ANOTHER VIEW OF TRANSPORT SYSTEM ANALYSIS

Martin Wohl

February 1968

P-3785

ANOTHER VIEW OF TRANSPORT SYSTEM ANALYSIS

Martin Wohl*

The RAND Corporation, Santa Monica, California

I. INTRODUCTION

With the interest in and use of system analysis increasing so enormously and with the development of analysis and problem-solving capabilities improving so rapidly, one must wonder -- properly -- about the necessity of yet another discourse on transportation system analysis. Certainly some analysts remain convinced that they already are conducting appropriate transport system analyses and that from here on out the task is simply one of obtaining the requisite project funding and "getting on the job." Others, though, would assert that even our statement of the conceptual framework and elements of transport system analysis -- much less our understanding of the working details -- needs modification, amplification and clarification prior to embarkment on large-scale systems studies or projects. The latter stance I find more realistic and must conclude that additional "talk" about transport system analysis will be helpful if not necessary.

As a consequence, in this brief paper an attempt will be made to structure transport system analysis more carefully, to characterize the inter-workings of appropriate system models, and to indicate positive directions for improving the quality of our analysis.

---

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors.

This paper to be published in the Special Issue of the Proceedings of the IEEE on Transportation.

II. A SYSTEMS APPROACH: ITS OBJECTIVE AND ESSENTIAL ELEMENTS

In barest form, to take a "systems approach" is to conduct the analysis and design (or synthesis, if you prefer) systematically and rationally, is to insure to every extent possible that all reasonable technological and policy alternatives are considered, and is to observe and measure in quantitative and objective terms those beneficial and adverse consequences stemming from alternative designs or policies. In this fashion it is to better enable decision-makers to select the "best" design or policy.[1]

Of course, to talk about a systems approach in this broad, comprehensive fashion is one thing, but to carry it out is indeed another. The disparity may not be so much a matter of lack of either understanding or desire on the analyst's part as it is a problem stemming from system scale and complexity and from lack of knowledge of the behavioral relationships. More specifically, as systems become larger, as the number of possible events increase significantly, and as the number of linkages magnify almost beyond comprehension, the task of characterizing the actions and interactions of the system elements -- that is, the task of identifying "what will happen if" -- is complicated enormously. Add to this complexity the risks and uncertainties of estimating the

---

[1] Realize, though, that the "best" design or policy may and probably will involve either political or equity considerations as well. Note, further, that system analysis is not to be confused or equated with sheer mathematical rigor, nor are mathematical techniques or tools even a necessary requirement for a systems approach. Rather than to imply a false sense of sophistication, a systems approach is to describe the general character, content and quality of the design or policy analysis, and it is to connote the manner in which the analysis is conducted.

future and the virtual infinity of alternatives open for discussion and one can more easily understand our failure to hardly feel other than a sense of frustration.

Amid these constraints and difficulties, and the simple realities of large-scale analysis, let us examine in detail some of the more important characteristics of a systems approach.

The first crucial issue involves one's point of view; that is, it involves designating whose interests and welfare are of concern and at stake. In more direct terms, the analyst must specify with some considerable care the particular group of people for whom the "best" plan is being sought; for example, is the analyst to adopt the point of view and to select the "best" subset of alternatives for city residents, or that for county, state or national residents? Is he to employ the value scales of the residents of the city, or those of the state, county or nation? And so forth. The necessity for precision on this score should be obvious (but apparently is not, judging from past experience); two points are appropriate. First and most obvious, the balance of gains and losses resulting from any specific design or policy will usually differ from one political or geographical unit to another, thus altering the desirability to the various units. Second, delineation of the particular group of people for whom the analysis is being conducted will provide the boundary which serves to indicate which gains (benefits) and losses (costs) are internal and which are external to the unit.[1]

---

[1]External benefits or costs are those gains or losses stemming from a design or policy which do not affect and thus which are (generally) not recognized by those people for whom the "best" design or plan is being

Second, the analyst must designate the particular consequences which are to be measured and used (collectively) to gauge the effectiveness of alternative designs; the measures of effectiveness for these consequences should be quantifiable, objective and representative of the entire system.

Third, the analyst must be able to evaluate or place objective values upon the various consequences such that the overall effectiveness can be judged; to satisfy this particularly difficult (though important and unavoidable) task does require that objective and commensurate value scales be used but in no sense implies that both pecuniary and non-pecuniary consequences and measures of effectiveness cannot be dealt with. Regarding the latter point, an insistence upon use of commensurate value scales for all measures and consequences, whether pecuniary or non-pecuniary, is simply to insure that non-pecuniary objectives are properly considered and "weighted" by the designer, analyst and decision-maker and that they are neither overvalued nor undervalued in the system planning process. And it is to recognize explicitly that tradeoffs between money and non-money values, between economic and social consequences, and between present-day and future generations are at stake and cannot be avoided simply by one's refusal to be more explicit about the process.

---

sought. Nevertheless, externalities <u>are</u> part of the <u>system</u> effects and should be identified and measured by the system analyst. Of course, whether or not the first party does take them into account depends on the extent to which they can be internalized (by income transfers, pricing mechanisms or laws of one sort or another) or perhaps on the beneficence of those affected advantageously.

Fourth, the analyst must be able to model the system and to predict the consequences stemming from the imposition of specific technological or policy alternatives.[1] Certainly there can be no more important task than that of developing the models, techniques and procedures which will permit (sufficiently) reliable measurement of the detailed consequences of different actions (on the part of operators, designers, etc.). And of equal certainty, there can be no more difficult task -- given the risks and uncertainties stemming from both human and natural events and from preferences, all of which defy prediction, and given the complexities and scale of transport systems and the human and social environment with which they interact.

Fifth, the analyst must: (1) create or search for alternative transport designs, technological devices and operating (or regulatory or pricing, etc.) policies; and (2) select a subset of alternatives (from a virtual infinity of possibilities) for more intensive investigation and closer analysis. In carrying out this two-step process -- of searching among all alternatives and of selecting that subset which probably contains the "best" design -- the analyst ideally (though not realistically) would be able to foresee all future inventions, technologies and events which might affect the system effectiveness, and would have at his disposal a "search" or alternative selection methodology which would insure that the "best" design was included within the smaller subset. In a certain sense, this is the key step in system engineering -- for it is evident that the "best" or even a

---

[1] The requirements and prospects for carrying out this task will be set forth in the section entitled, "The Technical Components and Requirements for Predicting System Consequences."

"better" plan will not be adopted if it is not even considered or offered for additional analysis and evaluation. Yet, even if the "best" plan is considered, and chosen for detailed analysis and evaluation, either the analyst or decision-maker must be able to ascertain that it is "best." This part of the system engineer's task bears an important relationship to the other phases of the process and has meaning only insofar as the remaining phases are properly and adequately handled. Without having specified appropriate measures of effectiveness, without reliable prediction techniques, and without appropriate evaluation methods and decision-making criteria, it is doubtful that much reliability can be placed in our judgment for choosing the "best" subset of alternative plans, designs or technologies, even if our ability to foresee future events is superlative.

In the process of meeting these tasks -- of searching for better technologies or operating policies, or predicting the consequences of the alternatives and of evaluating these consequences -- the system engineer should not fail to recognize that his job is not just (or even necessarily) one of building a more sophisticated technological system nor is it one of determining the best among the new systems improvements; that is, his responsibility is not to "get something done" or "to build something," regardless of all else. Rather, it is one of considering not only new or improved systems and technologies but also one of evaluating the wisdom of continuance, abandonment, or contraction of the existing facilities. Summarizing, the job of the system engineer is one of developing technologies and operations which will be consistent with man and his environment (both

present and future) and on certain occasions it is a task of recognizing that the most profitable system decision will be to do nothing, at least for the time being.

The above remarks seem particularly important in the face of increasing technology, of increasing knowledge and scientific activity since we seem more preoccupied with questions about how we can design and build this or that exotic transport vehicle or propulsion device than with questions about whether we should build them at all. In short, we may well be overlooking the possibility that the "ends do not justify the means," however glamorous or modernistic or advanced the latter may be.

