Resource Paper

PAUL O. ROBERTS
Massachusetts Institute of Technology

Travel demand forecasting has proved to be an elusive art, widely practiced, but increasingly difficult to believe as more and more projects are opened to traffic that exceeds all forecasts, sometimes even those of the design year. Other projects, such as the Massachusetts Turnpike or the Chesapeake Bay Bridge-Tunnel, have not reached their estimated volumes and consequent revenue generation. The next big project to open is the BART system in the San Francisco area. This project is one that will be watched carefully. Will actual volumes exceed the estimates, or fall short? If past prediction is any guide, there is little chance that they will come close to the mark.

This Achilles heel in the urban transportation planning process is a problem that is not likely to be solved by one conference. It will take research money, creative effort, and lots of hard work to bring long-range travel-demand forecasting to the state of being a science. What we must do are identify some of the reasons why we appear to have failed in the past and put forward ideas for future research.

NEED FOR LONG-RANGE FORECASTS

The first, and perhaps the most important, point I want to make is that there is a mistakenly perceived need for accurate long-range forecasts of traffic volumes. It is true that the transportation projects associated with most cities are large and expensive. The South Station section of the Boston central artery cost more than $75 million per mile. This is too large an investment and too important a facility to build without a careful estimate of demand. When one considers the effect on the area as a whole, including the secondary investments, the impact of this transport investment is enormous, affecting the land use patterns of the urban area literally centuries into the future.

How then can I say that accurate long-range demand forecasts are not needed? First, I believe that we are kidding ourselves to think that demand estimated by present methods will be accurate in the long range. Even if we were to develop very good demand models, travel is only one part of the complex overall urban process that includes changing land use and urban structure. Changes in the transport facilitie: system will themselves affect the equilibrium of the system. The new levels of the activity system that will result will influence the travel demand forecasts and render the inaccurate. Thus, I would argue that demand models considered in isolation of this process of change are necessarily short range in nature.

The conclusion that demand models are short-range prediction devices derives from the fact that the forecast volumes computed by the model depend on the validity of the inputs and, because these will be changing during time as the economic-activity system changes, the outputs will only be valid during that period for which we can accurately predict values for those input variables. At the moment we are not very facile at developing those inputs. One reason is that the urban transportation planning process has devoted very little time and effort to the construction of operational activity-system models of land use, housing, industrial location, and so on. Another reason is that the models are hard to develop. As a consequence, demand models will remain accurate only during the short range, not the long range, until these models have been implemented.

Beyond the practical realization that at the moment we do not have satisfactory activity-prediction models and therefore currently are not going to get long-range demand predictions, my original point was that there is a mistakenly perceived need for accurate long-range forecasts of traffic volumes in the first place. One current use of such demand forecasts is in the economic justification of a project. Long-range forecasts are needed because projects are typically capital intensive and require long lives for their economic amortization. This leads us to believe that accurate volumes are needed in the computation of the benefit stream. The fact is that in networks the flows over the system are changed by investments in new facilities in such a diffuse and subtle way that it is impractical to compute benefits external to the transportation/activity prediction system. Rather, what must be done is to integrate into the models the ability to compute the disutility incident to travelers or commodity movements as the result of each trip. Over time, as the activity system changes, those utilities will also change; and, as they do, accounting systems set up for the purpose will record them. One such measure is the gross national product. Urban equivalents to GNP could be formulated. Plans could then be evaluated by comparing time series of such measures. Volumes would be output merely as an afterthought.

An attractive alternative when we get good land use models will be the examination of the changes in land use that have occurred as the result of the changes in the transportation system. In fact, the goals of the urban transportation planning process might well include specifying the land use pattern being sought. Clearly, this would remove some of our current obsession with traffic-flow figures.

For the time being we should spend more time looking at the levels-of-service attributes that exist on important links in the system and between typical origins and destinations at equilibrium. The fixation on volumes is understandable, for they provide a readily comprehensible measure that is in many instances directly observable, at least in theory. In practice, they begin to get complicated: Hourly volumes, average daily traffic, thirtieth highest hour, practical capacity, and design volume all require elaborate explanation. With higher level objective functions, our concern can properly turn to other matters.

However, before we can do this we will have to address the concerns of our brethren whose job it is to design transportation facilities. They have for some time been looking to these forecasts for numbers to use in the sizing of facilities. I believe we should cut them off with nothing more elaborate than a statement that the figures we can furnish will not be appropriate to their use anyway without further discussion. Typically, the numbers needed for economic justification will not be the same as those needed for sizing and design. This is not to say that there is no relation between the two; however, the needs of the former are concerned with total flows, whereas the design concern is with conditions during the often-repeated peak period. The capacity-related figures frequently used for design are difficult if not impossible to get from present approaches. The design of the facility should probably be based not on traffic forecasts literally interpreted in any event but on consistency principles with as much future flexibility built in as each case will allow. For large projects it may be desirable to develop simulations for each of several designs to determine the consequences of each. In essence, I am suggesting that one first establish the facility design and then determine how traffic will use the facility provided.

Increasingly, the forecasting of demand will be in support of some notion of what one would like to see the city become in the future. This means that city goals must be well thought out. It also means that the task of designing future systems is greatly clarified. Designing systems without such a statement of desired goals is much more difficult, for there can be no measures of goal achievement except very narrow measures of traffic efficiency, which I contend are totally unsatisfactory.

There are at least two other reasons why I would like to debunk the importance of accurate long-range forecasts. The first is that the most important years from an economic point of view are the first years, not the later ones. The higher the appropriate rate of interest is, the truer this is. Our concern therefore should be primarily with the short-run effects of changes in the system. Many of these may be interim in nature. They may represent conditions before the system is "finished." As designers we have tended to be over-concerned with the design year, with the grand plan, and with symmetry of the finished system. Increasingly we have been frustrated by systems that, it has become embarrassingly obvious, will never be finished. We might be more convincing if we stuck to short-range forecasting and emphasized solutions that have more immediate payoff.

The second reason is closely related to the first. That is, long-range forecasts associated with a target year insidiously undermine rationality in the planning process. If the activity system changes in response to changes in the transport system, then it would appear to be patently absurd to attempt to jump ahead to some final year and predict activity levels at that time without working up to it gradually in steps that allowed one to make midcourse corrections. Actually, the target-year concept in conjunction with the needs approach to public works planning has been the device by which we have fooled ourselves for years. We have merely specified activity variables that would support the travel patterns for the facilities that we sought to supply in the first place. Whether these travel patterns could actually take place was not debated. We could then hide behind the comment, "We only build what the public needs."

This leads us to consideration of what is involved in forecasting long-range travel demand for urban transportation facilities. Clearly, the discussion thus far indicates that more than mere demand models are involved, though they are at the heart of the process. What is at issue is the whole urban transportation planning process and the entire system of models. Current methods do not recognize this, though there are many individuals around who do and have said so.

The problem as I see it is that the entire process must be revised to make it more interactive and interdependent and to integrate the currently missing elements. Others have criticized the faults with currently employed methods more extensively and carefully than I propose to do; nevertheless, I will point out those aspects that I find most in need of correction if we are to plan intelligently.

