PART II

DEVELOPMENT, TESTING, AND VALIDATION OF A WORK-TRIP MODE-CHOICE MODEL

CHAPTER 1

PRETESTING THE SAMPLE AND INITIAL MODEL SPECIFICATIONS

Introduction

The development of travel demand models requires substantial data collection and compilation efforts. It is wise to pretest not only the data collection and data compilation methods but also the model specification the analyst has in mind. In this way, important information is obtained regarding the potential value of variables and variable specification in addition to the help it provides to the data collection activities themselves. Pretesting is also likely to sharpen the analyst's intuition and give an indication of what important variables may be missing from the model.

The Urban Travel Demand Forecasting Project (UTDFP) was armed with this type of pretest data set for about 200 households. In addition, there existed considerable prior experience in the development of work-trip mode-choice models (McFadden, 1974; Ben-Akiva, 1973; Lisco, 1968; Stopher, 1969; Talvitie, 1971; and many others). Nevertheless, several issues required further examination, among them the following. First, additional knowledge was needed in modeling multinomial choice situations, particularly with regard to modeling access mode. In the previous studies the transit mode was always viewed to be a single alternative. However, even within a single linehaul transit mode, such as bus, there are two ways to reach the mode line: by foot and by car. A multinomial logit (MNL) model that considers bus-with-walk-access and bus-with-car-access as separate modes is estimated, and two problems associated with the variable definition and model specification are discussed. Second, the MNL model is the only way in which two or more access modes can be considered. The MNL model is valid only if the odds of choosing one alternative over another are independent of the existence of a third alternative (the assumption of the independence from irrelevant alternatives, or the IIA assumption). It may not be the case that the odds of choosing auto over bus-with-walk-access are the same whether or not the possibility exists of taking bus-with-auto-access. A choice model that does not incorporate the IIA assumption--the "maximum model"--is also estimated and discussed. Third, all the earlier models are based upon utility functions in which the attributes of alternatives are introduced directly as arguments. However, according to the neoclassical theory of consumer behavior, there exists a tradeoff between goods and leisure time. Models are then derived that incorporate this theory with the attributes of work-trip alternatives entering the constraints on utility maximization. The advantage of this approach is that it allows the functional form of the choice model to be related to the presumed structure of preferences between goods and leisure.

We examine these issues in the following three sections. All the models presented are based upon the sample of 161 individuals interviewed regarding their work-travel behavior; this sample is often referred to as the WTS sample.

The Basic Three-Alternative Model Specification

Table 2 gives estimates for a particular specification of the relative impedance of the three alternatives. The "representative" utility[1] function is assumed to be of the form $\beta' x$, where x is a vector of attributes of alternatives (or attributes of the individual interacting with attributes of the alternatives), and β is a vector of coefficients. The first column in Table 2 lists the elements of x and the second column gives the point estimates of the elements of β. The estimation is the method of maximum likelihood described earlier.

The cost and time variables are self-explanatory. The socioeconomic variables are alternative-specific variables and can be interpreted as proxies for unmeasured attributes of the alternatives in which the variable is entered; for example, "length of residence in community" enters the representative utility of the auto alternative, but not that of the bus alternatives because this variable assumes the value of zero for the bus alternatives. An explanation of its significantly positive coefficient could be that persons who reside for a long time in the same community happen to value the privacy that an automobile offers more than less sedentary persons. Or, that the length of residence is an indication of wealth not captured by the income variable.

[1] Representative utility is that part of the random utility function that is common to all individuals.

TABLE 2 Work-Trip Mode-Choice Basic Model

Mode 1:	Auto	Data:	Work Travel Survey, East Bay
Mode 2:	Bus, Walk Access	Model:	Multinomial Logit, Fitted by the Maximum Likelihood Method
Mode 3:	Bus, Auto Access		

Explanatory Variables	Estimated Coefficients	t-Statistics
Cost divided by post-tax wage, in cents/ (cents per minute) e/	- .0358	2.78
On-vehicle time, in minutes e/	- .0185	1.38
Walk time, in minutes a/ e/	- .0190	.972
Transfer-wait time, in minutes	- .0534	1.54
Number of transfers a/ e/	- .0723	.249
Headway of first bus, with a ceiling of 8 minutes, in minutes a/ e/	- .218	2.32
An index of distance to parking at home c/	- .318	.933
Family income with ceiling of $7000, in $ per year b/	.000434	1.51
Family income minus $7000 with floor of $0 and ceiling of $3000, in $ per year b/	.000785	1.66
Family income minus $10,000 with floor of $0 and ceiling of $5000, in $ per year b/	- .000617	2.47
Length of residence in community, in years b/	.143	2.90

Table 2, continued

Explanatory Variables	Estimated Coefficients	t-Statistics
An index of population density in neighborhood b/	- .741	2.58
Dummy if respondent is over 44 years of age	- .781	1.22
Dummy if there is child in household b/	-1.63	2.54
Number of persons in household who can drive c/	1.10	2.68
Auto alternative dummy d/	-5.49	2.27
Bus-with-auto-access dummy d/	-2.76	3.45

Likelihood ratio index: .5983
Log likelihood at zero: -176.9
Log likelihood at convergence: -71.05

Values of time saved as a percent of wage:

On-vehicle time: 52
Walk time: 53
Transfer-wait time: 149

One transfer with no waiting or walking is valued the same as 3.9 minutes of on-vehicle time.

