PART III

MODELING CHOICES OTHER THAN WORK-TRIP

CHAPTER 1

A STRUCTURED LOGIT MODEL OF AUTO OWNERSHIP AND MODE CHOICE

Introduction

A household owns some number of autos, perhaps zero. This number can be regarded as the outcome of a choice by the household members. This choice is called the auto-ownership choice.

The auto-ownership choice is interrelated with the work-trip mode. That is, the number of autos a household chooses to own depends upon the mode-choices of its workers, and each worker's mode-choice is affected by the number of autos available to the household.

The present study analyzes the choices of auto-ownership and work-trip mode. Models of these choices are developed and estimated that take explicit account of the interaction between the choices. The models are disaggregate in that the unit of analysis is the individual worker and household. As discussed below, the models are specified to be a structured logit model, which was shown before to be a generalization of the logit model.

We review the previous research on auto-ownership and work-trip mode-choice; present and discuss the model specification; present the work-trip mode-choice part of the joint model, and present the auto-ownership part of the joint model.

Previous Research[1]

Auto ownership research

Most research has confronted either the issue of auto-ownership or that of mode-choice. Of the auto-ownership models, most have been estimated on aggregate data. Such data present fairly severe problems in estimation. First among these problems is the paucity of observations that aggregate data afford. All the models using aggregate data have been based on annual time-series data. Consequently, the models were estimated with extremely few degrees of freedom. The model of Kulash (1975), for example, has only eleven degrees of freedom; that of RAND (1975) has only fourteen. Quarterly data or cross-sectional state data could perhaps have been used in estimation; none of the models chose these data, presumably because of the other problems with aggregate data: collinearity among variables and little variation in variables over time or across states. These problems are more severe with cross-sectional state data and quarterly national data than with annual national data, but are nevertheless nearly debilitating with annual national data. When coupled with the small number of annual observations available, these problems result in models with only a few significant explanatory variables. The problems of aggregate data point to the need to model auto-ownership choices with disaggregate data.

Only two models of auto-ownership have been estimated on disaggregate data (Johnson (1975) and Burns and Golob (1975)). While these models avoid the problems of aggregate data, they include very few explanatory variables. Only socioeconomic variables are considered to affect auto-ownership choices in Johnson's model; the costs and benefits to the household of owning autos were not included. Burns and Golob include the cost of auto-ownership and a variable representing the extra accessibility to shopping and other destinations which owning an auto allows; these two variables are the only explanatory variables in their model, however.

In all the models of auto-ownership, the effect of the worker's mode-choice is ignored.

[1] See Train (1977) for a more extensive review and critique of previous research.

Mode choice research

Previous models of mode-choice point to the need to use a disaggregate approach to modeling mode-choice and to model mode-choice and auto-ownership jointly. Aggregate mode-split models have been criticized by Stopher and Lisco (1970), Brand (1973a and b), and Domencich and McFadden (1975). The major problem with these models is that they generally describe correlative relations that exist between variables at the time and for the place in which the model was estimated, rather than describe causal relations. Furthermore, it seems that the aggregate data are not rich enough to allow a model to capture the true causal relations. Hence, a disaggregate approach is necessary.

A considerable amount of mode-choice research has been conducted with disaggregate data. In all of these models, however, the issue of auto-ownership is a source of concern. Warner (1962), Lisco (1967), and Quarmby (1967) include auto-ownership variables in their models with the (implicit) assumption that the number of autos owned is exogenous to the mode-choice. Lave (1969) recognizes that auto-ownership is endogenous; hence, he does not include an auto-ownership variable and states that his model formulation can be considered a reduced-form equation. McFadden (1974) and Train and McFadden (1976) do not include any auto-ownership variable in their models because of the econometric problems incurred by inclusion. Train (1976b) compares models with and without an auto-ownership variable and discusses the problems incurred with each specification. These problems point to the need for a joint modeling of mode-choice and auto-ownership.

Auto-ownership and mode-choice research

To our knowledge, the only joint econometric model of auto-ownership and mode-choice is that of Lerman and Ben-Akiva (1975). The model is useful in being the first of its kind and in pointing the direction for further research. Several problems with the model, however, inhibit its usefulness.

First, the number of alternatives designated to be available to households is quite restricted. Only two modes are available to the worker: auto and transit. As a result, policies directed toward shifting commuters from the auto-alone to the carpool modes cannot be analyzed with the model. Furthermore, only three auto-ownership levels are considered in the model, those being no auto, one auto, and two or more autos. Because many households today own three and more autos, and because many policies can be expected to have a greater effect on the choice between two and three autos than on the choice between none and one auto, the restricted choice set inhibits the plausibility and usefulness of the model.

A second problem with the model is the small number of explanatory variables included in the model. The only attributes of the work-trip that are considered explanatory are cost, on-vehicle time, out-of-vehicle time, and distance. Because out-of-vehicle time is not decomposed into time spent waiting for transit and time spent walking to and from transit, the effect of policies trading-off these two components cannot be analyzed. An example of such a policy is to place more buses on fewer bus lines, thus decreasing wait-time and increasing walk-times. The number of variables explaining the level of choice-of-auto-ownership is also fairly restricted. The cost of auto-ownership is included, as is a variable reflecting the ease of accessing shipping destinations by auto relative to that by transit. The effect of family structure on auto-ownership, however, is not considered. Thus, according to the model, a household does not have a higher probability of owning an auto when it gains a member. Nor does the probability of ownership rise in this model when a household member learns to drive and obtains a license. These limitations diminish the plausibility of the model.

Model Specification

Consider a household with one worker.[1] The worker can choose any one of several modes of travel to work; label the total number of modes available to him as J. The household is assumed to have a choice among I levels of auto-ownership (that is, it can choose to own none, one, two, etc., autos). The household and worker are assumed to choose an auto-ownership level and work-trip mode so as to maximize utility.

Adopting the sequential logit model (Part IV, Chapter 2), the utility is expressed as

$$(1) \qquad U_{ij} = Y_i + W_{ij}(1 - \theta) + \varepsilon_{ij} \quad ,$$

where Y_i is the "representative" utility of auto ownership level i, W_{ij} is the representative utility of mode j, θ is the similarity index discussed on page 301, and ε_{ij} is the unobserved component of utility. The combination of auto ownership level i and work-trip mode j is called alternative ij.

