the collection of these data and especially of land value information in a research study is fraught with difficulty, but

I believe that there is a more serious reason why these values have been neglected in spite of strong theoretical reasons for their inclusion.

If values (prices) are made explanatory variables leading to changes in the behavior of decision units, then future applications of the same theory and its derivative models require that these values be projected under new circumstances. The theorist then faces an ugly dilemma. If he chooses to predict future prices by means of proxy variables, he must build a purely descriptive model for this purpose which contains no ideas about cause and effect; and this being the case, he might just as well have left prices out of the original analysis and included the same proxy variables, admitting from the outset that his theory was in part purely descriptive. If on the other hand he takes the importance of these economic variables seriously, he must face the difficulty of reconstructing a complete market through some form of simulation. This reconstruction is complicated by the existence of submarkets, institutional stickiness, imperfect dissemination of information, and probable lags in equilibrium. If economic considerations were largely peripheral to the theory of land use and transportation systems, there would be less objection to taking the easy way out of this dilemma. I believe, however, that these considerations are so central that economic models must in the future be added to the implementation of transportation and location theory at full scale. This approach will involve much deeper consideration of equilibrium tendencies than was previously suggested, and perhaps a much more serious look at some aspects of the behavior of decision units.

Before turning to a discussion of the theory of the behavior of decision units, I must emphasize a vital distinction between the study of that theory and its application. To a very large extent, the study of the behavior of decision units can be undertaken independently of the simulation of system and subsystem performance which has been the subject of the prior discussion. This is true because at the moment when we examine the actions of decision units, the systems in which they are embedded have already performed their functions, interacted with each other, and thereby generated the environmental conditions and information of which the decision unit has knowledge and on which it acts. In this analysis, the experimental approach consists of searching out instances in which the environment and its informational content differ significantly from other environments, or the decision units differ significantly from other decision units, so that the general application and fruitfulness of the theory may be examined. When, however, the behavior of decision units as understood on the basis of such an analysis is to be explored experimentally under changed assumptions as to policies and technology, an entirely new situation arises. We can no longer assume that various sets of decision-makers are independent of each other. Each reacts with the environment and creates changes which result in messages reaching other decision-makers. This interaction, which is irrelevant to some analyses, becomes critical in system simulation. I thus assume that system simulation and decision analysis interact strongly with each other and that each is necessary for the other. But as a matter of research emphasis, I would give short-term priority to system simulation on the grounds that relevant experiments to test our understanding of the behavior of decision units probably cannot be performed without it.

Engineers and planners are vitally concerned with the behavior of households and business establishments in making use of the transportation system and in making locational decisions. Such behavior is the source of transportation demand. Private decisions in respect of automobile ownership, location, and new construction in the aggregate greatly influence the development of cities and regions. Finally, I am sure that if we understood thoroughly the whole constellation of decisions made by individuals and firms, we could understand at the same time the extent to which various urban arrangements satisfy their needs. Such an understanding is a vital key to producing plans and policies which will best serve the public interest.

Some of the differences between practicing planners and engineers can be traced to their different approaches to decision-makers' needs and preferences. The planner typically approaches the problem from the viewpoint of normative standards of behavior and social welfare. This is in part based on notions of minimum socially acceptable levels of welfare and in part upon an emphasis on the externalities of individual behavior—i.e., on the effects of one's behavior on others. These notions are linked with strong ideas of social control. The engineering approach tends to be more adaptive. Individual behavior is regarded as being largely self-motivated and not widely amenable to control. In dealing with supposed patterns of behavior as necessary inputs to engineering estimates, the engineer approaches the problem with naive concepts of motivation and of measurement. Neither planners nor engineers are in general well trained in the intricate issues of choice behavior, and present-day economics, sociology, and psychology offer little which is of general applicability to the problems which they face. The following remarks are therefore observations on a dilemma which will ultimately be resolved only by training people and developing methods which embody a combination of all of these disciplines in a new format.

The basic theory of choices by individual decision units deals in terms of alternatives and trade-offs, yet if we examine transportation and locational theories, or models, we find that these trade-offs are very deeply buried, if indeed they may be presumed to have been considered seriously at all. Since the same thing is true of econometric models in related fields, this is not a particularly telling criticism in terms of past performance, but it is clear that it may constitute a barrier which will have to be removed before a great deal of progress can be made.

Much of the difficulty concerns observation and measurement, and perhaps this may best be illustrated with reference to the theory of industrial and commercial location. Industrial location in particular has long been very carefully studied by locational economists and regional scientists, and interregional locational theory is a particularly well developed field. In this location theory three factors are particularly important: internal economies of scale which depend on the size of establishment; external economies of scale or agglomeration economies which depend on the sizes of the geographical assemblages of activities in which the establishment is located; and locational costs which depend mainly on the cost of land and the costs of interaction. In the complicated urban metropolitan scene, these economic variables turn out to be very difficult to define, more difficult to measure, and still more difficult to value. While it may be well known, for example, that the garment industry has large agglomeration economies and is sensitive to its accessibility to a particular labor force and to the cost of industrial space, these variables and their relationships are not well defined. The interaction requirements of offices and the agglomeration economies of retail trade establishments are also imperfectly understood. While these ideas enlighten a good deal of research design, anyone who has tried to set up an industrial or commercial survey knows that it is very difficult to tie them down specifically. Because of this situation and for allied reasons, it is beginning to appear that in spite of the much more sophisticated work over many decades in industrial location, the problems of residential location are more tractable and amenable to sound solution.

In the area of consumer behavior, some difficulty is introduced by the fact that certain decisions are made by individuals, others by households, and still others by individuals in a household context. These difficulties must be faced in research design, but they are relatively minor compared with other more obvious problems. One which has been both recognized and ignored (often simultaneously) is that of aggregation. Some researchers, perhaps moved by data difficulties, are inclined to deal with the means and medians of areal aggregations of data. This method of work is almost enforced by the form of the availability of published census data in certain cases. There is clearly here a latent conception that area averages represent some sort of aggregation of behaviors, but the implicit rules of this aggregation are not explored, and frequently the assumed behaviors are not fully defined. The gravity model of trip distribution clearly takes this approach at a descriptive level, while multiple regression models of modal split may be but a step closer to postulating real cause and effect. The Schneider model of trip distribution postulates more explicit behavioral patterns and works with areal aggregate data. In practice, however, this model reveals unexplained variations in decision behavior because it requires an area-specific determination of the proportions of long and short trips. This specification amounts to a statement that the behavior in the model is incompletely defined.

Those who avoid the implicit assumptions of working with grouped area data by using individual or household observations encounter another level of difficulty which helps to elucidate the problem of aggregation. Behavioral models of individual and household choice invariably produce tests in which only a small part of choice behavior is adequately explained. Typical coefficients of determination are in the range of 0.15 to 0.30. While this may mean in some cases that the models and theories employed are inadequate, it is more likely to imply that behavior is influenced by more or less unobservable cultural and psychological factors which (at least at any one time) may be statistically distributed in the population. The delicate problem of research design is to know when to stop trying to identify these factors and when to introduce assumptions about the statistical nature of their distribution in the population. After the first recourse is exhausted or while it is being further developed, it is apparent that the nature of the assumptions about the statistical distributions of behavior around their observed statistical means may strongly affect the characteristics of their aggregation. As a simple example, I have demonstrated elsewhere that if Schneider's L parameter (7) is assumed to have a certain statistical distribution rather than being fixed, his model converts readily to a modified gravity model or to a combined model. Certainly considerable statistical expertise will be required to explore this problem further.

One of the more subtle and neglected aspects of the analysis of decision units is the role of the history of the unit in its behavior. To some extent, the history of certain units is implicit in their state description—a family head aged twenty is probably recently married. But other and more subtle historical aspects may be overlooked. It is quite clear, for instance, that the history of industrial establishments is related to their tendencies to relocate, and the ethnic background of many population groups is related to their choice of residence. It has even been reasonably suggested that consumer choice of mode of travel is related to the individual's history in learning to drive. These historical aspects of the behavior of decision units have two very important relationships with more general systems analysis. The historical aspects of decisions are closely tied to the extent of the lags in movement toward system equilibrium, and only systems in which the history of decision units is unimportant will rapidly achieve equilibrium. At the same time, the introduction of these histories is a means of dealing quite explicitly with trend data, without at the same time building into the theory an assumption that trends will indefinitely continue. It should be apparent that this historical approach does not lend itself to easy application to aggregate data, at least in analysis. And if the histories to be considered become very complex, then Monte Carlo methods are almost required for any projection simulations.

In the light of the foregoing incomplete review, we may justifiably conclude that a theoretically sound and scientific approach to systems simulation of transportation and land use will require a great deal of rethinking of our theory of decision-units' behavior.

Let us now take a brief final view of the workaday implications of the type of program that has been sketched. The essential elements of this approach are six in number.

1. Since sound theory has so much to offer for practical progress, the work should be organized on a scientific rather than a mission-oriented or technological basis. We would thus also avoid the dangers implicit in harnessing these activities to suboptimal policies, and rely on the social and policy motivations of the scientists to maintain a well-directed drive toward ultimate application.
2. We would view these problems as related to certain real world systems and would deepen our efforts to achieve successful theories of the operation of those systems.
3. We would give appropriate recognition of the need for the study of the behavior of decision units in the context of larger systems which create their environment.
4. We would give explicit recognition to the theoretical problem of communication between the systems, subsystems, and decision units.
5. We would recognize that the scope of these investigations will require the unification of parts of different disciplines in institutions and in individuals.

6. We would recognize that in a specially defined sense this work is experimental in the best traditions of experimental science and that the experimental method will require special conditions for success.

It would seem that the scale of these problems and their importance for long-term policy development tends to argue against scattered research in connection with specific projects. Such projects in any case tend to impose their mission orientation on individual researchers. The resulting tension between the desire of the researcher to satisfy his scientific conscience and the desire of the management to get the job done sometimes borders on the tragic, or the comic. In any event, the problems are of general and national significance, and if worthy of consideration should not be charged to local or to special-purpose studies. It is also apparent that the variety of ability and knowledge required for an assault on these problems can rarely be assembled even in a large study of an ad hoc nature. Consequently, many such studies are repeating the work and perpetuating the errors of other studies for lack of resources to go further and try new methods. Finally, there are serious difficulties of communication within this scientific community which result from the excessive fragmentation of effort.

Special attention should be devoted to the requirements of the experimental method in this field. Consider, for example, designing a laboratory for social, engineering, and planning research. Instead of white mice, our experimental material is extensive data about metropolitan areas and regions. These data must meet certain rigorous standards and be well organized and accessible. Our main experimental tool is probably the computer, but this will include the software or operating programs which embody many or most of the elementary processes of simulation and analysis which we have discovered. Our experimental design is a model or group of models based on theory and using experimental material (data) and experimental tools (computers and software). In any good experimental design, we are apt to discover that some special-purpose tools will have to be made—in this case new programs will have to be written and in some cases new data collected. The essential aim of an experiment will be to make critical tests of theories by good experimental design and thus to decide, for example, on a clear definition of the relative merits of the gravity model, the Schneider model, the Tomazinis model, and the Harris model of trip distribution (2, 7, 8). The essential ingredient for progress in addition to all the niceties so far discussed is quick turnaround so that experiments may be rapidly executed once they are designed. I would estimate that under current conditions, with practically no standing stock of data and widely diversified programs, the turnaround time on experimental work of this type is roughly three to five years. This time should be reduced by a factor of three or more, and the content of the experiments should be far more conclusive than it is today.

I believe that these standards, both of theoretical excellence and mechanical performance, are achievable and that if achieved they will have tremendous payoffs in improved planning.

REFERENCES

1. Carrothers, Gerald A. P. An Historical Review of the Gravity of Potential Concepts of Human Interaction. Jour. Amer. Inst. of Planners, Vol. 22, 1956.
2. Harris, Britton. A Note on the Probability of Interaction at a Distance. Jour. of Reg. Sci., Vol. 5, No. 2. In press.
3. Harris, Britton, ed. Urban Development Models: New Tools for Planning. Special Issue, Jour. Amer. Inst. of Planners., Vol. 31, No. 2, May 1965.
4. Lowry, I. S. A Model of Metropolis. Memorandum RM-4035-RC, The RAND Corp., Santa Monica, California, August 1964.
5. Lynch, Kevin. The Image of the City. Joint Center for Urban Studies, Cambridge, Massachusetts, 1960.
6. Poincaré, Henri. Science and Hypothesis. Dover Press, New York, 1952.
7. Schneider, Morton. Gravity Models and Trip Distribution Theory. Papers and Proceedings of the Reg. Sci. Assoc., Vol. 5, 1959.

8. Tomazinis, Anthony R. A New Trip Distribution Model. Paper #347, Highway Research Board, Washington, D. C., 1963.
9. Weinberg, Alvin. But Is the Teacher Also a Citizen? Science, Vol. 149, 601-6, August 6, 1965.

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.

A Recursive Programming Theory of the Residential Land Development Process

KENNETH J. SCHLAGER
Southeastern Wisconsin Regional Planning Commission, Waukesha, Wisconsin

•INTEREST in the land development process has increased in recent years with the growth of urban and regional planning. Planning, by reason of its primary objectives, must influence the land development pattern. An understanding of the land development process is a necessary prerequisite for such an influence since it is difficult to control a process the characteristics of which are not understood.

Attempts to understand the land development process have generally taken form in mathematical models which have been formulated to represent and to simulate over time this process. Although such models have taken a wide variety of forms, and a significant amount of experimentation using digital computers is currently under way, the theoretical basis for these models has often been elusive, and the need for a comprehensive rationale has become increasingly apparent.

Any model is an artificial representation of the real world. This artificiality is as true of a physical model of a ship as it is of a mathematical model of a transportation system. Both represent the real world in an imperfect way.

Any mathematical model in representing the real world implies a theory of behavior of the real world. Such a model is comprised of variables and relationships, and the relationships between the selected variables in the model automatically imply a theoretical construct of the manner in which the system modeled in the real world operates. In most cases, the theories behind current land development models must be deduced from their variables and relationships since extensive discussion of the theoretical framework of the models is rare.

MACROSCOPIC AND MICROSCOPIC THEORIES

Examination of current models reveals two primary classes of land development theory that may be implied from the nature of the models themselves. The first class of models is macroscopic in outlook since it deals with aggregate variables and relationships. In this approach, the theory relates to an explanation of gross land development patterns in terms of variables such as accessibility to employment and shopping centers.

An example of such a model is the gravity model, originally applied to retail trade (1), which distributes residential development in direct proportion to the size of the center and some inverse power of the distance (or travel time) from the center. This gravity model in various forms has been the best known of the macroscopic class of models and typifies the macroscopic approach in its emphasis on abstract "forces" as the cause of developing land patterns. Because of the obvious parallel of this force concept to the physical sciences, macroscopic theories are usually classified as part of the field of social physics. Macroscopic theory is closely related to geography and demography and macroscopic economics in its emphasis on aggregate relationships.

A second class of models is formulated at a more microscopic level of relationships that attempts to describe the actual decision processes of individuals influential in land development. Such models explicitly consider the goals, information availability and choice selection patterns of decision-makers.

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

Examples of this second class of model are rare in land use development but quite common in microscopic economics where the classic theory of the firm is based on the optimizing concept of the business entrepeneur. A contrasting but still microscopic view of the firm is represented by the recent research of Cyert and March (2) centering on a behavioral explanation of decision-making in a business firm. This behavioral theory of the firm and its related models are based on detailed examination of organizational goals, expectations and choice processes and is representative of the microscopic approach to model-building.

The differences between macroscopic and microscopic models are really more basic than the choice of a level of aggregation. Macroscopic theory emphasizes the omnipotent forces structuring the patterns of land development independent of the decisions of individual households or businessmen associated with land development. Such a viewpoint parallels certain laws of thermodynamics that state aggregate relationships that prevail in the midst of the apparent random motion of the millions of molecules making up the gas under study. In such a framework, study of the movement patterns of the individual molecules is futile since such study does not reveal the important aggregate structure of the process.

Microscopic theory employs an opposite viewpoint that considers the key relationships to be at the level of the individual action event. To understand the process, intensive study of these individual activities is vital to a theoretical representation of the system.

Experience in land development theories and models has been too limited for even a preliminary evaluation of the merits of the macroscopic and microscopic approaches to land development modeling. Realizing this current state of the art in the field, this paper presents a particular microscopic theory of land development which has provided the basis for an operational land use simulation model in current use and offers, at the very least, a point of departure for further development of microscopic theory.

A DECISION THEORY APPROACH

A decision theoretic approach to a microscopic theory of land development involves the following sequence of research activity:

1. The identification of the key decision-makers in the land development process;
2. An examination of the goals of these decision-makers that provides a measure of the values of alternative courses of action;
3. A description of the type and quality of information available to these decision-makers and the procedures used in translating this information into expectations associated with alternative decisions; and
4. An analysis of the logic used by these decision-makers in relating goals and expectations in the selection of a desired course of action.

Successful completion of the above sequence provides the basis for both a theory of land development and a framework for a model to realize this theory in the simulation of the land development process.

DECISION-MAKERS IN RESIDENTIAL LAND DEVELOPMENT

Since residential land development occurs through a series of transactions in an economic market, the primary decision-makers must be the principals in these transactions: the buyer and the seller. In residential land-housing transactions the buyer is either a household purchasing shelter or a business investor securing a rental property. The seller is either a land developer (assuming that some development of raw land is necessary before an area is suitable for residential housing) or a builder who also serves as an intermediary in the sale of developed land.

Other persons and institutions are influential in land development even though they are not direct participants in the final transaction. Financial institutions such as commercial banks, savings and loan associations and insurance companies, play a vital role since they supply money in the form of loans to both the land developer seller and the household (or business) buyer. Indeed, such financial institutions are often in a decisive position in such transactions which could not take place without their cooperation. This influence is particularly important in a tight money market where mortgage money is in short supply. It is less influential when the money supply is adequate since competitive forces in the financial industry effectively neutralize money constraints on development.

The role of government (federal, state and local) on land development is secondary but powerful. Governmental influence is a necessary prerequisite for planning since urban plans could never be implemented if there were no governmental influence on the emerging land patterns. The governmental role is two-sided. Its most direct influence results from public works programs which provide the primary transportation, water, sewer and educational facilities necessary for any residential development. Whether the construction of such facilities tends to lead or lag land development is, of course, a key test of the real influence of these facilities.

A secondary role of government relates to the application of certain legal controls on land development such as subdivision regulations and zoning. Subdivision regulations affect the costs of the land developer by requiring him to provide certain improvements at his own expense. Zoning provides a constraint on certain transactions to the extent that it is able to withstand market pressures.

An important but more indirect influence on residential land development is exerted by private business institutions providing shopping facilities and opportunities for employment. Here the key factor is accessibility since such facilities tend to serve a limited retail or labor market area. Again, the time phasing of such development is an important consideration. If such facilities tend to follow residential development, they only reinforce existing trends. If they lead residential development into new areas, their influence can be crucial.

Since the planning, by its very nature, is expected to recommend plans and policies for government, the appropriate theory of land development for planning is one that is able to explain and predict the behavior of private (nongovernmental) decision-makers under a variety of governmental policies. The theory must be focused on the behavior of the private decision-makers in land development. Public decision-making is to be determined, not explained or predicted, based on its effect on private decision-making and the desirability of the resulting land pattern.

ECONOMIC AND BEHAVIORAL THEORIES OF RESIDENTIAL LAND DEVELOPMENT

With the decision-makers identified, the sequence of theory formulation now calls for the examination of the goals, the description of the information available to and the decision logic used by these decision makers. Two basic frameworks for a residential land development theory, the economic and the behavioral, are analyzed. Both of these frameworks involve assumptions relating to goals, information and decision logic.

A pure economic theory of land development involves the classic assumptions of the theory of the firm. The goals of land development are maximum profit for the developer and maximum utility for the household or business buyer. Perfect information relative to all variables affecting these goals is assumed so that the seller or buyer is aware of the value of each alternative. Finally, the choice of an alternative is based on maximum profit or maximum utility. Ignoring the formidable problems associated with the definition of the utility functions of the buyer and concentrating our attention on the more tangible profitability goals of the businessman-developer, the pure economic theory does offer a well-defined and logical approach to land development decision-making. Whether this well-defined and logical approach bears any resemblance to real life activity is, of course, a very different question.

The general criticisms originally voiced against the economic theory of the firm apply equally to a pure economic theory of land development. These criticisms encompass all aspects of the decision-making problem—goals, information and decision logic. The concept of a single economic goal, maximum profitability, as the guideline for all land development decisions, has been considered naive and oversimplified.

Human motivation, it is said, is much more complex, and the myth of the economic man has long since ceased to bear any relation to reality.

The criticism continues with questioning about the informational assumptions. Perfect information concerning the set of alternative decisions is rarely present in a real life situation. Such information is usually either too expensive to collect or even impossible to obtain at any price.

The final assumption, given the goals and information, is criticized in a less fundamental way. The sheer size of the computational and logical effort required in many decisions to select an optimal alternative tends to cast doubt on its practicality. Even with clear-cut goals and perfect information, the mathematical effort needed to uncover the true optimum is often staggering. That such a computational effort could be performed subconsciously by all land developers is not easily reconciled with practical experience.

Behavioral theories of the firm stress the existence of multiple goals, imperfect information and non-optimal decision logic. A land developer may well have goals relating to volume of business as well as the profitability of this business. The information available to the developer is imperfect particularly since much of this information is based on forecasts of future demand for land. Selection of the final alternative is influenced as much by the desire to avoid radical change as it is to select an optimum even within the limits of imperfect information. Some of this natural conservatism results from the realization that the information, particularly the forecasts, are imperfect. In such an atmosphere of uncertainty the desire to make haste slowly is a natural one.

While the behavioral theory is an improvement in its added degree of realism, it quickly tends to become vague and subjective, and it lacks the operational simplicity of the economic theory.

What is needed, to be sure, is a theory combining the best features of both the economic and the behavioral viewpoints. To the simplicity and common sense appeal of the economic theory, must be added the behavioral limitations brought about by the uncertainty of the forecasts of future land requirements and the natural resistance to radical change in land development trends. Such a composite theory will now be developed within the framework of a new modeling technique, recursive programming.

RECURSIVE PROGRAMMING

A Decision Framework for a Residential Land Development Theory

Recursive programming (3) is a decision simulation technique that provides for ". . . optimized decision-making over a limited time horizon on the basis of knowledge gained from past experience." (4) The limited time horizon is a direct result of the uncertainty of forecasting future land requirements. This uncertainty forces the decision-maker to base his decisions on short-term forecasts of relatively higher accuracy. Past experience continually updates these short-term forecasts over time. This same experience also alerts the land development decision-maker to the uncertainty of even these short-term forecasts and thereby discourages rapid changes in land development trends.

The analytical nature of recursive programming in its linear form is best described in the words of its originator (4):

> Recursive linear programming is a sequence of linear programming problems in which the objective function, constraint matrix and/or the right hand side parameters depend upon the primal and/or dual solution variables of the preceding linear programming problems in the sequence.

Recursive programming is a combination then of recursive simulation over time and linear programming. Each linear programming solution provides parameters for the next linear programming problem in the recursive time sequence. The recursive programming relationships for a combined economic-behavioral theory of residential land development take the following form:

$$\begin{aligned} f(x) &= \min cx(t) \\ A'x(t) &= \hat{b}'(t) \\ A''x(t) &\leq (1 + B)(t - 1) \\ A'''x(t) &\leq b'' \end{aligned}$$

where cx (t) represents the cost function minimized by the land developer with the row vector c representing the costs of developing land in different areas in the region and the x (t) column vector representing the amount of land developed in each area. The second relationship represents the forecast of land requirements with the total land developed in each period equal to the forecast of land requirements, $\hat{b}'(t)$, for that period. The forecast $\hat{b}'(t)$ for each type of residential land is based upon actual demand in previous periods.

$$\hat{b}'(t) = \lambda b'(t - 1) + \lambda^2 b'(t - 2) + \dots + \lambda^n b'(t - n)$$

where

$$\lambda \leq 1 \text{ and } \sum_{i=1}^{n} \lambda^i = 1$$

The land demand in each past period is weighted to arrive at a forecast of future land requirements in the new time horizon. More recent periods are weighted more heavily depending on the value of the λ parameter. Higher values of λ place more weight on the recent past. Lower values conversely are more affected by the distant past. This simple form of time series extrapolation based on a "smoothing" of historical experience was extended to a more elaborate model based on household types to be described.