Table 2, continued

All cost and time variables are calculated round-trip. Dependent variable is alternative choice (one for chosen alternative, zero otherwise). Sample size: 161 .

a/ The variable is zero for the auto alternative, and takes the described value for the other alternatives.

b/ The variable is zero for the bus alternatives, and takes the described value for the auto alternative.

c/ The variable is zero for the bus-with-walk-access alternative, and takes the described value for the remaining alternatives.

d/ The variable is one for the bus-with-auto-access alternative and zero otherwise.

e/ Sum of home-to-work and work-to-home.

The alternative-specific dummy variables also reflect the impacts of the alternative's unmeasured level-of-service attributes not captured in the included variables; their omission could bias the coefficients estimated for the observed and included variables. Consequently, alternative-specific dummy variables appear in most contemporary disaggregate models.

Three income variables were included to allow for a non-linear relation between income and representative utility of the auto alternative. These variables can be understood most readily by reference to Figure 4. The positive coefficient for the first income variable indicates that the representative utility of the auto alternative increases with income up to an income level of $7000 per year. The second income variable also has a positive coefficient, indicating that representative utility of auto increases with income for incremental income above $7000, up to a total increment of $3000 (total income of $10,000).

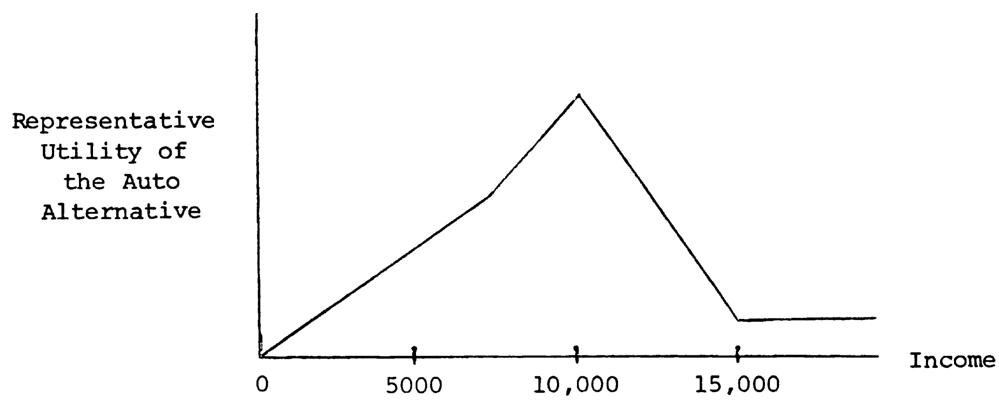

Figure 4: The Relation of Income to the Representative Utility of Auto

The slope of this segment is greater than that of the first segment because the coefficient of the second income variable exceeds that of the first. The negative coefficient of the third income variable indicates that representative utility of auto decreases with income for incremental income above $10,000, up to a total increment of $5000 (total income of $15,000). No fourth income variable being included indicates that representative utility does not vary with income for incremental income over $15,000. (A fourth income variable was included originally and found to have a very small coefficient and a t-statistic below 0.1 .)

The values of on-vehicle and walk times conform to previous estimates (Quarmby, 1967; Thomas and Thompson, 1971). Because walk time was calculated at the rate of two minutes per block, the walk and on-vehicle time coefficients being similar indicates that the representative individual is practically indifferent between walking an extra block and riding for two extra minutes. The value of transfer-wait time, which is much larger than the values of walk and on-vehicle times, conforms to the authors' expectations.

The estimated coefficient for "headway of first bus" is difficult to explain. Headway is the minutes between buses for the particular bus route at the time the respondent would travel. If the average amount of time a person waits for a bus is half of its headway with a maximum wait of four minutes, the estimated value of time spent waiting for the first bus is 1,190 percent of wage.[1] Apparently the headway variable is capturing something more than time spent waiting. It is possible that individuals dislike long headways not so much for the waiting time involved as the rigidity that infrequent service places upon their schedules. This does not seem to be the case, however, unless individuals think that headways of seven or eight minutes impose significant scheduling rigidity.

Three additional headway variables were included in a model in a manner analogous to the income variable explained above. This specification allowed for a non-linear relation between headway and representative utility of the bus alternatives. The additional headway variables affected the coefficients of many variables. Naturally, most affected were the the first headway variable and the transfer-wait time variables whose coefficients and t-values were cut by fifty percent or more. An example of the model estimated with two headway variables is in Table 3. (Note that the model is estimated with only 142 data points.)

---

[1]This large value of first wait time is partially confirmed by Algers, *et al.* (1974) on Stockholm data. They found a ratio of twelve between the values of wait and on-vehicle time. However, Algers, *et al.* were using a wait time formula from British studies (O'Flaherty and Mangan, 1970) that gives approximately half our value of wait time (measured as one-half headway) at the mean of the headway values for our sample.