It is shown that the sequential (structured) logit model of joint choice is a random utility model and is a generalization of the MNL model. In the sequential (structured) logit model, the choice probabilities are expressed as:

$$(2) \qquad P_{ij} = P_{i|i} \cdot P_i \quad ,$$

$$(3) \qquad P_{ji} = \frac{e^{W_{ij}}}{\sum_{k=1}^{J} e^{W_{ij}}} \quad ;$$

$$(4) \qquad P_i = \frac{e^{Y_i + (1-\theta)Q_i}}{\sum_{\ell=1}^{I} e^{Y_\ell + (1-\theta)Q_\ell}} \quad ;$$

---

[1]The problems of introducing two or more workers are discussed later.

and

$$(5) \qquad Q_i = \log \sum_{j=1}^{J} e^{W_{ij}} \quad ;$$

where $P_{j|i}$ is the conditional probability of the worker choosing work-trip mode j given auto-ownership level i ; P_i is the marginal probability of the household choosing auto-ownership level i (the marginality being over all work-trip modes); Q_i is called the aggregate work-trip utility associated with auto-ownership level i (this term is also called an inclusive value).

Note that the conditional probabilities (3) and the marginal probabilities (4) are both in the form of a logit model and hence are easy to estimate. The conditional probabilities can be estimated by specifying W_{ij} up to some parameters and estimating the parameters by the maximum likelihood method developed by McFadden (1973) for logit models. The choice being analyzed in this estimation is the worker's choice of mode given the household's observed choice of auto-ownership level. Once the conditional probabilities for work-trip mode are estimated, the aggregate work-trip utility can be calculated by (5) for each auto-ownership level. Then, the marginal probabilities of auto ownership levels, expressed in (4) can be estimated: the term Y_i is specified up to some parameters and these parameters are estimated along with θ by McFadden's maximum likelihood method for logit models. The aggregate work-trip utility term simply enters as another explanatory variable in the logit estimation; its estimated coefficient is the estimate of $1-\theta$.

Estimating the probabilities P_{ij} in this step-wise fashion allows one to use existing maximization programs and reduces the amount of data required for estimation. (Data on the attributes of the work-trip mode for non-chosen levels of auto-ownership are not required with the step-wise method.) Furthermore, breaking the probabilities P_{ij} into their conditional and marginal components facilitates analysis of the restrictions implied by the logit model. It was shown above that the structured logit model does not possess the IIA property for each pair of alternatives.

It is clear from (3), however, that the conditional probabilities for a given auto ownership exhibit the IIA property:

$$\frac{P_{j|i}(D)}{P_{k|i}(D)} = \frac{e^{W_{ij}}}{e^{W_{ik}}} \quad ,$$

where D is a set of work-trip modes available to the worker. This ratio is constant for any D that contains modes j and k and for any attributes of modes other than j and k in D .

From (4) it is clear that the marginal probabilities of auto ownership level exhibit the IIA property:

$$\frac{P_i(A)}{P_m(A)} = \frac{e^{Y_i + (1-\theta)Q_i}}{e^{Y_m + (1-\theta)Q_m}} \quad ,$$

where A is the set of auto-ownership levels available to the household. This ratio is constant for any A containing i and m and for any attributes of auto-ownership levels other than i and m in A . In conclusion, the structured logit model is less restrictive than the logit model because each pair of alternatives does not exhibit the IIA property; the structured logit model does, however, imply some restrictions on the choice probabilities, namely, that the marginal and conditional probabilities exhibit the IIA property.

In the following two sections, the structured logit model of auto-ownership and work-trip choice will be presented. First, we present the estimated model of the conditional probabilities, that is, the estimate of (3). Some tests of the specification of W_{ij} are applied, and the validity of the IIA property of the conditional probabilities is tested using methods introduced above. We next present the estimated model of the marginal probabilities, that is, the estimate of (4). Some tests of the specification of Y_i are applied, and the validity of the IIA property of the marginal probabilities is tested. In estimating (4), an estimate of θ is obtained. As explained above, when θ equals zero, the model is logit. The hypothesis of θ being zero is tested, and the effect on parameter estimates of restricting θ to zero is examined.

Conditional Probabilities of Work-Trip Mode-Choice

This section presents the estimated model of the conditional probabilities of the work-trip mode-choice given the household's auto-ownership level. The model was estimated and the tests applied on a sample of workers taken in the San Francisco Bay Area during 1975. The sample and data are described in detail in Train (1977).

Workers were considered to have a choice among seven alternative modes of travel to work:

(1) auto-alone;

(2) bus-with-walk-access-to-bus;

(3) bus-with-auto-access-to-bus;

(4) BART-with-walk-access-to-BART;

(5) BART-with-bus-access-to-BART;

(6) BART-with-auto-access-to-BART;

(7) carpool.

The model estimates the probability that a particular worker will choose each of the seven modes for his work-trip, given the observed auto-ownership choice of his household.

In the subsections below, we present and discuss the conditional model of work-trip mode-choice. Specification tests are applied to the model, and tests of the IIA property of the conditional probabilities are applied.

The conditional work-trip mode-choice

Table 21 presents the model estimated on the post-BART sample. The model is a particular specification of the probability of choosing a particular mode of work-trip travel, conditional upon the household's auto-ownership level. The model takes the form of (3). The term W_{ij} is called the representative utility of the work-trip mode and is specified to be linear-in-parameters:

$$W_{ij} = \beta' w(x_{ij},s) \quad , \tag{6}$$

where s is a vector of attributes of the household and worker (such as income); x_{ij} is a vector of attributes of mode j and auto-ownership level i (such as work trip cost); w is a vector-valued function of x_{ij} and s ; and β is a vector of parameters to be estimated.

Estimation of β was performed by the maximum likelihood method described in McFadden (1973). For each person in the sample, the available choice set was considered to be the seven alternatives listed above with the following exceptions. The auto-alone alternative was considered unavailable to a worker if the worker's household had chosen to own no autos. Any of the transit alternatives was considered unavailable to a worker if going to work by that alternative entailed more than three transfers either to or from work, total weighted time exceeding four hours in either direction, or other excessive attributes.

