CHAPTER 3

# A Theory of Individual Travel Demand

## 3.1. Introduction

This chapter explores the underlying theory of consumer choice behavior as it relates to urban transportation demand decisions. The purpose of exploring formal theoretical models of consumer choice is to make explicit the considerations that will guide the selection of variables to be included in the empirical analysis of transport demand, and to investigate possible restrictions on the demand function that aid in the empirical estimation of its parameters. The objective is to base restrictions on as few arbitrary assumptions as possible and to make these assumptions and their relationships to the underlying theory of consumer choice explicit.

The exploration of the underlying theory of consumer choice in the context of travel behavior in this chapter has two parts. The first consists of the development of a derived demand model for urban transport in which the choice of transportation alternatives is determined by the consumption activities available to the consumer. This analysis is primarily useful for providing a general framework for demand modeling rather than for providing specific demand theorems and functional forms.

The second major effort in theory development is concerned with the possibility of deriving useful restrictions by factoring the demand function into component parts. Rather than considering the demand for transportation as a set of simultaneous decisions about mode, destination, time of travel, and trip frequency, the implications of analyzing these choices separately are considered. It is found that the assumptions needed to factor the demand model into these components are quite plausible as a working approximation, and that this procedure helps considerably in reducing the estimation problem to manageable proportions.

While the basic approach of factoring the demand function into separate parts may initially sound very much like the procedure that conventional transport demand models follow, it will be seen that it is actually quite different. The fundamental difference is that in this approach the individual parts of the model are interrelated through the attributes of the trip—the time and cost variables—so that the separate parts link together into an overall *demand* model. Moreover, separately and together, the parts of the demand model are responsive to changes in the times and costs of travel. Both theoretically and practically this provides a more satisfactory treatment of travel demand for policy analysis.

## 3.2. Rational choice behavior

The theory of rational choice behavior asserts that a decision maker can rank possible alternatives in order of preference, and will always choose from available alternatives the option which he considers most desirable, given his tastes and the relevant constraints placed on his decision-making, such as his level of income or time availability. Suitably modified to take account of the psychological phenomena of learning and perception errors, this theory has been used successfully in analyzing and forecasting economic consumer behavior in a wide variety of applications, and it forms the foundation of modern economic analysis.

In the model of consumer behavior that follows, this theory is elaborated to focus on the relationship between consumer behavior and transportation. The consumer is assumed to have a utility function defined on both consumption and transportation attributes. The set of alternatives available to the individual is determined not only by the usual budget constraint, but also by the "household" technology for carrying out work and consumption activities in various locations, and the attributes of transport modes to these locations. Because transport often appears as a "fixed charge" concomitant of consumption activities and involves discrete choices, the set of available alternatives will not be a simple "budget set" of the type ordinarily encountered in consumer theory. Consequently, we will not obtain the usual consumer theory implications drawn from marginal analysis. (In the following chapter on behavior of populations of consumers, we reintroduce marginal concepts by considering the *extensive* margin in the population.)

Our initial model will consider consumer choice in the abstract, without specific reference to transport. The reader may, however, find it useful to keep in mind the range of transport-related decisions made by the consumer:

(1) The locations of residence and job;
(2) Sales of labor and purchases of commodities, including vehicles;
(3) Frequency of work, shopping, recreation, and other trips;
(4) Destination of trips;
(5) Time of day of travel;
(6) Mode of travel.

To encompass these decisions, which involve short- and long-run choices and the dynamics of consumption activities, it is in general necessary to consider a fully intertemporal theory of behavior.

We formulate our description of the economic consumer within the framework of the Court–Griliches–Becker–Lancaster consumption-activity–household-production model. In this model it is assumed that the individual has a series of basic wants, or drives, as for example "hunger", "thirst", and "rest", and the consumer is assumed to have a "utility" function defined for levels of satisfaction of these wants which "summarizes" his sense of well-being.

Over his lifetime, the individual has available a set $A$ of mutually exclusive alternative choices, with each choice representing a lifetime program of activities, or acts. Each choice of a lifetime consumption activity determines the levels of satisfaction of wants the individual will experience.[1] On the other hand, each lifetime consumption activity determines a vector of attributes describing market commodities purchased, trips taken, work performed, etc. The individual chooses an activity from $A$ which maximizes the derived utility; the corresponding vector of attributes defines his observed demands. In particular, this vector of attributes will specify transport demand behavior along the dimensions listed above.