TABLE 3 <u>Work-Trip Mode-Choice Maximum Model</u>

Mode 1: Auto
Mode 2: Bus, Walk Access
Mode 3: Bus, Auto Access

Data: Work Travel Survey, East Bay
Model: Multinomial Logit, Fitted by the Maximum Likelihood Method

Explanatory Variables	Estimated Coefficients	t-Statistics
Cost divided by post-tax wage, in cents/(cents per minute) e/	- .0448	3.1
On-vehicle time, in minutes e/	- .0303	1.8
Walk time, in minutes a/ e/	- .0185	0.90
Transfer-wait time, in minutes a/ e/	- .00202	0.04
Number of transfers a/ e/	- .312	0.90
Headway of first bus, with a ceiling of 8 minutes, in minutes a/ e/	- .138	1.1
Headway of first bus exceeding 8 minutes a/ e/	- .0639	1.7
An index of distance to parking at home c/	- .382	1.0
Family income with ceiling of $7000, in $ per year b/	.000343	1.1
Family income minus $7000 with floor of $0 and ceiling of $3000 in $ per year b/	.00107	1.9
Family income minus $10,000 with floor of $0 and ceiling of $5000, in $ per year b/	- .000748	(2.6)

Table 3, continued

Explanatory Variables	Estimated Coefficients	t-Statistics
Length of residence in community, in years b/	.176	(2.8)
An index of population density in neighborhood b/	- .713	2.1
Dummy if respondent is over 44 years of age b/	- .714	1.0
Dummy if there is child in household b/	-1.679	2.4
Number of persons in household who can drive c/	1.251	2.6
Auto alternative dummy b/	-5.190	1.8
Bus-with-auto-access dummy d/	-2.730	3.0

Likelihood ratio index:	.613
Log likelihood at zero:	-156.0
Log likelihood at convergence:	-60.4

Values of time saved as a percent of wage:

On-vehicle time:	68
Transfer-wait time:	5

One transfer with no waiting or walking is valued the same as 10.3 minutes of on-vehicle time.

Table 3, continued

All cost and time variables are calculated round-trip. Dependent variable is alternative choice (one for chosen alternative, zero otherwise). Sample size: 161 .

a/ The variable is for the auto alternative, and takes the described value for the other alternatives.

b/ The variable is zero for the bus alternatives, and takes the described value for the auto alternative.

c/ The variable is zero for the bus-with-walk-access alternative, and takes the described value for the remaining alternatives.

d/ The variable is one for the bus-with-auto-access alternative, and takes the described value for the remaining alternatives.

e/ Sum of home-to-work and work-to-home.

Thus, it appears that the headway coefficients are capturing many effects. The most plausible explanation for the headway coefficients is that they capture not only the waiting time effects but also some other unobserved effects correlated with headways. Clearly, the headway variable needs more attention and study. We will return to the problems associated with the headway variable on several occasions in this volume.

Examination of an Empirical Issue with the Basic Model

In the basic model (Table 2), variables were calculated for the bus-with-auto-access alternative under the assumption that individuals would drive their auto to the bus stop, park the auto, and then take the bus. If, instead, a member of the individual's family were to drive the individual to the bus stop and return home in the auto, then the actual cost and time variables would be different from those calculated. To determine the extent to which this problem could affect the estimates in the basic model, the following steps were taken. First, respondents whose households had more drivers than autos were identified. Second, the cost and time variables for the bus-with-auto-access alternatives were recalculated for these respondents under the assumption that these respondents would be driven to the bus stop. Thus, new estimates were made incorporating these variables. This method provides an upper limit to the actual difference between the estimate in Table 2 and the estimates based on exact knowledge of whether a person would drive to the bus stop and park or be driven. The actual number of respondents who would be driven is probably less than the number of households in which there are fewer autos than drivers. (Respondents whose households have more autos than drivers have no incentive to be driven and thus were always assumed to drive-and-park.)

The results of this recalculation are presented in Table 4. The estimates are close to those presented in Table 2. The indication, then, is that the choice of access mode can be included in an MNL model without materially affecting the coefficient estimates. Whether the IIA assumption is violated is not known; nevertheless, the stability of the coefficient estimates is a good omen. We will return to both the diagnostic tests for the violation of the IIA property and the choice of access mode later in greater detail.

TABLE 4 Work-Trip Mode-Choice Model

Mode 1:	Auto	Data:	Work Travel Survey, East Bay
Mode 2:	Bus, Walk Access	Model:	Multinomial Logit, Fitted
Mode 3:	Bus, Auto Access		by the Maximum Likelihood Method

Explanatory Variables	Estimated Coefficients	t-Statistics
Cost divided by post-tax wage, in cents/(cents per minute) e/	- .0364	2.77
On-vehicle time, in minutes e/	- .0166	1.31
Walk time, in minutes a/ e/	- .0240	1.22
Transfer-wait time, in minutes a/ e/	- .0539	1.55
Number of transfers a/ e/	- .0957	.326
Headway of first bus, with a ceiling of 8 minutes, in minutes a/ e/	- .235	2.54
An index of distance to parking at home c/	- .314	0.910
Family income with ceiling of $7000, in $ per year b/	.000426	1.48
Family income minus $7000 with floor of $0 and ceiling of $3000, in $ per year b/	.000790	1.66
Family income minus $10,000 with floor of $0 and ceiling of $5000, in $ per year b/	- .000635	2.51
Length of residence in community, in years b/	.146	2.94

Table 4, continued

Explanatory Variables	Estimated Coefficients	t-Statistics
An index of population density in neighborhood b/	- .736	2.56
Dummy if respondent is over 44 years of age b/	- .845	1.31
Dummy if there is child in household b/	-1.67	2.56
Number of persons in household who can drive c/	1.11	2.68
Auto alternative dummy b/	-5.74	2.36
Bus-with-auto-access dummy d/	-2.93	3.66

Likelihood ratio index:	.5976
Log likelihood at zero:	-176.9
Log likelihood at convergence:	- 71.18

Values of time saved as a percent of wage:

On-vehicle time:	46
Walk time:	66
Transfer-wait time:	148

One transfer with no waiting or walking is valued the same as 5.8 minutes of on-vehicle time.