The first column of Table 21 lists the elements of $w(x_{ij},s)$ in equation (6). The second and third columns list the estimates and t-statistics, respectively, of the elements of β in equation (6) .

The model of Table 21 is similar to that of Table 8, page 89. The main difference is that the model of Table 21 includes the variable "number of autos in household" instead of the variable "number of drivers in household". The model of Table 8, page 89, included the drivers variable as a proxy for the autos variable to avoid including an endogenous variable as an explanatory variable in a model of mode-choice that is not joint with a model of auto-ownership choice. In the present study, auto-ownership and mode-choice are modeled jointly and, consequently, including the endogenous autos variable does not present econometric problems.

TABLE 21 A Work-Trip Mode-Choice Model Conditional on Auto Ownership Level

(Mode 1 - Auto Alone; Mode 2 - Bus, Walk Access; Mode 3 - Bus, Auto Access; Mode 4 - BART, Walk Access; Mode 5 - BART, Bus Access; Mode 6 - BART, Auto Access; Mode 7 - Carpool)

Model: Multinomial Logit, Fitted by the Maximum Likelihood Method

Independent Variable (The variable takes the described value in the alternatives listed in parentheses and zero in non-listed alternatives)	Estimated Coefficient	T-Statistic
Cost divided by post-tax wage, in cents divided by cents per minute (1-7)	-.0263	3.92
Auto on-vehicle time, in minutes (1,3,6,7)	-.0473	3.54
Transit on-vehicle time, in minutes (2-6)	-.0192	2.01
Walk time, in minutes (2-6)	-.0861	3.28
Transfer-wait time, in minutes (2-6)	-.0479	1.99
Number of transfers (2-6)	.137	0.986
Headway of first transit carrier, in minutes (2-6)	-.0259	2.41
Family income with ceiling of $7,500, in $ per year (1)	-.000282	1.73
Family income minus $7,500 with floor of $0 and ceiling of $3,000, in $ per year (1)	.0000496	0.345
Family income minus $10,500 with floor of $0 and ceiling of $5,000, in $ per year (1)	-.0000474	0.828
Number of autos in household (1)	1.93	5.57
Number of autos in household (3,6)	2.08	5.51

Table 21, continued

Independent Variable	Estimated Coefficient	T-Statistic
Number of autos in household (7)	1.86	5.79
Autos per driver with a ceiling of one (1)	1.21	2.71
Dummy if person is head of household (1)	.605	3.01
Employment density at work location (1)	-.00166	3.50
Auto-alone alternative dummy (1)	-1.01	0.795
Bus-with-auto-access dummy (3)	-5.17	7.81
BART-with-walk-access dummy (4)	-2.25	3.35
BART-with-bus-access-dummy (5)	-.524	0.789
BART-with-auto-access dummy (6)	-4.22	6.74
Carpool alternative dummy (7)	-2.81	4.35

Likelihood ratio index	.4542
Log likelihood at zero	-964.4
Log likelihood at convergence	-526.4
Percent correctly predicted	67.65

Table 21, continued

Values of time saved as a percent of wage (t-statistics in parentheses):

Auto on-vehicle time	180	(2.55)
Transit on-vehicle time	73	(1.82)
Walk time	327	(2.42)
Transfer-wait time	183	(1.78)

Value of initial headways as a percent of wage: 98 (2.03)

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

Number of people in sample who chose:

Auto-alone	378
Bus-with-walk access	68
Bus-with-auto-access	9
BART-with-walk-access	4
BART-with-bus-access	6
BART-with-auto-access	33
Carpool	137
Total Sample Size	635

The variables in Table 21 are similar to those in the model of Table 8, page 89, and hence do not require further explanation. The qualitative discussion of that model applied equally well to that of Table 21.

Specification tests

Tests of non-genericity and non-linearity. The results of the tests of non-genericity and non-linearity applied to the model of Table 21 are the same as those of the tests applied to the model of Table 8, page 89. The hypothesis was not rejected that the various restrictions implied by the model of Table 21 are accurate. The details of these tests are presented in Train (1977).

Tests of taste variations. The model of Table 21 implicitly assumes that the coefficients of times and cost do not vary in the population. That is, it is assumed that all people have the same values of time. A more general model is one in which tastes are allowed to vary systematically with some observed socioeconomic characteristics.

Table 22 presents a model in which the coefficients of the variables "cost divided by wage", "auto on-vehicle time", "transit on-vehicle time", and "walk time" are allowed to be different for suburban dwellers and urban dwellers. The model is significantly better than that of Table 21 in that the hypothesis of equal coefficients for suburban and urban dwellers is rejected. The log likelihood value for the model of Table 22 is -521.0 . That of the model of Table 21 is -526.4 . Therefore, the test-statistic (from page 25) is 10.8 . The critical value (at the .05 significance level) of chi-squared with four degrees of freedom is 9.5 . Therefore, the hypothesis that the coefficients of the four variables are the same for suburban and urban dwellers is rejected at the .05 significance level.

TABLE 22 <u>Work-Trip Mode-Choice Model, Conditional on Auto-Ownership Level, with Different Values of Time for Urban and Suburban Dwellers</u>

(Mode 1 - Auto Alone; Mode 2 - Bus, Walk Access; Mode 3 - Bus, Auto Access; Mode 4 - BART, Walk Access; Mode 5 - BART, Bus Access; Mode 6 - BART, Auto Access; Mode 7- Carpool)

Model: Multinomial Logit, Fitted by the Maximum Likelihood Method

Independent Variable (The variable takes the described value in the alternatives listed in parentheses and zero in non-listed alternatives)	Estimated Coefficient	T-Statistic
Cost divided by post-tax wage, in cents divided by cents per minute, for urban dwellers; zero for suburban dwellers (1-7)	-.0163	2.17
Cost divided by post-tax wage, in cents divided by cents per minute, for suburban dwellers; zero for urban dwellers (1-7)	-.0412	4.40
Auto on-vehicle time, in minutes, for urban dwellers; zero for suburban dwellers (1,3,6,7)	-.0465	2.92
Auto on-vehicle time, in minutes, for suburban dwellers; zero for urban dwellers (1,3,6,7)	-.0606	3.76
Transit on-vehicle time, in minutes, for suburban dwellers; zero for urban dwellers (2-6)	-.0103	0.978
Transit on-vehicle time, in minutes, for suburban dwellers; zero for urban dwellers (2-6)	-.0306	2.31
Walk time, in minutes, for urban dwellers; zero for suburban dwellers (2-6)	-.0698	2.25