III. THE TECHNICAL COMPONENTS AND REQUIREMENTS FOR PREDICTING SYSTEM CONSEQUENCES

The system consequences of interest and concern are the costs (or resource commitments, whether human or material) which will be forfeited and the benefits (or gains, whether tangible or emotional) which will be received if some specific design or alternative is adopted. However, to predict and measure these consequences -- the costs and benefits which will result from adopting some particular design or policy -- requires knowledge about cost, price, performance and demand functions and requires that they be specified and interrelated on a realistic network and time-period basis. In more specific terms, it is necessary first to determine what the equilibrium flow or actual tripmaking will be *if* a specific facility or system is built and operated in a specified manner and second to determine what resources must be committed in order to provide for that amount of tripmaking, whether or not the travelers perceive or are aware of the expenditures.

The above statement of prediction and measurement differs markedly from the more common view of the forecasting problem, particularly from that implicit in many operations research-type models and analyses. Specifically, the system analyst should not be seeking to determine the resources required to "meet a need" nor trying to determine the "best" system to "meet the demand" wherein "the need" or "demand" is regarded as fixed in the all too common requirements sense.[1]

---

[1] A characteristic example of the incorrect "fixed requirements" approach may be seen in the illustration of the AFT input-output model (developed by BAARINC), as shown in W. S. Beller, "Attacking the Systems Problem," Space/Aeronautics, September 1967.

Rather, the problem is a more fundamental one in which it should be recognized that system performance is dependent both on the facility design and the extent to which the system is used, that (simultaneously) the extent to which the system is used is dependent on the system performance or service levels (as well as on price) and on the socio-economic characteristics of travelers and of the region, and thus that equilibrium performance and flow or volume levels are highly interdependent.[1]

Space does not permit a complete treatment of the substantial and important topic of travel forecasting -- that is, of determining what the actual or equilibrium travel volumes and service (or performance) levels will be for specified design and operating conditions. However, it will be useful to at least sketch the overall structure of this problem and to note the most important aspects of forecasting.[2] For this purpose, the discussion will be viewed in three parts: (1) Demand Functions; (2) Price and Performance Functions; and (3) Equilibration of the Price, Performance and Demand Functions, such that equilibrium flows and performance levels can be determined.

---

[1] If, for illustrative purposes, we cast this problem within a linear programming format, the link or unit costs (or prices) for travel networks are not uniform for all flow levels, the incoming or outgoing shipments or numbers of trips from each zone are not fixed regardless of the interzonal cost, and the interzonal transfer or trip distribution problem is not simply one of cost minimization; all in all, the determination of the actual equilibrium flow (and price) levels cannot be determined by simply another large-scale application of the Hitchcock-Koopmans Transportation Problem program.

[2] A more complete but still introductory article on this subject is "New Directions for Passenger Demand Analysis and Forecasting," by G. Kraft and M. Wohl, in Transportation Research, Vol. I, No. 3, 1967.

Demand Functions

In brief, travel demand functions are conditional tripmaking models which relate the amount of tripmaking to those service and price variables which influence the number of trips, as well as their mode, the time at which the trip is made, its route, and so forth. To characterize demand (that is, the dependent relationship between quantity of tripmaking and the independent service or price variables as a function[1] rather than as a "fixed point estimate or requirement for some future design year" and as an estimate to be regarded as invariant with the system design and its service or price clearly is a more realistic manner of representing the tripmaking, price and service interactions. Moreover, demand functions not only should be stratified according to modal or route preferences but they also should permit time-of-day preferences to be reflected; more directly, demand models (which presently are partially imbedded within most trip generation, trip distribution, modal split and route assignment models and

[1]The intercity passenger demand models that have already been developed for the Northeast and California corridors serve as important forerunners (as do the commodity demand models of similar form developed by E. D. Perle, Demand for Transportation, University of Chicago, 1964); an early model of this type was that developed and reported in Part V (authored principally by G. Kraft), Demand for Intercity Passenger Travel in the Washington-Boston Corridor, report submitted to the Office of the Under-Secretary of Commerce for Transportation, U.S. Department of Commerce, by Systems Analysis and Research Corporation, 1963. A follow-up demand model of similar characteristics and form was developed by R. Quandt and W. Baumol, Abstract Multi-Mode Travel Model: Theory and Measurement, Technical Report No. 4, Northeast Corridor Project, U.S. Department of Commerce, 1966. Also, this is not to overlook the pioneering development of urban passenger demand models recently completed by Charles River Associates Incorporated and Traffic Research Corporation for the Division of Bay Toll Crossings, State of California.

procedures)[1] should condition the amount of tripmaking during hour H_1, say, not only on the travel price and service during hour H_1 but also on the travel prices and services which will be afforded at other times of the day. In brief, to incorporate this additional time-of-day preference feature within our demand models would permit explicit treatment of the important institutional arrangements which affect the hour in which we travel (such as starting to work times, store opening and closing times, etc.) as well as the full range of choices which the traveler considers and expresses preferences for when deciding whether or not to travel, where to travel, what mode and route to use and when to travel; quite specifically, we could explicitly treat questions such as: "Would you rather travel during your preferred hour of the day and endure heavy traffic congestion, or would you rather arrive at your destination later than preferred but endure only light congestion?"

Also, among other important features, two items seem worthy of particular mention. First, tripmaking -- and its time-of-day, modal, route, and zonal characteristics -- should be looked upon as a highly interacting if not simultaneous decision-making calculus on the part of the traveler. That is, the decision about whether or not to travel is not independent of the modal possibilities, of the time-of-day travel conditions, and so forth, and by the same token the modal choice is not made independently of the other modal and time-of-day choices

---

[1] Among many descriptions of these types of procedures, see N. A. Irwin, N. Dodd, and H. G. Von Cube, "Capacity Restraint in Assignment Programs," Highway Research Board Bulletin 297, 1961; and M. Wohl and B. V. Martin, Traffic System Analysis, McGraw-Hill Book Co., 1967, Chapter 5.

open to the traveler. Second, in characterizing the price and service (or performance) conditions which face the traveler and influence his tripmaking choices (including that of deciding whether or not to travel), most travel forecasting procedures presently fail to take into account other than enroute travel conditions; that is, most travel forecasting models presently condition trip distribution, modal or route choices simply upon the elapsed enroute travel time (to include any waiting, transfer, or walking time or that stemming from car pooling or parking arrangements, etc.). By contrast, though, the travel forecasting procedures and demand models should incorporate another type of travel time variable, one which might be termed "schedule delay." Schedule delay can be defined as the difference between the traveler's actual arrival time at his destination and his preferred arrival time and can result from either a late or early arrival. While schedule delay can result both from infrequent schedules and from low capacity (relative to equilibrium travel volumes) can arise both for transit and auto travel,[1] one should not infer that the introduction of all stochastic processes and queuing aspects will properly account for schedule delay in all cases.

Clearly, the formulation of a travel demand function which will incorporate some or all of the factors noted before can take many forms, and no attempt will be made herein to suggest which seems most appropriate. Nevertheless, for the purpose of illustrating certain structural qualities and interrelationships, two different types of

---

[1] See Kraft and Wohl, op. cit., for elaboration on this point.

tripmaking models will be discussed in some detail. The first type is drawn from micro-economic theory and takes the form of a multivariate exponential function; the second is commonly referred to as a "gravity model" though more properly should be termed a relative gravity model or trip proportioning model.

The Exponential Form of Demand Model

This type of demand model embodies the following essential characteristics:

1. Passenger tripmaking can be determined for any set of specified service, price, and socio-economic conditions;

2. Changes in tripmaking which result from changes in trip price or service or from those in income, population or industrial growth, and so forth, can be easily identified and measured;

3. The amount of tripmaking or its change can be determined in both absolute and relative terms, thus permitting simultaneous determination of the aggregate amount of travel and its distribution among zonal pairs and modes, etc.

One such equation for expressing $V(ij|p,h)$, the quantity of tripmaking (for purpose p and at time-of-day h) which will be made between a pair of origin (i) and destination (j) zones under a set of specified trip price, service and socio-economic conditions, might be represented as follows:[1]

$$V(ij|p,h) =$$

$$\alpha_{ph}(Y_i)^{\beta_{ph}}(P_i)^{\gamma_{ph}}(E_j)^{\delta_{ph}}\prod_y c(ij|p,y)^{\theta_{phy}}\prod_y f(ij|p,y)^{\varphi_{phy}}\prod_y s(ij|p,y)^{\psi_{phy}} \quad (1)$$

[1]While this demand function could be made more realistic by the inclusion of other stratifications to represent the mode and route of

where $V(ij|p,h)$ is the trip volume which will move between zones i and j for purpose p and during time-of-day h under the given price, service, and socio-economic conditions; Y_i is an income measure for residents of origin zone i; P_i is a population measure for zone i; E_j is an employment measure for destination zone j; $c(ij|p,y)$ is the enroute travel time or congestion for purpose p and time y trips being made between zones i and j; $f(ij|p,y)$ is the schedule delay for purpose p and time y trips between zones i and j; α_{ph} is a parameter for purpose p and time-of-day h travel; the exponents β_{ph}, γ_{ph}, δ_{ph} are the demand elasticities[1] for purpose p and time-of-day h travel with respect to income, population and employment, respectively; and the exponents θ_{phy}, φ_{phy} and Ψ_{phy} are the demand cross-elasticities for purpose p and time h trips with respect to the enroute travel time, travel fare and schedule delay, respectively, for purpose p and time y trips.