Table 4, continued

All cost and time variables are calculated round-trip. Dependent variable is alternative choice (one for chosen alternative, zero otherwise). Sample size: 161 .

a/ The variable is zero for the auto alternative, and takes the described value for the other alternatives.

b/ The variable is zero for the bus alternatives, and takes the described value for the auto alternative.

c/ The variable is zero for the bus-with-walk-access alternative, and takes the described value for the remaining alternatives.

d/ The variable is one for the bus-with-auto-access alternative, and zero otherwise.

e/ Sum of home-to-work and work-to-home.

<u>An Examination of a Specification Issue with the Basic Model</u>

Estimation of the basic model was performed on a sample that contained persons who do not have an auto or do not have a license. How these persons should be treated depends upon how the choice to have an auto and license or not is made vis-à-vis the choice of work-trip alternative. The decision could be made recursively, as represented in Figure 5.

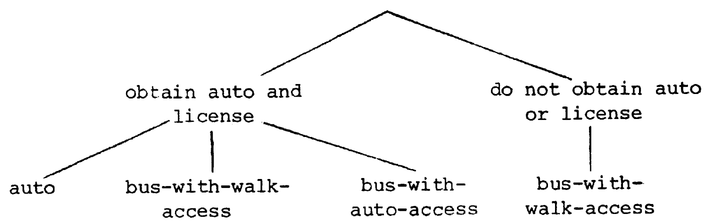

Figure 5

The person chooses to obtain an auto and license or not and then chooses a work-trip alternative. On the other hand, the decisions could be made jointly, as represented in Figure 6.

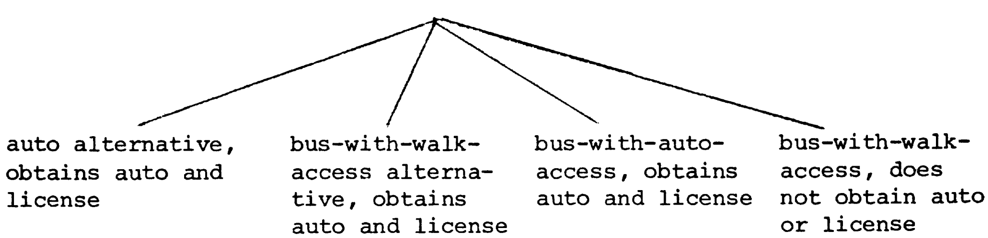

Figure 6

If the decisions are made jointly (and the four possible choices satisfy the assumption of the independence of irrelevant alternatives property), then the representative utility function can be estimated in either of two ways: (l) a four-alternative model could be estimated, with alternatives as the nodes in Figure 6 and attributes of auto and license ownership included as arguments of the representative utility function, or (2) a model similar to the basic model with the three work-trip alternatives can be estimated using the subsample of persons who have an auto or license. Neither of these approaches is feasible for the present study. Attributes of auto and license ownership were not obtained in the interviews in the first place, and, furthermore, the sample is too small to permit the estimation of coefficients.

If the decisions of auto and license ownership and work-trip alternative are made recursively, and the two decisions are independent, then the representative utility function can be estimated by adding to the basic model a dummy variable that assumes the value of one in the auto and bus-with-auto-access alternatives for all persons who have both an auto and a license, and zero otherwise. Table 5 presents estimates of such a model. The outstanding item of the model of Table 5 is that the values of time are about fifty percent higher than those of the basic model, and the estimated coefficient of "number of transfers," though insignificant, has the wrong sign. Whether or not specification error is involved, however, is unclear.

TABLE 5 Work-Trip Mode-Choice Model

Mode 1:	Auto	Data:	Work Travel Survey, East Bay
Mode 2:	Bus, Walk Access	Model:	Multinomial Logit, Fitted
Mode 3:	Bus, Auto Access		by the Maximum Likelihood Method

Explanatory Variables	Estimated Coefficients	t-Statistics
Cost divided by post-tax wage, in cents/(cents per minute) e/	- .0308	2.29
On-vehicle time, in minutes e/	- .0226	1.61
Walk time, in minutes a/ e/	- .0231	1.16
Transfer-wait time, in minutes a/ e/	- .0689	1.74
Number of transfers a/ e/	.0330	0.106
Headway of first bus, with a ceiling of 8 minutes, in minutes a/ e/	- .249	2.55
An index of distance to parking at home c/	- .222	0.461
Family income with ceiling of $7000, in $ per year b/	.000470	1.62
Family income minus $7000 with floor of $0 and ceiling of $3000, in $ per year b/	.000529	1.11
Family income minus $10,000 with floor of $0 and ceiling of $5000, in $ per year b/	- .000481	1.88