The approach that I will take is to address 4 interdependent topics: problems with the current UTP process, problems with the travel demand forecasting models of the UTP process, some proposals for improving the forecasting models, and some proposals for improving the overall UTP process. It will be difficult to separate my discussion of the nature of the problems with current methods from the prescription for their solution, but I will attempt to do this to achieve what clarity I can.

PROBLEMS WITH UTP PROCESS

It may be useful to clarify what I mean by the current UTP process. Basically, I am addressing the process of urban transportation planning and the associated travel demand forecasting models endorsed by the Federal Highway Administration and embodied in its package of urban transportation planning programs (1). Although these are not the only programs in current use or the most advanced ones, they do represent a sort of current state of the art and for a variety of reasons they must be "eclipsed" before they can be replaced.

The overall planning process for a given urban area is shown generally in Figure 1. I have the following problems with this overall approach:

1. Insufficient recognition of the interdependence of the elements in the process,
2. Absence of transportation-responsive activity-system models,
3. Use of a target-year approach to planning,
4. Obsession with predicting traffic volumes,
5. Seeking of accuracy when what is needed is a range of uncertainty,
6. Lack of appropriate objective measures with which to measure the effectiveness of the process, and
7. Insufficient integration with the decision-making process.

Some of the topics have already been alluded to, but it is useful to focus the arguments for each.

Insufficient Recognition of Interdependence

The basic problems here are a lack of explicit simultaneity in travel demand determination, an apparent reticence to compute equilibrium between supply and demand, and an almost complete failure to model the feedback between the transport system and the socioeconomic activity system. The real-world system clearly involves a high degree of interdependence; yet, the UTP process attempts to deal with it as a sequential process. It fails to recognize transportation as part of an essentially self-regulating equilibrium process. Thus, we have the thrust to provide enough capacity in the new freeway system, the attempt to answer the problem of growing congestion, and the failure to consider the do-nothing alternative.

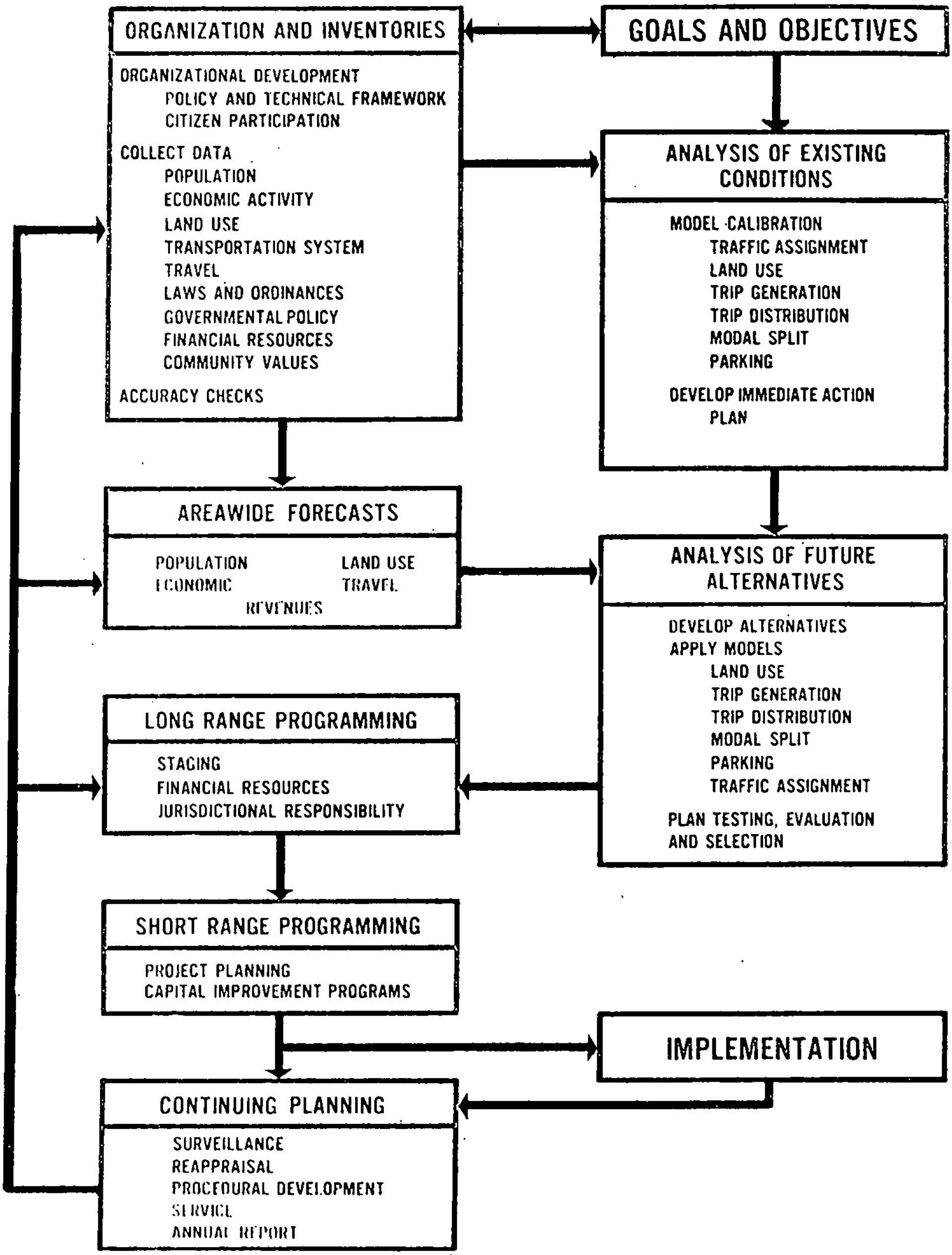
Absence of Transportation-Responsive Activity Models

The typical UTP process forecasts or projects area-wide population, land use, and economic activity. In only a few cases have socioeconomic-activity models of any type been used, and in almost no case could they be considered to be fully responsive to changes in the transport system. This may be too harsh a judgment to levy against all studies, for the National Bureau of Economic Research has reported on land use models developed and used in 5 different urban areas (2). It is no exaggeration, however, to assert that adequate models of economic activity for use with the transportation planning process still do not exist.

Use of Target-Year Approach to Planning

I feel that the target-year approach is the single most detrimental aspect of current planning. It basically ignores the question of how we get from here to there. It also ignores both the problems and the benefits associated with the staged introduction of improvements into the system. Benefit streams are badly misrepresented by this process, which assumes that all projects will be on-line by the target year. The benefits achieved in this year must be linearly applied from "zero" at the outset to "full" in the target year. This is clearly unrealistic. If, in the real world, it is possible to bring an important project on-line in the early years, the discounted present value of the benefit stream is greatly enhanced. Yet, this is ignored by the existing approach. Improper consideration of intermediate years also means that disruption during constructuion is not addressed directly. Finally, the shortcomings of the forecasting process are placed in the most unfavorable position because in essence the whole value of the project is determined based on extremely tenuous estimates of the far distant future.

Figure 1. Urban transportation planning process.