While no attempt is made to defend the propriety of this particular model form, certain of its structural qualities and implications with respect to the tripmaking process deserve emphasis. First, note that the total amount of tripmaking (i.e., the trip generation, to use the most common terminology)[2] is permitted to vary according to the

---

travel, the simpler form was chosen to make the exposition less cumbersome. The exclusion does not affect the underlying principles, however.

[1] The elasticity of demand with respect to some variable x is defined as the percentage change in trip volume (V) which results from a one percent increase in the variable x or as $(\partial V/V) \div (\partial x/x)$. For this particular model form, the demand elasticities will be constant over all service levels -- an assumption which probably is not valid other than for limited portions of the entire range of values.

[2] Trip generation is generally used to indicate the quantity of trips generated by a particular type and amount of land use at some geographical location; most traffic forecasting processes regard trip generation and thus the total amount of tripmaking as fixed or invariant with respect to the transport system and its service and price levels.

quality and cost of travel service in addition to the level of population and income; thus, trip generation is not assumed to be fixed at some arbitrary level regardless of locational, technological and service qualities of the transport system and regardless of the native preferences of travelers. Second, trip distribution, modal split and route assignment (i.e., trip-splitting procedures for proportioning trips among different destination zones, modes and routes, respectively) are not treated as travel choices which are made or exercised independent from each other and from the decision about whether to travel at all; rather, they are regarded and dealt with as a simultaneous tripmaking package (that is, one does not travel regardless of what modes are available, of how bad travel conditions are and of how expensive the trip will be). Third, this model incorporates two additional features over those in more common use, both of which seem to be important and necessary to fully explain tripmaking behavior; they are: time-of-day travel preferences and schedule delay. The first of these will permit proper prediction of the peaking characteristics of a transport system and the second will permit recognition of an additional service variable which influences tripmaking behavior and causes shifts to other modes or times of day, etc.

Again, this type of demand model can hardly be regarded as "perfect" but does embody and thus serve to highlight certain important structural qualities and to provide a useful starting point for development of more fundamental demand functions. Further, it is worth mentioning that transformation of the expression shown in equation (1) to its log/log form (i.e., by taking the logarithm of both sides of the

equation) places it in a form that is readily adaptable to multiple linear regression analysis for determination of the exponent or elasticity values; while this particular mathematical quality is most useful for parameter estimation, one should realize, however, that this type of reduction requires one to assume that the exponent or elasticity values remain constant over the entire range of parameter values, an assumption of doubtful validity.

Gravity-type or Trip Proportioning Demand Model

The so-called "gravity model" is used in conjunction with a series of factoring and trip-splitting procedures to determine $V(ij|p,h,r,m)$, the quantity of trips moving between zones i and j for purpose p, at time-of-day h, by route r and by mode m. For this type of process, the quantity of trips to be demanded under any set of specified price, service and socio-economic conditions are determined by making use of a four-step sequence involving trip generation, trip distribution, modal split, and route assignment.

Two important tasks will be carried out within the trip generation step. First, the number of daily trips of each purpose which originate from each zone will be calculated according to the particular land use activities and locational qualities of the zones in question. Second, the daily trips will be converted to peak- and off-peak period trips produced or generated by each zone, and stratified by trip purpose, then serve as inputs to the trip distribution step.

The trip distribution model serves to distribute the total number of trips emanating from any origin zone i among all other $(n-1)$ destination zones. One frequently used procedure for distributing

trips among the zones employs the so-called "gravity model," or what more appropriately should be termed a relative gravity model or trip proportioning procedure. As usually formulated, the model is as follows:

$$T_{ij}^{p} = T_{i}^{p} \frac{K_{ij}^{p} A_{j}^{p} / t_{ij}^{P}}{\sum_{k=1}^{n} K_{ik}^{p} A_{k}^{p} / t_{ik}^{P}} \qquad \text{for all } k \neq i, \tag{2}$$

where T_{ij}^{p} = interzonal trips from zone i to zone j for purpose p,

T_{i}^{p} = interzonal trips originating in origin zone i for purpose p

$= \sum_{k=1}^{n} T_{ik}^{p}$ for all $k \neq i$,

A_{j}^{p} = interzonal trips ending in destination zone j for purpose p,

K_{ij}^{p} = zone-to-zone "adjustment" or "calibration" factor,

t_{ij} = travel time or travel resistance for trips moving from zone i to zone j,

P = exponent of travel time or resistance for trips of purpose p,

n = number of zones in region.

Customarily, the K factors and P values are empirically derived for a particular region and are assumed to remain constant in the future. Also, it should be evident that this distribution model merely serves to split or proportion a constant number of trips (T_{i}^{p}) among the other (n-1) zones and does not allow the total number of trips to change as the travel time (or price) is increased or reduced.

Once all interzonal transfer volumes (or T_{ij}^{p} for all i and j) are determined, these trips must then be split among the modes serving each

origin and destination pair. While most procedures accomplish the modal split by making use of transit versus highway diversion curves in which the percentage of trips using transit varies with the ratio of the origin-to-destination travel time by transit to that by highway, some more sophisticated ones split the interzonal trips between auto and transit on the basis of relative travel time, relative cost, relative service (or waiting-walking-parking) time and income level of the origin zone residents.[1]

The fourth or route assignment step can be accomplished in a number of ways, though most often is carried out by assigning interzonal trips to that route offering the minimum enroute travel time.

Collectively, these factoring and trip-splitting procedures provide an approximate type of demand function, though one having distinctly different structural qualities than that illustrated in equation (1).[2]

Price and Performance Functions

The second set of technical requirements necessary for the overall travel forecasting process is that which permits description of the relationship between travel service (or, say, the trip price and performance conditions) and system usage or volume levels. That is, to

---

[1]For an example of the latter type of modal-split procedure, see T. Deen, W. Mertz, and N. Irwin, "Application of a Modal Split Model to Travel Estimates for the Washington Area," Highway Research Board Record 38, 1963.

[2]The capacity-restraint and feedback features that are used in conjunction with the trip generation-distribution-modal split-route assignment process outlined above should not be regarded as part of the demand function, but should be looked upon as a representation of the "supply function" and equilibration, respectively. In brief, it is the interrelation of the demand and "supply" functions which permits determination of equilibrium trip volumes and performance levels as distinct from the description of the functional relationship between tripmaking and the transport system performance.

what extent and in what fashion does travel time increase as link volume increases? Does the traveler actually perceive these changes in travel service? Does the travel price change as the volume or usage changes? And so forth.

One serious attempt to introduce these types of functional relationships between travel performance and link volume into the forecasting process is illustrated by the use of capacity-restraint functions;[1] these so-called capacity-restraint functions (which express travel time as a function of the hourly link vehicular volume and of the facility type) are used to insure that the travel time used to determine interzonal transfers and to assign these volumes to modes, routes and links is equal to that travel time which would result if assigned volumes of the predicted magnitude did actually use the facilities in question. In a very real sense, these capacity-restraint functions are entirely analogous to the steady-state waiting time expressions which can be found in the queuing theory literature. However, while these expressions do permit the analyst to determine speed or travel time conditions under different levels of vehicular volume, and for different facility designs or controls, they fail to be realistic or complete on at least two grounds: (1) the formulations are characterized solely in terms of volume rates (e.g., vehicles per hour) -- and without regard for the length of time over which the volume rate is sustained; and (2) the formulations do not permit

---

[1] See for example: N. A. Irwin, N. Dodd and H. G. Von Cube, "Capacity Restraint in Assignment Programs," Highway Research Board Bulletin 297, 1961, Figs. 2 and 3.

detailed evaluation of the performance or delay when flow rates exceed capacity (or, in the jargon of the queuing theorist, when λ exceeds μ). More specifically, queuing models and capacity-restraint functions -- as usually formulated -- apply only to steady-state or stable flow conditions and do not permit treatment either of transient queuing behavior or of exploding queue situations.