Table 5, continued

Explanatory Variables	Estimated Coefficients	t-Statistics
Length of residence in community, in years b/	.136	2.70
An index of population density in neighborhood b/	- .716	2.44
Dummy if respondent is over 44 years of age b/	- .663	0.990
Dummy if there is child in household b/	-1.55	2.26
Number of persons in household who can drive c/	.272	0.578
Auto alternative dummy b/	-6.92	2.72
Bus-with-auto-access dummy d/	-3.56	4.05
Dummy if respondent owns an auto and a license c/	2.59	3.14

Likelihood ratio index:	.6315
Log likelihood at zero:	-176.9
Log likelihood at convergence:	- 65.17

Values of time saved as a percent of wage:

On-vehicle time:	73
Walk time:	75
Transfer-wait time:	224

Table 5, continued

All cost and time variables are calculated round-trip. Dependent variable is alternative choice (one for chosen alternative, zero otherwise). Sample size: 161 .

a/ The variable is zero for the auto alternative, and takes the described value for the other alternatives.

b/ The variable is zero for the bus alternatives, and takes the described value for the auto alternative.

c/ The variable is zero for the bus-with-walk-access alternative, and takes the described value for the remaining alternatives.

d/ The variable is one for the bus-with-auto-access alternative, and zero otherwise.

e/ Sum of home-to-work and work-to-home.

The inclusion of a dummy variable for owning a car and license is only one way to deal with alternative availability. Another would have been to exclude auto and bus-with-auto-access choices from those individuals who do not own a car or a license to drive. If this course of action was chosen then another alternative involving auto, such as shared-ride and drive-drop to the bus stop, should have been included in the choice set. However, the WTS sample of 161 workers was not rich enough to permit the investigation of such issues.

The fact that the dummy variable for owning a car and a license obtains a significance coefficient of substantial magnitude points toward the importance of including the considerations of auto ownership and driver-license status in both the model building and travel forecasting work. Later in the sequel, models will include such alternatives as shared-ride and specific models will be developed for auto ownership, though no attempt will be made to model the choice of obtaining or not obtaining a driver's license. Perhaps in this day and age all eligible persons will know how to drive.

Testing of Hypotheses in the MNL Model

McFadden (1973) shows that the following statistic is distributed as chi-square with degrees of freedom equal to the number of restrictions implied by the null hypothesis:

$$(1) \qquad -2[L(\hat{\theta}^H) - L(\hat{\theta})]$$

where $L(\cdot)$ is the log likelihood function, $\hat{\theta}$ is the unconstrained maximum likelihood estimate of the vector of coefficients, and $\hat{\theta}^H$ is the maximum likelihood estimate of the vector coefficients under the null hypothesis. Using this statistic we obtain the following results.

(1) We accept at the five percent level the hypothesis that auto and bus on-vehicle times are weighted the same. The test-statistic is 0.70 , which is below the critical value (with one restriction) of 3.84 . The point estimates imply that bus on-vehicle time is weighted 2.3 times as much as auto on-vehicle time.

(2) We accept at the five percent level the hypothesis that auto and bus costs are weighted the same. The test-statistic is 0.24 , which is below the critical value (with one restriction) of 3.84 .

(3) We accept at the five percent level the hypothesis that auto mileage costs, tolls, auto parking costs, and bus costs are all weighted the same. The test-statistic is 6.5 , which is below the critical value (with three restrictions) of 7.81 .

(4) We accept at the five percent level the hypothesis that walk, on-vehicle, and transfer-wait times are weighted the same. The test-statistic is 1.02, which is below the critical value (with two restrictions) of 5.99 .

A Non-Logit Choice Model: The Maximum Model

In the logit model, the probability that an individual will choose the first of three alternatives is:

$$P_1 = \frac{e^{U_1}}{e^{U_1} + e^{U_2} + e^{U_3}} \quad ,$$

where U_1 , U_2 , U_3 denote the representative utilities of the three alternatives. The odds of choosing the first alternative over the second alternative are:

$$\frac{P_1}{P_2} = \frac{e^{U_1}}{e^{U_2}} = e^{U_1 - U_2} \quad .$$

Thus, the odds of choosing the first alternative over the second are independent of the existence of the third alternative (i.e., the magnitude of U_3). For choice situations in which this independence does not exist, the logit model is not the appropriate model.

A plausible model that does not entail the assumption of the independence of irrelevant alternatives is the maximum model. With this model a person is assumed to choose between auto and bus and then, if bus is chosen, to choose between walk- and auto-access. The auto-bus decision is based on attributes of the particular bus alternative (either bus-with-walk-access or bus-with-auto-access) that gives the higher representative utility. Thus the probability of choosing the auto alternative is:

$$P_1 = \frac{e^{U_1}}{e^{U_1} + e^{U_4}} \quad ,$$

where $U_4 = \max [U_2 , U_3]$ and 1 denotes auto, 2 denotes bus-with walk-access, and 3 denotes bus-with-auto-access.