Obsession With Predicting Traffic Volumes

Overconcern with traffic volumes could be construed to be an indication that objectives have not been well thought out. Traffic volumes on the links of a network (especially a complex network) bear little or no direct relation to benefits. It is difficult to look at the volumes on most networks and infer anything about objectives. Our interests in flows come about because of the correlation with equilibrium levels of service on the network. Level of service is a legitimate concern, but it is hard to determine the incidence of changes in level of service from those on individual links. The aggregation that is of more interest is the equilibrium level of service between zones and the consequent impact of the activity system.

Seeking of Accuracy Instead of Range of Uncertainty

We are not now, nor will we ever be, capable of making accurate forecasts. Accuracy in this undertaking is a myth. We should understand, however, that there is uncertainty associated with all forecasts. The real point is that we should understand the range of uncertainty implied by a particular forecast. It would also be useful if we also knew how the range of uncertainty changed with each choice variable. This will take a change in our current thinking, but it is essential to proper assessment of the impact of individual projects.

Lack of Appropriate Objective Measures

We really have done very little to define suitable objectives and to program the capture of the appropriate measures into the existing models. The plan and the traffic flow volumes on it have been viewed as the final answer. For those studies that have attempted to do more, a variety of approaches have been used in practice with obviously different objectives. This is probably as it should be. There will always be a variety of objectives that one would like to observe in making the decision, and a decentralized decision authority will want to see all of them. I feel, however, that at the moment in the planning stage we show the decision-maker almost nothing except the plan and the flows. I will have some suggestions as to the objective measures we should be using at a later point.

Insufficient Integration With Decision-Making Process

It is difficult, given the present structure of urban government, to interface with the decision-maker, for there is typically not one person or even one decision-making body but literally dozens. The UTP process as currently structured is ponderous. It cannot respond quickly to requests by local mayors or citizen groups to investigate local changes or to develop scenarios that could be presented to the people involved to show them what it will be like after the change is made. This lack of responsiveness to participatory planning makes the process less useful than it would be if it were less expensive, more quickly done, and more illustrative in output.

PROBLEMS WITH FORECASTING MODELS

Difficulties with the process as a whole are repeated in the models making up the system; however, I will make every effort not to repeat myself as we turn to the models themselves. The particular focus of this section will be on those components of the overall system that deal directly with the demand and network equilibrium portions of the process. I am acutely aware of the very fine job that others (3, 4) have done in summarizing the faults with the process; I will merely review the basic faults that I find with the models.

Although it cannot be considered to be a complete statement of the details of the UTP process, Figure 2 shows the basic thinking underlying the process (1). The 4 basic steps are trip generating, trip distribution, modal split, and traffic assignment. Economic activity and land use are essentially projected into the future without feedback from the transportation system, though feedback to future land use is shown here with a dotted line, indicating that "though we now know there should be interconnections, they have not been routinely implemented to-date." Trips are generated without concern for the supply of transportation or its effect on the level of service offered. Trip distribution is typically constrained by its calibration to maintain the existing trip-length distribution whether or not the network can support it or the land uses have changed to accommodate it. And, neither generation nor distribution is typically brought into the equilibration process with network flows. Finally, the future-system or target-year approach is indicated as the recommended approach.

The problems with these models can be listed as lack of policy responsiveness, improper selection of attributes for modeling demand, inadequate determination of equilibrium flows, and importance of activity-system models to long-run demand.

Lack of Policy Responsiveness

The most obvious problem with the models is that they are not policy responsive. That is, they are not designed to answer the questions posed by a particular agency or to understand the response of the system to particular controls held by that agency. The urban transportation system in a large metropolitan area is rarely under the control of a single authority but typically jointly controlled by a variety of transportation agencies and an equally large number of nontransportation agencies. One cannot overly criticize the designers of the models for failing to identify a particular decision-maker. The major problem here, however, is that the current model design does not properly reflect the trip-making response of the system to changes made in the system itself. As pointed out in another report (3), the models are nonbehavioral and noncausal as well. The model system also suffers from lack of ability to account for transportation-related features such as differentiated tolls, parking fees, and bus schedules; and it is unresponsive to the possible changes in public transportation offerings, vehicle exclusion, parking restriction, and signalization that might be imposed by a policy-maker.

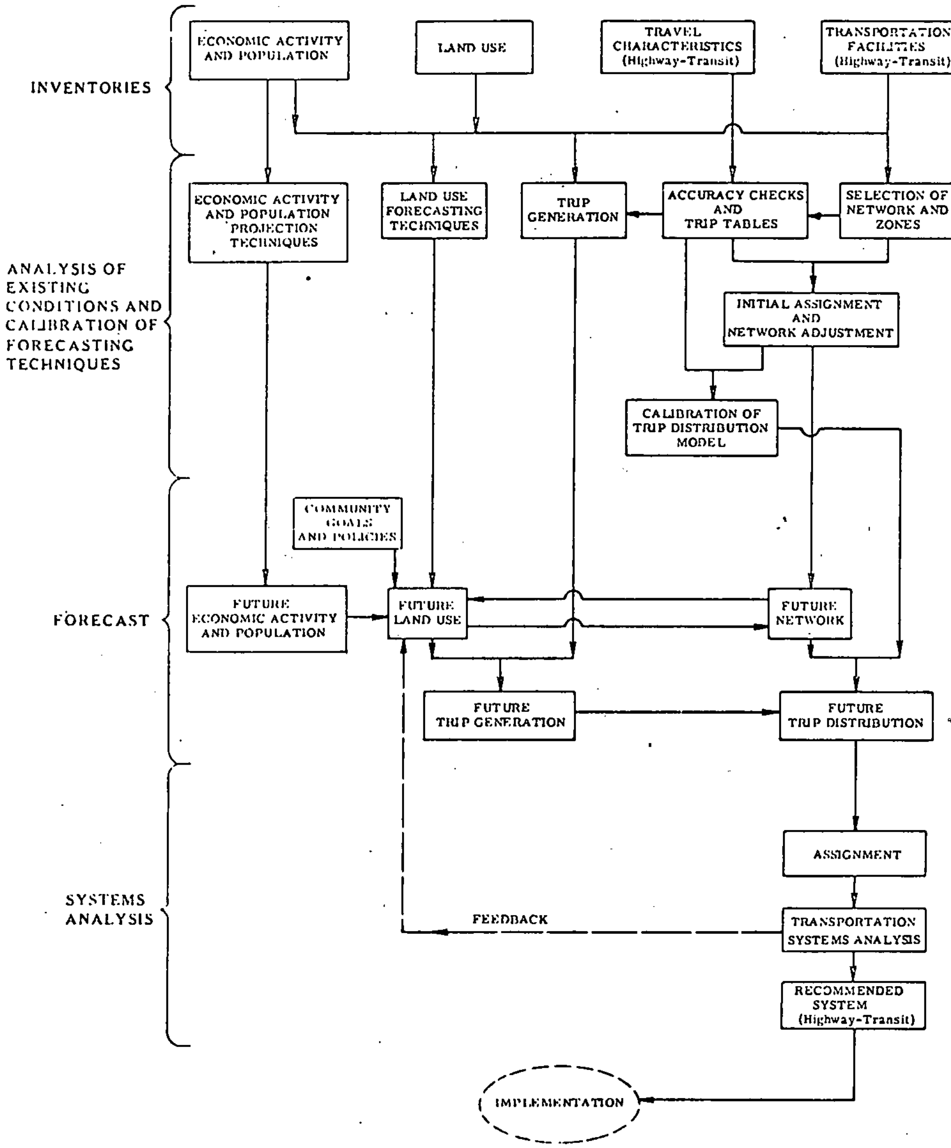
The most basic problem, therefore, is failure of the process to consider trip-making as responsive to travel conditions. This is a direct consequence of the fact that, in current versions of the travel demand forecasting package, trip generation is accomplished prior to and separately from both trip distribution and capacity restraint. The decision on where to travel, and, in fact, whether to travel at all, cannot be separated from the travel time, cost, and other travel consequences. Yet, existing programs do not even iterate on trip distribution, much less on the whole process.