That non-steady-state and exploding queue conditions are of interest and of relevance to travel forecasting should be self-evident. First, it is obvious that arrival rates for many travel facilities vary during the day and over time and that arrival rates on some facilities do exceed the service rates during peak periods of the day (i.e., the approaching volume exceeds the facility capacity during one- to two-hour peak periods).[1] The fact that capacity can be approached and even exceeded during peak periods and that congestion can thereby result clearly will influence the amount of travel which can (if not will) take place during peak periods of the day; further, it seems likely that the circumstances during peak periods will also affect the amount of travel during non-peak hours. (In simplest terms, for example, is it uncommon to expect at least some travelers to wait until after the peak period to make their trip in order to avoid the heavier congestion occurring during the peak period?)

---

[1] In assessing the validity of the latter comment some care must be taken in evaluating the evidence. For example, surely if the analyst observes flow on the downstream side of a facility bottleneck, then the arriving volume will not exceed capacity; similarly, if the arrival volume is observed even on the upstream side of the bottleneck, the arrival volume will not exceed capacity if flow is observed over long enough periods of time during which queues will be worked off.

Before dealing with the problem of determining the effect of peak period congestion on tripmaking (and the time of day at which it takes place), it will be useful to outline the form of some example capacity-restraint functions which will characterize the travel time for non-steady-state flow and exploding queue situations (i.e., for cases having arrival rates greater than capacity or service rates). The capacity-restraint functions included for this purpose are shown in Figs. 1 and 2, the first being applicable to facilities and situations having constant service times and equally spaced vehicular arrivals, and the second being applicable, more generally, to random service time and arrival cases. Also, in both instances the initial queue (at start of the time period over which arrivals are being sustained) is assumed to be zero.

Expressions of the sort illustrated in Figs. 1 and 2 are analogous to the usual "capacity-restraint" function which is increasingly employed in travel forecasting models that feature feedback between link capacity and trip assignment, modal split and distribution phases. However, the formulations included herein differ from those usually (if not always) employed in travel forecasting and assignment models in that explicit recognition is made of the dependence of delay on the time period over which arrival rates are sustained. Or, to put the matter in other terms, recognition is made of the fact that queuing and travel delay is dependent upon the length of the peak period as well as the rate at which traffic volume enters or attempts to enter the roads and streets during that period.

In those cases where the pricing policy dictates such, appropriate functions must also be developed to relate travel prices (as perceived

NOTE: $\bar{t} = t_0 + \bar{t}_q$

$\bar{t}_q$ = TIME-IN-QUEUE $\doteq \left(\frac{\lambda}{\mu} - 1\right)\frac{T}{2}$ FOR $\lambda/\mu > 1$

μ = 600 VEHICLES PER HOUR

T = TIME PERIOD OVER WHICH λ IS SUSTAINED; ALSO, INITIAL QUEUE IS ZERO

t_s = AVERAGE SERVICE TIME AT TOLL BOOTH (OR SIGNAL, ETC.)

t_{da} = AVERAGE TIME LOSSES FOR DECELERATING AND ACCELERATING ON APPROACHING AND LEAVING THE TOLL BOOTH (OR SIGNAL, ETC.)

L = LINK DISTANCE (IN MILES)

v = UNIFORM SPEED ON LINK L (IN M.P.H.)

FOR THE CURVES SHOWN, L=1, $t_s = t_{da} = 10$ AND v = 40

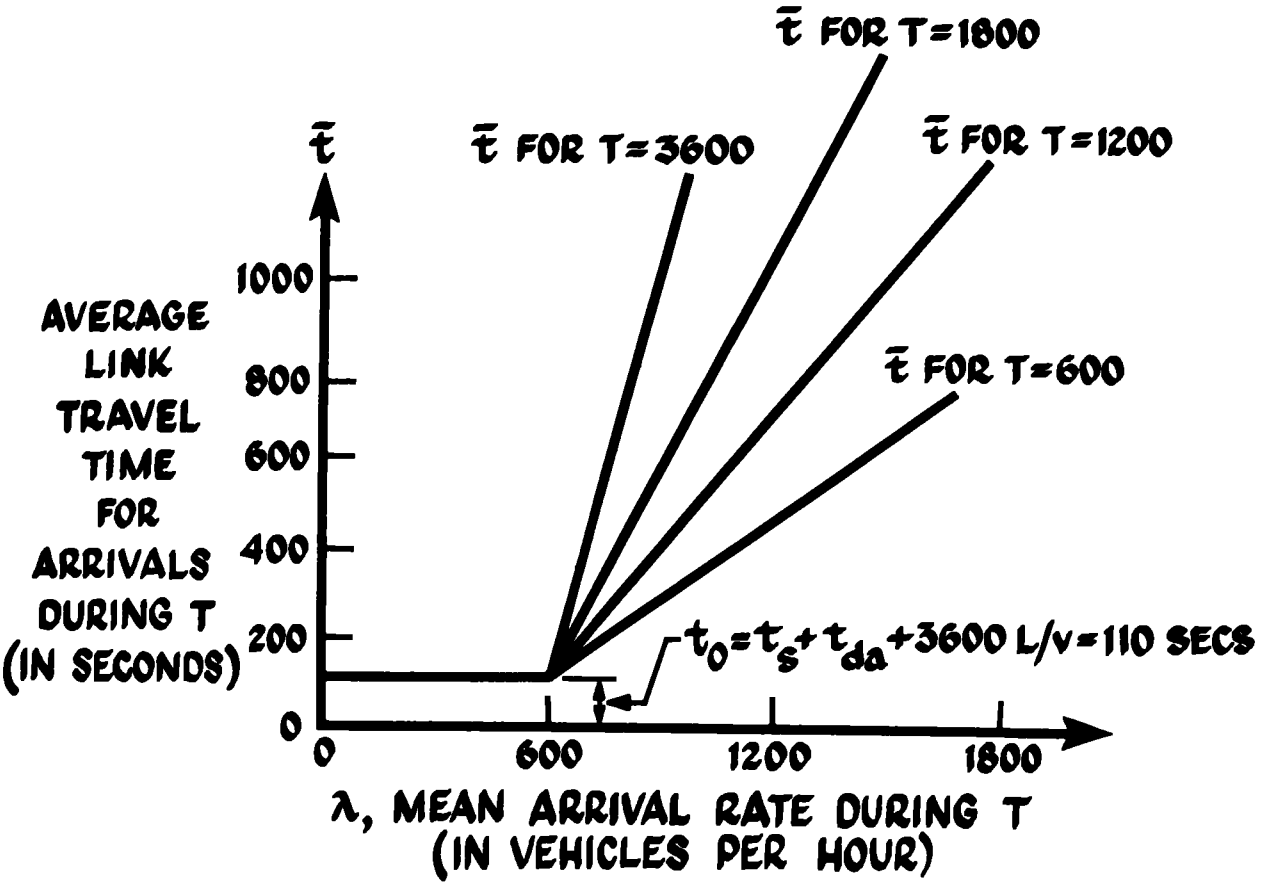

Fig. 1—Travel time versus volume rate functions for constant service times and equally spaced vehicular arrivals