In addition to fulfilling the forecasts of land demand, short-term land development optimization is also restricted by the third set of relationships which tend to limit rapid changes in land development. In effect, the third class of relationship states that the new land developed in any area in any period cannot exceed some proportion of the previous development in that area. The B parameter determines the rate of development in an area that has favorable development in that area. The B parameter determines the rate of development in an area that has favorable development costs compared to rival areas in the same region. A higher value of B (B will always be less than one) will permit more rapid development of favorable areas and indicate a bolder and more risk-taking attitude on the part of developers. A lower value of B will restrict rapid development in the most economic areas and transfer development to less efficient areas. To provide for the early development in a given area, the initial condition of the constraint is established at a small value.

The third set of relationships embraces all constraints on the solutions that do not change with time. These might include land availability (capacity) restrictions in each zonal area and accessibility relationships to employment and shopping facilities in other zones. Other technical and behavioral constraints may be desirable in certain applications.

Goals, Expectations and Choice

How has the recursive programming framework provided answers to the questions of the decision-maker's goals, expectations and decision logic? The goals of the land developer decision-maker are encompassed in the recursive programming objective function. Land development is viewed as a business and the land developer as a businessman seeking to advance his own fortunes. This aspect of the theory in its assumption of the goal of maximum profitability (or minimum costs) is identical to the classic economic theory of the firm. Although it would be possible to use the developer's estimate of costs rather than the true costs, this approach was not taken here, and the costs are assumed known. The real uncertainty of these costs, however, is reflected in the restraining effect of the B parameters.

Information availability and accuracy together with the expectations associated with alternative decisions are reflected in the uncertain forecast and the resistance to radical changes provided for in the model. The informational assumptions of the model are behavioral in form and depart radically from the perfect information assumptions of the economic theory of the firm.

Within the limitations of less-than-perfect forecasts and general distrust of data accuracy, the decision logic is optimal in structure in that it selects the lowest cost combination of land development. The optimum selected, and the resulting land development sequence will differ significantly from a pure linear programming optimum given a perfect forecast of land requirements.

The composite theory reflected in the recursive programming construct is a combination of the economic and behavioral approaches to decision theory. The economic goal of profitability, as reflected in minimal costs, remains. The behavioral concepts of limited information, however, are the key to the operation of the model.

The Household Decision

The recursive programming relationships previously discussed have stressed the role of the land developer as a decision-maker. The developer has been placed in the key role of determining the location of new land development. How do the goals, expectations and decision logic of the household home buyer or renter relate to the recursive programming theoretical structure?

Inherent in the theory is the concept that the land developer determines the location of new land development based on the demand generated by the desires of households. The actual land demand in each period $b'(t)$ is determined by the households in the market for new housing. These households include households relocating within the region, new household formations, and inmigrating households. The $b'(t)$ vector of land demand in each period provides for a demand value for each type of residential land and is the result of a transformation of the household vector (number of households, by type, seeking housing) into the land demand vector. In this way, the household determines the demand for each type of residential land which is then distributed spatially by the land developer. The household population is classified into types because of the different land needs and desires of different age, income, and family-size groups within the household sector.

Household goals and decision logic also reflect in the third set of recursive programming relationships relating to accessibility constraints. Household goals relating to accessibility restrict the area of choice to those areas accessible to employment and shopping. Other household contraints on land use relationships may also be reflected in this set.

Government Plans and Controls

The third major participant in the land development process, the government, is known to exert a major influence on the emerging land use pattern. This influence is represented in all components of the recursive programming framework. The cost function is influenced by the legal codes (such as subdivision regulations) which usually determine the proportion of land development costs to be borne by the private developer. Public works programs (such as highway and transit systems) determine the accessibility of potential residential areas to employment, shopping, and recreational facilities. Accessibility constraints are represented in the third set of constraint relationships in the recursive programming model. These accessibility constraints prevent land development in areas outside of specified travel-time limits from employment and commercial centers. Other public works (for example, sewer systems in combination with the legal framework) provide other constraints on certain types of residential development in some areas.

Since the primary purpose of a theory of land development for urban and regional planning is to provide the knowledge necessary to implement certain plan designs, the usual approach to application of the theory will be to test the effects of government policies (such as those related to land development costs) and public works on the course of land development. Experimental test and modification of these policies allows for the determination of a set of policies consistent with the target plan design.

MODEL IMPLEMENTATION

Variables and Relationships

The implementation of the recursive programming theory of land development involves the detailed specification of a set of variables and the estimation of a set of associated parameters. This implementation is best presented in terms of the general matrix of variables and parameters previously discussed. If each element of each matrix and vector can be identified as to physical meaning, the application will become more real and vital. The variable vector and each parameter vector and matrix are presented sequentially.

1. The x variable vector. The elements of this vector represent the land developed by residential-density class in each area of the zonal grid defined for the model. Because the type of soil provides the basis for the estimation of costs, subareas within each area will exist depending on the number of types of soil in that area. The size of the vector will depend on the number of soil groups in all of the areas, the number of zonal areas, and the number of residential-density classes. If each of 50 zonal areas had 3 soil types and 3 density classes represented, there would be 450 elements in the x vector.
2. The c cost parameter vector. The elements of this vector represent the cost of developing an acre of land of a given density on a given soil type in a specified zonal area. The number of elements will correspond to the x variable vector since a cost element is required for each x vector element.
3. The A' design standards parameter matrix. The elements of this matrix represent the amount of primary and service land required to provide an acre of residential land of a given residential-density class. Service land requirements include land for streets, local shopping centers, local school sites, and local recreational areas. An element value of 1.5 would indicate that one-half acre of service land is needed for each acre of residential land. The number of rows in this matrix depends on the number of residential-density classes. The number of columns depends on the size of the x variable vector.
4. The $\hat{b}'(t)$ land demand forecast exogenous variable vector. The elements of this vector represent the forecasts of regional demand for each residential density class. The size of the vector depends on the number of residential-density classes. This vector is exogenous to the main model and is externally programmed to change every recursive time interval. A number of methods might be used to prepare this forecast. In the current application, a land use demand forecasting submodel based on a household typology is used to generate the forecast. A transformation matrix is used to convert households by type into land demand by residential-density class.
5. The A'' behavioral constraint matrix. The elements of this matrix are all unity. The number of rows in the matrix depends on the product of the number of zonal areas and the number of residential-density classes. The number of columns correspond to the x variable vector.
6. The $(1 + B)(t - 1)$ variable vector. The elements of this vector represent the permissible development of each residential-density class in each zonal area based on some proportional increases over previous development. The value of the B coefficient depends on the behavioral resistance to change among local land developers. The number of elements is identical to the rows of the A'' matrix.
7. The A''' accessibility and zonal capacity matrix. A variety of constraint relationships is represented by this third matrix set. In the accessibility submatrix, all zonal areas accessible to a given shopping or employment center are represented by unity element values. Other zones are given zero element values. One row in the matrix will be required for each center. The zonal capacity submatrix is an extremely low-density matrix with a maximum of three non-zero elements (one each residential-density class) for each row of the matrix. This matrix and its associated right-hand

side constrain the total land developed to zonal land capacity. One row of the matrix is required for each zonal area.

8. The b'' accessibility and zonal capacity vector. In the accessibility subvector the element values represent the amount of residential land within the travel time service area that can be supported by each center. One element is needed for each employment and shopping center. The zonal capacity subvector elements represent the land capacity of that zone. One element is needed for each zone.

Parameter Estimation

Model parameters were estimated from sources independent of the land development history in the region. Regression analysis of variable histories was not used for parameter estimation. Such an approach allows for meaningful historical tests of the theory and model since parameters are estimated independently.

Cost parameters were based on engineering estimates of land development costs depending on the type of soil in the area. With this cost-soil type relationship established, the regional soil survey allowed for the estimation of development in all areas of the region. Raw land values were based on equalized assessed valuations of the land.

The design standards relating to the amount of service land required to support residential land developed were based on both historical ratios and normative design standards. The regional land use inventory was used to determine existing relationship between primary and service land in residential areas.

The behavioral B coefficients were the only parameters estimated by analysis of historical time series. Since these coefficients determine a nonlinear upperbound constraint, linear regression analysis is not suitable. A partially experimental method of nonlinear regression may be used to estimate these parameters.

Accessibility criteria for the household types represented in the model were determined from an analysis of the travel-time habits of sampled households. This analysis established the market areas of the shopping and employment centers.

Zonal capacities were determined from the "developable land" indicated in the land use inventory.

Early experimental experience with the model is more fully described in a technical report by the Southeastern Wisconsin Regional Planning Commission (5).

Plan Design and Public Policy

Residential land development theory must be related to the overall urban-regional planning problem in order to be meaningful and useful. The urban regional planning problem solution may be viewed as a three-stage sequence:

1. Determining the current state of the region;
2. Designing a plan for a desired future state of the region that satisfies the plan objectives; and
3. Developing public policies and programs necessary to transform the region from its current state into the desired future state.

Residential land development theory is related almost exclusively to the third stage of this planning sequence. Its relationship to the current state of the region is important only as a means to explain the historical decision-making that led to the current pattern. An understanding of this historical development, aside from its more academic interest, is important only insofar as it leads to a clearer understanding of the third stage relating to future land development.

The theory should not be used to design plans for the desired future state of the region unless these plans are intended only as a means of perpetuating existing trends. The most significant confusion concerning the application of land development theory seems to revolve about the approach to plan design. Since land development theory is essentially positivistic (what is) rather than normative (what should be) in nature, its use in plan design usually represents a perversion of the worst kind.

The final stage of urban-regional plan development is closely linked with an explicit or implicit theory of land development. The transition to a desired future land pattern involves a complex decision-making process only partially controlled by the planner. Attempts to influence or modify this process will be abortive unless they are based on a soundly conceived and experimentally verified residential land development theory. The recursive programming theory of residential land development may well provide at least a partial answer to this need.

REFERENCES

1. Reilly, W. J. Methods for the Study of Retail Relationships. Univ. of Texas Bull. No. 2944, Nov. 1929.
2. Cyert, Richard M., and March, James G. A Behavioral Theory of the Firm. Prentice-Hall, Englewood Cliffs, N.J., 1963.
3. Day, Richard H. Recursive Programming and Production Response. North Holland Publishing Company, Amsterdam, 1963.
4. Day, Richard H. Linear Programming and Related Computations. Dept. of Agriculture, Washington, D.C., 1964.
5. Technical Report No. 3. The Land Use Plan Design Model and the Land Use Simulation Model. Southeastern Wisconsin Regional Planning Commission, Waukesha, Wis., Jan., 1966.

Discussion of Land Use Forecasting Concepts

IRA S. LOWRY, The RAND Corporation. —The papers presented here form a rather neat progression from the general to the specific:

1. An essay on scientific method and its application to model-building;
2. An essay on some recurrent strategic problems in land use and transportation models; and
3. An exposition of an explicit model of residential land development.

I will comment on the first of these papers and on the last.

It is always a pleasure to read an essay by Britton Harris. This one is no exception. The ideas are well formulated, the illustrations are both lucid and engaging, and the prose style is first-rate. The essay is much too rich to be profitably summarized by a discussant, so I have picked out several of its themes for comment.

Simulation. I am mildly distressed by the use of the term "simulation" as it appears in Mr. Harris' title and repeatedly through the paper. He tells us immediately what he means by the term: ". . . the reproduction in some recognizable form of certain aspects of human behavior or of the performance of mechanical systems, or of a combination of these two." In other words, "modeling a system." Elsewhere, he speaks of a simulation model, a phrase surely redundant in view of his definition.

Such equivocal usage is unfortunately general among model-builders. The difficulty is that there also exists a particular class of quantitative models commonly called simulation models; these are most clearly distinguished from other quantitative and mathematical models of the same phenomenon by the method of solution. If we think of a set of models of a given phenomenon or real-world system as each identifiable with a set of simultaneous equations, the simulation model is one which is meant to be solved by numerical substitution rather than by analysis. The use of this brute-force technique is most often associated with large models whose logical closure is not self-evident, and also with Monte Carlo methods of generating inputs or intermediate values for certain of the variables.

While simulation models in this sense have in fact been applied to conventional forecasting problems, this is not, I submit, their proper métier. Simulation is a very useful if inelegant technique for exploring the homeostatic properties of a complex system, and is really most serviceable when applied to problems of system-design in which homeostasis is an important design criterion.

System Mechanics and Decision Theory. Among my colleagues, both in urban systems research and in economics, there is a persistent myth to the effect that each real-world system can in principle be reduced to a set of interacting elements which are themselves irreducible. These elements are called "decision units," and we are warned that unless we understand the behavior of the decision units, we can never successfully model the system. The first error in this argument is the notion that the decision-unit—say, a household, an automobile driver, a corporation—is irreducible. It is rather a system in its own right, also composed of elements, ad infinitum. The argument also ignores the fact that the only reasonably successful models so far produced for large human systems do not deal in decision units, but rather with much more aggregative systems-mechanics.

But the myth persists, and macro-analysis has acquired a bad name in the social sciences. When it comes to modeling urban systems, nearly everyone except J. Douglas Carroll apologizes for resorting to a macro-analytic approach which fails to exploit the infinite variety of behavioral possibilities at the level of decision-units.

I think that Mr. Harris gives an excellent account of the dilemma of urban systems research with respect to the level of analysis. The theme is perhaps the most persistent of his essay. If we approach the problem at the most aggregative level, the case material on which a model of systems mechanics may be based is "limited in extent, and experimental manipulation is both extremely slow and vastly expensive." On the other hand, if we tackle the decision units, we not only find that their behavior is more erratic than that of the system, but we are also unable to work out the mechanics of aggregation which will get us back to the system level. And it is, after all, the performance of the system in which we are primarily interested.

Heretofore, Harris has always given priority to micro-analytic research as the key to success in modeling urban systems. For the first time, to my notice, he has now reversed his priorities, arguing for the study of aggregative systems mechanics "on the grounds that the relevant experiments to test our understanding of the behavior of decision units can probably not be performed without it."

My own view is that success is where you find it. The construction of a truly comprehensive theory of the behavior of a large system may indeed lead to greater and greater disaggregation, further reduction of the irreducible. But building a model whose solution can be identified with a system's behavior within the environment of interest and with an acceptable level of verisimilitude may be much less taxing at higher levels of aggregation. As a matter of personal style, I tend to look to behavioral models of individual decision-units mostly for clues and hunches about the nature of highly aggregated system mechanics. The method is fallible, but I think, fruitful.

Mappings and Multiplications of Mappings. I like the description of the relationship between a theory and a phenomenon—that is between a theory and a real-world event or system of events. He calls this relationship a mapping, in which explicit correspondences are postulated between the elements of the theoretical system and the elements of the real-world system. The exactness of correspondence may vary; he distinguishes the crudest case as that of analogy, and he calls the complete one-to-one mapping of a theory onto a real-world system an isomorphism. In between, he recognizes explicit but imperfect mappings which he calls homomorphisms.

It is perhaps unfortunate in the essay that the passage just summarized does not lead directly to the discussion of the relationship of theories and models—unfortunate because this relationship also can be profitably conceived as a mapping. With some hesitation, Harris defines a model as "an experimental design based on a theory." We expect to find, then, correspondences between the elements of the theory and those of the model. There should also obviously be correspondences between the elements of a model and those of the real-world system which is being modeled. In experimental science, we typically thus have a multiplication of mappings—from theory to model to phenomenon. The direct mapping from theory to phenomenon is the device of the armchair or speculative scientist.

I think the point is important because it reveals two quite disjointed contexts for error. Our mapping of a theory into a model may be incorrect or ambiguous; and our mapping of the model into the real world may also be incorrect or ambiguous. On the other hand, this perspective also reveals the possibility that we may, by accident or intuition, have a good model along with a bad theory, or even without theory.

Let me add parenthetically that of my colleagues in the model-building field I can readily say that some are extremely proficient at mapping theories into models, and some are extremely proficient at mapping models into phenomena; but I do not number among my acquaintances anyone who is competent at both.

These last remarks provide a convenient point of departure for my reactions to Kenneth Schlager's paper. He also makes distinctions between a theory, a model, and a real-world phenomenon. He complains that "the theories behind current land development models must be deduced from their variables and relationships since extensive discussion of the theoretical framework is rare." And, in opposition to the dominant style of macro-models for land use prediction, he proposes a micro-model with an explicit basis in "decision theory."

> A decision theoretic approach to a microscopic theory of land development involves the following sequence of research activity:
>
> 1. The identification of the key decision-makers in the land development process;
> 2. An examination of the goals of these decision-makers that provides a measure of the values of alternative courses of action;
> 3. A description of the type and quality of information available to these decision-makers and the procedures used in translating this information into expectations associated with alternative decisions; and
> 4. An analysis of the logic used by these decision-makers in relating goals and expectations in the selection of a desired course of action.
>
> Successful completion of the above sequence provides the basis for both a theory of land development and a framework for a model to realize this theory in the simulation of the land development process.

Mr. Schlager goes through these steps in an exceedingly casual way for someone concerned with theoretical rigor. At all events, he comes out with three propositions about the behavior of land-developers:

1. They may have goals other than profits,
2. Their information is imperfect and they know it, and
3. Their decision-logic has a conservative bias.

His model, which is alleged to correspond to this theory, does not include a single variable which can be identified with those of the decision-paradigm for an individual land-developer. His model, as he explained, is a recursive linear program whose solution is that distribution of newly developed land which results in the minimum total cost of land development for a region, given aggregate demand forecasts for K types of housing. The solution is constrained by limits on zonal capacity, and by a binary rule governing colocation of households and shopping facilities. It is also constrained by upper bounds set on the rate of development in each zone. The solution vector of his model corresponds in the real world to the outcome of a market process involving many land-developers, many households, and assorted other parties. If this market process can be mapped onto the decision paradigm which Schlager presents as his "theory," it is only in the sense of analogy; certainly we are far from isomorphism.

What are, in fact, the correspondences? He emphasizes three which he considers strategic. The theoretical principle of multiple goals is represented in the model by the use of an objective function which minimizes cost rather than maximizes profit. The principle of imperfect information is represented by a demand forecast constructed from an exponentially-weighted time series of actual demands. The principle of conservatism in decision-making is represented by the local rate-of-development constraint which is based on past local rates of development.

If, in Schlager's mapping, the individual land-developer as decision-maker is blown up into a market system by scalar multiplication, the other decision-makers are very nearly relegated to the null space of the mapping. None of those listed by him as relevant—households, local governments, financial institutions—draws any information from the market, and there is no suggestion in the model that they too have decisions to make, alternatives, and trade-offs to consider in response to "moves" by the land-developer.

In fact, if I may put Mr. Schlager in bad company, he has done very much what I might have done under the same circumstances: pondered the paradigm of individual decision for hunches as to the kind of model that might work. But he has not in any sense established in his paper a close and rigorous correspondence between his model and his decision-theory.

I suppose it is apparent that I find fault with Schlager's mapping of his theory into his model. Let me now turn to his other mapping, from the model into the real world. This is the section of his paper where he identifies development costs with soil types, the demand for each type of housing with a linear combination of household types, accessibility of a zone with travel-time to the nearest shopping center, etc. The really crucial fact about this mapping is that he identifies the solution to his recursive linear program with a pattern of land development over time.

I think that there is a good chance that he will be able to show a reasonable correspondence between the solution to his model and the actual development pattern so long as he sticks to ex post facto prediction. Schlager makes the point that his parameters are estimated "from sources independent of the land development history of the region. . . . Such an approach allows for meaningful historical tests of the theory and model." But a few paragraphs further, he allows one exception, the B coefficients which are to be estimated from historical time series, apparently the same time series which are to provide the test of his model.

Because of his condensed notation, it is not clear to me whether B is a single number, a drift parameter whose value changes systematically with time, or a vector of numbers each specific to a local area within the region. If the last interpretation is correct, B is our old friend the k-factor under an assumed name. The expression containing B will very nearly be the error-term of the forecast which would have been made by a similar model lacking the B-type constraint.

If, on the other hand, B is a drift parameter or even a single number fitted from the time series which is then used to test the model's solution, it still offers some help in constraining the solution to match the historical pattern of development. He says that "the value of the B coefficient depends on behavioral resistance to change among local land developers." It is obvious that the variable thus defined is not measurable by any techniques at his disposal. Rather he will measure the serial correlation of development rates in local areas.

This is to my mind a better approach than the naive trend forecasting method not unknown in land-use planning. It is better because trend is here used as an upper bound rather than as a best estimate. The solution to Schlager's linear program also has certain cost and accessibility constraints which in principle make sense, although I am surprised to learn that in Wisconsin, soil type is so important a determinant of land-development costs.

In summary, I think Schlager's model might meet the tests of ex post facto prediction because (a) it contains some quite relevant variables, and (b) it is constrained by observed events of the back-casting period. I do not think, however, that it owes much to decision theory except inspiration, and on this ground I take exception to the lecture on theoretical rigor with which he introduces his model.

BRITTON HARRIS, Closure—One of the pleasant aspects of discourse with Mr. Lowry is the fact that our meetings frequently uncover rewarding and interesting differences, but seldom lead to sharp disagreements.

On the subject of simulation there could be endless debate, largely over semantics. I think the kernel of the problem is to be found in the fact that even many deterministic system descriptions resist analytic solution and hence, while analytic in nature, appear to be solved by simulation. As I tried to imply in my paper, some less clumsy methods should be sought for exploring not only the homeostatic, but also the more general dynamic properties of the systems in which we are interested. I agree that when we have analytic means of doing this, we may be forced to abandon the term "simulation," even though in the most literal etymological sense it may be correct.

I am inclined to disagree with Lowry that system elements occur in infinitely nested sets. In any case, however, there are levels of relevance. Systems theory surely can be used to demonstrate that a sulfur atom in my left toe-nail is hardly a suitable system for study in relation to metropolitan affairs. We can stop disaggregating our systems at the subsystem which we call a man and leave his internal functioning to medicine and psychology. Here, Lowry's attribution to me of a change in viewpoint is only partly correct. In the past while I may have preached micro-analysis, I have practiced macro-analysis. Conversely, while at present I acknowledge my emphasis on macro-analysis, I still maintain that we will ultimately make progress only by finding the appropriate bridge from man to the social system and the metropolitan system. Success may indeed be where you find it, but one of the purposes of theory is to suggest where to look. I still think we ought to look. I still think we ought to look under this bridge.

I like Lowry's comments on the multiplication of mappings, but I cannot go all the way with his method of dichotomizing theory and models. The interesting case which he identifies of a good model and a bad theory indeed gives us reason to ponder the whole situation. This is indeed armchair theorizing, but it does not correspond to my concept of theory at all.

KENNETH J. SCHLAGER, Closure—There are three basic objections raised by Jack Lowry that I will comment on in my rebuttal.

1. The first objection relates to the mapping of theory into models and the mapping of models into real-life.
2. The second relates to the estimation of the B factors in the model.
3. The third relates to the use of soil data for the estimation of residential land development costs.