The argument usually advanced for treating trip generation and distribution as given is that work trips are inelastic with respect to travel conditions. This may be true to a large extent. However, in the short run, the workbound traveler may vary his time of departure, his routing, and his travel mode. Travelers with nonwork purposes can also change their destinations and their frequency of travel.

Time of departure is not treated by any of the currently operational models. Yet, earlier departure times are clearly one way in which individual travelers in the system continue to cope with the capacity bind. For shopping and recreational trips, the opportunity to shift destinations, frequency of trip-making, and departure time increases with the construction of each new suburban shopping center. If travel demands are to be properly predicted by demand models, the equilibrium computation must be responsive to short-term changes in mode choice, routing, departure time, trip destination, and frequency of trip-making.

The activity system must likewise be responsive to a larger number of intermediate- and long-term variables. In the longer run, the traveler may decide to purchase a motor vehicle (or a second one). Over a still longer term, he may decide to change his

Figure 2. Urban travel forecasting process.



place of residence or even his job. These trade-offs are not appropriately modeled at the level of the daily travel equilibrium. Yet, they might be reflected in the changing income of an area, the aging of its population, or a change in its housing market.

Improper Selection of Attributes for Modeling Demand

The UTP demand forecasting models are extremely limited in the variables available. The trip generation and attraction phases can, and typically do, make use of a variety of socioeconomic-activity variables. It apparently was not convenient to use transportation level-of-service variables for either trip generation or attraction, for none is ordinarily used. This is equivalent to saying, "It doesn't matter how bad the traffic gets; I'm still going to the ball game." Once trips have been generated and attracted, they are typically pushed around by use of the distribution model, ordinarily the gravity model. This model typically uses travel time as the only variable. It is possible to weigh travel time by its average value and to add certain out-of-pocket costs such as parking fees or tolls and to use this as the variable affecting distribution of trips. This is sometimes done. The modal-split models have made use of a few more variables, but even here the number of variables that can be used to influence modal choice is limited.

Obviously time is an important factor, and it has been used in almost every study. However, not everyone has recognized the variety of time-related factors there are and their relative importance. Time is frequently separated into travel time and access time. It has been further differentiated by mode into walking time, wait time, line-haul time, transfer time, parking time, time variability, interarrival time, and schedule delay. Recent research has shown that these variables should have different values (5). For example, walking time appears to be 3 to 4 times more onerous than in-vehicle travel or transfer time.

A variety of other variables, including cleanliness, comfort, convenience, out-of-pocket cost, ability to carry packages, safety, schedule reliability, ability to read, privacy, and ease of carrying wife or children all may be important for the various modes. Very few can be routinely handled by the present process. A principal reason is that, although the attributes associated with the various modes may be available at the link level, they are difficult to determine from origin to destination over the network. This could be easily overcome by using the minimum path tree-tracing algorithms in more creative ways. For example, it is possible to use the concept of a travel resistance or impedance where

$$R = \sum_i c_i L_i$$

and where L_i is the level-of-service attribute found on the link, c_i is the cost per unit associated with encountering attribute i, and R is the consequent disutility of traveling over the link (6). Manheim (7) shows this relation in terms of utility functions, and Blackburn (8) refers to this concept and uses utility theory as the "inclusive price" of travel on the link. If R is determined at the link level, minimum R paths can be computed or minimum time paths can be traced and their consequent summed attributes determined simultaneously. None of this is typically attempted in the present programs.

There has been, as well, a notable lack of interest on the part of researchers working with the UTP package to develop more extensive supply attributes or supply models at the link level. I would reason that this is because there did not appear to be a way in which the information could be used in the process even if it were to be developed. Based on the methods described above, the information flows of such extensions become perfectly clear. Traffic volumes determined during a previous assignment could be used with variables describing transportation supply to determine level of service that, through valuation by an inclusive price scheme, is then used for tree-tracing of attributes in the next demand computation. Such a scheme does not limit the number of attributes to be used in describing transportation supply or demand functions as do the present UTP models.

Inadequate Determination of Equilibrium Flows

For practical purposes, there is almost no feedback in the present system of models. The modal-split portion is the most interactive, and equilibrium between modes joining the same origin and destination is achieved in some cases. The same cannot be said for network travel conditions and trip generation, attraction, and distribution. Obviously, the more simultaneity that can be reflected in the models, the better. In the real world, one observes a certain equalization of impedance over the network, at least as reflected by travel times. There are typically a larger number of possible paths between origin and destination than can be conveniently modeled. As the network becomes increasingly loaded, more and more of these paths are used. Travel times on the expressway approach those on city streets. Where there is disparity between the two, queues may build on the expressway because travel times on the freeway are still better than those on city streets—especially for the long-distance travelers or those without perfect information about the local street system. The transit system may share in this impedance equilibrium, if value of time to the travelers could be properly evaluated.

Achieving this equilibrium in the models appears to be a very difficult task. Even this understates the problem, for trip production is still intimately related to level of service on the network in spite of my attempt to present the parts as independent. A first attempt to solve the problem of simultaneity might do it by considering the parts independently and iterating. This has been done on occasion but not routinely. The criterion for equilibrium is typically stability on the network. Figure 3 shows a scheme for improving this equilibrium computation, but it should be viewed only as an improvement, not the answer.

Although current practice is variable, the FHWA package suggests that speeds measured on the present network should be used in building the minimum path trees used in trip distribution. Trip travel time distributions are then built up by using the origin-to-destination pattern actually observed and travel times computed from the minimum path trees. For the future, travel times on the network are assumed for the purpose of building trees for use as input to the distribution model. Then, the iteration between capacity restraint (if it is used) and travel conditions on the network is assumed to affect only choice of mode and routing. This is obviously simplistic.

A major problem facing those of us who are attempting to explain travel behavior is the sheer size and complexity of the network. Early models used several hundred nodes. For large urban areas, several thousand nodes are currently being tried, and the desire for more grows with every increase in computer capacity. The fact of the matter is that we will probably never have enough. We must somehow be able to model the volume-delay function of the corridor as a whole, while it is flowing in all directions at once, for it is apparent that, as major transport links become congested, flow is diverted to facilities of ever-decreasing levels of service. Perhaps the answer is the use of a spider-network plan that is rather different from the type we have tried to date or, alternatively, the consideration of the local streets as a sort of plain of impedance. My view is that bigger networks are not the answer. They merely lead to bigger computers, longer computer times, and increased complexity and expense.