Table 22, continued

Independent Variable	Estimated Coefficient	T-Statistic
Walk time, in minutes, for suburban dwellers; zero for urban dwellers (2-6)	-.106	3.25
Transfer wait time, in minutes (2-6)	-.0570	2.28
Number of transfers (2-6)	.187	1.31
Family income minus $7,500 with floor of $0 and ceiling of $3,000, in $ per year (1)	.0000559	0.393
Family income minus $10,500 with floor of $0 and ceiling of $5,000, in $ per year (1)	-.0000425	0.733
Number of autos in household (1)	1.81	5.25
Number of autos in household (3,6)	2.01	5.29
Number of autos in household (7)	1.73	5.39
Autos per driver with a ceiling of one (1)	1.24	2.76
Dummy if person is head of household (1)	.640	3.04
Employment density at work location (1)	-.00162	3.40
Auto-alone alternative dummy (1)	-1.04	0.830
Bus-with-auto-access dummy (3)	-4.99	7.45
BART-with-walk-access dummy (4)	-2.20	3.22
BART-with-bus-access dummy (5)	-.595	- 0.882
BART-with-auto-access-dummy (6)	-4.08	6.43
Carpool alternative dummy (7)	-2.26	3.30

Table 22, continued

Likelihood ratio index	.4598
Log likelihood at zero	-521.0
Log likelihood at convergence	-964.0
Percent correctly predicted	66.30

Values of time saved as a percent of wage (t-statistics in parentheses):

	Urban dwellers		Suburban dwellers	
Auto on-vehicle time	285	(1.72)	147	(3.10)
Transit on-vehicle time	63	(0.92)	74	(2.29)
Walk time	428	(1.53)	257	(2.69)
Transfer-wait time	439	(1.57)	138	(2.04)

Value of initial headways as a percent of wage: 162 (1.58) 64 (2.13)

All cost and time variables are calculated round-trip. Dependent variable is alternative choice (one for chosen alternative, zero otherwise). Number of people in sample who chose:

Auto-alone	378
Bus-with-walk-access	68
Bus-with-auto-access	9
BART-with walk-access	4
BART-with-bus-access	6
BART-with-auto-access	33
Carpool	137
Total Sample Size	635

The estimated values of time and headways are higher for urban dwellers than suburban dwellers, except for the value of transit on-vehicle time. This result is expected. There is, generally, a trade-off between cost of housing (normalized for quality) and the time necessary to spend traveling to work, shopping, and recreation destination. Housing units in urban areas are generally more expensive than housing units of comparable quality in suburban areas, yet travel times to shopping, etc., are generally shorter in urban areas than in suburban areas. Given this tradeoff, people with high values of time are expected to locate their homes more in urban areas than persons with low values of time. This expectation is confirmed by the estimates of Table 22.[1]

The slightly higher estimated value of transit on-vehicle time for suburban dwellers than urban dwellers does not fit into the pattern expected from the above argument. It might be the case, however, that suburban dwellers dislike the discomfort of transit and fear crime sufficiently (relative to urban dwellers) to offset the effect of locational patterns.

Diagnostic tests of the IIA property of the conditional probabilities

The conditional model of Table 21 seems likely to violate the IIA property. For instance, because some of the alternative modes are similar, unobserved attributes of each mode are perhaps correlated across modes. In particular, the comfort of on-vehicle travel is similar for bus-with-walk-access and bus-with-auto-access, and yet no comfort variable is included in the model. Similarly, the reliability of BART is the same whether a person walks, drives or takes a bus to BART; however, no reliability measure was included in the model. The question of whether these and similar problems are severe enough to reject the IIA property of the conditional probabilities is explored by applying diagnostic tests developed above.

The first tests employ the observation demonstrated by McFadden (1975) that any choice model can be expressed in logit form. Specifically, it is shown below that any choice model of work-trip mode given auto-ownership level can be written as a logit model that contains the variables listed in Table 21 plus other variables.

---

[1]The values of time in Table 22 are normalized by the person's wage rate. If each person were able to optimize the number of hours which he worked, then the result obtained above (that people with higher values of time locate in urban areas) would not hold for values of time expressed as a percentage of wage. Given, however, that the number of hours a person works is not completely flexible, the above argument is plausible.

Following McFadden, any model of the probability of choosing mode j given auto-ownership level i can be written as:

(7)
$$P_{j|i} = \frac{e^{Z_j(x_{i1},\ldots,x_{iJ},s)}}{\sum_{\ell=1}^{J} e^{Z_\ell(x_{i1},\ldots,x_{iJ},s)}} \quad .$$

Define

(8)
$$T_j(x_{i1},\ldots,x_{iJ},s) = z_J(x_{i1},\ldots,x_{iJ},s) - \beta' w(x_{iJ},s) \quad ,$$

where the last term is the representative utility for mode j in the model of Table 21 (see (6)). If the domain of T_j is closed and bounded and T_j is continuous on its domain, then T_j can be approximated to any degree of accuracy by a function that is linear in parameters:

(9)
$$T_j(x_{i1},\ldots,x_{iJ},s) = \delta' t(x_{i1},\ldots,x_{iJ},s) \quad ,$$

where t is a vector-valued function of the data and δ' is a vector of parameters. Using (9) and (8), the functions $z_j,\ldots,z_J$ in (7) can be written as:

(10)
$$z_j(x_{i1},\ldots,x_{iJ},s) = \beta' w(x_{ij},s) + \delta' t(x_{i1},\ldots,x_{iJ},s) \quad .$$

This shows that any conditional mode-choice model (that satisfies the weak topological assumptions mentioned above) can be written as (7) and (10). For the case of $\delta = 0$, the model expressed in (7) and (10) is the same as that expressed in Table 21. Because the model of (7) and (10) does not entail the IIA property, a test of the hypothesis that $\delta = 0$ is equivalent to a test of the IIA property of condition probabilities implicit in the model of Table 21.