In reference to the question about the relationship between the proposed model and the theory from which it is introduced, the concept of an analogy is characterized as crude. Lowry perhaps would like to see a direct, or more specifically, a detailed isomorphic correspondence between the elements of the theory and the elements of the model. If such a correspondence should take the form of an objective-by-objective, information-by-information, and decision-by-decision correspondence, he knows as well as I do that such a model is not now possible within the computational state of the art. The many types of households and land developers make such a model computationally, if not conceptually and statistically, an impossibility. But does that mean a microscopic model is impossible? Not at all, since the model proposed is microscopic in the sense of an analogy. The development of land in each zonal subdivision of the model is depicted as a "flow" controlled by the decisions of the land developer. No apologies need be made for analog models since almost all the advances in the application of science and technology to engineering, which is quite microscopic because engineering problems are rarely solved in the aggregate, have resulted from the use of hydraulic-electrical-mechanical analogies. To ignore the role of analog models is to ignore the whole history of the application of scientific theory. Land development models, as well as physical models in the microscopic sense, can benefit from the application of analogy. Isomorphic models may be of aesthetic interest to the theoretician but rarely have important meaning in application.

The estimation of the B factors used in the model was not the result of a time series analysis of the history of the system being studied, but the result of investigations of growth rates in areas at the subdivision level. For this reason the implication that the B factors were used as disguised k factors to improve the accuracy of the model is without foundation.

The objection that questions the relationship between land development costs and soil resources is a rather sad commentary on the tremendous gap that exists between the theoretical model builder and the realities he is trying to model. Anyone who has had any contact, direct or indirect, with land development will testify as to the importance of soils in development costs. These soil differences can change the cost of development by as much as 100 percent or more in a particular area. It is true that some areas will have greater variations in the types of soil than others and that the variations in development costs will fluctuate correspondingly. But the fact remains that insofar as soil differences do exist, they provide a sound method for the estimation of land development costs in any area.

In summary, the response to the three objections are:

1. An analogy may and usually does provide a basis for a microscopic model of a real-life phenomenon.
2. The B factors in the model were not estimated from historical output with which the model was being compared.
3. Soil resource data provide a sound and perhaps the only basis for the estimation of land development costs.

A Test of Some First Generation Residential Land Use Models

CARL N. SWERDLOFF and JOSEPH R. STOWERS, Highway Engineers, U.S. Bureau of Public Roads

•THIS paper reports on a comparative evaluation of five operational residential land use forecasting techniques, four of which have been previously used in urban transportation planning studies. These techniques are representative of the earliest of efforts in the development of operational urban activity simulation models and continue to serve, either in their original or in modified form, a great number of transportation planning organizations. Urban activity simulation models currently under development, while in most cases considerably more complex and, hopefully, more accurate, in many instances draw upon notions and fundamental concepts which either originated with or were adapted to these early techniques. Improvements being introduced in these later, second generation models include more complex statistical estimating procedures, the stratification of residential locators into several distinct groups, and the incorporation of behavioral relationships in the model formulation. These newer techniques may require several years of research, evaluation, and refinement before they become fully operational. Meanwhile, the less sophisticated approaches evaluated in this report should continue to be useful to smaller metropolitan areas lacking the resources for developmental research.

The primary objective of this project was to compare the relative accuracy of these approaches through a series of ex post facto tests, holding all conditions constant except the interrelationships among variables, so that differences in "forecasts" would be a function only of inherent differences in models.

There is a temptation to interpret a study of this nature as a contest of sorts and to turn to a table of results for the proclaimed "winner." Any such evaluation of the results is unwarranted for several reasons. First, the contestants are not all of the same class. Some are more truly "forecasts," and some are merely data fitting problems. The latter involve fitting different numbers of parameters. More information is used in some than in others. Perhaps most important, the results represent a sample of one, out of a rather large universe of possible test conditions. Entirely different results might occur in other cities, at other time periods, by other forecasters, working with other data problems.

GENERAL PROCEDURES

The five residential land use forecasting procedures are each variants of work done by others. The only innovations introduced here are the authors' simplifications and modifications to suit peculiar test conditions—apologies are made to the progenitors of these models for possible misrepresentations of their original work. In any realistic planning application, more care would necessarily be given to the particular forecasting tool used. Trends would be more carefully analyzed, the forecasters would be more familiar with the area, and output of models would be scrutinized in detail and modified as judgment indicated. In contrast, the authors have applied the models coldly and crudely, accepting the immediate output in an attempt to make objective comparisons.

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

The techniques used were (a) the density-saturation gradient method, (b) accessibility model, (c) regression, and (d and e) two intervening opportunity models.

The density-saturation gradient method (DSGM) is a simplification of the approach used by the Chicago Area Transportation Study (1, 2). Of the five techniques, the DSGM is least computer oriented, more demanding of subjective inputs, and therefore least suitable for objective comparison with other approaches, particularly when the forecasters are not intimately familiar with the area. The method is based essentially on the regularity of the decline in density and percent saturation with distance from the CBD, and the stability of these relationships through time.

The simple accessibility model is based upon the concept formulated by Walter Hansen (3, 4). Growth in a particular area is hypothesized to be related to two factors: the accessibility of the area to some regional activity distribution, and the amount of land available in the area for development. The accessibility of an area is an index representing the closeness of the area to all other activity in the region. All areas compete for the aggregate growth and share in proportion to their comparative accessibility positions weighted by their capability to accommodate development as measured by vacant, usable land.

The third method used in this study, multiple linear regression, is a popular approach because of its operational simplicity and ability to handle several variables (5, 6, 7). The proportion of total regional growth which locates in a particular area is assumed to be related to the magnitude of a number of variables which in some manner are measures of geographic desirability as viewed by those making the locational decision. The procedure is to determine those factors, and their weights, which in linear combination can be related to the amount of growth which has been observed to take place over a past time period. These factors (called independent variables) and their weights (regression coefficients), in linear combination (the regression equation) can then be applied to the individual analysis areas to forecast the magnitude of growth (the dependent variable).

Although more commonly applied to the problem of trip distribution, the intervening opportunities models can be used in simulating the distribution of urban activity. Two separate and distinct formulations were applied in this study, both based upon the general notion that the probability that an opportunity is accepted decreases as some function of the number of opportunities ranked closer to a central distributing point. The Stouffer formulation was originally applied to intra-urban migration (8). A related formulation has more recently been investigated as a trip distribution technique (9). Schneider's formulation was originally applied to trip distribution (10) and is currently being used in distributing urban activity (11, 12).

The test area used in this study was Greensboro, North Carolina. This city was chosen for a number of reasons. First and most important, a rather extensive information file on a small area basis for two time periods (1948 and 1960) was available. Secondly, it was felt that Greensboro was representative of the kind and size city for which forecasting techniques of the kind being examined would still be most appropriate after the development of more sophisticated models in the largest metropolitan areas.

The data for the study came from two major sources. The data obtained from the University of North Carolina contained a wide variety of information for the Greensboro area coded to 3,980 grid cells, each one 1000 ft square, for a circular area of about 7-mi radius. These data included quantitative measures of land use, population, residential density, proximity to various activities and to the CBD, and certain environmental measures (13). With certain exceptions, these data were available at the grid level for two time periods, 1948 and 1960.

The data supplied by Alan M. Voorhees and Associates included 1960 population, employment, accessibility to shopping, and accessibility to employment, for each of about 250 zones. These latter accessibility measures were computed from zone-to-zone traveltimes over the highway network.

A number of problems were encountered in combining the data from these two sources in a form suitable for testing of the models. Principal among these were the following.

1. The aggregation of grids to zones. Since it was felt desirable to work at a level of aggregation more typical of transportation studies, it was necessary to define new zone boundaries following grid lines approximating the irregular old zone boundaries. No important error was introduced since only accessibility scores from the original zone file were used in subsequent analyses—all extensive quantities used were grid aggregates.
2. Estimation of 1948 dwelling units. Consideration of all data sources and the purpose of the study led to the decision to use dwelling units as the item to be predicted. However, 1948 dwelling unit data were not directly available. Estimates were made and various checks applied by using 1948 land area, a 1948 USGS map for suburban areas, 1950 census block statistics for the central city (changes were not large for the inner area from 1948-1960), and the 1960 land area and dwelling unity densities.
3. Estimation of accessibility measures for 1960 for certain zones at the fringe. The area covered by the zone file did not extend to the boundaries of the grid coverage area in all directions. Rather than eliminate this area entirely, estimates of accessibility measures were made for about one-half of the outer ring of zones by examining contours of iso-accessibility lines, which follow fairly regular patterns in the fringe area.

MODEL DESCRIPTION AND METHODOLOGY

Density-Saturation Gradient Method

The DSGM is the least formally structured forecasting procedure of the five. No formal theoretical statements or mathematical hypotheses are required, although the staff of the Chicago Area Transportation Study have presented excellent conceptual explanations of their empirical findings and rationale for their projections (1). This theoretical development, however, is not essential to the purpose of this paper.

Before discussion of the actual application of the DSGM to the Greensboro area, mention should be made of certain reservations which existed prior to the testing. The only known previous application of this approach was for the Chicago area. There was some initial fear that the regularities in activity distribution about the central place, which is axiomatic to the method, would not be manifest for a city of the size of Greensboro. The declines in density and percent capacity result from the operation of the competitive land market, a mechanism which might not exert the dominating influence upon spatial organization in a city of Greensboro's size. It will be seen that these fears were unwarranted, and that in fact the distribution of residential activity was markedly structured about the CBD.

Two semi-independent forecasts were made using the DSGM in order to determine the sensitivity of the results to variations in the critical assumptions made. A principal distinction was that the first trial was made using air-line distance from the high value corner (HVC) as the key spatial variable, whereas traveltime to the HVC was used in the second trial. (The HVC is a point representative of the hypothetical activity center of the CBD).

Figure 1 shows the relationship between 1948 dwelling unit density and air-line distance from the HVC. Each point on this plot represents the gross residential density (street area included) for a ring around the HVC. Each ring is defined by the boundaries of all zones whose centroids fall within $\pm\frac{1}{2}$ mile of the nominal distance of the ring from the HVC with the exception of the first or CBD ring. The plot indicates a surprisingly regular decline in residential densities with distance from downtown in Greensboro in 1948. This was encouraging since the reliability of the DSGM depends greatly on the strength and stability of this relationship.

The method depends equally upon the relationship between distance and percent saturation. To compute the latter, residential capacity must be defined. Mathematically capacity is defined as existing dwelling units plus the product of vacant available, suitable land, and expected residential density. A decision had to be made at this juncture as to the density values to be used in the computation. Theoretically this should be the anticipated average density at which all future residential development will occur.

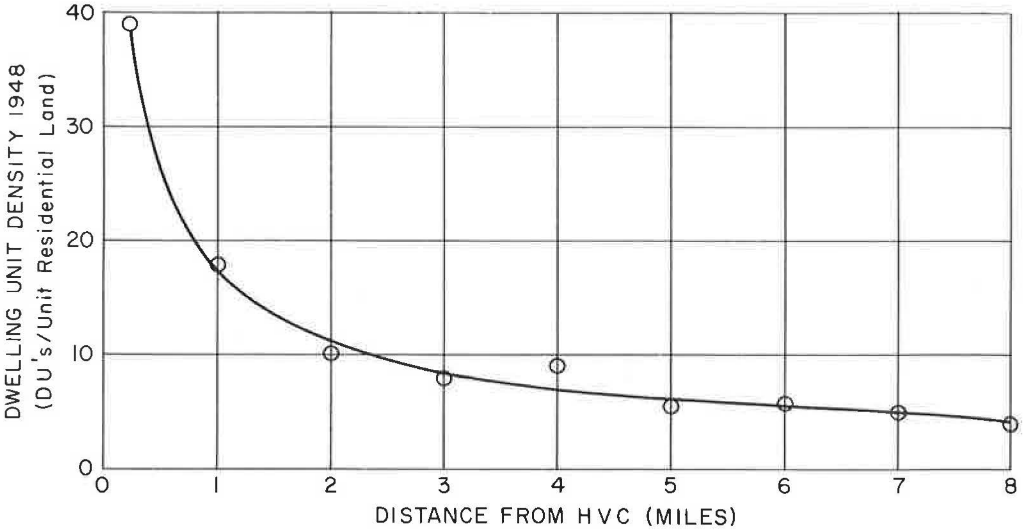

Figure 1. Dwelling unit density by distance bands—1948.

These values should be developed from an intensive analysis of trends in residential density patterns and zoning policies. For purposes of this study, however, future densities for each zone were assumed to be those given by the smooth hand-fitted curve of Figure 1. Prior to the acceptance of this single curve for the density gradient, gradients were plotted for each of five sectors. Although these plots exhibited less regular relationships, no significant variation between sectors was noted.

Vacant, suitable land for residential development was estimated by subtracting marginal land and land zoned for nonresidential uses from 1948 nonurban land. A systematic, but subjective procedure was used in the treatment of zoning: land was weighted by factors ranging from 0 for grids zoned only for industry to 1.0 for grids zoned only for residential use; land in grids zoned for mixed uses and other nonindustrial uses was weighted subjectively on a scale from zero to unity.

Having future residential development densities and vacant available land, it was possible next to compute both the residential saturations, in dwelling units and existing percent saturation, for each distance ring from the HVC. The latter values, resulting from the division of saturation into 1948 dwelling units, were then used to construct the percent saturation gradient. Figure 2 conforms very well with the plot expected for an urban area. The rather distinct and sharp transition between the $3\frac{1}{2}$- and $4\frac{1}{2}$-mi points indicates a transition from the area of urban character into the predominantly rural portions of the study region. The almost negligible slope of the curve beyond the $4\frac{1}{2}$-mi point is indicative of agricultural development and the absence of any strong competition for location with reference to central Greensboro.

The next step involved the 1960 projection of the percent saturation curve, also shown in Figure 2. (Percent saturation gradients by sector for 1948 were also plotted; however, as in the case of the density gradient, there was some additional scatteration of points, but no basis for using sector-specific gradients.) This is the most critical and subjective step in the forecasting process, the only restraint on the projected curve being that the area under the new curve must account for the projected regional growth. The number of dwelling units in the study area grew from a 1948 total of 27,191 to 41,250 in 1960 or a growth of 52 percent. One can proceed in almost an infinite number of ways insofar as establishing an acceptable projection of the percent saturation gradient. It was, however, found useful to first develop a feeling for the overall scale of the problem, that is, the area under the final curve which would be commensurate with the required final regional population. As a first approximation to the 1960 gradient each ordinate value was raised a distance equivalent to 52 percent of the 1948 value.

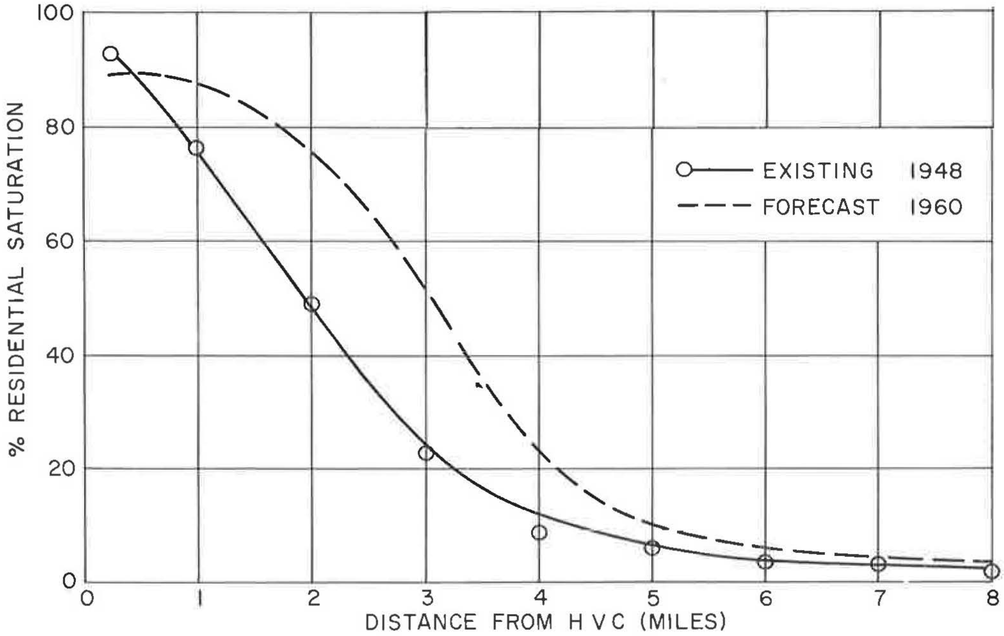

Figure 2. Residential saturation by distance bands.

The resultant curve then approximated the forecast condition under the assumption of uniform growth over the entire region. The following general criteria were then introduced to modify the naive first approximation of the shape of the gradient in 1960:

1. The bulk of the residential growth would occur in the 2-, 3-, and 4-mi rings.
2. The inner ring would suffer a slight decline.
3. The shape of the gradient would tend to bow out in the 1- to 3-mi range.
4. The sharp transition in slope of the 1948 saturation gradient observed at about the 4- to 5-mi point would become less abrupt in 1960.
5. The areas 5 miles and beyond would show some exurban growth, but the general flat slope would remain.

Relatively few attempts were necessary to arrive at a solution which was of satisfactory shape and which conformed with the actual 1948-1960 increase in total dwelling units.

Multiplying the appropriate ordinate value from the forecast percent saturation gradient (Fig. 2) by the ring saturation quantities established the forecast dwelling unit totals by analysis ring.

The projected growth of each ring was distributed to zones in a two-step process following the logic of CATS. The allocation to districts (defined by ring-sector boundaries) was handicapped by a lack of historical data. Ideally the trends in land use composition and growth rates between sectors should be studied in detail. For trial one, however, the simple assumption was made that sectors would share growth in proportion to available residential capacity.

The final distribution to zones was based on a systematic, but subjective linear weighting of the following factors:

1. Distance to convenience shopping,
2. Available residential capacity,
3. Distance to the major street system,
4. Percent of industrial development in the zone, and
5. Percent of residential development in the zone.

Trial two, which was conducted independently of trial one, differed from the above procedure in two principal ways:

1. Traveltime to the HVC was substituted for airline distance as the major independent variable. Zones were aggregated into 1-min interval rings for all analyses.
2. Ring growth was allocated to sectors (i.e., the district-level forecast) in proportion to the product of each sector's available residential capacity and the number of existing (1948) dwelling units.

Otherwise, the process followed that of trial one, including the method of estimating density and holding capacity, the sector definitions, and the allocation of growth from districts to zones.

Figure 3 shows the dwelling unit density gradient as determined from the ring analysis for trial two. As expected the same general shape is observed as for trial one. Figure 4 shows both the percent saturation curve calculated for the 1948 base period, and the forecast of the 1960 percent saturation curve. The shape of the latter gradient is quite similar to that for trial one except for a slight decrease in the growth allocated to the inner rings, resulting in a lessening of the bowing effect and a reduction in the slope of the gradient in the intermediate areas.

Accessibility Model

The generalized form of the accessibility model is as follows:

$$G_i = G_t \frac{A_i^a V_i}{\sum_i A_i^a V_i}$$

where

G_i = the forecast growth for zone i;
G_t = total regional growth $= \sum_i G_i$;

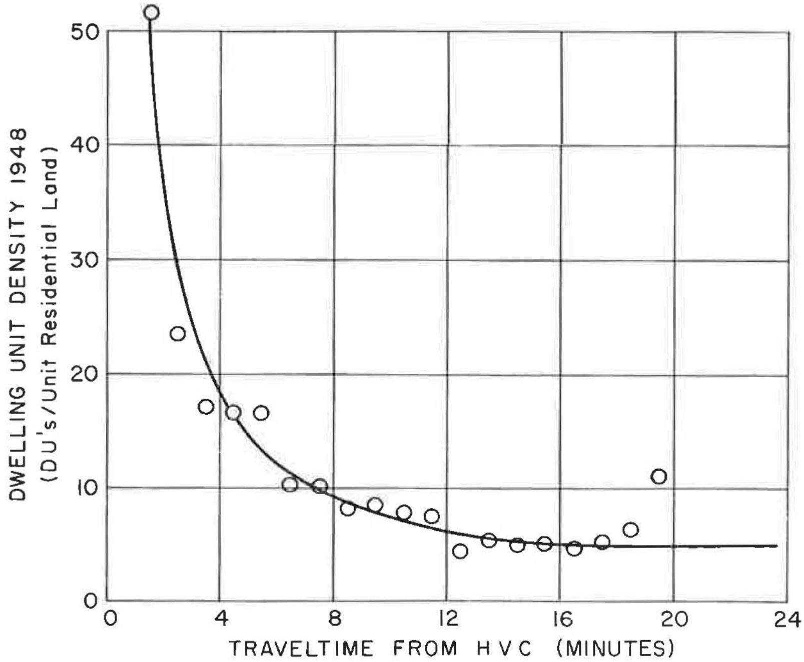

Figure 3. Dwelling unit density by time bands—1948.

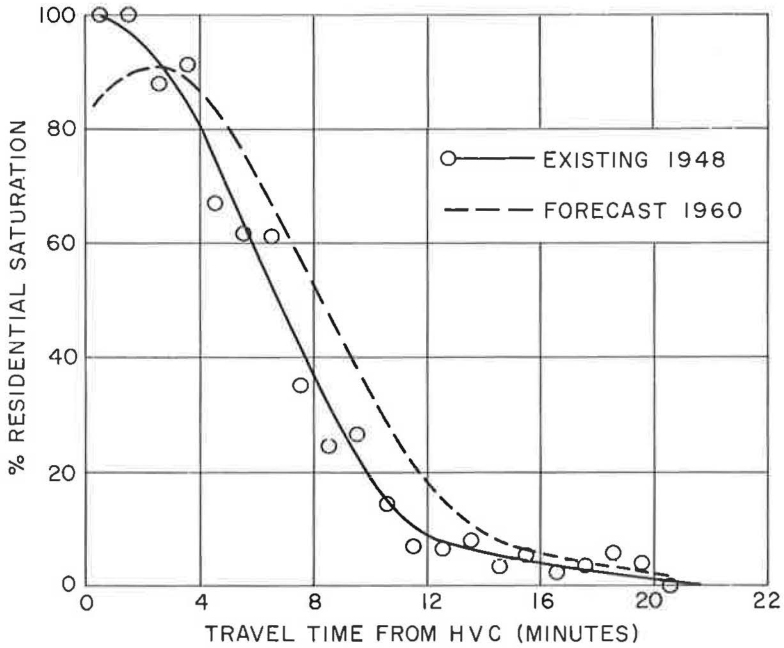

Figure 4. Residential saturation by time bands.

A_i = accessibility index for zone i;
V_i = vacant available land in zone i; and
a = empirically determined exponent.

The computation of the accessibility index traditionally is as follows:

$$A_i = \sum_j \frac{E_j}{T_{ij}^b}$$

where

E_j = a measure of activity in zone j (total employment used in this study);
T_{ij} = traveltime from zone i to zone j; and
b = an empirically determined exponent.

However, "friction factors" developed in the gravity model calibration by Alan M. Voorhees and Associates were actually used in the computation of accessibility:

$$A_i = \sum_j E_j \, F_{ij}$$

where F_{ij} is the friction of time separation of zones T_{ij} minutes apart. The F_{ij} values are approximately proportional to the actual number of trips T_{ij} minutes long per trip-end in each pair of zones T_{ij} minutes apart. In practice the computation of F_{ij} is considerably complicated by a desire to have the F_{ij} values form a smooth monotonic relation to T_{ij} yet maintain approximate equality between the resulting mean trip length and the actual mean trip length.

With the above definition of the model only one parameter, a, need be estimated to make the forecast. Two options were open:

1. Make a judgment of the value of a from previous work in other cities, and forecast 1960 zonal growth to have an independent test of the model; or
2. Fit a "best" value for a using the actual 1948-1960 changes in dwelling units.

Both options were actually used. For option 1 a value of 2 was assumed for a. (Hansen found that a value of about 2.7 was optimal for Washington, D.C.; the presumption that accessibility would have less influence in shaping growth in a smaller city is substantiated by the subsequent results in fitting values for a.) Methods used in fitting a to the 1948-1960 data are described in the Appendix.

Regression

For several reasons it was felt desirable to express the dependent variable of the multiple regression formulation as some function of the 1948-1960 growth rather than as some function of the absolute amount of cumulative development at a single point in time. The latter option was open, and has been used by others (13, 14); however, it was rejected to maintain comparability with the dependent variables of the other models, as well as to conform to standard practice in transportation planning models. As has been pointed out by the Traffic Research Corporation (15), there is good reason to expect greater accuracy for relatively short-range forecasts when predicting increments of growth.