Actually, travel-making behavior is considerably more complicated than the process shown in Figure 3 suggests. The complicating factor is related to the phenomenon known as peak-spreading. For inelastic trips, the one dimension of flexibility is time of travel. Everyone is familiar with the statements: "Better leave early to get ahead of the peak" or "It's too late now—might as well have another cup of coffee and wait for the traffic to clear." In fact, travel times may be relatively consistent throughout the peak.

Peak-spreading is partly due to the diversity of starting times and appointment hours and the randomness of schedule that may be found in any urban area. It is also the result of travelers having a choice in their schedules so that their travel does not coincide with peak-hour travel. For those employees who have starting times at 8 or 9 o'clock, there is still another factor influencing their decisions on when to travel. That is the variability of travel time as measured by the cumulative probability of travel time. If it is extremely important that you be some place on time (for example, at work), then you must allow enough time not only for average travel time but also for the extreme tails of the distribution to the desired probability. It is probably this additional factor that pushes the cautious employee to commute to work early and have breakfast downtown.

All of these factors make the prediction of the full set of network conditions affecting travel much more difficult and involved than present methods would even suggest, much less replicate. Obviously, if the previous discussion is to be believed, there is a need to predict peak-hour as well as average daily flows. And perhaps even more important there is a need to know the length of the peak as well. Our ability to do these things is currently extremely limited.

Importance of Activity-System Models to Long-Run Demand

The difference between short-run and long-run demand prediction is a distinction that was made in a before-and-after study (3) and it is useful to extend here. Because demand models are a function of both equilibrium level-of-service and activity-system variables,

$$V = d(L, A)$$

it is useful to ask how the FHWA process will deal with each of these inputs over time. If one feels that the activity-system variables are also a function of transport level of service,

$$A_t = a\ (L, A_{t-1})$$

then a basic problem exists with how to predict these variables several time periods out. One way is to assume that there is no relation between transportation and economic activity or that it is not significant and to merely project these variables by extrapolation. This is what has been done in the FHWA process. To do otherwise requires the use of an activity-system model. For the transport level-of-service attributes, there is conceptually no problem with keeping up with L because it is a direct function of the transport-system variables and the volumes, both of which are available from the models in the process.

$$L = s(T, V)$$

This poses a problem only in the sense that the FHWA process does not now do it—not that it could not be rather easily done.

For the short run, both level-of-service variables and activity-system variables are directly observable from the real world. For the long run, we have to either construct activity-system models or be prepared to make heroic assumptions.

PROPOSALS FOR IMPROVING FORECASTING MODELS

It is far easier to criticize the present set of forecasting models than to offer constructive proposals for improvement. I do believe, however, that it is possible to make a quantum jump in our ability to make short-run forecasts by merely implementing a number of the research advances that have been made during the past few years and by following up suggestions made by others. Furthermore, I am confident that a greatly improved short-run forecasting capability will carry us a long way toward being more effective in our advisory roles to decision-makers. The specific recommendations are summarized as follows: use the available demand-model knowledge to develop policy-responsive, behavioral, causal, short-run demand models; integrate supply and demand models in better equilibrium computations; develop activity-system models that can be used to support longer range use of demand models; and develop and incorporate performance measures useful in decision-making.

1. Use the increased knowledge of demand models to develop better models.

It is not difficult to improve demand models dramatically by merely implementing research already completed to date. This will ultimately improve long-range as well as short-range forecasting. Brand (9) has summarized the state of the art. The first, and perhaps the most, significant change that could be made is to a direct demand model instead of the multistage process in the FHWA package. Because a direct demand model handles generation, distribution, and modal-choice simultaneously within a single model, there is no need to think in terms of 3 separate steps. A variety of these models exist including those of Kraft (10); McLynn and Watkins (11); Quandt and Baumol (12); and Domencich, Kraft, and Valette (13). Another advantage is that the parameters of those models can, for the most part, be interpreted fairly easily as elasticities and therefore have an intuitiveness and a meaning that are useful in their own right.

Manheim (7) discusses the interrelation of these models and his general share model and shows how consistent direct models can be "disaggregated" into indirect models without loss of generality. It appears that, with the expenditure of a bit more time and effort in attempting to clarify the approaches that are available as well as their advantages and disadvantages, significant understanding will occur.

In listing problems with the current models, I mentioned the need to expand the set of variables available for use in the demand models. The use of the inclusive price technique appears to be the way to do that efficiently. Not only does it tie in the concepts of utility theory but it is useful from a computational point of view, for it can be readily implemented with existing minimum path algorithms. It is also useful from the standpoint of calibrating the model, reducing the difficulties caused by multicollinearity, and so on. The variables ordinarily used in CRA's disaggregate behavioral demand model (5) were greatly expanded, including several types of time and a variety of others expressing comfort and convenience characteristics.

That study also explored the possibility of calibrating a disaggregated or sequential model that first determines frequency of travel and then mode, time of travel, and destination. The model form used was the logit model calibrated by maximum likelihood techniques. The results were exciting. It appears that the methods for handling both choice of destination and time of day were satisfactory and that the results, particularly the elasticities, were quite significant. Although time of day was divided into only peak and off-peak travel, this may be the breakthrough needed to consider the equilibrium-over-time computation more fully.

2. Integrate supply and demand models in better equilibrium computations.

The subject of equilibrium in networks is one that has been badly neglected. The FHWA package makes only a limited effort at producing equilibrium flows through the use of a capacity-restraint routine. Conceptually the idea is that, if flow on a given link exceeds the "capacity" of that link, then some of the flow must be diverted to other routes or modes. To think this way, one must carefully define the time domain during which the model simulation is representative, the demand for flow during that period, and the capacity (or, better still, level of service) defined during that same period. The problem has been that, if overall levels of service get too bad, demand will decrease (or shift to the off-peak). Typically, the models have ignored this point and have simulated either peak-hour or all-day flows. Both have problems. Peak-hour figures are hard to estimate, and all-day capacities are meaningless.

Even ignoring these problems, we have had problems with equilibration procedures that we have wished would go away. The most difficult has been oscillation of flows on the network instead of convergence of our iteration procedures. This has made the assignment algorithms long-running, which has been the reason why we have failed to feed back all the places in the computation we should have. Weighting schemes where flow from the last iteration is weighted with flow from this iteration have been tried with only palliative success.

The most promising work in the equilibrium area has been that of Manheim and Ruiter (14) on DODOTRANS. This very flexible demand-computation and network-assignment package can do anything the FHWA package can do plus incremental assignment in conjunction with direct demand computation. The incremental-assignment technique, developed in 1962 (15), assigns an increment of travel demand over an increasingly loaded network with periodically updated computations of level of service. The result is an exact equilibrium of demand with supply done in somewhat the same way as the real world.

Integrating peak and off-peak travel into this system is something that will still require some effort, but a really computationally efficient version of this system should be programmed and implemented in a test project that could also include the best available demand techniques.

If this were done, there are still 2 areas in the equilibrium computation that need more conceptual work. The first is in the area of supply models. Almost no attention has been given to constructing simple supply models for use with the demand computation and assignment package. These could include the effects of number of lanes, curb parking, edge effects, traffic lights, left turns, and pedestrians.