[1] In the most general case, each choice of a consumption activity will determine a *lottery* over levels of satisfaction whose outcome is determined by chance. It will be unnecessary to consider behavior under uncertainty explicitly in our analysis.

The model above can be stated formally, although we will not make direct use of this formalism in the future. Assume the individual to have a lifetime extending over a finite sequence of short periods, indexed $v = 1, 2, \ldots, H$. Let $w_v$ denote a finite vector of levels of satisfaction of wants in period $v$, and let $w = (w_1, \ldots, w_H)$ denote the lifetime vector of want satisfaction levels. Individual utility is a function $u = W(w, s)$, where $s$ is a vector of individual social and demographic characteristics influencing tastes. A consumption activity is assumed to be a finite vector $a$ contained in a universe $\Omega$. Associated with each $a \in \Omega$ is a vector of want satisfaction levels $w = Ma$, and a finite vector of attributes $x = Na$, where $M$ and $N$ are taken to be linear transformations. Given a set $A \subseteq \Omega$ of available actions, the consumer solves the problem

$$\operatorname*{Max}_{a \in A} W(Ma, s), \tag{3.1}$$

with observed demands satisfying $x^* = Na^*$ for the maximand $a^*$. Of particular interest is the case in which each attribute vector $x$ is identified with a unique activity vector $a$; i.e., $N$ is a square non-singular matrix. Then, the set of available alternatives can be expressed directly in terms of observed attributes,

$$B = \{Na \mid a \in A\},$$

and utility can be written

$$u = U(x, s) \equiv W(MN^{-1}x, s),$$

so that (3.1) becomes

$$\operatorname*{Max}_{x \in B} U(x, s). \tag{3.2}$$

The framework set out above is still far too general to be useful in formulating econometric models of transport demand. In order to proceed further, we exploit the structure that is likely to prevail for transport preferences.

## 3.3. Substitutes, complements, and independence

The consumer chooses among lifetime actions which generally have different patterns of costs and ultimate results. Considering consumer choice in full generality involves considering the attributes of every conceivable lifetime alternative to a specific choice as part of the determination of that choice. The number of distinct alternatives may be enormous, presenting a problem not only to the investigator but also to the individual faced with the decision. Studies of decision behavior suggest that in circumstances such as these the individual is likely to follow a "tree" decision structure, behaving "as if" his utility can be factored in such a way that it is necessary for him to consider relatively few alternatives at each decision point. A good behavioral model should not only parallel the individual's decision tree, but should also exploit the separability of decisions implicit in this tree to make empirical analysis practical.

In conventional econometric demand analysis, commodity demands are assumed to depend only on the prices of close complements and substitutes, the effect of remaining commodity prices being assumed negligible. Analogously, one may postulate that the choice of a particular transport alternative depends only on the attributes of this alternative and those that are its close complements or substitutes.

In the context of urban transportation demand, substitute alternatives are trips by alternative modes, at alternative times of day, to alternative destinations, etc. Generally, substitutes can be thought of as different ways of accomplishing the same objective. For example, if the substitutes are two trips from the same origin to the same destination, at the same time of day, one by car and the other by bus, then the substitutes are expressed as the attributes of the auto trip (auto in-vehicle time, walking time, parking charges, etc.) versus the attributes of the transit trip (fare, walk, wait, and transfer time, line-haul time, etc.). This is an example of close substitutes, where the attributes of both will influence the choice between them. If, on the other hand, these alternatives corresponded to trips on different days, they would not be close substitutes and the demand for one would not be closely influenced by the attributes of the second.

In the context of transport demand, a complement to a choice is a second choice which tends to be tied to or induced by the first. For example, the return trip from a given destination is complementary to the outbound trip in that the choice of outbound mode and time of travel restricts the return alternatives. A second example is the complementarity between choices of residential and work location and auto ownership, and the attributes of available trips.

Implicit in a pragmatic selection of "close" substitutes and complements is a theoretical assumption of separability of preferences. We next explore the implications of such assumptions.

## 3.4. Factoring the demand model

When one considers the full range of decisions involved in the travel demand function and allows for widespread substitutes and complements, it is clear that there are an enormous number of potential variables in the demand function. In this section we develop a procedure for reducing the complexity of the empirical analysis by factoring the demand function into its component decisions. Under plausible working assumptions, this factorization will coincide with an actual separation of decisions. This can occur if the individual's preferences in fact conform to these assumptions, or alternatively, if the individual behaves "as if" his preferences satisfied these conditions because he follows a "myopic" or "local optimization" decision rule to economize on information and computation costs.