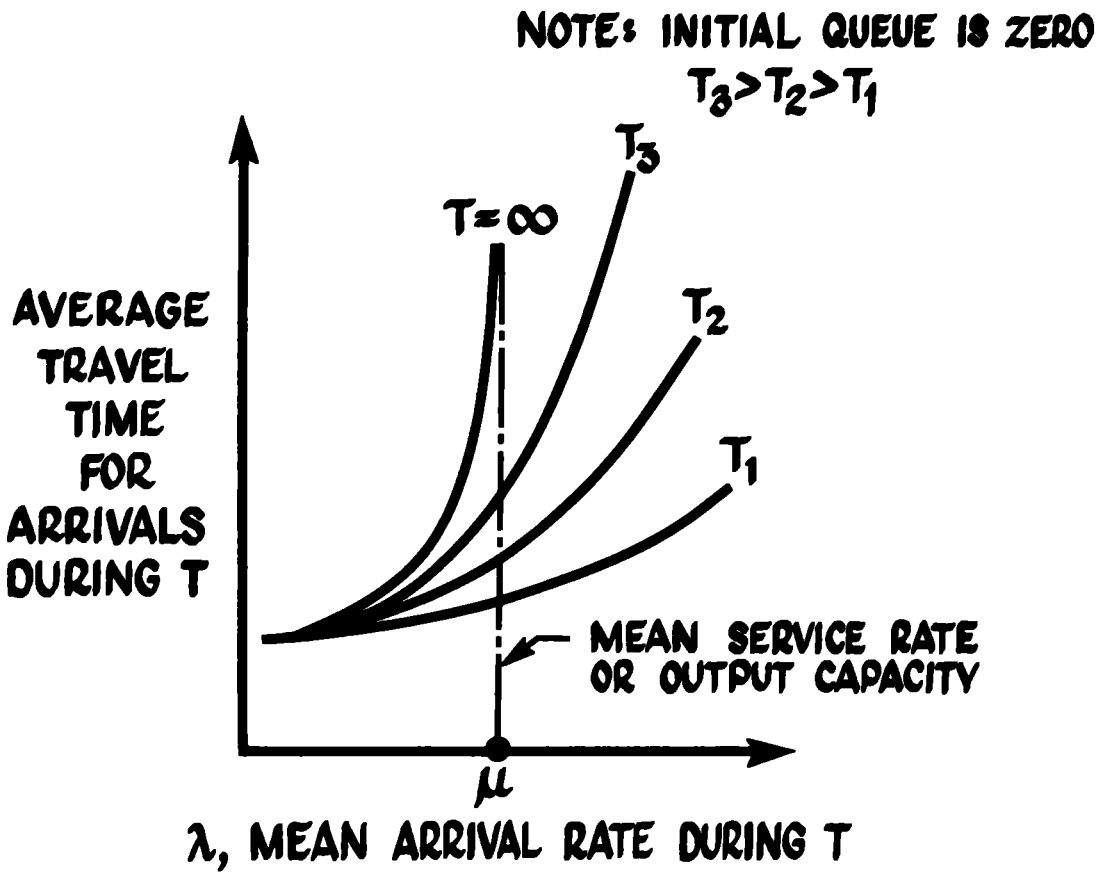

Fig. 2—Generalized travel time versus volume rate functions

by the traveler) to the volume of tripmaking.[1] More specifically, for an average cost pricing policy the trip price will be volume dependent and thus functional relationships will be required; on the other hand, in those cases when the traveler does not perceive changes in the travel costs (for whatever reasons) -- and thus the price stays constant -- or when the price is uniform regardless of the volume level, the problem is simplified considerably. Examples of the former (in which the travel price is volume dependent) would arise in those instances when toll authorities attempt to establish the uniform daily toll rate so as to provide sufficient revenues to just cover their bond payments. For the latter (or uniform prices) it might seem appropriate to characterize the travel price for most urban (but non-toll facility) trips as uniform and invariant with link volumes, at least to the extent perceived by travelers.

Cost Functions vs. Price Functions and Use of Former

Above it was noted that prices must be related to costs in those cases when the pricing policy called for cost-related prices; either average cost pricing or marginal cost pricing would be appropriate examples for indicating the necessity of having proper knowledge of costs. However, it is of some considerable importance to emphasize that the discussion thus far was limited to determining equilibrium travel volumes, performance levels and prices and thus was concerned with costs only to the extent that prices were cost related or that travelers considered costs in their trip decision-making calculus.

---

[1] In this discussion, of course, a distinction must be made between trip price and trip travel time.

This does not mean, of course, that we are disinterested in or unconcerned with actual costs or resource commitments; rather it simply means that for forecasting the extent and character of tripmaking we may (in certain instances but not in others) not need such cost information.

On the other hand, it is perfectly clear that cost data is required for any full-scale system analysis and evaluation, thus dictating that cost relationships must be known in some detail and must be used in conjunction with the data on equilibrium volume levels in order to determine the actual costs which must be foregone to provide the service that is demanded for the alternative in question. Also, it is clear that this costing should deal with both private and social (or internal and external) costs and should pinpoint their incidence (that is, denote who pays how much of the costs).

Equilibration of Price, Performance and Demand Functions

To equilibrate the price, performance and demand functions is, in equivalent terms, to interrelate these functions such that we can determine or forecast the (equilibrium) price, service and volume levels which will satisfy simultaneously the three sets of contraints. Symbolically, forecasting or equilibration may be defined (somewhat) as follows:

Find non-negative $\underline{V}$, $\underline{p}$, $\underline{s}$ and $\underline{t}$ such that

$$V = f(p;\ s;\ t;\ \text{socio-economic factors}) \qquad (3)$$

$$p = g(V;\ \text{facility design and operation;\ pricing policy}) \qquad (4)$$

$$s = h(V;\ \text{schedules and capacity;\ institutional/socio-economic factors}) \qquad (5)$$

and

$$t = k(V;\ \text{facility design and operation}). \qquad (6)$$

The resultant price ($\underline{p}$), flow level ($\underline{V}$), schedule delay ($\underline{s}$) and travel time ($\underline{t}$) which satisfy equations (3) through (6) represent the equilibrium price, flow, schedule delay and travel time conditions and, thus, are a forecast for the particular facility or system and pricing policy being analyzed.

To place this simple view of forecasting on an operational basis, particularly to the point at which multi-link and multi-modal transport systems can be dealt with, is clearly a task far beyond this brief summary paper. Even so it will be useful, first, to elaborate on some of the important aspects of large-scale system equilibration and, second, to illustrate some of these considerations with an example (though admittedly oversimplified) forecasting program.

(1) A Discussion of Large-Scale Equilibration Problems. To begin, specification of consistent and compatible input and output units is of crucial importance to determination of cost, price, performance and demand functions and to the equilibrating process. The task becomes particularly difficult and complex because most network links serve particular interzonal trip movements for only a portion of the entire door-to-door trip and because network links are utilized by travelers moving between a wide variety of origin and destination zones.

To illustrate these types of considerations, the region shown in Fig. 3 and represented by a series of nodes (or origin and destination zones) and connecting transport links or facilities can be utilized; essentially, we will assume that this set of zones and links is independent of and isolated from outside or external effects and activities. To specify the output units, at least four aspects must be

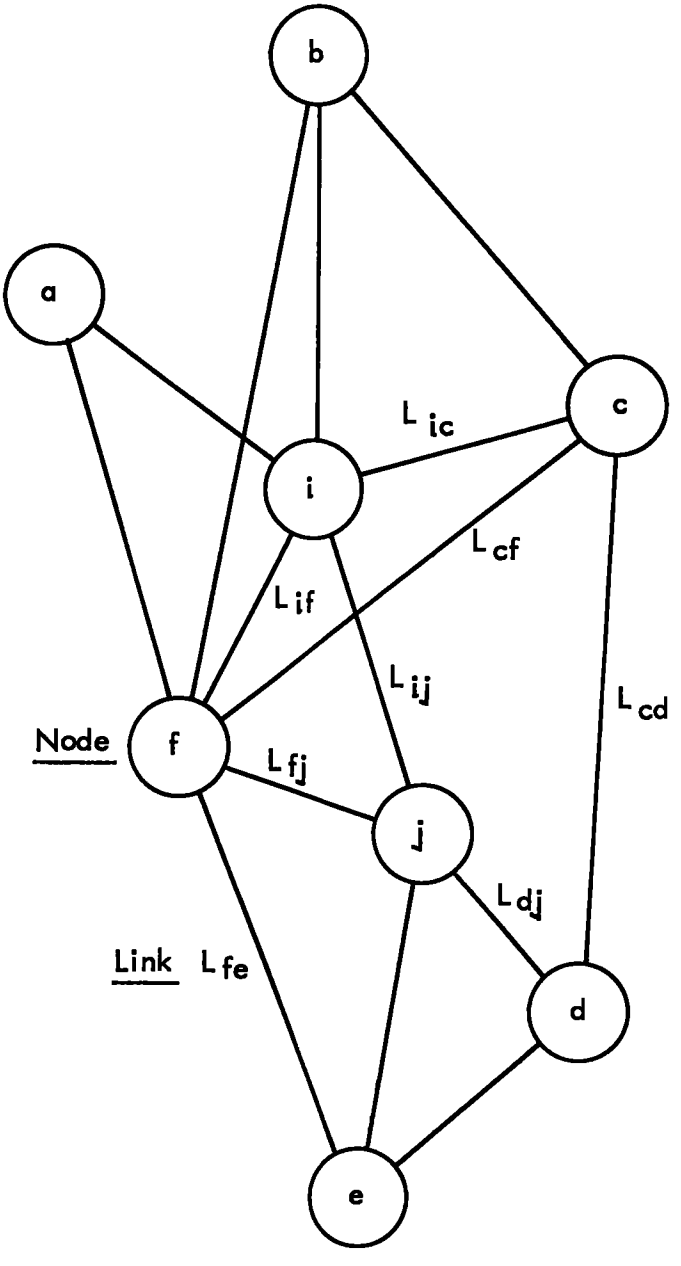

Fig. 3—Isolated link and node travel system

accounted for: (1) the trip unit (i.e., persons, vehicles and their classification, tons, or combination thereof, etc.); (2) the origin and destination of the trip; (3) the trip route or links over which the trip is made; and (4) the price and travel time for the route over which the trips are made, as well as the time-of-day and time period or year during which they are made.