The second area is that of working with big networks. I discussed this previously and will not belabor the point, but I think that our ability to model real situations is badly impacted by the fact that we seem to have to model the whole world before we can adequately represent the ability of the local street system to absorb or give up flows. Perhaps the answer is the use of aggregate models such as those of Koppelman (16). We can explore the use of spider-background networks that represent in a very general way the volume delay characteristics and volume response of the local street system in all areas except the one on which we are currently focusing our attention. This will require more research, however, for we do not now know how to do this adequately.

3. Develop activity-system models.

As I indicated earlier the entire travel demand forecasting process is incapable of moving beyond short-run demand prediction unless or until activity-system models are integrated into it. This is probably not the forum for a full-scale discussion of the development of these models. Yet, their importance must be immediately apparent. The expectations are not for immediate fulfillment of this need. There has been steady progress in this field since Lowry, however, and recent work may be imminently useful.

In general, the process of predicting the activity system is closely related to predicting the progress of the economic system. There are, of course, stocks of both transport facilities and buildings. The construction process because of its commitment of resources and time is a good place to separate the long-term changes in these stocks and the short-term use of these stocks by the economic system. The process shown in Figure 4 attempts to set out in a general way the relations between these stocks and the processes that decide their fate (the various demand processes and markets). The actors in these processes are individuals, firms, and governments. As in travel demand forecasting, the viewpoint of the individual and his choice is a good place to start. The family income sets an upper limit on the overall budget, but the amount used for items such as housing, clothing, food, and transportation is essentially flexible. The higher the income is, the more trade-off possibilities there are. The utility concepts applied in demand forecasting also can be applied here.

Obviously, the housing stocks and the housing market determination are the largest single group accounting for the largest amount of land and dealing with the most number of persons. A recent and landmark contribution to this area is Ingram's work (17), which is part of an overall urban simulation effort undertaken during the past few years by the National Bureau of Economic Research (18). The larger effort involves also industry and commercial location studies, but models that can be implemented are not yet at hand. Much research remains to be done in this area.

We should note in passing that we have not yet really addressed the problems of predicting urban freight flows. In view of the impact of those flows on urban street

Figure 3. Network equilibration process.

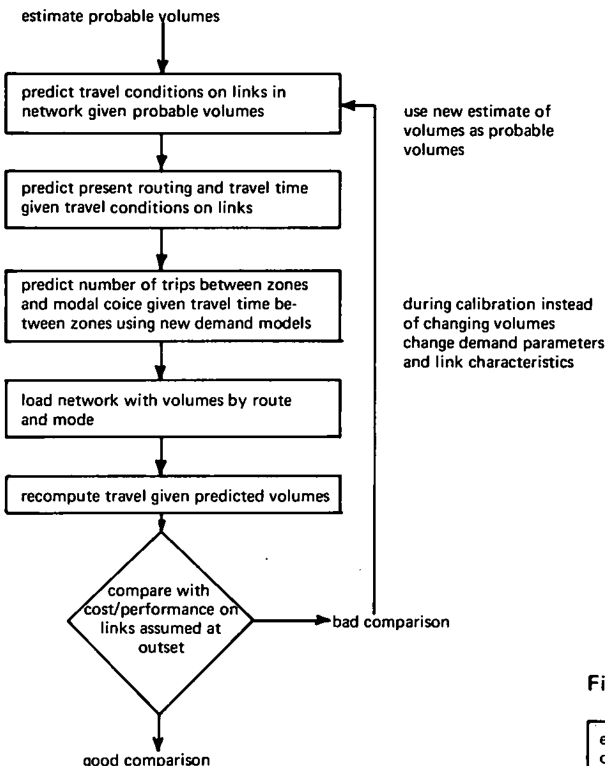
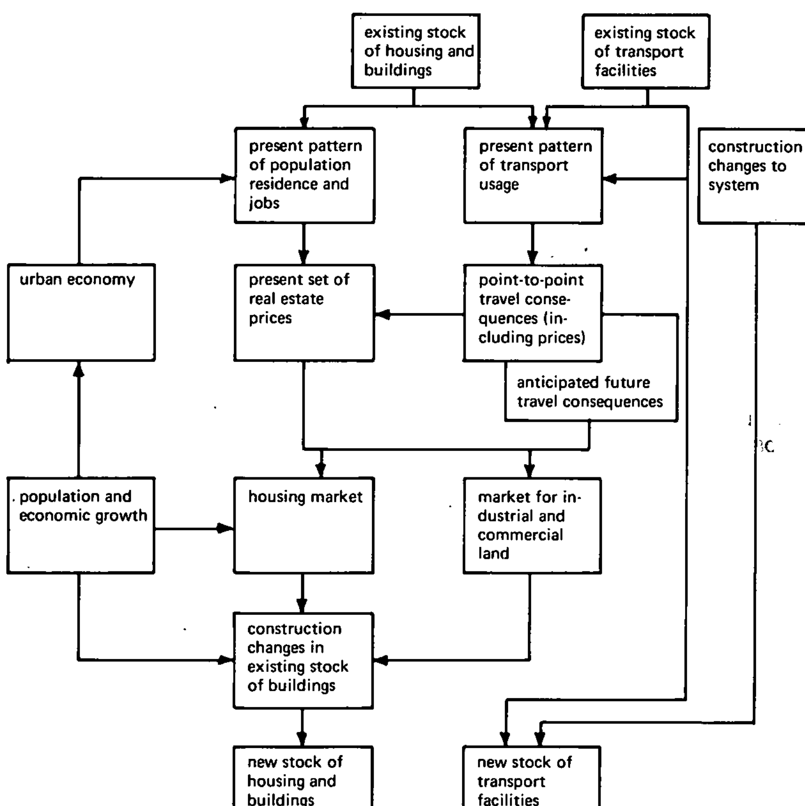


Figure 4. Urban growth process.



movements, it seems to me extremely important to begin to incorporate freight-movement factors into the system as well.

4. Develop useful performance measures.

The most important thing to keep in mind in the evaluation of transportation-system changes is that the final decision is basically a political and not a technical decision. Therefore, it is essential that performance measures properly reflect this. The actors in the political arena are likely to group themselves in any one of literally thousands of different ways. However, there are some basic building blocks that are common to a variety of groupings. These are firms and governments. They are typically referenced by resident location, affiliated locations, income, transport expenses, housing expenses, and mode. They may be aggregated by mode, industry, area, and income category. The relevant performance measures may be sorted or aggregated in a variety of ways for display purposes.

A major design issue, however, is the set of basic performance measures that will be developed by the models for use in evaluation. Clearly, costs of all types fall into this set. The factors in the inclusive price vector are all candidate costs, and the demand model weights indicate the way in which these attributes are traded off against money. This establishes their relative marginal values for use in evaluation. Within the modes, there are short- and long-range costs that are typically passed on to the shipper or passenger through the transport price or tariff. Therefore, they are captured in the inclusive price vector. Costs to the government or to industry for constructing facilities are also relevant and useful. Other, nonuser impacts will probably have to be developed in each particular case. The demand forecasting models may not be the place to gather the basic information for evaluation. For many variables, however, it will be the basic source.

Another major source of performance measures can be developed from the output of both the land use models and the model of the urban economy. Clearly, these, particularly the economic measures, will be very important outputs for use in evaluation. They are the overall measures that are most "macro" in outlook and must be considered to reflect the desirable overview from the standpoint of the overall economy (particularly for equitable economies).