The test is conducted as follows. A model is specified that includes all the variables in Table 21 plus some variables defined as such that attributes of an alternative are allowed to enter the representative utility function of another alternative. The hypothesis that the coefficients of all the extra variables are zero is tested. If the hypothesis of zero coefficients is rejected, then the hypothesis is rejected that the IIA property implicit in the model of Table 21 is correct.

The test of the model of Table 21 was performed against a more general model that includes the variables of Table 21 plus the following variables:

(1) Cost of auto-alone divided by post-tax wage, taking the described value in the bus-with-walk-access and BART-with-walk-access alternatives and zero otherwise.

(2) Cost of bus-with-walk-access divided by post-tax wage, taking the described value in the auto-alone and BART-with-walk-access alternatives and zero otherwise.

(3) Cost of BART-with-walk-access divided by post-tax wage, taking the described value in the auto-alone and bus-with-walk-access alternatives and zero otherwise.

(4) - (6) Variables defined as (1) - (3), respectively, but with "total weighted time" rather than "cost divided by post-tax wage", with total weighted time being the sum of on-vehicle time, walk time multiplied by 2.5, transfer-wait time multiplied by 2.5, and first headway multiplied by 1.25.

The log likelihood at convergence for this more general model is -521.9 . The log likelihood at convergence for the model of Table 21 is -526.4 . Therefore, the test-statistic (page 25) is 9.0 . The critical (.05 significance level) value of chi-squared with six degrees of freedom is 12.6 . The hypothesis that the IIA property implicit in Table 21 is correct is not rejected at the .05 significance level. The hypothesis is rejected only at significance levels exceeding .17 .

The more general model that includes the six extra variables described above does not necessarily represent accurately the dependence structure among the alternatives. As a result, the test of the model of Table 21 against the more general model might not indicate whether or not the IIA property implicit in the model of Table 21 is correct. An even more general model, which necessarily reproduces the dependence structure among the alternatives, is the "saturated" model. The saturated model is one that has the same number of parameters as there are free probabilities in the data set; that is, the saturated model is such that each free probability is essentially a parameter. For a data set consisting of 100 persons each of whom has a choice among three alternatives, the saturated model has 200 parameters (a parameter for each of the two free probabilities for each person is not free because it is determined by the first two and the constraint that probabilities sum to one).

The model of Table 21 has twenty-two parameters. The saturated model on the same data set has 2420 parameters. The test of the MNL assumptions implicit in the model of Table 21 is equivalent, therefore, so that the extra 2398 parameters of the saturated model equal zero.

Because in the saturated model each free probability is essentially a parameter, the log likelihood of the saturated model attains the value of zero (with no repetitions; for a detailed discussion, see McFadden, Tye, and Train (1976)). The log likelihood for the model of Table 21 is -526.4 . Therefore, the test-statistic (page 25) is 1052.8 . The critical (.05 significance level) value of chi-squared with 2398 degrees of freedom is slightly more than 2398. The hypothesis that the MNL assumptions implicit in Table 21 are correct is not rejected at the .05 significance level.

The next set of tests directly employs the IIA property of the conditional model. This property implies that the parameters of a model estimated on a subsample of choices conditioned on choice from a subset of modes are the same, asymptotically, as the parameters estimated on the full sample. A test of the hypothesis that the parameters estimated on the subsample are the same as those estimated on the full sample constitutes, therefore, a test of the IIA property.

The test is conducted as follows. Estimation is performed on the subsample of individuals who chose an alternative within the subset of alternatives relevant for the test. The model parameters are estimated on the subsample and the log likelihood at convergence is calculated; in addition, the log likelihood is calculated on the subsample with parameters restricted to the values of Table 21. Using the test statistic from page 25, the hypothesis that the parameters estimated on the subsample are the same as those of Table 21 is tested. If the hypothesis is rejected, then the hypothesis is rejected that the IIA property implicit in the model of Table 21 is correct.

The results of the tests for various subsets of alternatives are shown in Table 23. The subsets chosen for testing were those that seemed most likely to result in rejection of the hypothesis of equal parameters. For example, models similar to that of Table 21 estimated on a sample taken before BART was providing service greatly overpredict the use of BART-with-walk-access (see Train (1976a)); hence, the subset consisting of all alternatives except BART-with-walk-access seemed particularly relevant for testing based on conditional choice. As Table 23 shows, the hypothesis that the parameters are equal is not rejected at any reasonable significance level for any subset of alternatives.

Several other test were proposed in the section. Many of these tests are not applicable to data sets without repetitions (that is, data sets in which no two persons have exactly the same observed attributes and face alternatives with exactly the same observed attributes). Other of the tests are applicable only when two data sets are available. The only test that could be applied and is not is the test of association applied to the residuals and estimated probabilities of an MNL model. This test, however, is relatively expensive and time-consuming. Furthermore, the test of association was found not to be powerful in detecting violations of the MNL assumptions. Considering these disadvantages, the test was not applied.

TABLE 23 Results of Tests Based on Conditional Choice

Alternatives included in the subset of alternatives	Log likelihood at convergence for subsample choosing an alternative within subset of alternatives	Log likelihood with coefficients restricted to values of Table 4.1	Test Statistic	Degrees of Freedom	Critical (.05 level) value of chi-squared with appropriate degrees of freedom	Significance level above which the hypothesis is rejected
All alternatives except the three BART modes	-408.0	-410.7	5.6	19	30.1	.99
All alternatives except the two bus modes	-381.3	-384.7	6.7	20	31.4	.99
All alternatives except the bus-with-auto-access and BART-with-auto access modes	-411.2	-413.8	5.2	19	30.1	.99
All alternatives except the BART-with-walk access mode	-502.0	-502.9	1.8	21	32.7	.99
All alternatives except carpool	-197.2	-205.3	16.2	20	31.4	.70

Marginal Probabilities of Auto-Ownership Levels

This section presents the estimated model of the conditional probabilities of auto-ownership levels. The following subsections present and discuss the model. Specification tests are applied to the model. The tests concern (1) the structure of the joint model of auto-ownership and work-trip mode-choice, (2) the problem of households with more than one worker, (3) different methods of allowing the price of autos to enter the model, and (4) the possibility that urban and suburban dwellers have different model parameters. Tests of the IIA property of the conditional probabilities implicit in the model discussed below are applied.