The above distinctions are made necessary because: (1) the cost, price and performance functions will vary from link to link and will vary with changes in the composition of trip units and with changes in the time interval; and (2) demand will vary for different origin and destination pairs, for different times of day (and time periods), as well as for other factors. Equilibration will therefore require that the cost, price, performance and demand functions are consistently stated in terms of equivalent input and output measures accounting for these four aspects.

Equilibration will also be complicated in two other major respects ignored thus far. One, the demand for travel between any two nodes or zones $\underline{i}$ and $\underline{j}$ is not or may not be dependent solely upon the equilibrium travel time and price for link $\underline{L_{ij}}$ but may also be dependent upon the equilibrium travel times and prices (etc.) for other alternative routes connecting nodes $\underline{i}$ and $\underline{j}$ (such as route $\underline{L_{if}} + \underline{L_{fj}}$ or $\underline{L_{ic}} + \underline{L_{cd}} + \underline{L_{dj}}$). Two, the equilibrium volume, travel price and travel time on each of the links connecting nodes $\underline{i}$ and $\underline{j}$ (including $\underline{L_{ij}}$) is dependent both upon the demand for travel between nodes $\underline{i}$ and $\underline{j}$ and upon the demand for travel between other pairs of nodes which can be made over the same links; for instance, the equilibrium flow

on link L_{ij} can consist of trips being made between i and j, between a and d, between c and j, between a and j, and so forth. For other than those circumstances where a link and the demand for travel over that link can be isolated from all other portions of the network or system, a situation which is rare, it will be necessary to use iterative computational procedures for determining the equilibrium flow, travel time and price conditions (regardless of the pricing policy). These iterative procedures will have to simultaneously take account of the demand functions for all pairs of nodes and of the cost functions for each of the links, all as they relate to each other.[1]

To repeat, in making equilibrations of the sort outlined above, proper specification of the input and output units is of key importance because of their relationship to demand, cost, performance and price functions. Thus far, all these functions were presumed to have output units of vehicle trips per hour; furthermore, the output was assumed to include only passenger trips, each vehicle trip was assumed to be identical in terms of the number of passengers per vehicle, and all travelers were assumed to perceive identical private travel costs (for travel time, effort, accident hazards, tolls, vehicle ownership and operations, etc.).

To correctly handle these aspects (as well as intra-temporal or time-of-day and inter-modal demand cross-elasticities and interactions) will markedly increase the difficulties and complexities of the equilibration procedure and probably will alter significantly the final

---

[1] For elaboration on these techniques, see articles by T. Deen et al. and by M. Wohl in Highway Research Board Record No. 38, Washington, D.C., 1963; see also M. Beckmann et al., Studies in the Economics of Transportation, Yale University Press, 1956, Chapters 3, 4 and 5.

outcome (i.e., the design choice, the operating and pricing policies and staging plan). Needless to say, the data requirements will also be greatly increased by replacement of the above assumptions (regarding homogeneous output) with more suitable representations. In the following paragraphs, some guidelines for integrating these more practical considerations will be outlined; clearly, though, their specific form and detail will depend on the outcome of more intensive data analysis.

First, it is evident that the cost, travel time and price functions should not be restricted to output stated simply in terms of passenger tripmaking; that is, congestion and travel time on specific links of a highway transport system depend both on cargo and passenger tripmaking and thus on tripmaking in trucks, buses and passenger vehicles. As a consequence it will be necessary to jointly consider the demand for cargo movement and that for passenger movement and to carry out equilibration while considering the interaction of cargo and passenger vehicles which use and congest the same facilities and links; in simple terms, this will require that demand be stratified by type of movement (or, say, by vehicle type -- heavy truck, light truck, bus, taxi and passenger car). Furthermore, since these different types of movement will be made over the same facilities or links, the cost, travel time and price functions for each movement type will be dependent on the joint output, whose units must therefore be expressed in terms of the composition and amount of all types of vehicle units; in essence, this means that it will be necessary to determine cost, performance and price relationships for <u>each</u> type of

movement as a function of the total movement volume and the proportion to account for these interdependencies will necessarily be most complex, particularly if the inter-modal and time-of-day demand cross-elasticities are also to be handled. (Clearly, the cross-elasticities should be characterized and integrated since passenger shifts among bus, taxi and auto modes will occur in practice, as will both passenger and cargo shifts from hour to hour during the day.)

Second, it is important to characterize the differences among individual travelers with respect to the way in which they perceive and evaluate their travel effort, auto ownership and operating costs, car pooling and other significant preferences and tradeoffs. Without question, there are wide variations in the manner and extent to which travelers consider and are affected by travel time and congestion, by vehicle ownership and accident payments, by the inconvenience of car pooling, and so forth.[1] While it is clear that travelers whose trip value (i.e., the value of making a specific trip to some particular destination) is high will be willing to endure more effort and higher private travel prices in order to make the trip than will those with lower trip values, it cannot be assumed that those with higher trip values will necessarily regard the private inconvenience or discomfort of car pooling or of congestion as being more "costly" than will those travelers having low trip values. Thus, since the output consists of travelers having different trip values and having different travel

---

[1] See, for example, G. P. St. Clair and N. Lieder, "Evaluation of Unit Cost of Time and Strain-and-Discomfort Cost of Non-Uniform Driving," Highway Research Board Special Report 56, Washington, D. C., 1960.

service preferences with respect to auto quality, parking convenience, car pooling and traffic congestion, and since the trip values and travel service preferences are only partly dependent on income level, equilibration cannot be accomplished accurately simply by stratifying demand, performance and price functions by income level. However, one might attempt to equilibrate demand, travel time and price functions stratified by income level as a first approximation.

Satisfactory treatment of these sorts of variations in travel time and price which are (or may be) dependent both on the level of output and on the particular groups of people and goods involved in tripmaking will require highly complex simulation models to accomplish equilibration; and it would appear that iterative trial-and-error procedures offer the only hope for simultaneously satisfying the intricate demand and service (or performance) conditions. However, even these sorts of procedures may lead to multiple solutions and ambiguities. Considerable research and investigation must be undertaken before any firm conclusions can be reached on their dimensions, reliability or accuracy; the major difficulty appears to rest with the problems of gathering and processing the data requisite to the formulation of multi-dimensioned and compatible demand, cost, travel time and price functions and with the enormous intellectual and electronic data processing resources necessary to simulate the realistic conditions for regions and networks.

(2) An Example Equilibration or Forecasting Program. The kinds of interactions discussed in the previous sections and, more specifically, the process of equilibrating capacity-restraint, price and

demand functions and of determining the actual equilibrium volume, travel time and price conditions can be described more explicitly by examining the problem of determining an appropriate toll rate for a toll facility. For this example, we will consider a case in which price is partially dependent on facility cost and on equilibrium flow and in which demand varies from one time period (i.e., time-of-day) to another; however, this analysis will be restricted to a single-mode and single-link transport system and thus can only be regarded as an illustrative but overly simple example.

Specifically, assume that a certain toll authority seeks to determine the toll rate which will provide sufficient revenues to cover the capital and interest costs, i.e., the bond payments, for the facility; also, assume that the same toll will be applied during all hours of the day and to all vehicle types. For these conditions, the toll authority's criterion for setting the toll rate may be stated in either of two equivalent ways:[1] (1) find the lowest rate which provides sufficient revenues to cover bond payments; or (2) find the rate which provides sufficient revenues and which maximizes usage of the facility.