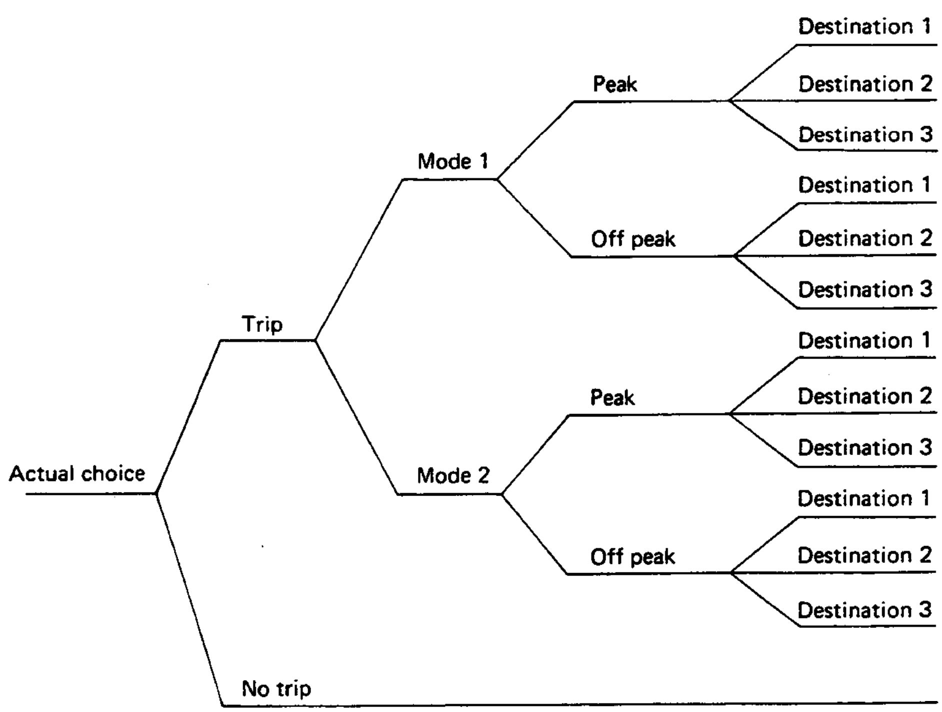


Fig. 3.1.

One need only examine the dimensions of the travel demand analysis to see why it is desirable to factor the demand function. Suppose there are only two modes to be considered, two alternative times of day (e.g., peak and off-peak), and three alternative destinations for each trip-maker. In addition, the traveler can choose between taking or not taking a trip. The options are shown in fig. 3.1.

Each terminal branch of this "tree" represents an activity available to the consumer; hence the overall demand function in this simple case involves thirteen different options. Moreover, since each of the twelve trip options involves at least two or three modal attributes, 25–35 variables would have to be included to analyze this fairly simple case, and as the number of options increases, the number of variables escalates. If there were three modes, three times of day of travel, four destinations per traveler, three trip frequencies, and three modal attributes per option, 324 variables would have to be considered, and the introduction of inter-temporal repetitions of this choice over the lifetime of the consumer makes the number of variables astronomical. It is easy to see from these few examples that one could rapidly lose control of the empirical analysis unless the problem can be considerably simplified.

One way to do this is to factor the overall demand function into its component decisions of choice of mode, time of day to travel, trip destination, and trip frequency, and to use this device to place restrictions across the different aspects of this demand function so as to further reduce the complexity of the estimation problem.

To examine this method, let us consider the hypothetical question: "If you were making a round trip from home to $\gamma$ at time $\beta$ on day $\pi$, what mode would you use?" Our hypothesis is that the answer to this question depends only on the attributes of the different available modes and is independent of the time of day the trip is made, the destination of the trip or the number of trips that are made for this purpose each day. By this we mean that for the given purpose, the choice of mode depends on the comparison of times and costs of travel by the different modes (the attributes of the modes), while the evaluation of a given minute of travel time or dollar of cost is independent of when or where the trip is made or the number of trips that are made for that purpose each day. We hypothesize that it is also independent of other prices in the economy. (However, we would expect the answer to depend on socioeconomic characteristics like income, family size, and auto ownership.)