The marginal model of auto-ownership

Table 24 presents the model estimated on the sample described in Train (1977). The model is a particular specification of the probability of a household choosing to own a particular number of autos, aggregated over all possible modes for the work-trip. That is, the model specifies the marginal probability of a particular auto-ownership level, the marginality being over work-trip modes. The form of the marginal probability model is that of (4) . The term Y_i is called the representative utility of auto-ownership level i and is specified to be linear-in-parameters:

$$(11) \qquad Y_i = \psi_y(a_i,s) \quad ,$$

where a_i is a vector of attributes of auto-ownership level i (for example, cost); s is a vector of attributes of the household (for example, income); y is a vector-valued function of a_i and s ; and ψ is a vector of parameters to be estimated.

Estimation of ψ was performed by the maximum likelihood method described in McFadden (1973). For each person in the sample, the available choice set was considered to be the following four alternatives:

(1) own no autos;

(2) own one auto;

(3) own two autos;

(4) own three or more autos.

TABLE 24 A Model of Auto Ownership

(Alternative 1- No Autos; Alternative 2 - One Auto; Alternative 3- Two Autos; Alternative 4 - Three or more Autos)

Model: Multinomial Logit, Fitted by the Maximum Likelihood Method

Independent Variable (The variable takes the described value in the alternatives listed in parentheses and zero in non-listed alternatives)	Estimated Coefficient	T-Statistic
Annual auto cost divided by household income (1-4)	-2.26	2.23
Number of persons in household (2)	.573	2.90
Number of persons in household (3)	1.76	7.21
Number of persons in household (4)	2.89	10.1
Proportion of persons in household who are drivers (2)	4.24	4.68
Proportion of persons in household who are drivers (3)	9.77	8.45
Proportion of persons in household who are drivers (4)	16.9	10.8
Accessibility to non-work destinations by transit (1)	.279	0.884
Accessibility to non-work destinations by auto or transit (2)	.664	1.59
Accessibility to non-work destinations by auto or transit (3)	.745	1.73

Table 24, continued

Independent Variable	Estimated Coefficient	T-Statistic
Accessibility to non-work destinations by auto or transit (4)	1.04	2.16
Home location in or near CBD (2=in CBD, l=near CBD, 0 otherwise) (2)	-.829	1.76
Home location in or near CBD (2=in CBD, l=near CBD, 0 otherwise) (3)	-1.30	2.52
Home location in or near CBD (2=in CBD, 1=near CBD, 0 otherwise) (4)	-1.34	2.12
Household income (2)	.0000905	1.84
Household income (3)	.000197	3.76
Household income (4)	.000183	2.92
Dummy for one auto (2)	-2.10	1.36
Dummy for two autos (3)	-12.1	6.05
Dummy for three or more autos (4)	-23.6	8.42
Aggregate work-trip utility	.800	4.66

Table 24, continued

Likelihood ratio index	.3970
Log likelihood at zero	-909.4
Log likelihood at convergence	-548.4
Percent correctly predicted	62.80

Dependent variable is alternative choice (one for chosen alternative, zero otherwise).

Number of households in sample who chose:

No autos	45
One auto	269
Two autos	272
Three or more autos	70
Total Sample Size	656

The first column of Table 24 lists the elements of $y(a_i,s)$ in (11) plus the Q_i term of (6) (called "aggregate work trip utility"). The second and third columns list the estimates and t-statistics, respectively, of the elements of ψ and $(1 - \theta)$.

Many of the variables in Table 24 require explanation and interpretation. The first independent variable is the annual cost of the number of autos designated by the alternative divided by household income. For no autos, annual auto cost is zero. For one or more autos, the annual cost is assumed to be $1000 per auto. (This assumption, following Lerman and Ben-Akiva (1975) is cavalier; the effect of this assumption and alternative methods for treating auto cost are considered below). Auto cost was divided by household income so that the importance of auto cost is allowed to be less for high income households than low income households. The negative estimated coefficient of cost divided by income indicates that increasing the cost of autos decreases the probability of choosing one auto over none, two autos over one or none, and three or more autos over two or less.

The number of persons in a household affects the household's need for autos. The number of persons was allowed to enter separately the representative utility of each of the alternatives. It was expected that increasing household size increases the probability of choosing three or more autos over two or less, the probability of choosing two autos over one or none, and the probability of choosing one auto over none. This expectation is confirmed by the estimated coefficients of the "number of persons" terms, which take the increasing values for the first, second, third, and fourth alternatives of zero (by normalization), .573, 1.76, and 2.89, respectively.

Household size, in itself, does not explain all of the household's needs for autos. The composition of the household in terms of drivers and non-drivers is also a factor affecting auto needs. The proportion of persons in the household who are drivers was allowed to enter separately the representative utility of each of the alternatives. It was expected that increasing the proportion of drivers increases the probability of owning one or more autos over none, the probability of owning two or more autos over none or one, and the probability of owning three or more autos over two or less. This expectation is confirmed by the estimated coefficients of the "proportion of drivers" terms, which take increasing values for the first, second, third, and fourth alternatives.

The ease with which household members can travel to various destinations for shopping, recreation and other activities affects the auto needs of a household.

An "accessibility" term was constructed and allowed to enter the representative utility of each alternative. This accessibility term was constructed as a measure of the ease of travel to non-work destinations for the number of autos designated by the alternative. For the first alternative, the accessibility term was constructed under the assumption that no auto is available and hence transit must be used. For the other alternatives, accessibility was constructed with either auto or transit being possible.