Further, it will be assumed that the remainder of the total user price (or total user price minus toll) is equal to the private, short-run average variable travel costs, which we will assume to be constant (per vehicle trip) and equal to $\underline{P}_o$. Thus,

$$p = P_o + U \tag{7}$$

[1] Such an objective function is not argued to be the best one for either the bondholders or the general public, but is adopted for use in this illustrative example (though it is generally consistent with the policy of most toll authorities).

and

$$U = f(F_x;V_1,T_1;\ldots;V_k,T_k;\ldots;V_n,T_n) \quad , \tag{8}$$

where p is the total user price, U is the toll, F_x is the annual facility cost (i.e., annual capital plus interest payments) for facility x, V_k is the equilibrium flow during the kth demand period, T_k is the number of hours during the kth demand period, and n is the number of demand periods per day; P_o was defined previously. For these conditions

$$U = F_x \div (365 \sum_{k=1}^{n} V_k) \ . \tag{9}$$

Making use of the simplified but convenient travel time (or capacity restraint) function shown in Fig. 1, let us assume that $\bar{t}_k$, the average travel time experienced during the kth demand period, will be as follows:

$$\bar{t}_k = t_o + \tfrac{1}{2}(V_k/\mu - T_k) \text{ for } \lambda/\mu > 1 \ , \tag{10}$$

where μ is the mean service rate for the toll facility (in vehicles per second) and t_o is the average travel time required for mean arrival volume rates below the mean service rate (i.e., for λ or $V_k/T_k < \mu$). Clearly, for this simplified case,

$$\bar{t}_k = t_o \text{ for } \lambda/\mu \leq 1 \ . \tag{11}$$

Finally, full knowledge of the time-of-day demand relations and cross-relations will be assumed; thus the volume of travel during the kth time-of-day demand period will depend not only on the price and travel

time occurring during the kth period but also on those during the other time periods as well. Also, for this illustration assume that the demand function has the following linear form:[1]

$$V_k = a_k + b_{k,1}\bar{t}_1 + \ldots + b_{k,n}\bar{t}_n + c_k p \ , \qquad (12)$$

where $b_{k,k}\bar{t}_k$ expresses the direct demand relation with respect to travel time $\bar{t}_k$ during time period k, where $b_{k,j}\bar{t}_j$ (for all $j \neq k$) expresses the demand cross-relations with respect to average travel times during other time periods, and where $c_k p$ expresses the direct demand relation with respect to travel price p; as before, V_k is the equilibrium flow during the kth demand period.

The set of mathematical relationships necessary to describe the interaction of tolls, facility costs, travel times and volume levels and to insure that revenue will cover the fixed facility costs can be stated as follows:

$$\text{Maximize} \qquad Z = \sum_{k=1}^{n} V_k \qquad (13)$$

such that U, $\bar{t}_k$, and V_k (for $k = 1,\ldots,n$) are non-negative (14)

and that

$$p = P_o + U \ , \qquad (15)$$

$$\bar{t}_k = t_o \quad \text{or} \quad t_o + \tfrac{1}{2}(V_k/\mu - T_k), \text{ whichever is larger} \qquad (16)$$

$$U = F_x \div (365 \sum_{k=1}^{n} V_k) \ , \quad \text{and} \qquad (17)$$

---

[1] This expression, rather than the more complete one shown in equation (1), was chosen to simplify this illustrative example and to adapt more conveniently to a linear programming format; its validity is not argued as a general proposition, however.

$$V_k = a_k + b_{k,1}\bar{t}_1 + \ldots + b_{k,n}\bar{t}_n + c_k p \ . \qquad (18)$$

In turn, this formulation can be placed within a linear programming format simply by using r piecewise linear approximations for the toll function. The resultant linear programming formulation would be as follows:

Determine non-negative values of U, $\bar{t}_k$ and V_k for $k = 1,\ldots,n$ which will

$$\text{Maximize} \qquad Z = \sum_{k=1}^{n} V_k \qquad (19)$$

subject to the following constraints:

$$V_k = a_k + b_{k,1}\bar{t}_1 + \ldots + b_{k,n}\bar{t}_n + c_k P_o + c_k U \quad \text{for } k = 1,\ldots,n \qquad (20)$$

$$\bar{t}_k \geq t_o \quad \text{for } k = 1,\ldots,n \qquad (21)$$

$$\bar{t}_k \geq t_o + \tfrac{1}{2}(V_k/\mu - T_k) \quad \text{for } k = 1,\ldots,n \qquad (22)$$

$$\text{and} \quad U \geq h_e + w_e \sum_{k=1}^{n} V_k \quad \text{for } e = 1,\ldots,r \ . \qquad (23)$$

The last of these expressions, equation (23), is simply a linearization (in r pieces) of the non-linear toll function shown before in equation (17).

Finally, it may be useful to note that use of this particular linear program will not necessarily result in "minimizing the travel time" for the total volume of travelers among them.[1] To realize this,

[1] Do not infer, however, that this is an improper or unrealistic outcome.

consider the circumstances for travelers who would face lengthy queuing delays if they chose to travel during time period $\underline{k}$; also, assume that the volumes during the $\underline{k}$ + 1th time period are negligible (prior to any "excess" volume from period $\underline{k}$ shifting over to period $\underline{k}$ + 1). In this case, while the "excess" volume might find their total enroute travel time reduced (i.e., $\bar{t}_{\underline{k}+1} < \bar{t}_{\underline{k}}$) if they shifted their time period of travel (since less queuing would occur), they would suffer the inconvenience of having to travel during a less desirable hour (from their point of view) and thus some may regard a shift in their time of travel as unworthwhile; in short, some may choose not to "minimize their enroute travel conditions," and some may decide otherwise. In a similar vein, ours is not simply a problem of minimizing travel time during a preferred time of travel since some will shift over to other hours when enroute travel conditions do get bad enough. These propensities are embodied, in both absolute and relative terms, within the coefficients or exponents of the demand equations, depending on the type of demand function employed for the forecasting process. Moreover, in balancing or "trading off" the discomforts of more (or less) queuing delay against the inconvenience of a less (or more) preferred travel hour, travelers are faced with more than a dichotomous travel choice. That is, they may choose not to travel at all rather than either travel slower but more conveniently or travel faster but less conveniently.

Among the more important conclusions which stem from the remarks in this section on The Technical Components and Requirements for

Predicting System Consequences are the following:

1. Performance conditions on travel facilities, particularly those for which the demand leads to heavy peaking and travel congestion, cannot be suitably represented by steady-state travel time and delay functions; that is, travel conditions during peak- and off-peak time periods cannot (always) be properly represented and forecast unless the circumstances for non-steady-state or transient queuing behavior are included as part of the analysis. An example dealing with transient queuing behavior was included and it was shown that queuing delays were dependent not only on arrival volume and capacity levels but also on the length of time over which the arrivals were sustained.

2. The quantity of tripmaking to be demanded at any particular time is both time-of-day and travel service dependent; that is, both the volume of tripmaking and the time of day at which trips are to be made depend not simply on the travel conditions at certain preferred times of the day but upon those to be afforded at other less preferred times of the day. Furthermore, it was argued that tripmaking -- the choice to travel or not, to take this or that mode or route, to travel at this or that time of the day, etc. -- was an integrated traveler decision-making process with the trip, mode, route and time choices all being made simultaneously. These interdependencies and requirements for simultaneous tripmaking decisions were illustrated both structurally and mechanistically by making use of a simplified (yet hopefully informative) linear programming formulation.

While these conclusions and findings do appear to suggest some important new directions and aspects for incorporation into the travel forecasting process, they nevertheless can only be regarded as tentative and imcomplete. Particularly, it should be emphasized that high quality research will first be necessary: (1) to establish reliable and accurate capacity-restraint functions; (2) to develop compatible travel demand functions (to include consideration of the effects of technology, service and time changes); and (3) to perfect or at least render operational the programming techniques and procedures which will permit equilibration to be accomplished for multi-link, multi-modal and multi-time period transport systems.