We assume that attribute vectors of alternatives can be used to index consumption activities directly, and adopt the notation of eq. (3.2). Thus utility $U(x, s)$ is expressed as a function of a vector of socioeconomic characteristics $s$ and a vector $x$ of attributes of each alternative, with a set $B$ of available alternatives. Suppose the hypothetical question above has two possible responses, modes $a$ and $b$. The vector of attributes $x$ can be partitioned into subvectors $x_{(1)}$ and $x_{(2)}$, such that $x_{(2)}$ is the same for the choices $a$ and $b$, while $x^a_{(1)}$ and $x^b_{(1)}$ differ. The factorization hypothesized above requires

$$U(x^a_{(1)}, x_{(2)}, s) \geqq U(x^b_{(1)}, x_{(2)}, s), \tag{3.3}$$

if and only if

$$U(x^a_{(1)}, x'_{(2)}, s) \geqq U(x^b_{(1)}, x'_{(2)}, s). \tag{3.4}$$

A sufficient condition for this property is that $U$ have the additively separable form[2]

$$U(x_{(1)}, x_{(2)}, s) = \psi(\phi^1(x_{(1)}, s) + \phi^2(x_{(2)}, s)). \tag{3.5}$$

Then, $a$ is chosen over $b$ if

$$\phi^1(x^a_{(1)}, s) > \phi^1(x^b_{(1)}, s). \tag{3.6}$$

Since $x_{(1)}$ is a "relatively short" vector, and the set of alternative $x_{(1)}$ subvectors for $x$ in $B$ is "relatively small", this structure greatly simplifies econometric analysis.

We would not expect the additivity hypothesis above to hold strictly. For example, the relative utility associated with $a$ and $b$ may differ depending on the commitment of family vehicles specified in $x_{(2)}$, or on levels of fatigue resulting from other trips specified in $x_{(2)}$. Further, the effects of varying transport time and cost on the consumer's expenditure and time budget constraints will, in general, introduce interaction effects in the demand functions, even when the utility function is separable. Despite these exceptions, additive separability seems a good general working hypothesis.

The assumption that the marginal rates of substitution between modal attributes are constant for a given trip purpose, regardless of the time

[2] When there are three or more subvectors of $x$, each with the independence property above, additive separability is also necessary. See Debreu (1960b, theorem 3, p. 21). The function $\psi$ must be monotone increasing.

of day, trip destination or travel frequency seems quite plausible. One can, for example, envision a traveler evaluating the trade-off between a minute spent waiting for a bus and a minute of bus line-haul time differently in the peak as compared to the off-peak, but the difference would probably stem from the uncertainty associated with waiting and perhaps the greater need to be punctual in the peak. However, if this were the case, the list of modal attributes should also include measures of the need to be punctual (e.g., schedule delay), in which case there would no longer be reason to expect the weights assigned to each attribute to be different for the two time periods.

With regard to the effect of overall income and time constraints, we note that the importance of these effects depends on the proportion of total income and time spent on transport. These proportions are likely to be sufficiently low to allow us to ignore the overall effects of changes in the cost and time of trips. In the special case that utility can be written as a function linear in numériare commodity, this commodity will "absorb" all income effects and the additive separability will carry over to the demand functions, even if transport expenditures are a substantial proportion of income. This special case for the utility function will appear again in our discussion of the measurement of consumer benefits from transport projects.

A further observation on the structure of the utility function can be made by recalling that the function $U(x, s)$ of attributes of alternatives is derived from a more basic utility function defined over levels of satisfaction of fundamental wants. Consequently, it will usually be the generic attributes of an alternative that matter to the individual, and not the specific "labeling" of the alternative. For instance, walking time is walking time regardless of the principal mode used for the trip. This structure allows a further reduction in the complexity of the description of alternatives.

By contrast to the hypothetical question posed above, in which the mode choice is indicated ceterus paribus for the remaining environment, one could ask a question of the form: "If you had to make a round trip to $\gamma$ on day $\pi$, at what time of day $\beta$ would you go, given that you can choose your preferred mode ($a$ or $b$)?" The response to this question is conditioned on an optimal choice along another travel dimension. Thus, one would expect this choice to depend not only on the socioeconomic factors and destination attributes, such as work and child-care schedules and size of crowds, but also on the attributes of each mode at each alternative travel time.

We now assume, analogously to the previous case, that the consumer's "trade-off" between such factors is not influenced by the mode of travel or day of the trip. For example, the degree of additional crowding an individual would accept at a shopping destination in exchange for lower goods prices should be independent of the mode used to reach the destination when the attributes of the trip remain the same. One can envision a traveler finding a crowded store more objectionable because he rode a crowded bus to get to the store, but we assume such instances of interdependence are rare or that their effects are negligible. This assumption holds if utility is now assumed to have the separable form

$$U(x_{(1)}, x_{(2)}, x_{(3)}, s) = \psi\left(\sum_{i=1}^{3} \phi^i(x_{(i)}, s)\right), \tag{3.7}$$

where $x_{(1)}$ describes the attributes of the trip, $x_{(2)}$ the attributes of the destination at each time of day, and $x_{(3)}$ the remainder of the environment.