Using change in dwelling units, or some function thereof, as the dependent variable, it was not possible with the available data to produce an independent forecast to check against the 1960 data. The equation parameters had to be estimated from the full 1948-1960 data files. Hence, accuracy results are shown in the next section only for a fitted model, and not for a forecast, in contrast to the other 4 methods. Dwelling unit data for a third point in time would be required to examine the forecasting reliability of the calibrated regression equation.

The usual regression approach differs from the other models used in this study in two additional important ways:

1. Many, rather than one or two independent variables may be incorporated, and
2. Variables are related to growth only in linearly weighted combinations, although variables may be transformed prior to regression.

The latter restraint is imposed by the use of a standard regression program (the BIMD 34 stepwise multiple regression program developed by the UCLA Bio Medical Center for the IBM 7090/7094 was used in this work). Of course nonlinear regression equations may be developed, but different normal equations must be solved and standard regression programs may not be used.

Numerous equations were developed, each involving the testing of various hypotheses regarding the functional relationships between variables. A total of 44 independent variables plus certain selected nonlinear transformations were examined in all, including:

1. Measures of zone size and amount of land in different uses;
2. Accessibility to employment;
3. Time and distance to HVC;
4. Zonal employment, total and by major type;
5. Densities for 1948;
6. Vacant available land;
7. Zoning protection;
8. Land value; and
9. Proportions of total land and developed land in each major use.

Four definitions of the dependent variable were tested:

1. Increase in dwelling units (DU);
2. Log DU;
3. DU per unit of available land (DU/L); and
4. Log [DU/L].

The logarithmic transformations were employed to test certain hypotheses regarding exponential relationships, as for example, are expressed in the accessibility model. The growth-per-unit-of-available-land transformations were employed in an attempt to remove all measures of zone size from the equations, and thereby, to avoid the possibility of distorted relationships due to the pecularities of area definitions.

The final equation accepted after comparing the accuracy and reasonableness of all trials was

$$Y = -2.3 + 0.061\,X_1 + 0.00066\,X_2 + 1.1\,X_3 - 0.11\,X_4 - 0.0073\,X_5$$

where

Y = logarithm of growth in dwelling units 1948-1960 per unit vacant land;

X_1 = zoning protection, 1948;
X_2 = percent of total land area in residential use, 1948;
X_3 = logarithm of accessibility to employment, 1960;
X_4 = dwelling unit density, 1948; and
X_5 = percent of total use land in industrial use, 1948.

The coefficient of correlation is 0.61. Table 1 contains the t and beta (β) values (standardized regression coefficient) for each of the independent variables in the equation. All regression coefficients are significantly different from zero with 95 percent confidence. Having the greatest β value, the transformed accessibility variable is shown to exhibit the most influence upon the estimate of the dependent variable. Percent of urban land which is in industrial use has the lowest β values and, therefore, contributes least to the total equation estimate.

The zoning code was a value from 0 to 9, where a higher value indicated zoning control closer to single family residential only, and lower value marginal-to-no zoning control. The positive relationship then indicates the positive environmental influence of strict residential zoning policy. The positive contribution of accessibility to work areas is self-explanatory. Also, the positive contribution of percent of total area devoted to residential development is interpreted as a measure of residential clustering. The tendency for slow growth or even decline in the residential stock of the close in, old city areas, coupled with the rapid increase in the fringe and newly settled locations accounts for the negative coefficient for dwelling unit density. The negative contribution of percent industrial land is indicative of the restraint on new residential development in areas immediately adjacent to industrial areas.

Because the estimation was couched in both logarithmic and intensity units, several operational difficulties were introduced. The estimating equation was incapable of either accepting negative values for the dependent variable or estimating decline in any zone. All zones which suffered dwelling unit decline over the calibration period were approximated to have shown no change. An additional problem was encountered for several zones which experienced dwelling unit growth, but which had no vacant land available in 1948. Without some adjustment the growth intensity value becomes infinite. These few cases were handled by substituting large arbitrary values of growth intensity. Finally, there is no built-in provision, as there is for other models, to assure that the accumulated zonal estimates obtained from the regression equation solution will equal the actual total regional growth. All regression estimates had to be factored up to sum to the actual regional growth.

TABLE 1

RELATIVE SIGNIFICANCE AND EXPLANATORY POWER OF VARIABLES IN REGRESSION EQUATION

Independent Variable	t	β
Log accessibility to employment, 1960	4.30	0.321
Zoning code, 1948	2.89	0.213
Percent of total land residential, 1948	2.70	0.187
Dwelling unit density, 1948	3.28	0.177
Percent of urban land industrial, 1948	2.98	0.159

Two Intervening Opportunity Models

Although the two opportunity models tested are based on quite different initial assumptions and take on dissimilar mathematical form, nevertheless, both can be reduced to a simple general hypothesis. In the context of this problem, the probability that a suitable residential opportunity (a unit of available capacity) is accepted for development is hypothesized to be a monotonically decreasing function of the number of intervening opportunities, opportunities being ranked by time from the HVC.

Some improvement in these models could undoubtedly be made by allocating increments of growth from more than one point, perhaps from all major centers of employment in proportion to the amount of employment in each center. This would make the test of the intervening opportunities models more comparable to the accessibility model procedure.

Stouffer Formulation. The Stouffer model may be defined in the following manner:

$$g_p = \frac{k\,O_p}{O}$$

where

g_p = number of dwelling units forecast to be located in a particular area p;
O_p = opportunities in interval p;
O = total number of opportunities from central distribution point through interval p; and
k = constant of proportionality to assure that the total number of dwellings located equals the actual total growth.

As stated, the Stouffer formulation can be applied without the need for assuming any parameter values. However, it is an operational requirement that the study area be structured into a number of discrete geographic units which are then ranked from a central distribution point, the HVC in this case. One method of aggregating areas, which Strodtbeck' has shown to have some appealing properties, is to delineate a small number of rings containing approximately equal numbers of opportunities (16). For the initial application of the Stouffer model to the allocation of residential growth, the Greensboro study area was divided into 10 rings, each of which was composed of a whole number of zones and an approximately equal number of opportunities. Zones were assigned to rings according to their ranking in time from the HVC.

It was then possible to determine g_p, the forecast number of dwellings in ring p by direct substitution in the formula. The ring forecasts were then proportioned among the constituent zones on the basis of opportunities.

For an explanation of the fitting of the Stouffer equation to 1948-1960 data the equation must be converted into its continuous differential form as follows:

$$d(G_p) = \frac{kd(O)}{O}$$

By integrating

$$G_p = k \ln O + C$$

where

G_p = the total number of dwellings allocated to all opportunities from the central point up to and including opportunity interval p;
$d(G_p)$ = dwellings allocated to opportunity interval p;
$d(O)$ = opportunities in interval p; and
C = constant of integration.

This equation plots as a straight line of slope k where the ordinate, total allocated dwellings, is in linear form and the abscissa, total accumulated opportunities, is a logarithmic scale. As a test of the appropriateness of the Stouffer formulation in describing the spatial distribution of residential growth in Greensboro, the actual accumulated zonal dwelling unit growth 1948-1960 was plotted against accumulated 1948 opportunities, the zones being ranked by traveltime to the HVC. If the Stouffer model is valid the resulting plot should follow a straight line. It was immediately obvious that a single straight line could not be adequately fitted to the points, but rather that two distinct straight lines were necessary (Fig. 5). After hand fitting the two lines, 1960 growth estimates were made to the individual zones from the straight lines and the error computed. These results and those computed from the initial, noncalibrated test of the Stouffer formula are discussed later with the results of the other four models.

Schneider Formulation. As applied to the distribution of residential activity, the Schneider model takes the following form:

$$d(G_p) = g_t \left[e^{-\ell O} - e^{-\ell (O + O_p)} \right]$$

where

G_p = total number of locations in opportunity interval from the central point up to interval p.
g_t = total growth to be allocated;
ℓ = model parameter expressing the probability of an opportunity being accepted for location;
O = total number of opportunities ranked from the central point up to interval p.

As a necessary condition for applying the model the parameter ℓ must be stipulated. For the first trial of the model for a 1960 forecast without benefit of the 1948-1960 data, ℓ was estimated from the assumption that the actual dwelling unit increase within the study boundaries was 99 percent of the aggregate Greensboro oriented growth. (The theoretical model is based on a distribution to an unbounded area; applicaton to a finite area requires specification of the number of accepted opportunities being outside the

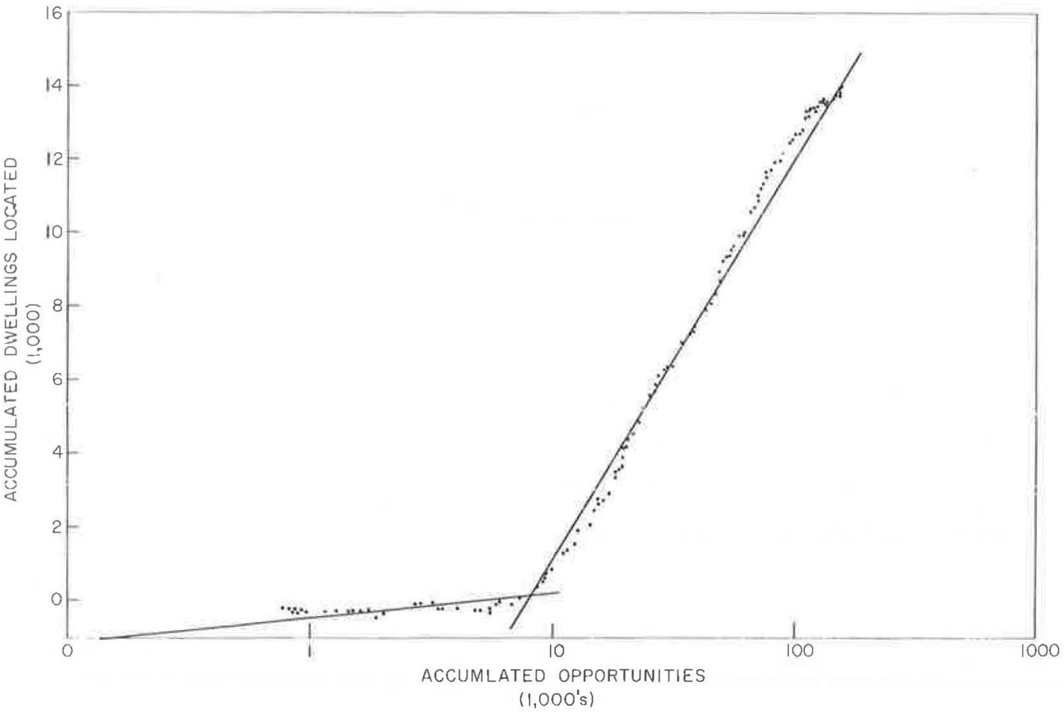

Figure 5. Test of Stouffer's formulation.

boundary, or equivalently, the percentage accepted up to the boundary.) The ℓ resulting from this assumption was 12.76×10^{-6}.

For an explanation of the fitting of the Schneider formulation to 1948-1960 data, the formula can be restated after integration as

$$G_p = g_t \left[1 - e^{-\ell O}\right]$$

Subtracting g_t from both sides and rearranging,

$$g_t - G_p = g_t e^{-\ell O}$$

or

$$\ln (g_t - G_p) = \ln g_t - \ell O$$

This relationship plots as a straight line where the ordinate, $(g_t - G_p)$, is in logarithmic scale and the abscissca, total accumulated opportunities from the central point (O), is in linear scale. The slope is ℓ and the intercept g_t.

If the Schneider formulation effectively replicates the spatial distribution of residential growth in Greensboro then plotting the actual quantity $(g_t - G_p)$ versus accumulated opportunities (O), in semilogarithmic forms, should yield a straight line (Fig. 6).

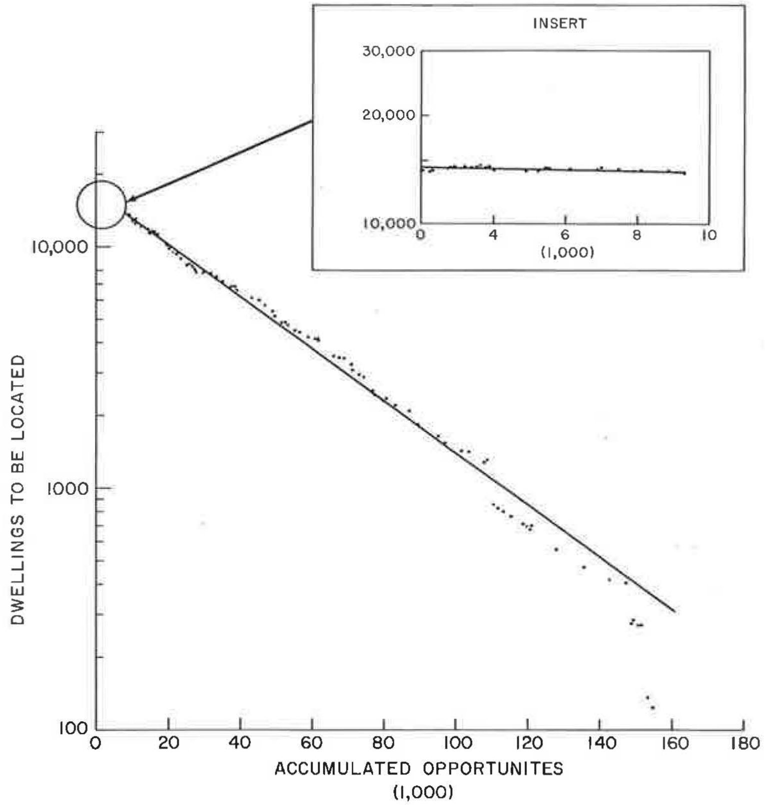

Figure 6. Test of Schneider's formulation.

As with the Stouffer formulation, the Greensboro data appear to exhibit two distinct straight line segments, rather than one, as required by the initial model formulation. The zones comprising the transition area between the two straight line segments (Fig. 6) are the same ones as those at the juncture of the two line segments for the Stouffer formulation (Fig. 5). The slopes of the fitted lines can be loosely compared to the short and long trip ℓ's which have become standard practice in applying the Schneider formula as a trip distribution model. The slope for the central city line segment is 1.707×10^{-6}, and that for the outer, suburban area is 10.9×10^{-6}.

The distribution of residential growth in Greensboro from 1948 to 1960 did not adequately conform to either of the intervening opportunities formulations over the complete range of opportunities. It is noteworthy, however, that the data plot as two straight lines in both Figures 5 and 6. It was also pointed out that the transition points in the vicinity of the intersection of the fitted straight lines in both figures were the same data points representing the same zones. Although a detailed examination of these zones has not been attempted it does appear that they approximate a transition ring in Greensboro which separates the "inner city," marginal growth area from the suburban, rapid expansion area. This band encircles the HVC at a radius of $1\frac{1}{2}$ to 2 miles. For a city the size of Greensboro, which in 1948, exhibited a leveling off in the percent saturation gradient at $3\frac{1}{2}$ to $4\frac{1}{2}$ miles from the HVC, the area circumscribed by this transition band probably was characteristic of similar areas in most cities—old and perhaps showing signs of blight with little available residential capacity.

The inner area straight line slopes drawn to the two plots are both very close to the horizontal. In contrast, there are quite steep slopes for the plots representing suburban areas. A hypothetical locator viewing the opportunity surface from the HVC in accordance with either of the two plots apparently assesses himself a greater penalty in passing up suburban opportunities as opposed to inner-city ones. That is, the inner-city opportunities are a less desirable subset of the total as evidenced by the significantly lower slope on the plots, hence a lower probability of accepting individual opportunities. One may conjecture that location choices from the inner-city opportunity subset are responsive more to the individual living qualities of the opportunities other than its accessibility, which may be extended to the notion that the inner-city opportunities are viewed more or less as of homogeneous access in opposition to the suburban subset where opportunity access is of greater import in the locational choice.

Of interest from a purely forecasting viewpoint is the question of the stability of the handfitted lines in Figures 5 and 6. Do the slopes remain more or less constant over time and how does the transition area behave in relation to the total opportunity surface? One may speculate, for example, that the straight line relationships fitted to the data will hold over time and that the diffusion in residential location observed in the past is merely a reflection of the diffusion in the opportunity surface; that is, a physical dispersion outwards occasioned by the filling in of less distant areas, rather than of an alteration in the location function. On the other hand, it is possible that over time the slopes of the plots may be flattening out which is symptomatic of a society less restrained by the impedance of travel. Clearly, answers to speculations of this nature are required before one can estimate the applicability of the fitted lines to forecasting to a future time point.

PERFORMANCE AND INTERPRETATION OF RESULTS

Performance

The single accuracy measure which was calculated for all trial forecasts was the sum of squares of dwelling unit forecasting error. These measures were computed at four levels of geographic aggregation: sector, ring, district, and zone, for all trials. A sixth forecast was made using the naive assumption of equal growth for all zones. The error sum of squares computed under this assumption, which will be referred to as the naive model, is (n - 1) times the variance in actual zonal residential growth. It will serve as a benchmark in evaluating the results of the five techniques listed.

Table 2 gives the computed error sum of squares for all of the forecasts and calibrations at each level of aggregation. For sake of complete comparisons, the results of zone level forecasts for each of the models (not for the DSGM) have been aggregated to districts and rings defined both by time and distance from the HVC. Trial one of the DSGM was based on analysis at the level of district as defined by distance from the HVC; therefore results are not shown for districts as defined by time to HVC, and vice versa for trial two of the DSGM.

The sums of squares of differences between estimated and actual are analogous to "unexplained" variances of a statistical model. However, since valid statistical inferences obviously cannot be drawn, this terminology should not be used. The error measurements of Table 2 do provide an index which can be used to compare results in any single column, that is, for the same level of aggregation. Comparisons between columns are meaningless, since different numbers of areas and different variances from mean growth rates are involved at different levels of aggregation.

To provide some degree of comparison between levels of aggregation, as well as between forecast techniques, Table 3 gives the ratio of each error to that for the naive model.

There are rather poor results at the zone level for all five methods. In some instances the naive model, assuming equal growth for all zones, actually exceeds the level of accuracy of forecasts. The particularly discouraging results of the DSGM at the zone level are evidence of poor choice of criteria by the authors in distributing growth from districts to zones. As pointed out earlier, this method requires historical data that were not available and requires intimate familiarity with the area, which the authors lacked. The technique itself should not be blamed.

Undoubtedly, a substantial amount of the error at such a fine level of detail as the zone can be attributed to inaccuracies in data—assumptions made in certain estimates, incompatibility of merged files, differences in definitions between time periods, etc. However, other factors are contributory. The average zone contained only 109 dwelling units in 1948 and increased 56 to 165 by 1960. These values are far too small to hope for reliable predictions with any model. Obviously, differences between zones

TABLE 2

ERROR SUM OF SQUARES FOR ALL TRIALS[a]

Method	Levels of Aggregation					
	Zone	Districts		Rings		Sector
		By Distance Ring	By Time Ring	By Distance	By Time	
DSGM						
Trial I	2.33	6.97	—	8.36	—	9.69
Trial II	2.41	—	4.43	—	4.07	3.02
Accessibility model						
Forecast	1.80	4.16	2.84	3.25	2.33	4.58
Fitted	1.79	3.98	2.76	2.18	1.99	4.46
Regression (fitted)	1.85	4.71	3.14	5.16	2.84	3.71
Stouffer model						
Forecast	2.21	6.45	4.22	5.57	3.48	11.25
Fitted	1.91	4.72	3.07	2.42	1.46	8.84
Schneider model						
Forecast	2.07	6.16	4.13	4.10	3.38	13.92
Fitted	1.95	4.65	3.08	1.91	1.65	10.18
Naive model	2.20	7.66	5.22	20.64	10.54	16.18

[a] All values have been multiplied by 10^{-6}.

TABLE 3

RATIO OF ALL ERRORS TO NAIVE MODEL ERROR

Method	Levels of Aggregation					
	Zone	Districts		Rings		Sector
		By Distance Ring	By Time Ring	By Distance	By Time	
DSGM						
Trial I	1.06	0.91	—	0.41	—	0.60
Trial II	1.10	—	0.85	—	0.39	0.19
Accessibility model						
Forecast	0.82	0.54	0.54	0.16	0.22	0.28
Fitted	0.81	0.52	0.53	0.11	0.19	0.28
Regression (fitted)	0.84	0.62	0.60	0.25	0.27	0.23
Stouffer model						
Forecast	1.01	0.84	0.81	0.27	0.33	0.70
Fitted	0.87	0.62	0.59	0.12	0.14	0.54
Schneider model						
Forecast	0.94	0.80	0.79	0.20	0.32	0.86
Fitted	0.89	0.61	0.59	0.09	0.15	0.63
Naive model	1.0	1.0	1.0	1.0	1.0	1.0

at this level are largely due to random variations not explainable by models. The districts represent a more reasonable level of detail at which to examine and compare accuracies. For the sake of comparison with transportation study practices, the average district (defined by distance rings) used in this study could be expected to have about 8,000 person trip-ends in 1948 (about 660 dwelling units with 3.2 persons per dwelling and 4 trip-ends produced per person).

Table 4 shows the relative accuracy of the accessibility model forecast at various levels in comparison to the size of the values being forecast. In this table the root-mean-square-error (RMSE) is used as the measure of error, since it can be compared with the magnitude of the forecast values: about two-thirds of the errors fall within RMSE values.

The RMSE is roughly half of the average 1960 dwelling units per zone, and about a third of the average 1960 dwelling units per district. Of course, these accuracies must be viewed in relation to the overall growth rate of 52 percent. Intuitively one would expect that the ratios of the RMSE's to the 1960 values might be nearly cut in half if the overall growth rate was half as large.

The accessibility model performed substantially better than other unfitted models at most levels of aggregation (Table 3); but the fitted Stouffer and Schneider models were

TABLE 4

COMPARISON OF ERRORS TO SIZE OF FORECAST VALUES
ACCESSIBILITY MODEL FORECAST

Levels of Aggregation	RMSE $\sqrt{\frac{s \cdot s}{n}}$	Average 1960 DU (per areal unit)	Average Growth 1948-1960	Number of Areas
Zone	85	165	56	249
District[a]	381	1,006	342	41
Rings[a]	600	4,580	1,560	9

[a] By distance.

quite comparable to the fitted accessibility model. Somewhat surprisingly, the addition of several other explanatory variables in linear regression form did not improve the accuracy.

Results at the sector level are of interest because of the implications for forecasting radial corridor movements. Here the intervening opportunity models yield comparatively poor results, perhaps because they were not made sensitive to the distribution of employment, as were the accessibility model and regression equation.

Trial one of the DSGM assumed relative growth by sectors in proportion to available capacity—a weak assumption judging by comparison with the error of trial two. The importance of residential character in attracting additional growth apparently holds at all levels—between sectors as demonstrated by comparison of the two DSGM trials, and as a factor at the zone level as demonstrated by the statistical significance of that factor in the regression analysis.

Examination of Actual Patterns of Growth

All forecasts of 1960 density were based on the assumption that development in any zone would occur at the density indicated by a smooth line drawn through the 1948 density vs distance (or time) from the HVC. Figure 7 compares the actual 1960 density-distance gradient with that for 1948. There appears to have been a rather uniform amount of decrease in density at all distances, except for the core area where the decrease was substantial. This obviously accounts for some error in the forecasts which required estimates of 1960 density (DSGM and the opportunity models), especially in the core area.

The actual 1960 and 1948 percent saturation gradients are compared in Figure 8, along with the forecast curve used for trial one of the density-saturation gradient method. Not surprisingly, the actual 1960 curve does not follow as smooth a curve as for 1948, since the plot represents percentage of 1948 capacity rather than 1960 capacity. The most significant errors in the forecast appear to be due to the unexpectedly large decline in the core and the amount of growth that occurred in relatively remote portions of the area, ring 5 and 6. However, the general shape of the forecast curve is appropriate.