IV. THE DISTINCTION BETWEEN DECISION-MAKING AND SYSTEM ANALYSIS

While the system engineer attempts to create and design the "best" technological system (along with its operating characteristics) and while he therefore must intuitively if not methodologically employ a decision-making scheme and relevant criteria, he usually is not the decision-maker and is not uniquely or even particularly well equipped for the decision-making task. That is, the goal of the system analyst and of the system engineer is to create the "best" design (to be sure), but his particular task is to create alternative designs as rationally as he knows how and, more importantly, to predict the consequences which will stem from the adoption of one design or another. Thus, his primary task and responsibility is to determine "what will happen if" and to pose what seem to be the "most reasonable set of alternatives" rather than to say "which design is best;" yet, in so doing he cannot avoid going beyond this task and cannot avoid evaluating as well as analyzing the consequences of alternate designs. (Otherwise, on what basis would "unlikely" or bad designs be dropped from further consideration.)

By way of emphasizing this distinction, consider some of the problems associated with the specification of appropriate decision-making criteria. Primarily, the problem resolves to that of deciding upon the appropriate philosophy of the collective society with regard to: (1) the risks and uncertainties of future consequences or benefits and costs, and the consequent investment strategy to be followed; and (2) the evaluation and distribution (or incidence) of system costs and benefits, the latter which involves resolution of

the questions, "who is to incur the costs" and "who is to accrue the benefits."

Turning first to investment strategy, the problem is not easily resolved -- as the individual and highly personal philosophies and moralities of the entire body politic must be "balanced" and considered collectively. For example, when the uncertain future is at stake, no one can say for certain which plan is "best;" we can only state which plan will (given our predictions of the probabilities of various states of nature and man, etc.) result in the highest expected net gains, or perhaps we can state what the probabilities are of returning higher or lower net gains, and so forth. However, with some people preferring to gamble on the future according to "expected value," with others preferring to adopt that plan which will probably return the highest minimum net present value (the MINIMAX solution), and with still others preferring to adopt that plan offering the highest maximum net present value (the MAXIMAX solution), and so forth, it is indeed difficult to judge which is the "best" strategy to follow. (As we will see from the discussion in the following paragraphs, one may well ask whether it would be appropriate to have the "losers" -- that is, those who wanted to follow a different strategy than was finally chosen -- compensated for their "mental anguish" by the "winners." Or, at what point does one abandon the entire set of alternatives because of the inability to resolve these conflicts?

The second decision-making issue to be examined, that involving the evaluation and distribution of system benefits (gains) or costs (losses), is even more difficult to resolve on any clear-cut, rational

basis. More directly, this issue required the introduction of moral arguments and ethical considerations, neither of which can be reduced to a value scale commensurate with other consequences. Nevertheless, these sorts of issues are important; they do differ from one design to another; they do affect decisions about which design is better or best and about whether a project should be undertaken at all.

Distributional problems usually arise or stem from involuntary situations in which an individual's payments (in whatever form) for a service are higher than the satisfactions or benefits which he receives from the service; that is, matters of fairness and of equity are at stake when one group receives benefits at the expense of another group. Typically, these sorts of incidence problems stem from public projects or programs which presumably are undertaken "in the public interest" and which generally fall into the public sector because of being describable as a "public good" (such as defense), of having scale economies, or of being faced with external costs or benefits; also, and of some importance, we should note that decisions to effectuate an income transfer from one group to another or to have one group subsidize another are often made by administrators who do not bear the brunt of the losses. One should ask, therefore, whether it really is all very well for highway departments to construct highways that inflict both money and non-money costs on abutting tenants and property owners in order that society at large can reap even larger net benefits -- particularly when the highway administrators do not have to bear the losses. And he should ask whether it is

all right for public utility commissions and other regulatory bodies to deem that it is in the "public interest" for a private firm to continue to provide public services at a loss -- particularly when the regulatory bodies and those who sit in judgment are not burdened with the loss. Similarly, we should ask whether it is proper for tunnel and bridge authorities to devote part of the funds collected from its users in order to provide services for non-users, again under the umbrella of the "greater good of society," particularly when the tunnel and bridge directors do not have to bear the burden of paying tolls over a longer period than would otherwise be necessary.

Clearly, the "greater good of society" is of interest and of importance, but it does not lighten the burdens, be they financial or social, of those who must suffer losses in order that "society" -- in the aggregate -- may be better off. Also, one cannot observe the transportation industry for very long without realizing that these sorts of issues are commonplace, particularly for publicly administered transport functions. Indeed, closer examination of the circumstances surrounding urban highway construction, the continuance of subsidized commuter railroad operations, and Federal involvement and support of supersonic transport development -- as but three examples -- should vividly point out the types of conflicts that often arise and should serve to emphasize the important role that matters of incidence play, first in the process of analyzing and evaluating consequences and their incidence and, second, in that of decision-making when the policy-maker must decide whether or not it is desirable to benefit one group at the expense of another. Again, the point is simply that matters

of equity, of one man's gain at another's expense, of who should pay, and so forth, are crucial to decisions and thus are of importance to the system analyst -- who must provide the decision-maker or policy-maker with information on both the differential and aggregate effects of alternate actions, and with details on who benefits and who suffers under different policies.

V. CONCLUDING REMARKS

While system analysis has been widely talked about and "seemingly applied" to transport problems, it seems fair to allege that most transport system analyses fail to account for certain important interactions and fail to view with sufficient breadth those system effects which should undergo analysis and measurement. As a consequence, the purpose of this paper has been to examine more closely and definitively the components and interworkings of transport systems and to indicate some directions for improving our analysis capabilities.

Certainly one of the more important emphases of this brief review paper is that the demand for travel should not be looked upon in the usual fixed "needs" or point estimate sense but should be regarded as a functional relationship between the quantity of travel to be demanded and the quality and price of the travel service. Further, the manner in which demand should be stratified (according to time-of-day and modal preferences, and so forth) and in which it should be related to the variable performance relationships for transport links was described in some considerable detail. In so doing, some fairly explicit directions were offered for improving our predictive or forecasting processes and thus for measuring the consequences stemming from alternative system designs.

Toward these ends, the weaknesses of most trip-splitting types of so-called demand models were explored and contrasted with other more adaptable demand functions. Specifically, it was emphasized that tripmaking (as viewed by the traveler) cannot always (and probably not often) be looked upon as a sequential type of process in which

tripmaking (i.e., trip generation) is treated as "fixed" and in which the forecasting problem is regarded merely as a task of splitting trips among the available zones of destination, modes and routes; to improve on the structural shortcomings of such procedures another type of demand model, one in which the decision-making calculus (about whether to travel, how to travel, where and when to travel) could be treated as a simultaneous process and could reflect the inherent service, time-of-day and other dependencies, was offered as a more suitable expression for transport system studies. In addition, some deficiencies with respect to our description of the performance capabilities of transport networks and of their links were outlined; in particular, the necessity for introducing non-steady-state and transient queuing conditions was pointed out and some preliminary relationships for handling this problem were spelled out. The interactions of these demand and performance aspects were explored in both qualitative and quantitative terms.

Buttressing the discussion of forecasting techniques -- of demand, price and performance functions, their character, development and interrelationships -- was a section which noted the importance of describing the differential as well as aggregate effects of alternative designs or policy, all as part of the system analysis. More specifically, an attempt was made to emphasize the necessity of recognizing and accounting for the distribution of the gains and losses associated with one design or the other. Emphasis was placed not only upon the way in which the matter of "who benefits" and "who pays" affect the individuals who gain or lose but also upon the character

of our administrative and decision-making process as it witnesses and takes account of these distributional effects. In brief, apparent deficiencies with respect to the resolution of these sorts of conflicts is felt to depend in an important way upon the kind and amount of information which is afforded the policy-maker by the system analyst.

Finally, it would be inappropriate not to mention the important way in which extensions of and improvements to our forecasting and system analysis procedures depend first upon our general understanding of the transport system and its interworkings -- much as described herein -- and second upon the quality and character of the requisite data to develop and test hypothesized demand, price and performance relationships; in addition, our abilities depend importantly upon the various types of programming techniques and data processing facilities which play a necessary role in the matter of equilibrating demand, price and performance functions on a multi-mode, multi-route, and multi-time period basis (even without considering the long-run interaction between the transport system, its usage and the location of businesses, residences and industries). Though the remarks herein tend to minimize -- or at least fail to emphasize -- the importance of the latter two aspects, no such implication was tended. Rather, it was felt that attention was more deserving with respect to the problem of formulating and structuring the overall transport system problem and that others having more formal training in the related fields of operations research and applied mathematics could then play a more useful role as members of an interdisciplinary transport system analysis team.