The maximum model can be estimated in a two-step procedure. First, the sample is limited to those respondents who chose one of the two bus alternatives. A logit analysis of the choice between walk access and auto access is performed on these respondents. Second, the logit function estimated in step one is used to determine whether walk access or auto access gives higher representative utility for each person in the sample. A logit analysis is performed on the choice between auto and the bus alternative which gives higher representative utility. The estimates obtained in Step 2 of this method are presented in Table 6. The estimated value of on-vehicle time is somewhat higher than in the basic model of Table 2, while that of transfer-wait time is smaller. What renders this model implausible, however, is the positive estimate for the value of walk time. The t-statistic, on the other hand, is relatively small.

A tentative conclusion can be drawn, nevertheless, that it is preferable to include the choice of access mode within the MNL structure rather than to use a non-logit model to accomplish the same purpose. Again, a caveat must be added to note that the violations of the IIA property have not been thoroughly investigated so far, and that some way must be found to deal with the many alternatives that can exist even within a mode-choice model. An example of this would be BART; four access modes (walk, bus, drive, and driven) and two egress modes (walk, bus) translates into eight BART modes. It would be highly presumptuous to believe that no violation of the IIA property would take place in using such a choice set.

TABLE 6 Work-Trip Mode-Choice Maximum Model

Mode 1:	Auto	Data: . .	Work Travel Survey, East Bay
Mode 2:	Bus, Walk Access	Model:	Multinomial Logit, Fitted by
Mode 3:	Bus, Auto Access		the Maximum Likelihood Method

Explanatory Variables	Estimated Coefficients	t-Statistics
Cost divided by post-tax wage, in cents/(cents per minute) e/	- .0319	2.42
On-vehicle time, in minutes e/	- .0233	1.65
Walk time, in minutes a/ e/	.0330	1.23
Transfer-wait time, in minutes a/ e/	- .0395	1.23
Number of transfers a/ e/	- .0551	.189
Headway of first bus, with a ceiling of 8 minutes, in minutes a/ e/	- .156	1.68
An index of distance to parking at home c/	- .243	.682
Family income with ceiling of $7000, in $ per year b/	.000360	1.36
Family income minus $7000 with floor of $0 and ceiling of $3000, in $ per year b/	.000768	1.63
Family income minus $10,000 with floor of $0 and ceiling of $5000, in $ per year b/	- .000540	2.15
Length of residence in community, in years b/	.138	2.80

Table 6, continued

Explanatory Variables	Estimated Coefficients	t-Statistics
An index of population density in neighborhood b/	- .803	2.86
Dummy if respondent is over 44 years of age b/	- .646	1.02
Dummy if there is child in household b/	-1.26	2.02
Number of persons in household who can drive c/	.448	1.34
Auto alternative dummy b/	-2.65	1.24

Likelihood ratio index:	.5170
Log likelihood at zero:	-111.6
Log likelihood at convergence:	- 53.90

Values of time saved as a percent of wage:

On-vehicle time:	73
Transfer-wait time:	124

One transfer with no waiting or walking is valued the same as 2.4 minutes of on-vehicle time.

Table 6, continued

All cost and time variables are calculated round-trip. Dependent variable is alternative choice (one for chosen alternative, zero otherwise). Sample size: 161 .

a/ The variable is zero for the auto alternative, and takes the described value for the other alternatives.

b/ The variable is zero for the bus alternatives, and takes the described value for the auto alternative.

c/ The variable is zero for the bus-with-walk-access alternative, and takes the described value for the remaining alternatives.

d/ The variable is one for the bus-with-auto-access alternative and zero otherwise.

e/ Sum of home-to-work and work-to-home.

The Tradeoff between Goods and Leisure

In the above analysis, cost is divided by wage, and the time variables do not interact with wage. Elsewhere (McFadden, 1976), the opposite procedure was used: times were multiplied by wage, and cost did not interact with wage. Because an individual's wage conveys information about his tradeoff between goods and leisure in the neoclassical model of consumer behavior, the manner in which wage is treated in the logit analysis has implications for the form of the goods/leisure tradeoff. In order to determine the extent of these implications and choose the most satisfactory way of treating wage, a utility maximization model is employed in which utility is a function of goods and leisure, and in which attributes of work-trip alternatives enter into the constraints on maximization.

The general procedure is as follows. Choose some specific functional form for the utility function, $U(G,L)$, where G is goods and L is leisure. Assuming the price index to be constant and normalized to one, the following identities must hold:

$$G = V + w \cdot W - c \quad ,$$
$$L = T - W - t \quad ,$$

where

$V =$ unearned income (given) ;

$w =$ wage rate (given) ;

$T =$ total amount of time (given) ;

$W =$ number of hours worked (a continuous variable, non-negative) ;

$c =$ cost of transportation to and from work (a discrete variable that can assume values $c_1,\ldots,c_n$) ;

$t =$ time of transportation to and from work (a discrete variable that can assume values $t_1,\ldots,t_n$) ;

$n =$ number of work-trip alternatives.

The individual chooses the work-trip alternative (and hence c and t) and the number of hours worked so as to maximize U, subject to the identities.