Suppose, as before, that two modes, $a$ and $b$, are available; and now suppose two choices of time of day, peak ($p$) and non-peak ($n$), are offered. The subvector $x_{(1)}$ will vary with both mode and time of day; e.g., $x_{(1)}^{ap}$ will describe the attributes of a trip by mode $a$ at the peak. The subvector $x_{(2)}$ will vary with time of travel, $x_{(2)}^p$ and $x_{(2)}^n$, but not with mode, while $x_{(3)}$ will be the same for all choices. The response to the hypothetical question above will be a choice of a peak trip if

$$\phi^2(x_{(2)}^p, s) + \operatorname*{Max}_{i=a,b} \phi^1(x_{(1)}^{ip}, s) > \phi^2(x_{(2)}^n, s) + \operatorname*{Max}_{i=a,b} \phi^1(x_{(1)}^{in}, s). \tag{3.8}$$

The important observation to be drawn from this formula is that all the attributes of the trip are summarized in the "index" of desirability $\phi^1(x_{(1)}, s)$, whose form could be determined from the first hypothetical question. Thus, the analysis of time of day of trip can proceed with the use of this previously determined "index" rather than with the much more extensive "raw" data on trip attributes. This saving can obviously be extremely valuable in facilitating econometric analysis.

Provided utility has the appropriate additive structure, we can repeat the arguments above to construct a "pyramid" of relatively simple choices corresponding to the decision tree illustrated earlier. In our empirical analysis, we shall make the explicit structural assumption that the decision sequence is as depicted in fig. 3.2. Individuals first choose work and residential locations, then vehicle ownership, then trip frequency, etc. At each decision level, choice can be viewed as being made conditioned on fixed preceding decisions and optimal succeeding decisions. For example, destination choice is made conditioned on fixed decisions on location, vehicle ownership, and trip frequency, with time-of-day of travel and mode assumed to be chosen for each alternative

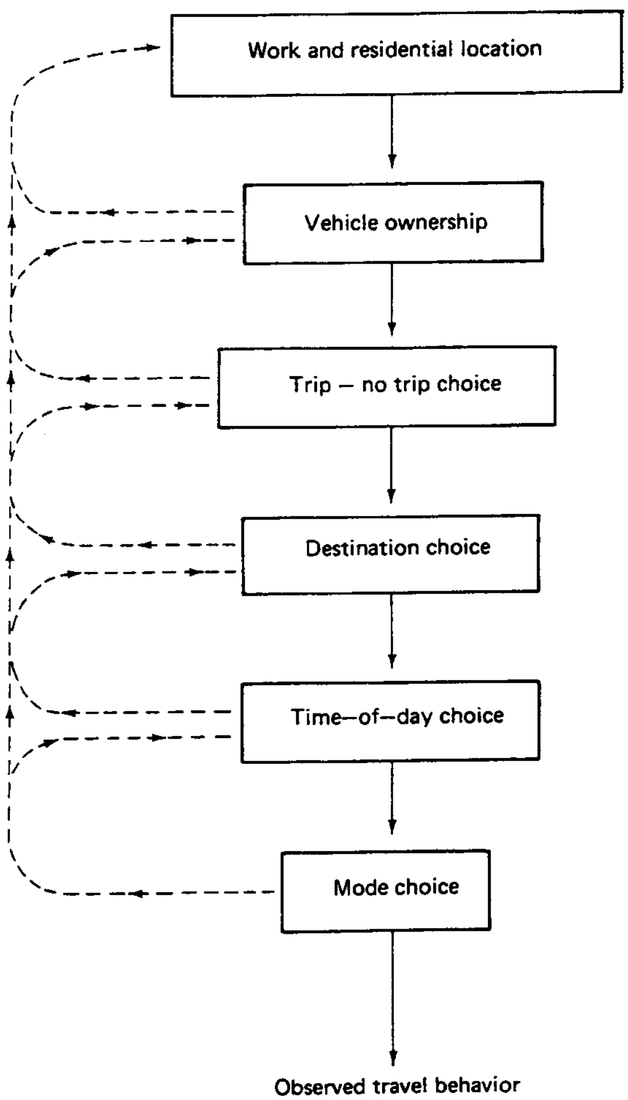


Fig. 3.2.

destination to be the most desirable. The solid arrows in the diagram represent the decision sequence, while the broken arrows represent the flows of information on optimal values of succeeding choices which enter the decision process at any level.