Figure 9 shows the same comparisons for the results of the accessibility and regression models. The agreement with the actual 1960 gradient is quite good, except for the obvious inability of these techniques, as used in this study, to predict decreases in the core.

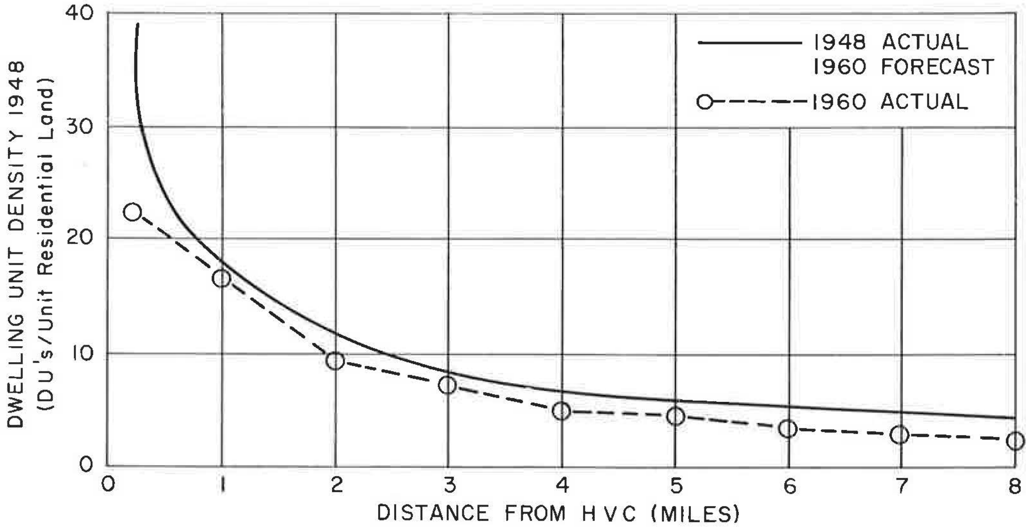

Figure 7. Dwelling unit density by distance bands.

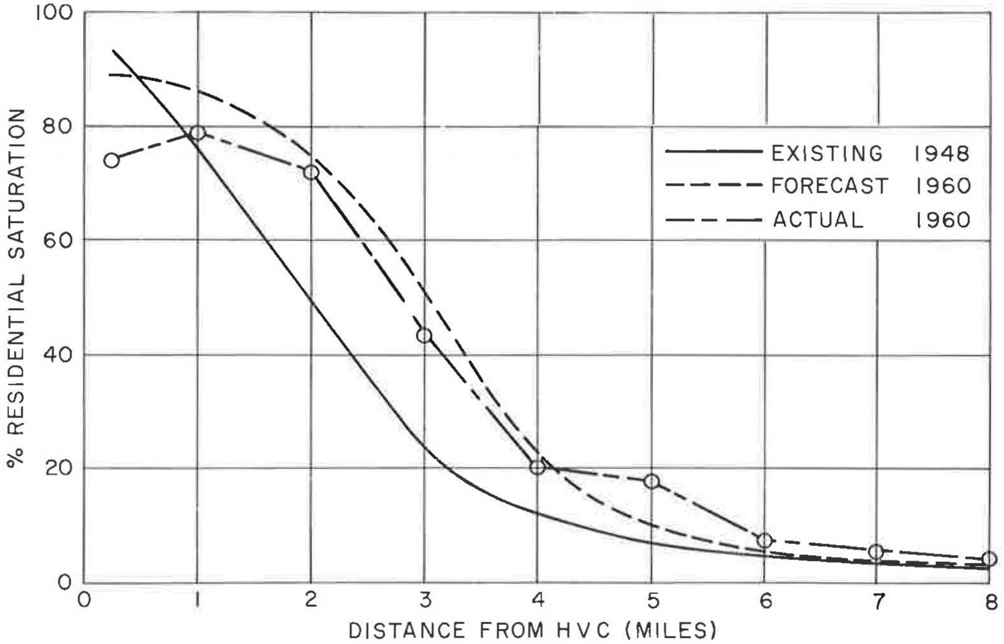

Figure 8. Percent residential saturation by distance bands.

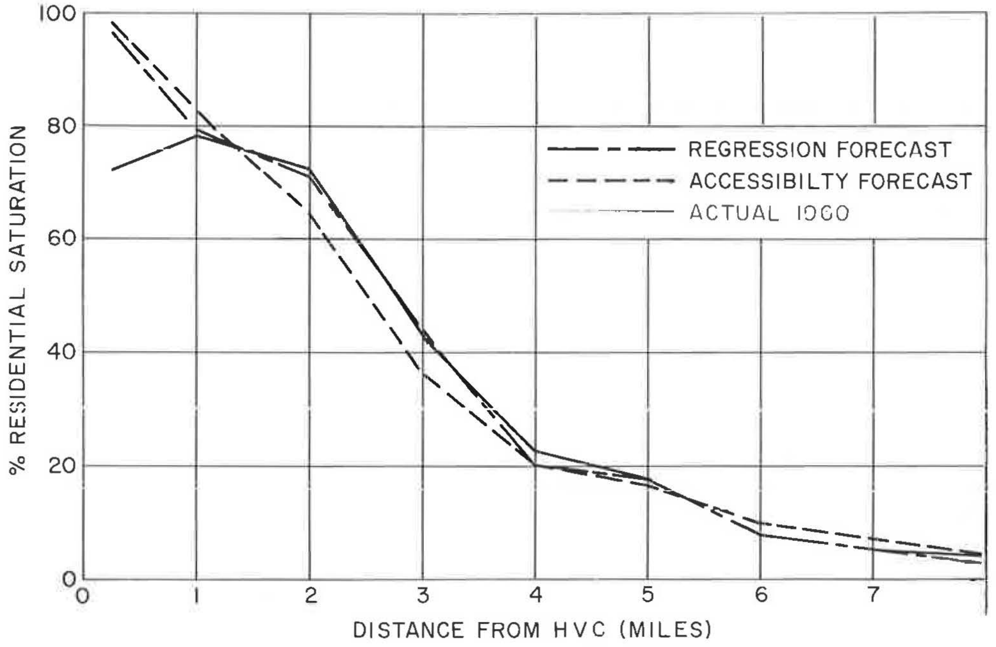

Figure 9. Percent residential saturation by distance bands.

In an attempt to picture how the residential density structure of the study region changed, Figure 10 was drawn. Using the data for total dwelling units and residential land area from the distance to HVC ring analysis, cumulative percent of total regional dwellings was plotted against cumulative percent total residential land area on a ring aggregate basis, proceeding outwards from the core ring. The plots for the actual conditions in 1948 and 1960 are shown. If smooth curves were drawn the slope at any point would represent the inverse of density for the marginal dwelling unit. A diagonal line drawn on Figure 10 would represent uniform residential density for the entire study area. The bowing of each of the curves below the diagonal indicates the decline in density as one proceeds outwards from the HVC. If densities in the inner area were to decline along with an increase in the dwelling unit densities in the outer rings, the region as a whole would be approaching a state of uniform density, and the curve would shift toward the diagonal. On the other hand, if the difference between inner and outer area densities were to increase substantially, then there would be a shifting of the plot down and to the right. Understanding that the plots in Figure 10 represent an overall increase from 1948 to 1960 of 52 percent, the rather minute change in the density structure of the study area as described by these plots is outstanding.

Although the two plots (Fig. 10) appear to coincide almost exactly, they should not be misread as indicating no change in the geographic distribution of dwelling units from 1948 to 1960. Each of the data points representing a distance ring has shifted downward and to the left from its 1948 position to 1960. That is, inasmuch as the majority of residential growth occurred in the suburban rings, the dwelling stock of the inner rings in 1960 represents a smaller proportion of the total region stock than in 1948 and also utilizes a smaller proportion of total residential land; hence, the shifting of the data points downward and to the left.

An interesting question is whether similar plots for other urban areas exhibit this same constancy as found in Greensboro. If this is found to be so, such plots could be quite helpful in residential forecasts.

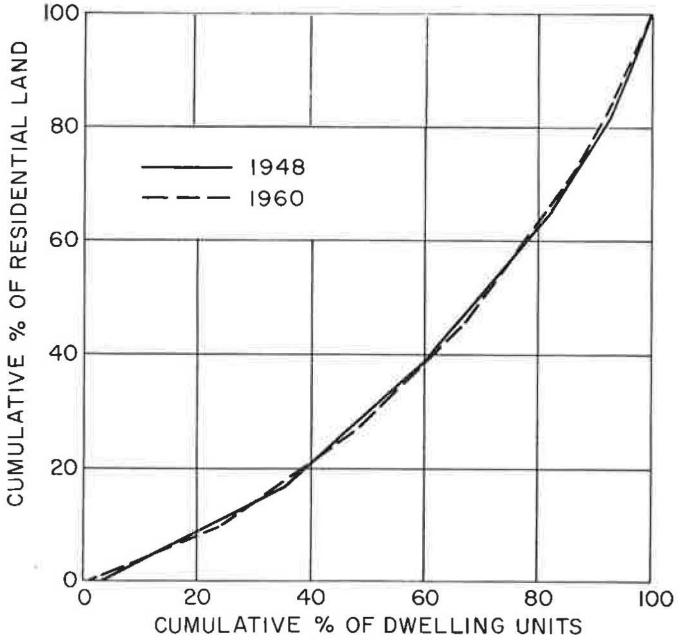

Figure 10. Cumulative dwelling vs cumulative residential land.

CONCLUSIONS

1. Simple, nonbehavioral residential land use forecasting models, which do not discriminate between the locational patterns of different types of households, are sufficiently accurate to be recommended for use in relatively small metropolitan areas of 100,000 population or larger. The Greensboro area's spatial structure and pattern of growth clearly demonstrates a degree of organization warranting analytical treatment in the planning process.

2. Land use forecasting with simple first generation models produced reasonably accurate results for levels of geographic aggregation where the average areal unit contained a population of about 2,000 persons. Efforts to forecast growth for much smaller areas may prove unjustified. At zone levels of about 300 population, these models appeared to offer little or no assistance in forecasting.

3. Differences in accuracy among the five forecasting methods are not large enough to warrant a strong recommendation for any single one in preference to others. Any of the methods would appear to be preferable to forecasting without the benefit of analytical techniques.

4. The simple accessibility model yielded the most accurate forecast of all methods used without benefit of calibration to time series data, for this one test. Errors in fitting were relatively insensitive to small changes in the exponent of accessibility.

5. None of the multiple linear regression models tested offered improvement over two-variable fitted models despite the fact that five or more factors were included in the regression equations.

6. Multiple regression models possess certain drawbacks. If the dependent variable is expressed as an extensive quantity (e.g., increase in dwelling units) then measured relationships with independent variables are influenced by pecularities of area definition and size, and may not conform satisfactorily with logical hypotheses regarding the land development process. Nonlinear transformations on the dependent variable such as logarithms or fractional power functions are unsatisfactory because the usual least squares criterion tends to bias the parameter estimates to produce good fits to small values and poor fits to large values. Expression of the dependent variable as an intensive quantity (e.g., dwelling unit increase per unit area) may be the most satisfactory operational solution except that relationships which are actually nonlinear may not be properly represented. Perhaps this might be handled by treating certain independent variables as sets of dummy variables.

7. Although the two intervening opportunity models performed satisfactorily as used in this study, some evidence pointed to the possibility of improvement by allocating growth from all major centers of employment rather than from just a single point, the CBD. In addition, each of the two models implies a different straight line plot on different semilogarithmic coordinates which did not hold true for Greensboro over the entire study area. Apparently the hypotheses are valid, but separate functions may be necessary for the built up, inner-city area, and the developing suburban area.

8. The forecasting approach used by CATS differed from the other models in important respects. It forces the analyst to become intimately familar with the study area before attempting to forecast. This is probably the strongest feature to recommend it. The graphical analyses that the method is based on represent excellent descriptions of the key spatial relationships of a metropolitan area—even for relatively small areas. The methods of analysis are useful tools regardless of the forecasting technique used. They can serve as checks on the reasonableness of forecasts made by less subjective models.

However, as applied in this study, the method is time-consuming, requiring considerable hand work and far more data manipulation. The method is less adaptable to the computer, and hence would be cumbersome for testing of alternative land use policies, or for recursive use in combination with other submodels.

9. The five techniques examined admittedly are far from representative of the extent of current land use forecasting research. They do represent the initial attempts and as such lack the sophistication and elegance of later thinking. These are descriptive models in that they do not involve themselves with the behavior of decision-makers;

nor do they possess any real theoretical content. It is highly probable that the key to increased forecasting accuracy for small subareas lies in the ability of the analyst to simulate the decision process of subpopulations of the region.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to both Professor F. S. Chapin, Jr., of the University of North Carolina and to A. M. Voorhees of Alan M. Voorhees and Associates for making available the data utilized in this study.

Appendix

CALIBRATION OF ACCESSIBILITY MODEL

Two procedures were used in the attempt to estimate the optimal exponent of accessibility: linear regression on transformed variables and an iterative, nonlinear least squares fit of the untransformed dependent variable.

Linear Regression on Transformed Variables

Three transformed versions of the standard accessibility model were tested:

$$\log G_i = \log a + b \log V_i + c \log A_i \tag{1}$$

which, in nonlogarithmic form is

$$G_i = a V_i^b A_i^c \tag{2}$$

$$\log \left(\frac{G_i}{V_i} \right) = \log a + b \log A_i$$

or in nonlogarithmic form

$$G_i = V_i A_i^b$$

$$\log G_i - \log V_i = \log a + b \log A_i \tag{3}$$

which is the same as Eq. 2 in nonlog form.

The nonlogarithmic forms of Eqs. 2 and 3 are essentially equivalent to the standard form of the model as stated in the body of this report. They would be identical if the normal equations contained the condition that

$$a = \frac{G_T}{\sum_i V_i A_i^b}$$

Since a standard regression program was used, this condition may be violated, and equation estimates must be factored to sum to actual total growth. This holds for all three of the transformed versions of the model.

TABLE 5

RESULTS OF THREE VERSIONS OF LINEAR REGRESSION ON TRANSFORMED ACCESSIBILITY MODEL

Item	Eq. 1	Eq. 2	Eq. 3
Accessibility exponent (b)	3.52	1.63	2.29
Log a	-8.0	-3.2	-4.9
Vacant land exponent (c)	1.51	1	1
Sums of squares of error (× 10^{6})	2.21	1.89	1.78

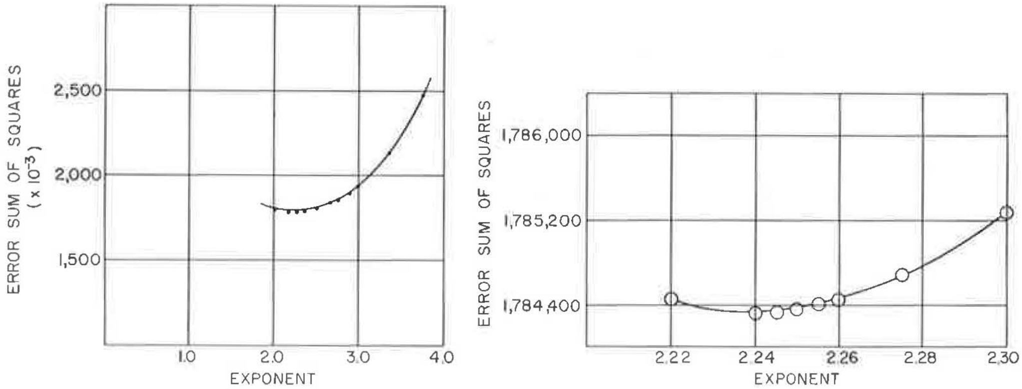

Figure 11. Accessibility model error vs accessibility exponent.

Eq. 1 also expresses vacant land as a power function in contrast to its linear form in the standard formula.

The basic problem, however, is that the least squares criterion is different for each version (the minimization of unexplained variance in the dependent variable) since the dependent variable is different for each. None is the correct criterion. The log transform tends to produce a bias toward better fits for small values of the untransformed dependent variable. Table 5 summarizes the results of the three versions.

The fairly wide variation in the accessibility exponent, as well as in the error term leads one to be suspicious of regression on transformed dependent variables.

Nonlinear Least Squares Fit of Exponent

A routine was programed to iterate toward the true least squares solution for the standard accessibility model

$$G_i = G_T \frac{A_i V_i^b}{\sum_i A_i V_i^b}$$

Figure 11 shows the results in the form of a plot of the sums of squares of error vs a range of exponents. A smooth curve with a minimum at b = 2.24 is apparent.

It is interesting to compare these results with the b value of 2.7 reported by Hansen for Washington, D.C. One might expect this value to increase with the size of the city.

REFERENCES

1. Hamburg, J. R., and Sharkey, R. H. Land Use Forecast. Chicago Area Transportation Study, 3.2.6.10, Aug. 1, 1961.
2. Hamburg, J. R. Land Use Projection for Predicting Future Traffic. Highway Research Board Bull. 224, pp. 72-84, 1959.
3. Hansen, W. G. Land Use Forecasting for Transportation Planning. Highway Research Board Bull. 253, pp. 145-151, 1960.
4. Hansen, W. G. How Accessibility Shapes Land Use. Jour. of the Amer. Inst. of Planners, May 1959.
5. A Projection of Planning Factors for Land Use and Transportation. Baltimore Regional Planning Council, Tech. Rept. No. 9, March 1963.
6. Graves, C. H. The Multiple Regression Models of Small Area Population Change. Highway Research Record 102, pp. 42-53, 1965.
7. Hartford Area Traffic Study Report. Volume I. Conn. Highway Dept., July 1961.

8. Stouffer, Samuel A. Intervening Opportunities: A Theory Relating Mobility and Distance. Amer. Soc. Rev., Vol. 5, No. 6, Dec. 1940.
9. Tomazinis, Anthony R. The Development of the Penn-Jersey Transportation Study Trip Distribution Model.
10. Chicago Area Transportation Study. Final Report. Vol. II. July 1960.
11. Hamburg, J. R., Lathrop, G. T., and Young, G. F. An Opportunity-Accessibility Model for Allocating Regional Growth. Highway Research Record 102, pp. 54-66, 1965.
12. Hamburg, J. R., and Lathrop, G. T. An Opportunity-Accessibility Model for Allocating Regional Growth. Jour. of the Amer. Inst. of Planners, May 1965.
13. Chapin, F. S., Jr., and Weiss, S. F. Factors Influencing Land Development. Univ. of North Carolina, April 1962.
14. Lowry, Ira S. A Model of Metropolis. The RAND Corp., Aug. 1964.
15. Traffic Research Corporation. Review of Existing Land Use Forecasting Techniques. Boston Regional Planning Project, July 1963.
16. Strodtbeck, F. L. Equal Opportunity Intervals: A Contribution to the Method of Intervening Opportunity Analysis. Amer. Soc. Rev., Vol. 14, No. 4, Aug. 1949.

Verification of Land Use Forecasting Models: Procedures and Data Requirements

DAVID E. BOYCE and ROGER W. COTE
Battelle Memorial Institute, Columbus Laboratories, Columbus, Ohio

A review of the available land use forecasting models suggests that verification and evaluation of these models is effectively limited because confidence statements on the forecast variables are not available from the models. This paper proposes possible procedures for providing confidence statements for two types of models. The data requirements for verification studies of the type proposed are described.

•DURING the past 5 years, model development in urban transportation planning has been oriented primarily to the problem of land use forecasting; that is, the allocation or distribution of urban activities to small analysis zones at future points in time. The publication of a special issue of the Journal of the American Institute of Planners, May 1965, on urban development models marks the completion of the first round of intensive model development. Reported in this Journal issue and elsewhere [e.g. Lowry (10), Hill et al. (5), and Donnelly et al. (1)] are the results of more than 10 separate studies whose objective was the formulation, development, and testing of land use forecasting models. These models represent a wide variety of approaches and orientations to model development and incorporate much of what is known and theorized about urban structure and growth. Although a great deal remains to be done before even crude forecasts of land use will be available on a production basis (in the sense that urban trip forecasts are now prepared), it is clear that a foundation for future model development and refinement has been provided.

This being the case, it is timely to ask

1. How well do the current models perform?
2. What procedures, tests, and criteria are available to evaluate their performance?

A review of current models indicates that the first question may be answered only superficially, for two reasons: (a) these models cannot be verified in a strict sense because their formulations do not provide a confidence statement about the relationship between the observed and predicted values; and (b) the data required for testing often exceed the data available, a condition that has perhaps properly reduced the priority of model verification in past studies. However, expected developments in data collection and manipulation (2) suggest that future studies will have at their disposal sufficient data for model testing and evaluation.

For these reasons, this paper does not attempt to evaluate these land use forecasting models, or to review the evaluation procedures in use. Rather, one requirement of verifiability is identified, and two techniques for satisfying the requirement are proposed. Then, data requirements for verification of models are reviewed. The proposed model verification and evaluation procedures should lead to an increased ability to verify forecasts and provide a basis for evaluating the model itself before test forecasts are prepared.

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

REQUIREMENTS AND PROCEDURES FOR MODEL VERIFICATION

Requirements for verification of forecasting models are discussed by Theil (11). Verifiability is defined as that characteristic of a model which makes it possible to conclude, after a certain time and in an unambiguous manner, whether the forecast has turned out to be correct or incorrect. It must be possible also to verify the procedure by which the prediction itself was derived. From this definition, Theil derives a series of requirements of forecasts and forecasting models for verification. These requirements are quite basic, and include such concepts as the explicit statement of the model, specification of the time to which the forecast refers, and internal consistency of multiple forecasts, requirements satisfied by nearly all land use forecasting models.

A requirement not satisfied by most land use forecasting models concerns the specification of the relationship between the forecast and the actual outcome. For the forecast to be verified at all, it is necessary that a probability statement relating the forecast and observed values be given. For point predictions, it is not expected that the prediction will coincide with the observed value, so it is necessary to specify the distribution of error around the prediction if an evaluation is to be made. For interval predictions, the probability that the observed value will fall within the predicted interval is required.

Inasmuch as the land use forecasting models already mentioned do not provide, as part of the forecast output, statements concerning the probability distribution of forecast variable, it is not possible to evaluate the performance of these models. It should be noted that measures of accuracy such as the coefficient of determination (R^2) and the root-mean-square error are often applied to compare observed outcomes with predicted values from these models, or to test how well the calibrating data fit the model. Definitions and analyses of these accuracy measures may be found in Irwin and Brand (6), as well as Theil (11). For comparing the performance of alternative models, these accuracy measures may be of limited value. For evaluating a single model, however, in the absence of any confidence statement about the forecast variable, there is little basis for evaluating the magnitude of these accuracy measures.

In addition to permitting model verification, the specification of confidence statements on the forecast variable permits the model to be evaluated in another sense. Since the land use forecast is an input to trip forecasting models, the distribution of the error about the forecast must be sufficiently small so that these inputs are meaningful. Generally, the acceptable level of error for land use forecasts as an input to the trip-forecasting procedures can be determined. As presently formulated, land use forecasting models do not provide for such confidence statements, so the forecasts cannot be evaluated, even on this basis.

Two procedures for providing confidence statements for land use forecasting models are described. These procedures have shown potential merit in preliminary investigations. The first procedure applies to the class of models for which the parameters are estimated outside the model and independently of each other. Such models are hereafter referred to as "nonstatistical" models. An excellent example is the Pittsburgh model developed by Lowry (10); another model of the same type is the direct trip allocation model developed by Lathrop et al. (9), which is used for explanatory purposes here.

Models in which the parameters are estimated simultaneously so as to maximize the goodness-of-fit to the data are a second class of models. These "statistical" models, such as the simultaneous equation regression model of Hill et al. (5), commonly employ one of the least-squares techniques to estimate the model parameters. The selection of the regression estimation technique to be used is not discussed here. The problem of interest is how to provide confidence statements for the forecasts.