PROPOSALS FOR IMPROVING UTP PROCESS

The suggestions of the previous section were largely design oriented. That is, they dealt with how an urban travel demand forecasting model system might be structured. This section will deal more with how such a system should be used, particularly in the interim period before such a system might be fully implemented.

My specific suggestions are be policy oriented, use the currently available parts of such a system to explore short-run consequences, attempt to integrate models more closely with the decision-making process, integrate supporting performance measures into the system, attempt to develop understanding of the degree of control that is possible in the real-world process, and use the model system in a time-staged planning process.

1. Be policy oriented.

Too often as model designers, academicians, and theorists, we are caught up in the model-design issues and not the real-world policy issues. Because it is, after all, the policy issues we are attempting to answer with the models we are designing, I suggest that we do our best to understand them and structure our models so that they can be addressed. A great deal of the criticism of the current set of UTP models centers on exactly this point. Let us not make the same mistake again.

2. Use available parts to explore short-run consequences.

Because a large portion of the demand model methodology already exists, at least in pilot form, I suggest that what is needed next are proposals for specific model designs and then identification of test sites, data collection, specification, calibration, programming, and testing of the prototype models that emerge. I feel that it would be a good idea to support 2 or 3 competitive designs with arrangements for comparison and evaluation. The model systems formulated at this time will necessarily be short run; that is, they will not embody operational activity-system models. I would suggest that pilot versions of activity-system models be undertaken simultaneously, but probably as a separate effort.

The important thing about the demand forecasting prototype models is that they will be short run. They should not be used for 1990 planning. They should instead be used to explore 2- to 4-year operational policy; i.e., What will happen immediately after the opening of the new expressway? They could also be useful in the analysis of transit, bus, or expressway-bus planning. To the extent possible the outcomes should be monitored carefully (19). Large differences between predicted and observed data will indicate either structural problems in the models or a need for recalibration by using the new data.

3. Attempt to integrate models with the decision-making process.

This point follows directly from the previous one. Elected decision-makers are extremely responsive to the political process. If our models are to be useful to them, they must predict the short-run outcomes. The system I have described would do just that. It should be possible to interface more closely with mayors, city councils, governors, and public works directors. An additional point is that demand models calibrated on individual rather than aggregate zone data would be much more general than our present models and would not require recalibration from use to use or city to city. Therefore, they should be less expensive to use and also less ponderous. Huge staffs will not be required. Instead, we will be able to work much more closely with public officials and respond more quickly to their questions or suggestions. Admittedly, in any given application, there are still data on the supply aspects of a given system to collect, code, and debug; and if equilibrium computations are required, we still may have lots of computing to do. The operational problems of performing analysis will not become simple overnight.

4. Integrate supporting performance measures into the system.

To be policy-oriented, short-run responsive, and "plugged into" the decision-making process, the system must have a carefully designed set of performance measures integrated into it during the formulation stages. Both physical and valued consequences should be available in hierarchical form. Aggregations and summaries should be available as well. In my estimation, this is a key factor in the utility of the final system. Although the implementation of this suggestion does not require a high degree of analytical skill, careful conceptual design is required. That is not a trivial job.

5. Attempt to develop understanding of the process.

Once the full system is available, the next step is to use it to develop a fuller understanding of the urban process and the degree of control that can be exerted over it. This will require an extensive effort with the possibility of false starts, failures, lack of understanding, and frustration at every step. This process has already begun. We can help push it along with integrative programs and financial support.

At this point, it is not clear what can be controlled. We may not want to control the process even if this turns out to be possible. More probable, we are not going to have the political or institutional framework needed to handle this control. If, for example, we found that land use could be controlled, what person or council would we trust to exercise the control? This appears to me to have been the reason highway planners have traditionally fought the notion that the construction of highway facilities was responsible for influencing traffic demand. They really had to believe they were building to match "needs." Anyway, it is institutionally convenient to believe that the part of the process for which you are responsible is not creating waves. The most likely happening, I believe, is that we will find that it takes a considerable amount of coordinated policy to achieve a highly predictable level of control of land use. These controls will almost surely involve elements of zoning, housing policy, tax policy, environmental controls, utilities control, and use policies in addition to coordinated transportation policy. A device that could represent the future direction of development is large-scale planned unit development by a consortium of private and government entities that would plan and develop large tracts on a modular basis. Another, and I believe more probable direction, is the continued growth of urban areas by innumerable small developments.

6. Use the model system in a time-staged planning process.

From my point of view, the most important thing to learn about implementation is that it must be done project by project over time. This is in almost direct contradiction to the traditional view of planning, which tends to focus on the ultimate system as it could someday be. It may be this dichotomy that leads to the frequent failure of planning. To be successful we must, I believe, organize the implementation of our projects one by one over time. If our predictive models are to be useful, they must be able to show us what happens as each project is either implemented, dropped, or delayed. The combinatorial possibilities of this are too many to conceive, but in practice we will be able to explore all the combinations we will need to analyze as long as we focus on possible alternatives to be tried on the relatively short run and make long-term runs to see their ultimate impact on socioeconomic activity and discounted present value of the benefits.

Each time-staged set of projects becomes a plan to analyze in the system model. As the years go by, the runs start a year later and run a year longer. The construction budget projected for each year into the future interacts with project costs to set an upper limit on the set of projects that can be implemented. Dropping a project out of the plan makes it possible to introduce new projects or advance other projects scheduled for later.

The modeling process that I am referring to produces time series as output (Fig. 5). The time series produced are the performance measures that I indicated previously. One such measure might be the total system inclusive price discussed earlier. Because this is one measure of total costs to the public of using the system, it would be interesting to compare the discounted present value of this figure to the discounted present value that represents the cost of constructing it.

This is not the only, or perhaps even the most relevant, output of such a model system. The discounted present value of any variable of interest can be presented. For example, the number of person-hours of noise higher than 30 perceived decibels might be shown visually by zone or the revenues presented by mode. The possibilities are endless. Design and implementation of the system will, however, depend in large part on the variables we have to predict.

Finally, it is useful to indicate how such a model system is calibrated (Fig. 6). Of course, the individual components will have been calibrated independently. The portion that remains to be calibrated might be viewed as the time constraints of the overall dynamic system. The lags and leads of the various time series must be adjusted so they will correspond to those found in the real world. Therefore, I conclude that the appropriate method of verifying these constants is to attempt to use the models to replicate history. As adjustments are made to the constants, one may view this as using up degrees of freedom. The constraints obtained in one metropolitan area may not be exactly the same as those in another. Comparisons between areas may then constitute a sort of cross-sectional analysis.

Figure 5. Some time series inputs and outputs.