The accessibility terms were constructed with the aid of Ben-Akiva's (1973) model of the choice of non work-trip mode and destination. Ben-Akiva estimated the representative utility, V_{md}, of going to destination d by mode m. This representative utility is a function of cost, time, employment at the destination, and socioeconomic characteristics of the household. Using the logit formula the probability, P_{md}, of going to destination d by mode m is specified. The "accessibility to non-work destinations by auto or transit" is defined as:

$$(12) \qquad \sum_{\substack{m=\text{auto,}\\ \text{transit}}} \sum_{d} P_{md} V_{md} \quad .$$

For the first alternative, no auto is available; hence, all trips must be taken on transit. The "accessibility to non-work destinations by transit" is defined as:

$$(13) \qquad \sum_{d} P_{td} V_{td} \quad ,$$

where t denotes transit.[1] Hence, the accessibility terms are simply weighted sums of the representative utility of going to various destinations by various modes, the weights being the logit probabilities of choosing the particular destination and mode.

The accessibility terms are *ad hoc* in that their definitions are not derived from reasonable behavioral assumptions. If the joint choice of auto-ownership, work-trip mode, and non-work-trip mode and destination satisfies the structured logit assumptions, then the accessibility measures could be defined in a rigorous

[1]Computing these sums over all possible destinations is quite expensive. For the purpose of conserving computer time, fifty representative transit zones were chosen in the San Francisco Bay Area and the sums were calculated for these destinations.

manner, analogous to that of "aggregate work trip utility". That is, "accessibility by auto or transit" would not be defined by (12) but rather by:

$$(14) \qquad \log \sum_{\substack{m=\text{auto,}\\ \text{transit}}} \sum_{d} e^{V_{md}} \quad ,$$

and "accessibility by transit" would be defined by an analogous term.

The problem, however, is that the joint choice of auto-ownership, work-trip mode, and non-work-trip mode and destination does not satisfy the structured logit assumptions. It is not reasonable to assume utility maximization for the joint choice of auto-ownership, work-trip mode and non-work-trip mode and destination. When the choice of non-work-trip mode and destination for a particular day is made the household owns a certain number of autos. If the household were to choose the number of autos to own on that day so as to maximize utility, the number might be more or less than the utility maximizing number for the previous day because the non-work-trip needs and opportunities change daily. Transaction costs prevent the household from owning the utility-maximizing number of autos on each particular day. Rather, the household owns the number of autos that maximizes expected utility, the expectation being over the possible non-work trips and destinations. Though neither (14) nor (12) is derived with reasonable assumptions from expected utility theory, (12) seems preferable because it is a weighted sum of representative utilities and is hence similar in form to the expression for the expected utility of non-work-trips, which is some weighted function of "true" utilities. This reason is weak, but no better criteria exist and calculation of several accessibility measures under various definitions is prohibitively expensive. It is doubtful that the model of auto-ownership would be substantially different if a different definition of accessibility were used.

The sign of the estimated coefficient of "accessibility to non-work destinations by transit (1)" is positive, indicating that, as expected, an increase in the ease of travel to non-work destinations by transit increases the probability of owning no autos (given that accessibility by auto or transit somehow remains constant). The estimated coefficients of "accessibility to non-work destinations by auto or transit" indicate that as the ease of travel by auto increases (and the ease of travel by transit remains constant) the probability of owning one auto over none increases, the probability of owning two autos over one or none increases, and the probability of owning three or more autos over two or less increases. These estimated relations conform to expectations.

The accessibility terms do not account for the difficulty of parking an auto near the home location. This difficulty is generally acute in the central business district (CDB) of a city, less severe in the area surrounding the CBD, and even less severe in the suburbs. The variable "home location in or near CBD" was intended to capture the difficulty of parking. It was allowed to enter the representative utility of each alternative. As expected, the estimated coefficients of these terms decrease for the first, second, and third alternatives. The estimated coefficient for the fourth alternative is not lower than that for the third, but the two coefficients are quite close.

Household income was allowed to enter the representative utility of each of the alternatives. These terms were intended to capture the differential value that households of different income levels place on the unmeasured attributes of auto ownership, such as prestige. The estimated coefficients of these terms indicate that, as expected, high income households are more likely to choose one auto over none than low income households, independent of the effect of income on the value of auto costs. Similarly, the estimated coefficients indicate that high income households are more likely to choose two or more autos over one or more than low income households. The effect of income on the probability of choosing three or more autos over two is estimated to be negative, but negligible in size.

Alternative-specific dummy variables were included for each alternative. The variables capture the common effect of unincluded variables.

"Aggregate work trip utility" is defined in (6). As discussed below, its coefficient is an estimate of $1 - \theta$, where θ is an index of the similarity of the unobserved components of utility of alternatives with the same auto ownership level and different modes. The estimated value of $1 - \theta$ in Table 24 is .80 . Hence, the similarity measure is estimated to be .20 .

When $1 - \theta$ equals one (that is, when θ equals zero) the joint model of auto-ownership and mode-choice is logit rather than structured logit. This hypothesis is tested below.

Discussion and tests of various specification issues

Restricted coefficient of aggregate work trip utility. The hypothesis that the joint model of auto-ownership and work-trip mode-choice is MNL was tested. Table 25 presents the model of auto-ownership estimated under the constraint that the coefficient of aggregate work-trip utility equals one. The log likelihood for this model is -549.0 ; that for the unrestricted model (Table 24) is -548.4 . Using formula (6) , the test statistic is 1.2 . The critical (.05 significance level) value of chi-squared with one degree of freedom is 3.84 . The hypothesis that this coefficient equals one is not rejected at the .05 significance level.

Amemiya (1976) has shown that the above test is not accurate because the test is performed under the assumption that aggregate work-trip utility is known exactly, whereas only an estimate of this variable is actually available. If, however, the hypothesis is not rejected at a given significance level with the above test, then the hypothesis would not be rejected at that significance level with a test that incorporated the variance of the estimate of aggregate work-trip utility. Hence, it is possible to state that the hypothesis that the joint model is logit cannot be rejected at the .05 significance level.

The estimated coefficients of the model of Table 25 indicate that the qualitative discussion of Table 24 applies equally well to Table 25. All further specification tests are conducted on the more general model of Table 24.

The number of workers in a household. The model of auto-ownership contains an aggregate work-trip utility term for only one worker in each household. This limitation was necessary because data were available on only one worker per household for the estimation sample.