Confidence Statements for Nonstatistical Models

The objective of this procedure is to provide confidence statements for the forecast variables. A complete specification of the probability distribution of the forecast variable constitutes the most desirable and complete type of confidence statement that can be made. If this is not possible, a specification of the variance of this distribution and the type of distribution will permit precise confidence statements to be made. A less desirable, but still quite useful, statement is the specification of a confidence interval and the probability that the observed value will fall within this interval. The procedure produces confidence statements of any of the above types, depending on the form of the confidence statements on the variables and parameters that are inputs to the model.

The direct trip allocation model (9), an example of a nonstatistical model, is used to illustrate the procedure. This model is defined by the equations

$$A_j = A\left[e^{-\ell O^{(j)}} - e^{-\ell(O^{(j)} + O_j)}\right], \; j = 1, 2, \; . \; . \; . \; , N \tag{1}$$

where

A_j = the amount of activity to be allocated to zone j;

A = the total amount of activity to be allocated within the study area;

ℓ = probability of a unit of activity being sited at a given opportunity;

O_j = opportunities in zone j;

$O^{(j)}$ = the opportunities for siting a unit of activity rank-ordered by access value and preceding zone j = $\sum_{i=1}^{j-1} O_i$; and

N = the number of zones in the study area.

On the basis that the parameters ℓ, O_1, O_2, . . . , O_N, are to be estimated outside the model and independently of each other, the following approach is proposed. First, it is asserted that it is unrealistic to attempt to produce point estimates of the parameters. Specification of point estimates requires more than is known about these parameters and at the same time discards valid information on the range and variance of the parameter value. One can, however, hope to obtain an interval of values for a parameter and a corresponding probability measure of that interval.

A further extension of this approach leads to a probability distribution on the range of possible values for each parameter. Evidently, the most meaningful way to express the available information about the parameters is to state that the quantities ℓ, O_1, . . ., O_N are independent random variables with distributions determined outside the model. These distributions are not necessarily independent, and in fact may be highly interdependent. Conceptually, this interdependence could be satisfied by specifying multidimensional distributions on the input variables. Operationally, however, this approach is not realistic for more than two or three variables. Therefore, as in the original model itself, the assumption of independence is necessary to the solution of the problem. The variables A_j, j = 1, . . ., N, being functions of the random variables ℓ, O_1, . . ., O_N, are also random variables with distributions made up of the same parameters because of the functional relations comprising the statement of the model.

In general, it is quite difficult to determine analytically the probability distribution of a random variable that is a function of several other random variables; however, one can develop an empirical solution using Monte Carlo techniques (7).

Empirical distribution of the random variables, A_j, can be constructed as follows:

1. Perform a sequence of independent sampling experiments that yields an observation $\left(x_0^{(1)}, x_1^{(1)}, \ldots, x_j^{(1)}\right)$ in which $x_0^{(1)}$ is a random number generated by the probability distribution of the random variable ℓ, and $x_i^{(1)}$, (i = 1, . . ., j) is a random number generated by the distribution O_i.
2. According to Eq. 1 compute

$$A_j^{(1)} = A\left[e^{-x_0^{(1)} \sum_{i=1}^{j-1} x_i^{(1)}} - e^{-x_0^{(1)}\left(\sum_{i=1}^{j-1} x_i + x_j\right)}\right] \tag{2}$$

3. Repeat these first two steps T times and obtain a set of T random numbers $\left(A_j^{(1)}, A_j^{(2)}, \ldots, A_j^{(T)}\right)$ from the distribution of A_j.

4. Organize these T random numbers into a frequency distribution and study its properties, in particular its mean and variance. This empirical distribution is an approximation of the exact, theoretical distribution of A_j.

5. Repeat Steps 1 to 4 for each $j = 1, \ldots, N$.

Following the development as outlined, one obtains probability distributions for the A_j's, rather than point predictions. In this manner, the amount of uncertainty associated with the individual parameters after they have been estimated is retained and translated into the uncertainty about the A_j's. One can evaluate the quality of the prediction in terms of the dispersions of the distributions of the A_j's, the dispersion or variance of the distribution being a measure of the precision of the forecasting model. A model that yields a predicted distribution for A_j with small dispersion permits one to establish upper and lower bounds on A_j that are fairly close with high probability. On the other hand, a model that results in a relatively flat distribution for the predicted variable restricts one to establishing upper and lower bounds that are close only with low probability; the higher the probability desired, the lower the precision obtainable.

The Monte Carlo approach also provides a means of tightening the forecast distributions. By carrying out this distribution sampling experiment under a variety of specifications of the distributions of the independently estimated parameters, it is possible to obtain an indication of how much the individual parameter distributions are contributing to the dispersion of the forecast variable. Thus, by working "backwards," one can obtain the required specifications for the input distributions. These specifications can then indicate goals to be met when estimating the parameters and determine where it is most economical to invest estimation effort.

Confidence Statements for Statistical Models

Confidence statements for the dependent variable in regression models (the only statistical models considered here) may be derived for the classicial treatment of univariate multiple regression models as follows. It is assumed that the mean of a random variable, $y_{\underline{x}}$, denoted by $m(y_{\underline{x}})$, is a linear function of the k elements of a nonrandom vector $\underline{x} = (x_1, \ldots, x_k)$. A sample of n independently observed points $(y_i, x_{1i}, \ldots, x_{ki})$, $i = 1, \ldots, n$, is used to obtain the maximum likelihood estimates of the parameters σ, α, $\beta_1, \ldots, \beta_k$, where the linear hypothesis about the relation of y to $(x_1, \ldots, x_k)$ is

$$m(y_{\underline{x}}) = \alpha + \sum_{j=1}^{k} \beta_j x_j \tag{3}$$

and where σ is the standard deviation of $y_{\underline{x}}$ about the mean function at any point $x_1, \ldots, x_k$.

The application of the general method of confidence limits to this model yields the interval (Y_T, Y^T), which is a γ percent tolerance interval of confidence level p/100, if the probability of this interval (Y_T, Y^T) containing γ percent of the possible values of $y_{\underline{x}}$ is equal to p/100. The interval (Y_T, Y^T) is determined by the formula

$$\hat{y}_{\underline{x}} \pm k S_{\hat{y}_{\underline{x}}} \tag{4}$$

where $\hat{y}_{\underline{x}}$ denotes the estimate of $m(y_{\underline{x}})$ and $S_{\hat{y}_{\underline{x}}}$ denotes the standard error of the estimate $\hat{y}_{\underline{x}}$. This interval specifies a range within which γ percent of the possible values of $\hat{y}_{\underline{x}}$ will occur with probability p/100. The number k depends on γ, p, and the sample size, n. Once the values of $\hat{y}_{\underline{x}}$, $S_{\hat{y}_{\underline{x}}}$, and $k(\gamma, p, n)$ have been determined and used in Eq. 4 to obtain the interval (Y_T, Y^T), one can state with p percent confidence that γ percent of the observed values of the random variable $y_{\underline{x}}$ lie within k standard errors of the prediction, $\hat{y}_{\underline{x}}$.

As an example, consider the problem of forecasting the number of households per zone, $m(y_{\underline{x}})$, and evaluating the forecasting model by comparing this predicted value with the observed outcome. If the model is valid, the observed outcomes will fall within the 95 percent tolerance limits (Y_T, Y^T) with probability 0.99, for example. An acceptable model is one for which observed values are so located, and the tolerance limits are sufficiently narrow, at an acceptable level of probability, to provide meaningful inputs to the trip forecasting methodology.

The confidence statements for this classical regression model are limited in that the problem formulation does not refer explicitly to forecasting future values of the dependent variable. An alternative approach to confidence statements, described by Zellner and Chetty (12), specifically considers this prediction problem. In this Bayesian formulation, the probability distribution of the next q observations is derived on the basis of the past n observations. Procedures for employing this distribution to make inferences about the future values of the dependent variable are developed. Specifically, the variances of the predictions is a function of the original n observations as in the classical treatment, plus the q observations on the independent variables for the forecast period. From these variances, confidence intervals can be constructed for each of the q predictions.

Data Requirements for Model Verification

In general, land use forecasting models use cross-section data, that is, observations on the model variables for each analysis zone for a given point in time.

Data requirements for model verification may be considered in terms of two classes of forecasts—incremental growth estimates and total estimates of activity. The incremental model estimates of the increment of growth by zone over a specified time period, whereas the total activity model estimates the total activity for a given point in time. Most land use forecasting models are of the incremental type in that they incorporate the land use pattern for some base year into the model. However, models that allocate the total activity, such as Lowry's Pittsburgh model, have been developed.

Verification studies of incremental models require a minimum of two sets of cross-section data, in addition to the data used for model calibration. These models may require data on specified activities for the second as well as the first time period. For example, an incremental residential and retail model might require inputs of the total activity patterns for the base year and the basic employment pattern for the forecast year. The ability to test the reliability of these models increases in proportion to the number of cross-sections of data available.

The total allocation models, on the other hand, require only one set of cross-section data for verification purposes. These models specify certain activities as inputs and estimate the total amount of the remaining activities. Although this type of model requires fewer data, it is much more demanding in that the total structure of activity location, which may reflect past location decisions made under outmoded technologies, is to be reproduced. In view of the vast changes in transportation and building technology over the past several decades, this requirement may be indeed difficult. Inasmuch as the principal interest in land use forecasting is the preparation of a land use forecast based on the existing land use pattern, it is expected that the incremental model will continue to dominate the course of model development.

Model verification studies of the types proposed depend on data inputs over and above the cross-section data described earlier. In particular, for the Monte Carlo studies, it is necessary to determine the probability distribution of the model parameters. In some cases, this may only require a specification of the measurement error for the data. If the data are based on a sample, as contrasted with an inventory, the specification of the sampling error is required. Several models also include parameters which cannot be measured directly, such as the probability of siting, ℓ, in Lathrop's direct trip allocation model. Nevertheless, the specification of the probability distribution of ℓ is required for these verification studies. A logical choice in this case is the β distribution because it has a range from 0 to 1, as does the probability ℓ. The specification of this distribution and its parameters should be regarded as an assumption about ℓ.

In summary, data requirements for model verification studies involve more than the availability of a sufficient number of cross-sections, a demanding requirement in itself. Also needed is the information necessary to characterize the distributions of the observed variables and parameters, and to support assumptions on other parameters. Fulfillment of these requirements will demand more and more concern with sampling methods, measurement of sampling error, analysis of data characteristics, and related data problems by urban transportation studies.

CONCLUSION

The preliminary research reported here suggests several procedures for evaluating land use forecasting models by providing for confidence statements on forecasts. Through the use of these procedures it can be determined whether these models are providing forecasts of sufficiently low variability, let alone accuracy, to be of value. In addition, the Monte Carlo technique provides a means of exploring the properties of the models themselves.

It is also clear from the brief analysis of data requirements that verification of these models depends heavily on many more data than are currently available. However, it is expected that this situation will be alleviated considerably in the next few years; hopefully, the verification techniques necessary to fully exploit these data will be developed in the interim.

REFERENCES

1. Donnelly, T. G., Chapin, F. S., Jr., and Weiss, S. F. A Probabilistic Model for Residential Growth. Institute for Research in Social Science, Univ. of North Carolina, Chapel Hill, 1964.
2. Garrison, W. L. Urban Transportation Planning Models in 1975. Jour. of the Amer. Inst. of Planners, Vol. 31, No. 2, 1965.
3. Hald, A. Statistical Theory With Engineering Applications. John Wiley, New York, 1952.
4. Harris, B., ed. Urban Development Models: New Tools for Planning. Spec. Issue, Jour. of the Amer. Inst. of Planners, Vol. 31, No. 2, 1965.
5. Hill, D. M., Brand, D., and Hansen, W. B. Prototype Development of a Statistical Land Use Prediction Model for the Greater Boston Region. Highway Research Record 114, pp. 51-70, 1966.
6. Irwin, N. A., and Brand, D. Planning and Forecasting Metropolitan Development. Traffic Quarterly, Vol. 19, No. 4, 1965.
7. Johnston, J. Econometric Methods. McGraw-Hill Book Co., New York, 1960.
8. Klein, L. R. A Textbook of Econometrics. Row, Peterson and Co., Evanston, Illinois, 1953.
9. Lathrop, G. T., Hamburg, J. R., and Young, G. F. An Opportunity-Accessibility Model for Allocating Regional Growth. Highway Research Record 102, pp. 54-66, 1965.
10. Lowry, Ira S. A Model of Metropolis. The RAND Corp., Santa Monica, 1964.
11. Theil, H. Economic Forecasts and Policy. North-Holland Pub. Co., Amsterdam, 1961.
12. Zellner, A., and Chetty, V. K. Prediction and Decision Problems in Regression Models from the Bayesian Point of View. Jour. of Amer. Stat. Assoc., Vol. 60, No. 310, 1965.

Methodology for Developing Activity Distribution Models by Linear Regression Analysis

DONALD M. HILL, Senior Research Analyst, and
DANIEL BRAND, Senior Project Engineer, Traffic Research Corporation

•A PROPOSED mathematical framework for developing urban activities distribution models is described. The models distribute forecast regional totals of socio-economic variables to small zones; for example, resident population by various income levels would be distributed to traffic zones. The distribution is carried out as a function of future public policies relating to highway and rapid transit improvements, public open space, etc.

To calibrate activities distribution models, information over a historic time interval on growths and declines of the activities to be distributed is needed. Thus changes in zonal values of activities, and similar changes in the policy variables to be tested are the information with which the models are calibrated.

This paper describes a methodology for developing an activity distribution model by linear regression analysis. A simple example of the regression model is the linear equation constructed with three variables

$$\Delta R = a + b_1 \Delta Z_1 + b_2 \Delta Z_2$$

where R is the measurement of growth or decline of a land use activity; ΔZ_1 and Z_2 reflect changes in measurable and causal factors; and a, b_1 and b_2 are parameters derived by application of the least squares principle. The best values of a, b_1 and b_2 are established to minimize the expected error of estimate of ΔR by solution of the equation with known values of ΔZ_1 and ΔZ_2.

However, by the use of linear regression analysis, it is frequently argued that the model builder is seriously limited in the flexibility of the model's construction. Critics of regression analysis are quick to point out the following troublesome restrictions of regression analysis.

1. Linear relationships must exist between the dependent variable ΔR and the independent variables ΔZ_1 and Z_2.
2. The effects of the independent variables are additive and the ΔZ_1 and ΔZ_2 variables must not be interrelated with one another. Furthermore, the errors of estimate of ΔR from values of ΔZ_1 and ΔZ_2, must be normally distributed with mean zero and constant variance.

In view of these restrictions, it is argued that the advantages of regression analysis are soon canceled by the violation of one or more of the above restrictions in using a particular data set.

Evidence is presented that the above restrictions are not insurmountable obstacles in the development of a linear regression model. If any of the restrictions are violated due to the nature of the data, which appear to invalidate the construction of a linear model, then the model can be reformulated to avoid such violations. For example, the following precautionary procedures are possible:

1. Nonlinear relationships between ΔR and ΔZ variables can be linearized by breaking up the single ΔZ variable into several ΔZ variables, i.e., ΔZ_1, ΔZ_2, ΔZ_3, etc. By

---

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

doing so, a linear relationship will exist between ΔR and each ΔZ. Transformation of the ΔZ variable by logarithms, cosines, etc., can achieve the same results.

2. The application of factor analysis techniques can create from highly interrelated ΔZ_1 (adj) and ΔZ_2 (adj) variables which are independent of one another. In so doing, the assumption of additive effects of independent variables is confirmed. If such techniques are not available for use or not preferred, then the expected errors of estimate of ΔR which have unequal variances can be dealt with satisfactorily by suitable transformations of the ΔR and/or ΔZ variables to insure constant variance for expected errors of estimate.

Explicit analysis of locational behavior can be incorporated in the model's design. Regression models do not have to depend primarily on a blanket interpretation of past events. The model's development can be shaped in accordance with a theory of allocation of growth of activities or urban development. The researcher in the development of the model will be back and forth between the theory of the model and tests of its behavior with data. Adjustments of the theory will result to improve the model's application with empirical data. However, the theory of the model should not be warped or distorted solely to achieve a best fit to the data.

Development of the model can be achieved by applying several types of regression analysis techniques; for example: (a) ordinary least squares, (b) indirect least squares, (c) limited information-single equation method, (d) 2-stage least squares, (e) simultaneous least squares, and (f) full information maximum likelihood method.

While method (a) deals with single equation models, methods (b) to (f) deal with models formulated as systems of simultaneous equations. If single equation models are formulated, method (a) is adequate and the one to use. However, most activity distributions require models formulated as systems of equations—methods (b) to (f). The relative efficiency of each of the methods for parametric estimation is discussed in the case of simultaneous equation models.

There are distinct computational and economic advantages associated with the use of linear regression analysis. Readily available analysis methods and economical computer programs can be used by the researcher for the model's development. Also, through the economies and flexibility of regression analysis techniques, several test models can be easily evaluated. In general a great deal of knowledge and modeling experience can be gained from constructing and testing regression models.

MODEL DESIGN BY LINEAR EQUATIONS

In the typical model design, one must choose a mathematical framework to describe a hypothesized set of structural relationships. This framework will comprise the variables chosen and specify the ways in which these variables are interrelated. A model framework convenient for use is a linear structural equation as follows:

$$\Delta R = b_1 \Delta Z_1 + b_2 \Delta Z_2 + \ldots + b_k \Delta Z_k + u \qquad (1)$$

Here ΔR is an urban activity variable dependent on the measurements of a number of independent variables, $(\Delta Z_1, \Delta Z_2, \ldots, \Delta Z_k)$. The parameter set $(b_1, \ldots, b_k)$, describes the relationship between the dependent variable and the independent variable set. The error term, u, occurs due to the imperfect fit of a mathematical equation to observed phenomena of urban development. It is the principle of model calibration to estimate the parameter set $(b_1, \ldots, b_k)$, so as to minimize overall the error terms, u, as well as to eliminate systematic bias in the error terms.

Eq. 1 accommodates adequately the situation where the dependent variables, ΔR, to be predicted, are not interrelated with one another. However, many model designs are premised on the occurrence of interrelationships between the dependent variables to be predicted. In accordance with this design requirement, it is desirable to formulate a framework of simultaneous linear equations; for example:

$$\begin{aligned}
\Delta R_1 + a_{12}\,\Delta R_2 + \ldots + a_{1m}\Delta R_m &= b_{11}\,\Delta Z_1 + \ldots + b_{1k}\,\Delta Z_k + u_1 \\
a_{21}\,\Delta R_1 + \Delta R_2 + \ldots + a_{2m}\Delta R_m &= b_{21}\,\Delta Z_1 + \ldots + b_{2k}\,\Delta Z_k + u_2 \\
a_{m1}\Delta R_1 + a_{m2}\Delta R_2 + \ldots + \Delta R_m &= b_{m1}\Delta Z_1 + \ldots + b_{mk}\Delta Z_k + u_m \qquad (2)
\end{aligned}$$

Within this framework, it is possible to account for the interrelationship between the dependent variables, $(\Delta R_1, \ldots, \Delta R_m)$, as well as accommodate the dependency of each ΔR variable on the independent variable set, $(\Delta Z_1, \ldots, \Delta Z_k)$. As in the case of Eq. 1 (which of course is a special case of equation system Eq. 2 where $a_{ij} = 0$ for $i \neq j$) the error terms, u, must account for the imperfect fit by the mathematical equation. The parameter sets $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ are estimated so that the overall errors (u_i) are minimized by the regression process of least squares.

The selection and formulation of variables in the model is critical in the model's design. The dependent variables should measure adequately the distribution which we propose to predict. The independent variables should provide adequate explanation of the distribution to be predicted, as well as retaining their separate identity with respect to one another. In particular, the following two criteria are suggested for the formulation of variables:

1. The variables formulated for incorporation into the model should be the same type. That is, variables which are changed in basically different ways by changes in definition of subregional areas and size should not be mixed in a single model. Variables will in general be of two types, i.e., point variables and aggregate variables. Point variables do not tell anything about area aggregates unless multiplied by some base quantity such as total land or total activity. Examples of point variables are densities, accessibilities, and area rate of growth. Area aggregate variables, on the other hand, refer to measurable magnitudes or quantities. Examples of aggregate variables are total population, and total employment or total land area.
2. The construction of the variables should be such that their interpretation is clear. The variables must be capable of being measured and named. Data categories assimilated to form a variable should furnish it with a logical name or explanatory description.

The formulation of variables should simplify the design of the model wherever possible. If two or more variables demonstrate similar locational characteristics and otherwise appear to cluster together due to a similarity in name and procedure of measurement, it is desirable to aggregate the variables into a single variable. Clustering or aggregating dependent variables will simplify the model design by reducing the number of estimating equations of the system. There must be one equation in the model for every dependent variable to be predicted. By aggregating dependent variables, it is possible (all else being equal) to increase substantially the predictive accuracy of the model over what might be achieved with a more complex model. Aggregation of independent variables which are highly interrelated is preferred for other reasons.

CRITERIA FOR APPLYING LINEAR REGRESSION ANALYSIS IN MODEL DESIGN

Linear regression analysis is simply defined as the estimation of the value of one variable (ΔR) from the values of other given variables (other ΔR and/or ΔZ) via a framework of some chosen linear equation. Descriptions of various regression techniques suggested for use in distributing urban activities are described hereinafter. Such regression analysis may be used provided the following criteria are met:

1. It is hypothesized in the construction of the activity distribution model that linear relationships exist between the dependent and independent variables.
2. It is hypothesized that the influences of the variables are additive. While the dependent variables are assumed to be interrelated with each other as well as being

related to the independent variables, it is desirable for the independent variables not to be interrelated with each other.

Linear Influences of Variables

In the application of regression analysis to estimate the parameters of a model it is essential that there is a linear relationship between the expected value of the dependent variables and the independent variables. Fortunately, even when this condition does not apply, it is often possible to modify the original variables in some way so that the new variables meet the requirement. The modifications or transformations of data most commonly applied are the logarithmic, the square root, or the reciprocal.

One of the assumptions of the linear model is the serial independence of the error terms, u, that is, covariance $(u_i, u_{i+j}) = 0$ for all observations i and j, where $j = 0$. However, there are circumstances in which the assumption of a serially independent error term may not apply. It is possible that one may make an incorrect specification of the form of the relationship between dependent and independent variables. For example, one may specify a linear relationship between the ΔR and ΔZ variables when the true relation is quadratic. While the error term in the true relationship may be non-autocorrelated, the new quasi-error term associated with the linear relationship must contain a term in ΔZ^2. If serial correlation exists in the ΔZ-values (i.e., characteristic of time series variables), then serial correlation will occur in the quasi-error terms.

In cases of autocorrelated errors, there are three main consequences of applying straight-forward regression processes without transforming the variables affected:

1. While the estimates of the parameters will be unbiased, their error variances could be larger than those achievable by applying suitable transformations in the estimation process.
2. The estimates of the error variances associated with parameters will be understated.
3. Inefficient predictions with large errors of estimation will be obtained.

The satisfactory manner of testing for linear relationships between the dependent and independent variables is by plotting the relationships between pairs of variables on graph paper. Based on the results, a decision can be made on the value of transforming variables, so as to linearize their influences.

Additive Influences of Independent Variables

Two variables exhibiting a high degree of interrelationship are said to introduce non-additive influences on the dependent variables. Unless interactance terms descriptive of the interrelationships are introduced in the model, there occurs serious ambiguity in the calibration process in separating the influences of the two variables. This ambiguity can be reflected in large fluctuations in the parameters associated with each model derived from calibrations with different aggregations of the subregions and variable sets, etc. Also, the signs associated with parameters of the affected variables may disagree from that expected from a priori reasoning.