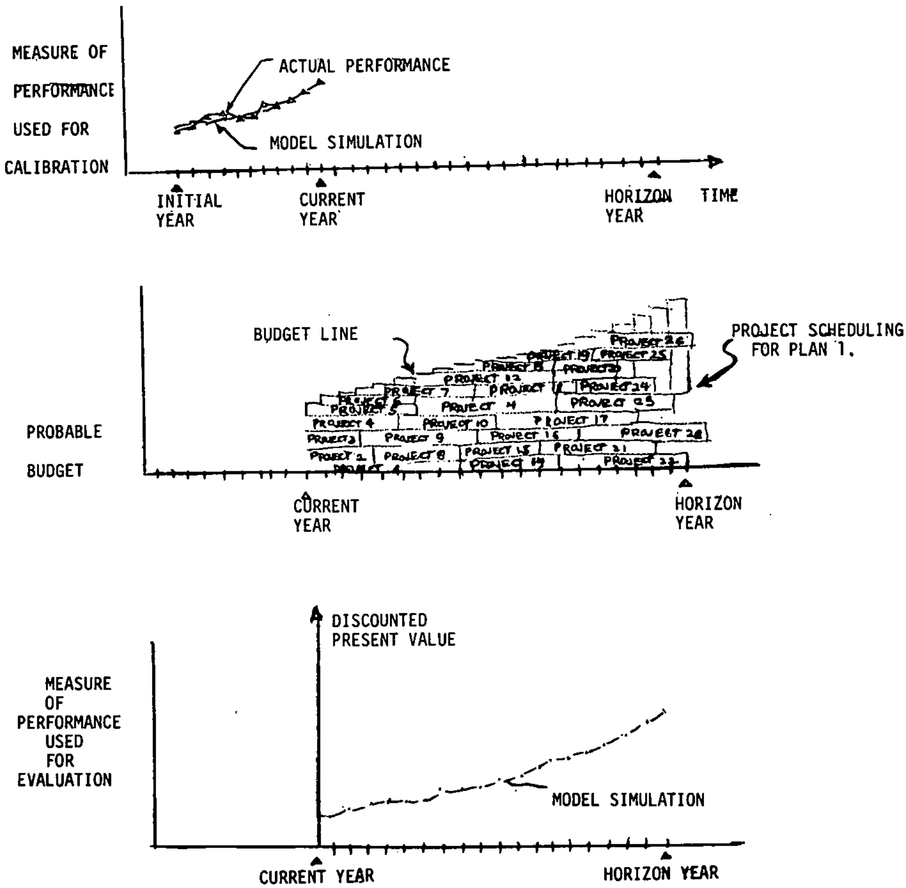
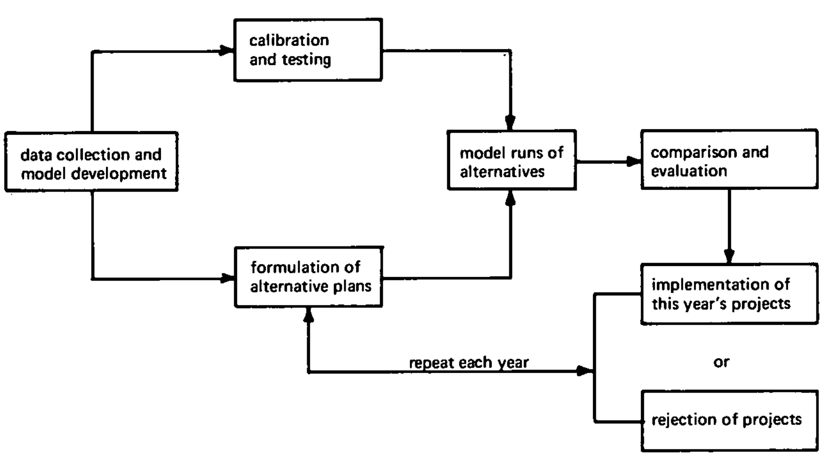


Figure 6. Model calibration and use for specific urban area.



SUMMARY AND CONCLUSIONS

The process I have described here is neither simple to develop nor simple to use. In fact, long-range forecasting is not simple to do. However, I do not believe that extreme accuracy is needed. What is needed is some indication of the range of values as well as the direction and magnitude of how the world will begin to change in response to the changes we are contemplating. We can use the models we now know about to forecast short-run outcomes of changes, and these should be of great benefit until we can get the other longer range activity-shift models developed and calibrated. After they are integrated into the system, it will take some time to develop the understanding of what the urban growth process is, what its controls are, and how we can manipulate the system to produce a favorable result.

The effort required to accomplish this will be major. The funds required will not be inconsequential, but let me hasten to add that, in light of the funds we will use during the next few years to build systems that are uneconomical, inefficient, and poorly planned, the funds will be trivial in the extreme. Furthermore, it is about time that we as human beings learned how our combined decisions are causing our cities to grow in ways that we do not want them to. We need the understanding that the development of this series of models implies. It is time we began in earnest.

REFERENCES

1. Urban Transportation Planning. Federal Highway Administration, June 1970.
2. Brown, J., Ginn, R., James, F., Kain, J., and Straszheim, M. Empirical Models of Urban Land Use: Suggestions on Research Objectives and Organization. Nat. Bur. of Econ. Res., New York, Rept. 6, 1972.
3. Measurement of the Effects of Transportation Changes. Charles River Associates, July 1972.
4. Manheim, M. Practical Implications of Some Fundamental Properties of Travel Demand Models. Dept. of Civil Eng., M.I.T., Cambridge, Res. Rept. R72-63, Oct. 1972.
5. A Disaggregated Behavioral Model of Urban Travel Demand. Charles River Associates, March 1972.
6. Kresge, D. T., and Roberts, P. O. Techniques of Transport Planning. In Systems Analysis and Simulation Models, Brookings Inst., Washington, D.C., Vol. 2, 1968.
7. Manheim, M. L. Fundamental Properties of Systems of Demand Models. Dept. of Civil Eng., M.I.T., Cambridge, 1971.
8. Blackburn, A. Unpublished PhD thesis. Harvard Univ., Cambridge, Mass., 1964.
9. Brand, D. Travel Demand Forecasting: Some Foundations and a Review. Paper in this Spec. Rept.
10. Kraft, G. Demand for Intercity Passenger Travel in the Washington-Boston Corridor. Systems Analysis and Research Corp., 1963.
11. McLynn, J. M., and Watkins, R. H. Multimode Assignment Model. In Approaches to Modal Split: Intercity Transportation, U.S. Dept. of Commerce, 1967.
12. Quandt, R. E., and Baumol, W. J. Abstract Mode Model: Theory and Measurement. Jour. of Reg. Sci., Vol. 6, No. 2, 1966.
13. Domencich, T., Kraft, G., and Valette, J. P. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. Highway Research Record 238, 1968, pp. 64-78.
14. Manheim, M., and Ruiter, E. DODOTRANS I: A Decision-Oriented Computer Language for Analysis of Multimode Transportation Systems. Highway Research Record 314, 1970, pp. 135-163.
15. Martin, B. V., Memmott, F. W., and Bone, A. J. Principles and Techniques of Predicting Future Urban Area Transportation. M.I.T. Press, Cambridge, 1965.
16. Koppelman, F. Research Into Development of Macro Travel Demand Model for Urbanized Areas: A Preliminary Study. Dept. of Civil Eng., M.I.T., Cambridge, unpubl. rept., Nov. 1972.

17. Ingram. A Simulation Model of a Metropolitan Housing Market. Nat. Bur. of Econ. Res., New York, Sept. 1971.
18. Ingram. The Detroit Prototype of the NBER Urban Simulation Model. Nat. Bur. of Econ. Res., New York, Feb. 1972.
19. Measurement of the Effects of Transportation Changes. Charles River Associates, July 1972.