We note that any joint or sequential process involving the dimensions of choice above can be cast in the framework we have just outlined. Structural properties of the decision process are reflected in the information flows up the decision ladder. If utility has no separability properties, then all the information on alternatives which enter succeeding decisions must be passed back up the ladder. If utility is separable, as we have assumed in this chapter, then a summary variable reflecting "desirability" or "inclusive cost" for succeeding decisions is sufficient to guide choice at each level. We emphasize that in any case the form and content of the information passed up to any decision level is determined within the utility maximization process. Thus, any plausible model of this process must at each level incorporate a realistic specification of the information available from succeeding decisions. The most straightforward way to ensure this plausibility is to formulate the decision model directly in utility terms.

It is instructive to compare the decision structure outlined above with conventional modal split and trip generation and distribution models. One may depict such models on the figure above, with the trip—no-trip choice corresponding to trip generation, and the destination choice corresponding to trip distribution. In this sense, conventional models are compatible with the utility maximization structure we have outlined. However, the information on succeeding decisions implicit in the conventional models is often either inconsistent with a theory of individual utility maximization or corresponds to an implausible utility structure. For example, conventional trip distribution models often utilize a summary variable on the trip attribute, "impedance", which is based solely on auto driving time. This implicitly requires either that individuals be irrational, failing to reconcile their actual mode choice with the hypothesized mode choice assumed at the level of the destination decision, or that mode choice depends solely on auto driving time independent of the attributes of transit. Neither implication is particularly plausible. We conclude that the framework of individual utility maximization provides restrictions on the information that should enter decisions at each level. One contribution of behavioral analysis of travel demand, even within the context of conventional UTP models, is to provide more plausible specifications of these information flows.

In algebraic terms, the structure we have assumed above for a given trip purpose and day of travel is an additively separable utility function,

$$U(x, s) = \psi\left(\sum_{i=1}^{7} \phi^i(x_{(i)}, s)\right), \tag{3.9}$$

where the subvectors of $x$ are defined as follows:

- $x_{(1)}$ Attributes of the trip mode choice, such as fares, tolls, in-vehicle time, walk time, and wait time.
- $x_{(2)}$ Attributes of the time-of-day choice, such as congestion at the destination, availability of services, convenience of scheduling other activities.
- $x_{(3)}$ Attributes of the destination, such as variety and cost of services available.
- $x_{(4)}$ Attributes of the trip—no-trip choice, such as levels of inventories of household goods, characteristics of recreational opportunities at home.
- $x_{(5)}$ Attributes of the choice of vehicle ownership, such as the availability, cost, and reliability of household autos.
- $x_{(6)}$ Attributes of the choice of residential and work location, such as neighborhood density and quality of public services.
- $x_{(7)}$ Attributes of all consumer choices other than the sequence above, including decisions on trips for other purposes and on other days.

Because of the limitations of the data utilized in this study, we do not attempt to estimate the parameters of vehicle ownership and location decisions.

## *3.5. Summary*

In exploring theoretical considerations of consumer choice behavior based on utility maximization subject to constraints, we find that the empirical analysis can be considerably simplified by making what seems to be a plausible assumption about the additivity of the consumer utility function. The assumption basically amounts to assuming that relationships between certain variables remain constant over different choice functions. This assumption can be used to combine a number of variables into a single "index of desirability", or inclusive price of travel, for a given mode. This inclusive price can then be minimized over modes, allowing a further reduction in variables.

In an analogous manner, the separate attributes of alternative destinations or times of day can be combined into a single variable by using parameter estimates from these choice models to aggregate over attributes. In estimating the trip frequency function for a given trip purpose then, a single minimum price of travel can be used to measure the disutility of travel. This price will be built up from modal attributes combined into an inclusive price for each mode and minimized over modes and destinations (and if desired, over times of day of travel). Thus, throughout the empirical analysis, estimates of relationships from earlier steps are used to reduce the complexity of the statistical estimation at later stages. The structure of the approach is such, however, that the effects of travel time and cost—the policy variables—are built into each separate choice model. Moreover, the separate models are tied together through the explicit incorporation of relationships from other decision functions. Thus, the approach can be used to measure separately or in combination the effects of changes in policy variables on the choices of how, when, and where to travel, and how frequently to make a trip.