Nonadditivity of a particular variable, unless previously eliminated, will frequently cause heterogeneity of error variance which is associated with the estimating equation for a particular dependent variable. This should not occur as it can have a serious effect on the parametric estimation achieved by regression analysis. Regression analysis may only be validly performed provided the error variance of the estimates of the expected value of a dependent variable is constant for all values of the independent variable (i.e., homogeneity of error variance is important).

The degree of interrelationship between variables can be measured in two ways: graphical analysis by plotting pairwise relationships on graph paper, and calculation of bivariate correlation coefficients. The value of the correlation coefficient will vary between minus unity and plus unity, and in either case as it approaches its limits, a high degree of interrelationship or correlation is indicated.

If two independent variables are correlated, one of three courses may be followed: (a) eliminate the one variable considered least important to the model design, or which one believes a priori to be less important; (b) combine the two variables, provided the new aggregate variable can be named and measured; (c) substitute a scale of a variable which is natural (i.e., which experience or theory suggests is additive) to reduce and even eliminate interactance between variables. Examples of transformation by logarithms or reciprocals have been shown to reduce interrelationships.

If it is considered important to include all variables in the model, then course (b) or (c) is preferred.

If course (b) is followed, factor analysis can be useful in aggregating variables into independent, and therefore, additive influences. The basis for conducting factor analysis is a matrix of correlation coefficients describing the pairwise relationships between all variables affected. Factor analysis processes will construct factors comprising a linear function or equation of the variables whose pairwise correlations are being analyzed. The principle for constructing these factors is such that the factors are statistically independent of one another. The factors should be able to be named and associated with an aggregate influence on urban development.

Heterogeneity of error variance, caused by nonadditivity, will usually be reflected by a relationship of the error variance to the mean (m) or expected value of the dependent variable for a particular independent variable. The choice of a suitable variable transformation will frequently depend on the relationship between the error variance

TABLE 1

SUMMARY OF TRANSFORMATIONS[a]

Variance in Terms of Mean m	Transformation	Approximate Variance on New Scale in Absence of Heterogeneity
m $\lambda^2 m$	$\sqrt{x}$ or $\sqrt{x + \frac{1}{2}}$ for small integers	0.25 $0.25\lambda^2$
$\lambda^2 m^2$	$\log_e x$, $\log_e (x + 1)$ $\log_{10} x$, $\log_{10} (x + 1)$	λ^2 $0.189\lambda^2$
$2m^2/(n - 1)$	$\log_e x$	$2/(n - 1)$
$m(1 - m)/n$	$\sin^{-1} \sqrt{x}$ (degrees) $\sin^{-1} \sqrt{x}$ (radians)	$821/n$ $0.25/n$
$km(1 - m)$	$\sin^{-1} \sqrt{x}$ (radians)	$0.25k$
$\lambda^2 m^2 (1 - m)^2$	$\log_e [x/(1 - x)]$	λ^2
$(1 - m^2)^2/(n - 1)$	$\frac{1}{2} \log_e [(1 + x)/(1 - x)]$	$1/(n - 3)$
$m + \lambda^2 m^2$	$\lambda^{1} \sinh^{-1} (\lambda\sqrt{x})$, or $\lambda^{-1} \sinh^{-1} (\lambda\sqrt{x + \frac{1}{2}})$ for small integers	0.25
$\mu^2 (m + \lambda^2 m^2)$	$\lambda^{-1} \sinh^{-1} (\lambda\sqrt{x})$, or $\lambda^{-1} \sinh^{-1} (\lambda\sqrt{x + \frac{1}{2}})$ for small integers	$0.25/\mu^2$

[a] Bartlett, M. S. "The Use of Transformation," Biometrics 3, 39-52, 1947.

and the mean of observations. This relationship is usually determined by empirical analysis with subregional data.

Table 1 gives transformations that have been found to have practical value.

SCOPE OF MODEL DESIGN

The design of the activity distribution model (1, 2) is based on a combination of deductive and inductive reasoning based on observations of urban development patterns. It represents an iterative procedure in which the analyst begins with general observations of subject matter; develops a hypothesis or theory of the causal system which explains the behavior of his subject matter; tests this hypothetical structure for its power to explain the observed data of his field, in this case urban development; studies carefully the discrepancies between the explanation provided by his hypothetical structure and the observed data; revises his hypothetical structure on the basis of these discrepancies; tests the structure again; etc. The analyst is thus back and forth between his theoretical explanation of the causal system and his observation of all possible aspects of the subject matter on urban development. His goal in this iterative process is to reduce the discrepancies between theory and observation to a minimum.

Identification of Equation Systems

The problem of identification in a system of causally interrelated variables is connected with making an empirical estimation of the system from observed data. The problem only exists for systems of simultaneous equations, and does not occur when the area of study can be fully explained by a single equation. Each equation in the system will be designed to explain one dependent variable of the system in terms of those causes which exert a direct or approximate influence on it. These causal variables include both other dependent variables, and independent variables.

The essential meaning of identification can now be stated. Any particular equation in our system is identified if it is sufficiently different from all of the other equations, i.e., in its form, the variables included in it, and any restrictions on the values which its parameters can take. By "sufficiently different" we mean that it must be impossible to arrive at an equation which "looks like" the particular equation we are testing by any linear combination of other equations in the system, or of all of the equations including the one being tested.

Sample Identification Problem. Suppose that our system consists of two dependent variables, ΔR_1, ΔR_2, and three independent variables, ΔZ_1, ΔZ_2, ΔZ_3. Suppose that we are assuming linear relations, and that we have as yet no clear ideas about structure specification. We might then simply put all variables in the system into each equation.

$$\begin{aligned} &(a)\quad a_{11}\Delta R_1 + a_{12}\Delta R_2 + b_{11}\Delta Z_1 + b_{12}\Delta Z_2 + b_{13}\Delta Z_3 = u_1 \\ &(b)\quad a_{21}\Delta R_1 + a_{22}\Delta R_2 + b_{21}\Delta Z_1 + b_{22}\Delta Z_2 + b_{23}\Delta Z_3 = u_2 \end{aligned} \tag{1}$$

The a's and b's are constant coefficients or parameters, and the u's can be treated here as either constant terms or as random disturbances. We can assume that Eq. (a) is supposed to explain ΔR_1 and that Eq. (b) is intended to explain ΔR_2. Let us further assume that the system we are analyzing is represented by a sample of observed data.

$$\left[\Delta R_{it}\right],\ \left[\Delta Z_{jt}\right]\ (i = 1,\ 2;\ j = 1,\ 2,\ 3;\ t = 1,\ 2,\ \ldots,\ T) \tag{2}$$

We now attempt to use these data to estimate the parameters of our system (1) above. But since the two equations look exactly alike, when we apply our observed data to the estimation of parameters we get exactly the same result for each equation. There is no way of distinguishing the behavior of one part of the system from that of the other using empirical methods.

Suppose, next that we do more work on the theory of our system, and arrive at a specification which excludes ΔZ_2 and ΔZ_3 as variables from (a) and ΔZ_3 from (b). Let us call the new equations (c) and (d). Now the two equations "look different" from each other. We have restricted b_{12}, b_{13} and b_{23} to zero. This is the most common kind of restriction which aids identification. But is there still any danger of getting those two equations mixed up in empirical estimation? Suppose we test by making a linear combination of (c) and (d). Thus suppose we form ℓ(c) + m(d) where ℓ and m are arbitrary multipliens. The resulting equation has the form

$$a_1\Delta R_1 + a_2\Delta R_2 + b_1\Delta Z_1 + b_2\Delta Z_2 = v \tag{3}$$

This is different from the new specification we have made for (c), for it excluded both ΔZ_1 and ΔZ_2, but it is no different from our new specification for (d). In our new system (c) is completely distinguishable empirically from the rest of the system, but (d) is not. Therefore (c) is identified, and (d) is not identified.

Now suppose that our theoretical specification had removed ΔZ_2 and Z_3 from (a) and ΔZ_1 from (b), giving equations (e) and (f). Suppose we make a linear combination ℓ(e) + m(f),

$$a_1\Delta R_1 + a_2\Delta R_2 + b_1\Delta Z_1 + b_2\Delta Z_2 + b_3\Delta Z_3 = w \tag{4}$$

This form does not look like either (e) or (f), and both equations in our system are fully distinguishable and hence identified.

In conclusion, the main basis for identification is the inclusion of only the main causal variables in each equation, and the exclusion of irrelevant variables, both dependent and independent. But there are other bases for obtaining distinguishability of one equation from all others, and these include cases like the following. It might be that there is a natural restriction that two parameters in the equation have a preordained ratio to each other, or that one or more parameters have preordained values, indicated by theory, or arrived at by separate studies. Sometimes a nonlinearity in an equation may insure identifiability, or even a specification of differences in the variances of the random components in particular equations may achieve this.

A necessary, but not sufficient, condition of an identified system of m equations is that in each equation, at least m - 1 of the variables are restricted, usually by setting them to zero. This is known as the "order" condition of identifiability. If fewer than m - 1 variables are restricted in any equation, the system is said to be under identified, and cannot be solved by the parameter estimation programs. If more than m - 1 variables in some equations and at least m - 1 variables in all equations are restricted, the system is said to be over identified. This will usually be the case with activity distribution models.

Methods of Identifying a Model. By and large, the identification of the system of simultaneous equations which comprise the model will be determined by a priori reasoning in support of a particular theory of urban development. These are, however, empirical tests which can be applied as a guide in choosing an appropriate identification for the model.

Tests of Model Design

The testing of the model is usually carried out by regression processes, such as least squares (LS) or maximum likelihood (ML). Their purpose is to make the best possible tests and estimates of the structural parameters associated with variables of the model. In doing so, a complete separation is sought between the systematic part of the relationships and the random part. Generally, testing can profitably begin with an examination of our estimates of the random component.

An examination is conducted of nonsystematic residuals of the equation which the estimation process may have produced. If these reveal any trend, cycle or sawtoothed behavior then the model design (i.e., its identification) is on this basis rejected. It is concluded that the model does not contain all of the systematic forces which affect the dependent variable being explained, or it may contain some forces which should not be there.

Next, one examines the standard errors of the parameters attributable to variances associated with the observed data and conducts accompanying t-tests of significance. Here one tests again the model design, this time to see which variables test out as significant and as causes affecting dependent variables. But these tests can only be suggestive rather than rigorous, if our residual has already tested to be nonrandom and containing systematic elements.

In making the tests of significance of parameters (and hence of the associated causes), the model design can be open to two types of error. First, the test may reject a design which is really appropriate. This is the well-known Type I error. It can arise because the source of data is not complete or adequately representative of subregional development patterns. Application of more representative data, with an appropriate level of significance can reduce this danger.

A second kind of error which one may make is to accept a design which is false. This is the Type II error. Some other identification of the model is correct, but the one chosen has produced estimates which happen to fall into the range of acceptance for the model. Here we have an identification error which could slip by the tests.

Finally, one tests the results at this stage through reapplying to them one's knowledge of the subject matter. On the basis of general observation of the pattern of development, and of the tests of the primary model design on this basis, one achieves concepts about the sizes and signs of the parameters associated with the variables of the model. If the regression tests produce results which are markedly different from expected, one must take this as a rejection of the model design, or otherwise as some combination of data error and error in the model's identification. Consequently, it is such rejections which lead the analyst forward in the iterative process of model testing.

During this process of iterative revisions of the identification, there is always the danger of warping the theory, and hence design, to make the model fit the particular data source. This is a real trap, and no doubt one could fall into it. But there is a defense against it. The defense lies in carefully preserving the strength, logic and realism of the model's design. It is only when the observed data, and the discrepancies or residuals between observed data and the systematic explanation, reveal some clearly relevant but hitherto unsuspected force or omitted force that the identification should be revised. Design should never be altered merely to get a good statistical fit when the theoretical underpinning of such alterations is weak, illogical, and unrealistic.

When the scientific process has reached a terminal stage, one should have minimal identification errors, and hence the estimates of standard errors of estimates should be realistic. During the process one has resisted rejecting a good theory on the basis of statistical tests, while at the same time one has been even more resistant to warping a design solely to get good statistical fits. The systematic model should be in agreement without general observations and knowledge about the subregional development. And finally, the residuals should be in a purely random sequence, with mean zero and constant variance.

The test of successful estimation of the true model comes partly in its explanatory power, and partly in its predictive power. If one has found satisfactory causal explanation of development, and if the model is performing in a known way, one should be able to make satisfactory predictions.

REGRESSION PROCESSES

Development of the model can be achieved by applying several types of regression analysis methods.

Ordinary Least Squares

One applies ordinary least squares to a single equation in a model (3, Chap. 4, pp. 106-138), i.e.,

$$\Delta R = B\Delta Z + u \tag{1.1}$$

where

ΔR = vector of dependent variables;
ΔZ = vector of independent variables;
B = parameter associated with independent variables; and
u = residual error.

If, however, there are two or more dependent variables in each equation one does not know which dependent variable to select as the primary dependent variable of an equation, i.e.,

$$A\Delta R = B\Delta Z + u \qquad (1.2)$$

where A = parameters associated with the dependent variables. The remaining dependent variables are always correlated with the error term in the equation because of the simultaneous nature of the equations in the model. Therefore, ordinary least square estimators are always biased (estimate does not equal true value) and they will also be inconsistent—for increasing numbers of sample observations, the estimates continue to be biased (3, Chap. 6, pp. 148-150).

For these reasons ordinary least squares is considered to be an unsuitable estimation method for dealing with systems of simultaneous equations. On the other hand, when dealing with a single equation containing one dependent variable, it is the method to use.

Indirect Least Squares

In the situation where a system of simultaneous equations is exactly identified, this is the proper estimation method to use. The other simultaneous estimation methods to be mentioned below always provide identical estimators to the indirect least squares method for the case of exact identification (exactly m - 1 of the parameters are set equal to zero where m is the number of dependent variables in total). The indirect least squares method is less complicated than the other methods, hence it provides definite computation economies.

The procedure (4, Chap. 4.4, pp. 135-137) is to estimate the parameters of the reduced form equations by application of the ordinary least squares method. A reduced form equation has only one dependent variable which is defined as the primary dependent variable, i.e.

$$\Delta R = D\Delta Z + u \qquad (2.1)$$

By deciding that a certain number of the parameters in each equation of the simultaneous equation system are zero, the reduced form equations are converted into a simultaneous system where each equation contains one or more dependent variables, i.e., multiply (2.1) by A to obtain

$$A\Delta R = AD\Delta Z + Au \qquad (2.2)$$

Write B = AD; therefore (2.1) is converted into a simultaneous system

$$A\Delta R = B\Delta Z + u$$

To recap, the exact number of parameters per equation which are set equal to zero is m - 1.

Limited Information Estimation Methods

Limited Information Single Equation Method (LISE) or Least Variance Ratio Method (LVR). This is a limited information maximum likelihood approach. It is a maximum likelihood approach (4, Chap. 6.2, pp. 166-167) in that the logarithmic likelihood function for the dependent variable is defined, i.e.

$$L(\alpha) = \tfrac{1}{2}\log AWA' - \tfrac{1}{2}\log \alpha M\alpha' + k - \tfrac{1}{2}\log \text{determinant } W \qquad (3.1)$$

where

$$\alpha = [A, B]$$

$$M = \begin{bmatrix} M_{\Delta R\Delta R} & M_{\Delta R\Delta Z} \\ M_{\Delta Z\Delta R} & M_{\Delta Z\Delta Z} \end{bmatrix}$$

$$M_{\Delta R\Delta R} = M = \frac{1}{T}\sum_{t=1}^{T} \Delta R^2_t \text{ etc., } (T = \text{number of observations})$$

$$W = M_{\Delta R\Delta R} - M_{\Delta R\Delta Z}\, M^{-1}{}_{\Delta Z\Delta Z}\, M_{\Delta Z\Delta R}$$

Next, the function is maximized to yield uniquely the ratios of the parameters associated with the dependent variables of each equation. By setting one of the parameters equal to unity the remaining parameters are defined. The parameters of the independent variables are determined by solving a mathematical identity of dependent variable parameters and values of both dependent and independent variables.

The application of the method requires the user to know the specification of the single equation being estimated (i.e., which parameters are zero), and the independent variables appearing in the remaining equations which are assumed to have non-zero parameters. The detailed specification concerning the parameters of dependent variables in remaining equations is assumed unknown. Hence, only limited information needs to be known to obtain the estimators.

Two-Stage Least Squares

The basic idea (3, Chap. 9.5, pp. 258-260) of the 2-stage least squares (TSLS) is to select a dependent variable in each equation of the system and set its parameter equal to unity; i.e., rewrite $A\Delta R = B\Delta Z + u$

$$\Delta R_1 = A_2\Delta R_2 + B_2\Delta Z + u \qquad (3.2)$$

Next replace the remaining dependent variables by their estimates based on ordinary least squares regression between each dependent variable and all independent variables in the model,

$$\Delta R_2 = \Delta Z(\Delta Z'\Delta Z)^{-1}\Delta Z'\Delta R_2 \qquad (3.3)$$

Finally ordinary least squares is applied to the selected dependent variable, the regression estimates of the remaining dependent variables, and the independent variables in each single equation.

There is a basic similarity between LISE and 2-stage least squares as they both make use of all the independent variables in the model in order to estimate the parameters of a single equation, but do not require a detailed specification of the dependent variables in the remaining equations of the model. Both methods are consistent estimating methods. For large numbers of sample observations, both methods provide unbiased estimates of the parameters. It is reported that for special cases with a small number of observations, the 2-stage least squares method may provide more efficient estimators than LISE—estimators with smaller limiting variance (5).

Full Information Method (FI)

This method implies the use of full information concerning the specification of the simultaneous equation system. The FI methods are anticipated to provide the most efficient estimators of all the methods. There are two different techniques which comprise the FI method, simultaneous least squares and maximum likelihood.

Simultaneous Least Squares (SLS)

SLS (6) is a distribution free method of estimation (no assumption is made about the distribution of residual error). The method is the simultaneous equation counterpart of ordinary least squares. It takes completely into account the simultaneous interactions of all dependent variables in the system:

$$A\Delta R = B\Delta Z + u \tag{4.1}$$

It is a least squares method in that the sum of the squared deviations between observed and estimated dependent variables are minimized; i.e. minimize

$$E^2 = \sum_{t=1}^{T} \sum_{i=1}^{N} u^2_{sit}$$

where

$$u_S = A^{-1} u$$

Maximum Likelihood Technique (ML)

Complete information on the simultaneous system is taken into account (4, Chap. 5, pp. 143-162). The likelihood function for the dependent variables, conditional upon the values of the independent variables, is determined for the complete model. By assuming that the residuals of the estimating equations are multivariate normally distributed, the logarithmic likelihood function is defined,

$$L(\alpha, \sigma) = \log \det B - \tfrac{1}{2} \operatorname{trace} (\alpha' \sigma^{-1} \alpha M) + k - \tfrac{1}{2} \log \det \sigma$$

where

det = determinant;
$\alpha = [A, B]$;
σ = non-singular covariance matrix of residual error u; and
trace matrix R = $\sum_i r_{ii}$ (sum of diagonal elements).

Maximizing the logarithm of the function with respect to the parameters of the model and its residuals lead to difficult estimating equations.

There are two assumptions involved in the use of ML, which may restrict its application. The first is the assumption that the residual errors are multivariate normally distributed. While the distributions of the residuals are probably bell-shaped and may be asymptotically normal (a property of large samples), the assumption of normality is not closely met with small samples of data. The second assumption (7) concerns the optimal properties of structure estimation. If the residual errors are normally distributed, both maximum likelihood and least square techniques lead to identical results which are linear unbiased estimates. However, where the residuals are non-normal then the ML and LS estimators are quite different. Nevertheless, the LS estimates are still the best linear unbiased estimates.

In conclusion, SLS is preferred to ML because of its distribution free properties and secondly because of anticipated computer economies. The computation economies are achieved by using a truncated procedure of SLS. This gain will, of course, be at the small expense of loss of accuracy in estimation. In truncated SLS the results are accepted after two or three stages of the recursive procedure of SLS estimation.

FACTOR ANALYSIS PROCESS

Variables which possess high statistical association are grouped together in clusters called factors. In particular, the intercorrelations among all the variables under study constitute the basic data for factor analysis (8).

All variables are assumed to be in standardized form, i.e., each has a mean value of zero and a variance of unity. It is the object of factor analysis to represent a variable in terms of several underlying factors, by a simple mathematical model of the linear form,

$$\Delta_{Rj} \text{ or } \Delta Z_j = a_{j1} F_1 + a_{j2} F_2 + \ldots + a_{jf} F_f$$

Naming the Factors

The factors are not named by the process, and this anonymity must be removed before the statistical association indicated by factor analysis can be evaluated against the planner's a priori knowledge of cause and effect relationships. The variables which are most closely associated with (those which supposedly make the most significant contribution to) each cluster should help in naming the factor.

Significance of Factors

The relative importance of each factor is indicated by its eigenvalue, which represents a measure of the total contribution of the factor to the variances of all the variables being analyzed. Eigenvalues for all factors are produced by the technique. An eigenvalue of unity or greater is considered to indicate a significant factor. Experience with our prototype activities distribution models (1, 2) has shown, however, that there are a small number of factors with very high eigenvalues, a few more with eigenvalues of unity or more, and a large number of factors which have eigenvalues less than unity. The latter, strictly speaking, are considered little more than statistical "noise," whose contribution to the variances of the variables will generally be insignificant.

Selection of Factors

The factor analysis process provides for specifying the number of factors to be used. Normally the process discards all factors with an eigenvalue less than unity. In some instances, the arguments for using unity as a cutoff are marginal, and in some cases a factor with an eigenvalue less than one may be significant. With a small number of input variables, a factor with an eigenvalue of less than one could make a significant contribution to the variance of the variables. In such cases, one may specify the number of factors required.

Regardless of which cutoff option is employed, the eigenvectors associated with eigenvalues of the factors are computed and are normalized so that the squared eigenvector coefficients associated with each factor add to unity or less. The normalized eigenvector coefficients associated with each factor (known as factor loadings) are produced in array form.

Structure of Each Factor

The construction of factors is established by a regression procedure, based on the array of factor loadings. Each factor is presented as a linear function of the variables. An array is produced which indicates for each factor the statistical importance of each variable in its construction.

Factor Rotation

There is a possibility that several factors will look very much alike, and possess similar eigenvalues. In order to sharpen the picture of the system as much as possible, a varimax method of rotation is utilized in factor analysis processes. The rotation should maximize large factor loadings and minimize small ones, and the distinction between factors should be much sharper in the rotated than in the unrotated case. Both the unrotated and rotated arrays are true shadows of the same shape taken in different lights. Traditionally, the multiplicity of true shadows offered by factor analysis has deterred investigators from using the method as a "proof." The cautious investigator has assured himself that he uses it (in moderation) only to stimulate his insight into a mass of data; that is, to prompt a review of his logic.

It is emphasized that use of factor analysis is always subject to demand for a logical explanation of clustering. However, it seems intuitively attractive with the large amount of data available in computer-size models to suppose that the surface of the factor shape is sufficiently regular that the maxima found by rotation, if not the best view, is at least one of the good views.

REFERENCES

1. Hill, D. M., Brand, D., and Hansen, W. B. Prototype Development of a Statistical Land Use Prediction Model for the Greater Boston Region. Highway Research Record 114, pp. 51-70, 1966.
2. Irwin, N. A., and Brand, D. Planning and Forecasting Metropolitan Development. Traffic Quarterly, Oct. 1965.
3. Johnston, J. Econometric Methods. McGraw-Hill, New York.
4. Cowles Commission for Research in Economics. Studies in Econometric Method.
5. Nagar, A. L. The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations. Econometrika, Vol. 27, pp.575-595, 1959.
6. Brown, T. M. Simultaneous Least Squares: A Distribution Free Method of Equation System Structure Estimation. International Economic Review, Vol. 1, No. 3, Sept. 1960.
7. David, F. V., and Neyman, J. Extension of the Markoff Theorem on Least Squares. Statistical Research Memoirs, Vol. II, London: Dept. of Statistics, Univ. of London, Univ. College, pp. 105-116, Dec. 1938.
8. Harman, H. H. Modern Factor Analysis. Univ. of Chicago Press, 1960.