Define:

$$G_i = V + w \cdot W - c_i \quad ;$$

$$L_i = T - W - t_i \quad ;$$

$$U_i = U(G_i, L_i) \quad .$$

Thus U_i is a function only of the variable W. Because W is continuous, U_i can be maximized in the normal way, setting

$$(1) \qquad \frac{\partial U_i}{\partial W} = U_{i1} \, w - U_{i2} = 0 \quad ;$$

or, as usual:

$$(2) \qquad w = \frac{U_{i2}}{U_{i1}} \quad .$$

Solve this for W and call the solution W_i^*. Substitute W_i^* into U_i and call the resulting value U_i^*.

Alternative i is chosen by the individual if and only if $U_i^* > U_j^*$ for all $j = 1,\ldots,n \, ; j \neq i$. Let U^* be the function of c and t, which assumes values U_i^*, $i = 1,\ldots,n$. This U^* is the function used in logit analysis. Thus, one can determine what functional forms of the goods/leisure tradeoff (that is, $U(G,L)$) produce particular forms of U^*.

An equivalent approach, which is less heuristic but easier computationally, is the following. Choose a specific $U(G,L)$ and derive the corresponding expenditure function:

$$(3) \qquad E = E(U,w) \quad ,$$

where E is expenditure(s) and prices are normalized to one. The following identity holds for expenditures when utility is maximized:

(4) $$E = V - c + w(T - t) \quad .$$

Substitute (4) into (3) and solve for U. The solution is U^*. The U^* obtained in this manner is the same or a monotonic transformation of the U^* obtained in the first approach.

Some specific examples follow.

Example A. Let $U = \alpha_1 \log G + a_2 L$. With this function the derivative of the utility-maximizing G with respect to income is zero: all extra income is absorbed in leisure. The resulting U^* is:

$$U^* = \alpha_1 \log\left(\frac{\alpha_1}{\alpha_2} w\right) + \alpha_2 T - \alpha_1 + \alpha_2 \frac{V}{w} - \alpha_2\left(\frac{c}{w} + t\right) \quad .$$

When comparing U_i^* and U_j^* $(i \neq j)$, all terms that do not contain either c or t drop out. Thus, operationally, U^* is

(5) $$U^* = -\alpha_2\left(\frac{c}{w} + t\right) \quad .$$

In this case, cost is divided by wage rather than time being multiplied by wage. There is only one parameter since c/w is in units of time, and the individual values work time and transportation time the same on the margin. The two variables can have different coefficients by specifying a model analogous to that in Example D below.

Example B. Let $U = \alpha_1 G + \alpha_2 \log L$. With this function the derivative of the utility maximizing G with respect to income is zero and all extra income is absorbed in goods. The resulting U^* is:

$$U^* = \alpha_1\left(V + Tw - tw - \frac{\alpha_2}{\alpha_1} - c\right) + \alpha_2 \log \frac{\alpha_2}{\alpha_1} w \quad .$$

Operationally,

(6) $$U^* = -\alpha_1(tw + c) \quad .$$

In this case, time is multiplied by wage rather than cost being divided by wage. Different parameters can be given to the two terms in a manner analogous to Example D below.

Example C. Let U be a Cobb-Douglas utility function:

$$U = AG^{1-\beta}L^{\beta} \, , \quad 0 < \beta < 1 \quad .$$

The expenditure function (with prices normalized to one) is:

$$E = Uk^{-1}w^{\beta} \quad ,$$

where k is a constant. Recall that

$$E = V - c + w\,(T - t) \quad .$$

Thus,

$$Uk^{-1}w^{\beta} = V - c + w(T - t) \quad .$$

And,

$$U^* = k(w^{-\beta}V - w^{-\beta}\,c + w^{1-\beta}T - w^{1-\beta}t) \quad .$$

Operationally,

(7) $$U^* = -k(w^{-\beta}c + w^{1-\beta}t) \quad .$$

When β approaches 0 , (7) becomes (6); and when β approaches 1 , (7) becomes (5). For values of β between zero and one, the derivatives of the utility-maximizing L and G with respect to income are greater than zero. The choice of β is an empirical issue.

Example D. The analysis in Example C can be extended so that the terms in (7) have different coefficients and components of time and cost enter each with different coefficients. Let U be the same as in Example C. The definitions of goods and leisure are re-specified:

$$(8)\qquad \begin{aligned} &\text{“effective” leisure} \quad L_i = \theta_{0i}T - W - \sum_{j=1}^{M} \theta_j t_i^j \quad ; \\ &\text{“effective” goods} \quad G_i = \gamma_{0i}V + wW - \sum_{j=1}^{N} \gamma_j c_i^j \quad ; \end{aligned}$$

where

- t^j is the j-th component of travel time (say, on-vehicle time);
- c^j is the j-th component of travel cost;
- M is the number of time components;
- N is the number of cost components;
- θ_j is the psychometric weight attached to a minute in travel component j , in work time units;
- θ_{0i} is the psychometric weight attached to the total time budget, in work time units, when mode i is used;
- γ_j is the psychometric weight attached to a travel cost component, in wage income units;
- γ_{0i} is the psychometric weight attached to non-wage income, in wage income units, when mode i is used.

The θ_j and γ_j reflect the relative onerousness or burden of different time or expenditure activities (associated, for example, with the exertion, fatigue, or bother involved). The parameters θ_{0i} and γ_{0i} reflect the value of added units of time or income (again evaluated in working units). They may differ with mode if choice of mode itself affects the types of consumption and leisure activities available to the consumer. "Effective" income is:

$$E_i = (\gamma_{0i}V + w\theta_{0i}T) - \sum_{j=1}^{N} \gamma_j c_i^j - \sum_{j=1}^{M} \theta_j w t_i^j \quad .$$

Substituting this expression into the expenditure function for Example C and solving for U obtains:

$$U_i^* = k\left[\gamma_{0i}Vw^{-\beta} + \theta_{0i}w^{1-\beta}T - \sum_{j=1}^{N} \gamma_j c_i^j w^{-\beta} - \sum_{j=1}^{M} \theta_j t_i^j w^{1-\beta} \right] \quad .$$

This formula was used assuming the γ_{0i} and θ_{0i} constant in i, so that operationally:

(9) $$U^* = -k \sum_{j=1}^{N} \gamma_j c^j w^{-\beta} - k \sum_{j=1}^{M} \theta_j t^j w^{1-\beta} \quad .$$

Example E. The analysis in Example D can be extended so that U^* depends on socioeconomic variables and alternative dummy variables. Redefine "effective" goods and leisure as:

"effective" leisure $L_i = \theta_{0i}T - W - \sum_{j=1}^{M} \theta_j t_i^j - \sum_{j=1}^{P} \theta_{j+M} u_i^j$;

"effective" goods $G_i = \gamma_{0i}V + wW - \sum_{j=1}^{N} \gamma_j c_i^j + \sum_{j=1}^{Q} \gamma_{j+N} v_i^j$;

where

- u_i^j is the j-th unmeasured time component of travel by mode i ;
- v_i^j is the unmeasured consumption of "good" j in traveling by mode i .

An example of a u^j is the length of time that one would usually arrive early to work so as not to be late for work if one's travel mode were delayed. If this variable were measured, it would assume a different value for each alternative, because the probability of being delayed a certain length of time varies across modes. An example of a v^j is the privacy that one "consumes" in a particular travel mode.

Because the variables are unmeasured, they are approximated by other, measured variables, such as socioeconomic variables:

$$\begin{aligned} \sum_{j=1}^{P} \theta_{j+M} u_i^j &= \sum_{j=1}^{R} \delta_j x_i^j + \varepsilon_i \quad , \\ \sum_{j=1}^{Q} \gamma_{j+N} v_i^j &= \sum_{j=1}^{S} \eta_j y_i^j + \mu_i \quad , \end{aligned} \tag{10}$$

where

x_i^j is the j-th measured variable used to approximate the unmeasured time components of travel by mode i ;

y_i^j is the j-th measured variable used to approximate unmeasured consumption from travel by mode i ;

δ_j and η_j are parameters; and

ε_i and μ_i are errors that allow equations (10) to hold exactly rather than approximately.

Substituting (10) into the definitions of "effective" leisure and goods and solving for U^* as in Example D obtains:

$$\begin{aligned} U^* = -k \sum_{j=1}^{N} \gamma_j c^j w^{-\beta} - k \sum_{j=1}^{M} \theta_j t^j w^{1-\beta} \quad , \\ -k \sum_{j=1}^{S} \eta_j y_i^j w^{-\beta} - k \sum_{j=1}^{R} \delta_j x_i^j w^{1-\beta} - k(\varepsilon + \mu) \quad . \end{aligned} \tag{11}$$

For logit analysis, it is assumed that $-k(\varepsilon + \mu)$ is distributed Weibull in the population and is independent across alternatives.

To estimate β and the coefficients, logit estimation can be performed for various values of β between zero and one, with utility defined as in (11). The value of β that results in the largest value of the likelihood function is chosen as the estimate of β .

Unfortunately, equation (11) does not help to identify which socioeconomic variables are to be considered y's and which ones x's . The choice of which way to treat a socioeconomic variable affects the likelihood function and hence the estimate of β . Several divisions of the socioeconomic variables were considered. When all the socioeconomic variables and alternative dummy variables were treated as y's , the estimated β was approximately 0.7 . For this model, with β equal to 0.7 , the estimated values of time were generally about half as large as those estimated for the basic model. For other divisions of the socioeconomic variables, the estimated β was generally between 0.7 and 1.0 , with the estimated values of time rising with the estimated β .

Thus, the results here suggest that dividing travel cost by wage is better than multiplying travel time by wage, and that the time variables do not, in fact, need to be interacted with wage. In some ways this result is very convenient. First, the specification of the model is simple and the practitioner has one less variable to worry about. Second, the assumption of neoclassical economic theory that individuals are trading off goods and leisure time, with the number of hours worked being a continuous variable, is not a fully satisfactory basis for model specification, for a number of reasons. It is unlikely that most people have an unrestricted choice of how many hours they want to work. The division of time into working time and leisure time is somewhat arbitrary; there are many other considerations which, in real life, enter the division of time (paid and unpaid) into various activities.

The empirical conclusions that $\beta = 1$ is consistent with a neoclassical theory of behavior. On the other hand, it implies a simple model structure which should appeal to pragmatists who do not accept or consider irrelevant the neoclassical formulation of the choice problem.

In conclusion, the issues warranting closer examination with the full sample can be identified as being the availability of alternatives and the choice set, especially with regard to the handling of access modes; the independence from irrelevant alternatives property of the MNL model; and, naturally, further testing of the specification itself.