A Sampling Technique for Updating a Quantitative Land Use Inventory

SAM M. HADFIELD, Survey Design Analyst, and
JOHN D. ORZESKE, Chief, Data Collection and Maintenance, Chicago Area Transportation Study

Transportation studies involved in the continuous planning process are required to update, periodically, their data inventories to fulfill their responsibilities for continuing review and evaluation of their transportation plans. One of the major inventories used in this process is that of the categories and quantities of land use and its relationship to trip generation. Basic detailed land use inventories have to be examined regularly and developmental changes analyzed for their effect on transportation plans. Because of the time and personnel requirements, it often is not economically feasible to perform a detailed resurvey of the area. This paper describes how an area sampling technique which had been employed previously in flood plain mapping and other geographical surveys was adapted to the land use updating problem at the Chicago Area Transportation Study. It illustrates how a stratified, systematic, unaligned pattern of points or "dots" was superimposed on aerial photography to yield reliable estimates of the proportions of various land uses within an area.

•AFTER completion of the original work program of data collection and analysis, publication of reports, and the development of a comprehensive transportation plan for an area, a transportation study group is charged with continuing the planning process by regular review and evaluation of its transportation plan. The study also must maintain the currency of the data on which the plan was based, and test proposals for modifying or expanding the plan. It is charged further with a continuing research program aimed toward improving study methods and the development of new analytical techniques.

In the performance of these duties, the staff of the continuing study is dependent on the basic inventories of travel, transportation systems, and land use, which are compiled at the earliest stages of the original study through a massive data collection program involving home and roadside interview surveys, etc. The land use survey is an integral part of the overall study design. The relationships between trips and land use are measured and these relationships then are applied to plans and estimates of the future distributions of land use in order to forecast future trip volumes and travel patterns.

TYPICAL LAND USE SURVEY METHODS

Customarily, the original land use inventory is the result of a lengthy process involving field listings and an extensive effort in identifying, classifying, coding, and measuring land areas on maps and aerial photos. This process, as employed in the 1956 Chicago Area Transportation Study land use survey, employed Sanborn fire insurance maps in the city of Chicago and aerial photography in the suburban portion of the study area as its basic source materials.

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

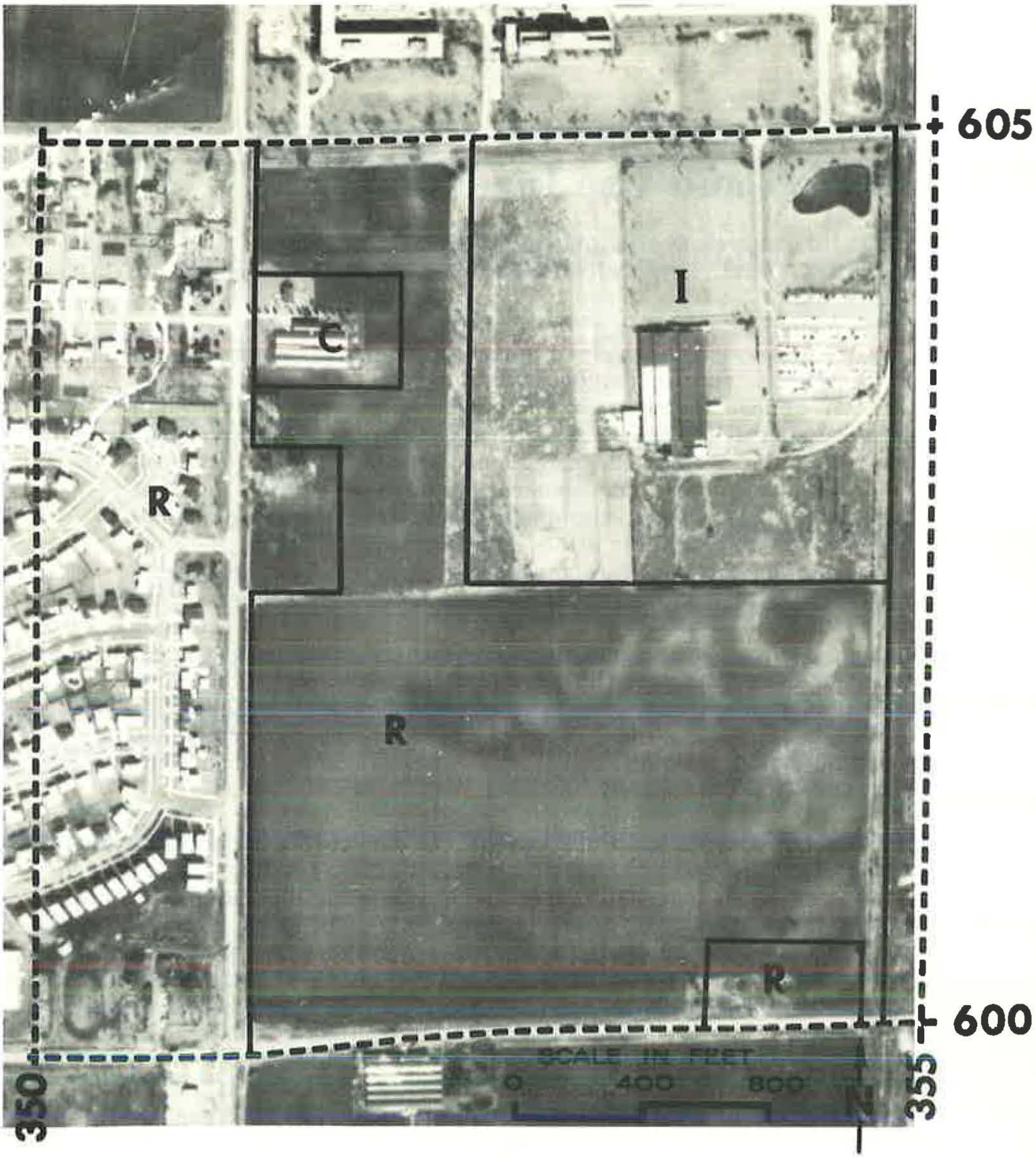

Figure 1.

Within the city generalized land use and floor area in establishments were measured and identified directly from Sanborn maps (1). For the suburban area, where Sanborn coverage was minimal, the procedure involved affixing a transparent acetate overlay, on which the quarter square miles were marked, over a sheet of aerial photography. Generalized land uses in the quarter square mile grid square then were identified and site boundaries marked (2). The next step was to measure the area of the entire grid square and of each land use within it by means of various map measuring devices (Fig. 1) (3). This procedure was repeated for more than 4000 quarter square mile units in the Chicago suburban area.

Thus, when the transition to the continuing phase occurred in 1962, CATS had available a very detailed quantitative land use inventory on which to build its continuing program. The problem was to maintain the currency of the data in order to be able to make accurate travel estimates for traffic assignments and to be in a position to evaluate, and if necessary, modify the Study's forecasts of growth in the area. As would probably be the case in most continuing transportation studies, staff, budget, and other considerations make it impractical to update by repeating the exhaustively detailed processes used in compiling the original inventory.

Ideally, of course, data inventories would be maintained by large computers fed by current tax assessor records, building permits, etc. Such a "data bank" program was investigated by CATS. It soon became evident, however, that there were enormous problems in coordinating the schedules and data formats for all the agencies who would be submitting data on the approximately 2,000,000 parcels of land in the Chicago area to the bank, as well as servicing its deposits and withdrawals. While such a program certainly is possible and probably will be begun in the Chicago area sometime in the future, the responsibility would better be assigned to an independent agency and not to an operating transportation study.

It was decided at CATS that sampling techniques should be employed as an immediate method for updating the land use inventory. Within the city, blocks in the Loop or central business district, and in areas where it was known that major change had occurred, e.g., new expressway routes and urban renewal areas, were completely remeasured using current Sanborn data. A systematic sample of 5 percent of the remaining blocks in the city was selected, measured, and compared to the 1956 record. Change rates were developed and applied to remainder of the blocks in Chicago.

A point sampling technique was selected to provide current information on suburban land use.

THE POINT SAMPLING TECHNIQUE

Again, aerial photography was the primary data source available for determining land use in the suburban portion of the CATS study area. A point sampling technique which had been used previously in flood plain mapping and other geographical surveys was adapted to the problem. In this technique, a transparent overlay on which a pattern of points has been printed is superimposed on an aerial photograph. The land use on which each point falls is identified and a count of points is made for each category of land use. The points counted are a sample of an infinite population of points covering an area, and yield estimates of the proportions of the actual population of points which would fall within areas of land devoted to each type of use.

The standard error of the sample proportions for an area of given size depends on the density of points. In the CATS survey, a density of 1 point per 640,000 sq ft of land area or approximately 44 points per square mile was used. This density was calculated to yield standard errors which were judged small enough to be usable in estimating trips and to yield adequate statistics on land use change for other purposes. For example, the standard error of the proportion for a township (36 sq mi), where a given land use (such as residential) occupies 50 percent of the land area, would be 0.013. Thus the chances would be two-thirds that the true population proportion would fall within plus or minus 0.013 of 0.50. This calculation is made simply, as follows:

$$\sigma_p = \sqrt{\frac{pq}{N}} = 0.013$$

where

p = the sample proportion of points in residential use (in this example, 0.50)
q = the proportion in other uses (also 0.50)
N = the total number of points falling in the township; at a density of one point per 640,000 sq ft, N = about 1570 points.

Similarly calculated, a type of use, such as commercial, which occupies 5 percent of the land area, would have a standard error of the proportion of about 0.005 (or 0.5 percent) for a township. For smaller areas such as CATS zones, which in the suburbs are 4 sq mi in size, the standard errors would be 0.038 and 0.017 for land uses occupying, respectively, 50 percent and 5 percent of the zones.

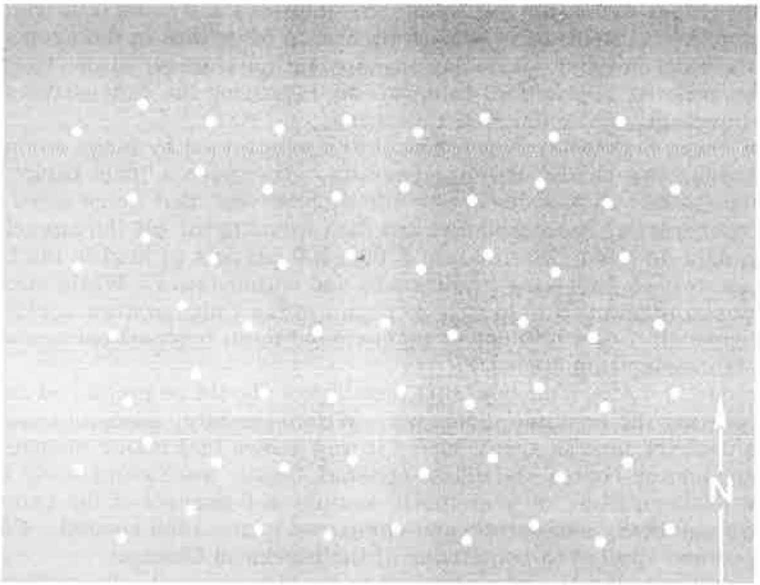

Figure 2.

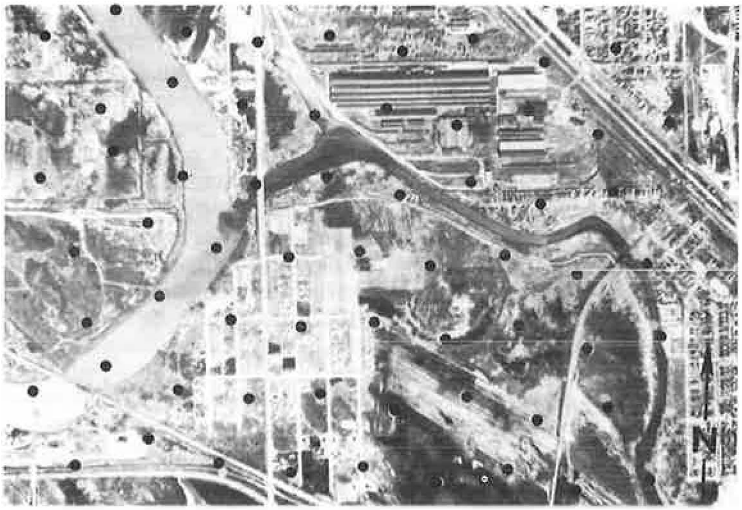

Figure 3. To show the dot pattern clearly, the individual dots are reproduced at a larger scale than the aerial photography.

The above estimates are conservative (maximum) estimates, based on an assumption of randomness. The pattern of points actually used was not random, however. It was a combination random and systematic pattern called a "stratified systematic unaligned" sample. It produced the benefits of systematic (checkerboard) samples: lower standard errors than either simple random or random-within-stratum patterns, without suffering from bias caused by periodicities in land uses which might occur on the ground.

A stratified systematic unaligned pattern is produced as follows: a plane surface is divided into squares of equal size which constitute strata. First, a point is placed in

SUBURBAN LAND USE UPDATING FORM FORM IV-10 4/64

LINE NO.	AERIAL PHOTO NO.	DATE CODE	COORDINATES		LAND USE IN 640,000'S OF SQ. FEET								
			X	Y	RES.	COMM.	MFG.	PUB. BLDGS.	PUB. OPEN SPACE	TRANS.	VAC. & UNUS.	STREETS & ALLEYS	TOTAL AREA
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													

Figure 4.

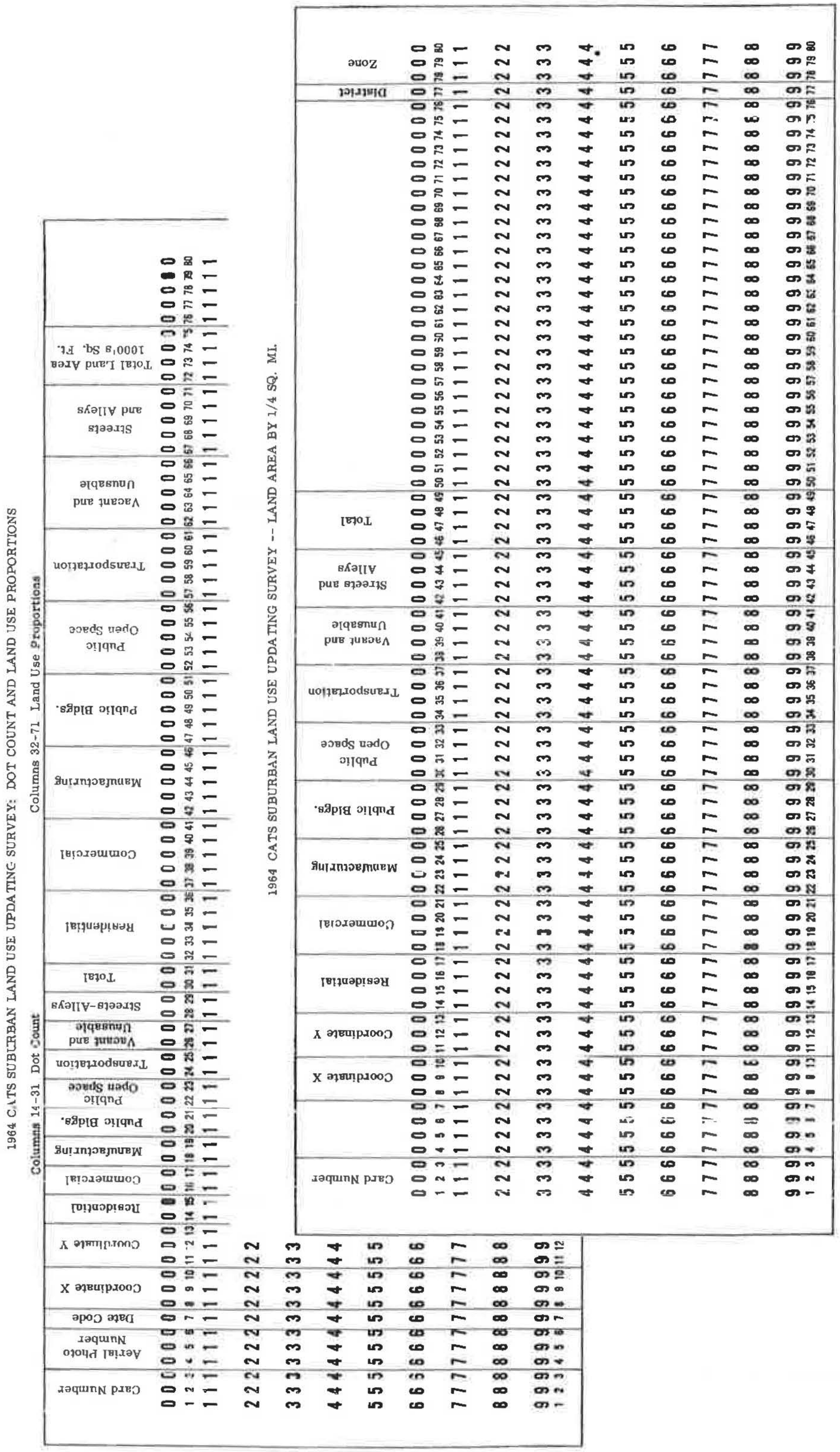
1964 CATS SUBURBAN LAND USE UPDATING SURVEY: DOT COUNT AND LAND USE PROPORTIONS
Columns 14-31 Dot Count
Columns 32-71 Land Use Proportions
Card Number
Aerial Photo Number
Date Code
Coordinate X
Coordinate Y
Residential
Commercial
Manufacturing
Public Bldgs.
Public Open Space
Transportation
Vacant and Unusable
Streets-Alleys
Total
Residential
Commercial
Manufacturing
Public Bldgs.
Public Open Space
Transportation
Vacant and Unusable
Streets and Alleys
Total Land Area 1000's Sq. Ft.
1964 CATS SUBURBAN LAND USE UPDATING SURVEY -- LAND AREA BY 1/4 SQ. MI.
Card Number
Coordinate X
Coordinate Y
Residential
Commercial
Manufacturing
Public Bldgs.
Public Open Space
Transportation
Vacant and Unusable
Streets and Alleys
Total
District
Zone

Figure 5.

each square in the left-hand column of squares, all points having the same Y-coordinate within each square (the first Y-coordinate is chosen at random), but different X-coordinates, each chosen at random. A table of random numbers is used to obtain the coordinate locations. Second, points are similarly placed in the bottom row of squares, this time using a constant X-coordinate but different Y's for each square. Finally, points are placed in the remainder of the squares, each point taking its X-coordinate from the left-hand column in the same row and its Y from the bottom row in the same column of squares (Fig. 2). The construction and benefits of such a pattern are further described by B. J. L. Berry of the University of Chicago (4).

The actual mechanics of the sampling procedure for determining the proportions of land use is as follows. First, the aerial photo is covered with a sheet of transparent acetate. Locational reference points then are drawn on the acetate. At CATS, quarter-square mile units defined by the CATS grid coordinate system were used. The next step was to insert randomly the dot sampling overlay between the photo and the acetate covering it. The sample dots then are marked on the acetate covering, and the dot sampling overlay is removed (Fig. 3). Finally, the land use under each dot is identified and recorded on a tally form (Fig. 4). Each line on the tally form represents a quarter-square mile and is recorded in one punched card. Land area proportions then can be determined by totaling the sample dot count and calculating the percentage of dots in each land use category. Actual land area estimates are determined by applying the derived percentages to the measured total area in each quarter-square mile unit (Fig. 5).

ACCURACY TEST

To attempt to check the accuracy of the point sampling technique, an empirical test was made at CATS, comparing the results of complete measurements as opposed to those from a point sample. In this test, land uses were marked off and identified in eight typical suburban ¼ sq mi. Complete measurements then were made of land use, by type, using a grid overlay. At least two measurements were made of land uses in each ¼ sq mi, and if differences in the total area for each ¼ sq mi exceeded 1 percent, or for the largest of the land uses, 2 percent, additional pairs of measurements were made until the differences were reconciled. Averages of the two final sets of measurements were computed and used as the check measurements.

Measurements then were repeated using the point area sample overlay. Seven independent measurements were made for each ¼ sq mi, using a random placement of the overlay for each remeasurement. This would be equivalent to making one measurement each for seven different ¼ sq mi, in all equivalent to measuring 56 different ¼ sq mi.

TABLE 1

LAND USE TOTALS

Land Use	Check		Point Sample	
	Sq Ft (1000's)	%	No.	%
Residential	17,945	33.0	200	33.8
Commercial	820	1.5	9	1.5
Manufacturing	5,840	10.7	67	11.3
Public buildings	875	1.6	8	1.4
Public open space	25,645	47.2	276	46.6
Transportation, etc.	460	0.8	4	0.7
Vancant	2,120	3.9	21	3.5
Streets and alleys	650	1.2	7	1.2
Total	54,350	99.9	592	100.0

TABLE 2

COMPUTED AND ACTUAL STANDARD ERROR OF LAND USE PERCENTAGES

Land Use	Actual Difference Pt. % - Ch. %	Standard Error of %
Residential	0.7	±1.9
Commercial	0.0	±0.5
Manufacturing	0.7	±1.3
Public buildings	-0.3	±0.5
Public open space	-0.5	±2.1
Transportation, etc.	-0.1	±0.4
Vacant	-0.3	±0.8
Streets and alleys	-0.1	±0.5

The measurements were repeated on the same areas to avoid multiplying the labor of making check measurements. A total of 592 points was counted for all seven measurements of the eight $1/4$ sq mi. The land use totals for all measurements are compared in Table 1. The percentage distributions agree quite closely. This would be about the agreement expected for a group of three or four CATS suburban zones.

Standard errors of the land use percentages were computed and are given in Table 2 in comparison with the differences actually obtained in the test. It is evident that the actual errors all are well within the calculated range, and usually are much smaller than would be expected from a random sample. In fact, this test probably overstates the obtained error because the check measurements themselves are not perfect, containing some random error of measurement.

CONCLUSIONS

Point sampling is an efficient, and low cost method of estimating proportions of land use. When a reasonable density of points is used, low standard errors will result. It has the additional advantage of avoiding bias more successfully than the complete measurement method. More time can be devoted to each sample point to obtain correct identification. Since much less care is needed in counting points than for making complete measurements, error which usually occurs in practice can be avoided, especially where the volume of work is large.

Additionally, point sampling is an excellent method for continuous updating of land use. The same point locations, recorded on acetate overlays, can be placed on new photography at future dates, providing an efficient way to measure changes. Random sampling differences would automatically be removed by placing each point on exactly the same spot at successive times.

Where the total areas already are known, as they were for CATS areal units from the 1956 survey, the land use percentages can readily be converted to areas of land. Even where total areas are not known and must be measured, the point sampling technique is much less costly and time consuming than complete measurement. This was demonstrated during the past year when CATS made available overlays, forms, and other details of the technique to the Northeastern Illinois Planning Commission. That organization, successfully applied the point sampling technique to the remainder of the Northeastern Illinois Metropolitan Area outside the CATS study area. These data have been turned over to CATS for integrated processing and the preparation of a land use inventory of the entire Northeastern Illinois Metropolitan Area.

It may be concluded that land use estimates may be made very accurately and economically by the point sampling technique. This method was tested and applied at the Chicago Area Transportation Study during 1964-65 and the resulting land use data have been used to make trip estimates for current traffic assignments.

REFERENCES

1. Chicago Area Transportation Study. Land Use Survey Manual. Chicago, 1956. 74 pp. mimeo.
2. Chicago Area Transportation Study. Geographic Reference System. Chicago, 1963. 111 pp. mimeo.

3. Chicago Area Transportation Study. Suburban Land Use Survey Manual. Chicago, 1956. 19 pp. mimeo.

4. U. S. Department of Agriculture. Sampling, Coding, and Storing Flood Plain Data. Agriculture Handbook No. 237, Farm Economics Division, Economic Research Service, Aug. 1962.