The omission of aggregate work-trip utility terms for other workers has two effects: (1) the model of Table 24 is less complete and hence its explanatory power is lower than it conceivably could be, and (2) the coefficients of variables which are correlated with the omitted aggregate work-trip utility terms could be estimated with bias. The latter problem, if it exists, is more serious than the former, but its existence is more easily tested.

TABLE 25 A Model of Auto Ownership, with Constrained Coefficient of Aggregate Work Trip Utility

(Alternative 1- No Autos; Alternative 2 - One Auto; Alternative 3- Two Autos; Alternative 4 - Three or more Autos)

Model: Multinomial Logit, Fitted by the Maximum Likelihood Method

Independent Variable (The variable takes the described value in the alternatives listed in parentheses and zero in non-listed alternatives)	Estimated Coefficient	T-Statistic
Annual auto cost divided by household income (1-4)	-2.30	2.24
Number of persons in household (2)	.582	2.93
Number of persons in household (3)	1.84	7.81
Number of persons in household (4)	3.01	11.1
Proportion of persons in household who are drivers (2)	4.28	4.66
Proportion of persons in household who are drivers (3)	10.1	8.90
Proportion of persons in household who are drivers (4)	17.4	11.5
Accessibility to non-work destinations by transit (1)	.202	0.642
Accessibility to non-work destinations by auto or transit (2)	.670	1.57
Accessibility to non-work destinations by auto or transit (3)	.760	1.73

Table 25, continued

Independent Variable	Estimated Coefficient	T-Statistic
Accessibility to non-work destinations by auto or transit (4)	1.06	2.16
Home location in or near CBD (2=in CBD,l=near CBD, 0 otherwise) (2)	-.888	1.87
Home location in or near CBD (2=in CBD, l=near CBD, 0 otherwise) (3)	-1.35	2.59
Home location in or near CBD (2=in CBD, l=near CBD, 0 otherwise) (4)	-1.40	2.20
Household income (2)	.0000859	1.75
Household income (3)	.000195	3.69
Household income (4)	.000181	2.87
Dummy for one auto (2)	-2.07	1.32
Dummy for two autos (3)	-13.0	6.85
Dummy for three or more autos (4)	-25.2	10.1
Aggregate work trip utility	1	---

Likelihood ratio index	.7470
Log likelihood at zero (and one for coefficient of aggregate work trip utility)	-2170.
Log likelihood at convergence	-549.0
Percent correctly predicted	62.35

Table 25, continued

Dependent variable is alternative choice (one for chosen alternative, zero otherwise).

Number of households in sample who chose:

No autos	45
One auto	269
Two autos	272
Three or more autos	70
Total Sample Size	656

A Chow-test was performed of the hypothesis that households with one worker have the same values of the model parameters as households with more than one worker. A model was estimated in which the two groups of households were allowed to have different parameter values. The log likelihood of this model is -532.3 . The model of Table 24 was estimated under the constraint that the two groups households have equal parameters. The log likelihood of the model of Table 24 is -548.4 . Using the formula of page 25, the test-statistic is 32.2 . The critical (.05 significance level) value of chi-squared with twenty-one degrees of freedom is 32.7 . The hypothesis of equal coefficients is not rejected at the .05 significance level. It seems, therefore, that omitting the extra aggregate work-trip utility terms does not produce substantial bias in the estimated parameters.

Issues concerning the annual cost of owning an auto. In the model of Table 24, the annual cost of owning an auto entered with the specification that: (1) annual cost is $1000 per car, and (2) the importance of cost varies in the population with the inverse of household income. Both of these specifications are arbitrary and require discussion.

The true annual cost of owning an auto depends upon, among other things, the type and vintage of auto. To correctly specify the cost of an auto, however, requires development of a joint model of the choice of number of autos to own, auto type, and vintage. Data limitations prevent such a joint model in the present study.

Given that the annual cost of owning an auto is not specified to vary across households, the figure that is used for annual cost is not crucially important because cost enters the model divided by income. This is because the value of "cost divided by income" multiplied by its estimated coefficient is the same for any constant value of cost. Suppose that the "true" value of cost is $1200 rather than $1000. If the model were reestimated with the true value then the coefficient of "cost divided by income" would be -1.79 rather than the present -2.15.

The implication of this fact for forecasting is that the estimated effect of a percentage change in the cost of autos is the same for any (constant) value of cost used in estimation. For example the change in the representative utility of owning one auto due to a ten percent increase in car cost is the same for the model of Table 24 and a similar model with cost set equal to $1200; in both cases the effect (for a household with $10,000 income) would be a decrease in representative utility of .0215 . Because representative utility in both cases changes by the same amount, the change in probability of choosing one auto is also the same in both cases.

The effect of a multiplicative change in the annual cost of owning an auto can be forecast without correctly specifying true cost. It must be noted, however, that the effect of an additive change cannot be forecast without correctly specifying true cost. This limitation of the model seems minor.

The annual cost of an auto enters the model divided by household income under the assumptions that the importance of cost decreases with income. Other methods for entering auto cost are possible. The importance of auto cost could be allowed to decrease with household income by entering a term defined as the log of remaining income, where remaining income is the household income minus the cost of the number of autos designated by the alternative. A similar approach was used by Lerman and Ben-Akiva (1975). This approach, however, does not seem as desirable as entering cost divided by income. First, with the log of remaining income, the correct specification of auto cost is quite important. The effect of a multiplicative change in cost could not be forecast accurately without correctly specifying cost in calculating the log of remaining income. With price divided by income, correct specification of cost is not necessary for such forecasting. Second, a model estimated with the log of remaining income does not fit the data as well as the model of Table 24. A model similar to that of Table 24 was estimated with "log of remaining income" replacing "cost divided by income"; the log likelihood of the model is -548.9 whereas that of the model of Table 24 is -548.4 . For these two reasons, it seems preferable to use "cost divided by income".

Dividing cost by income assumes that the importance of cost varies with the inverse of income. It seems reasonable to assume, however, that the importance of cost is greater for a household with three members than a household with the same income but with only one member. This relation could be due to the fact that three people require more income to attain the same standard of living as one person. What seems desirable, therefore, is to divide cost by some figure that reflects the standard of living that the household attains given its income and the number of members. This figure is probably not simply income per person, because two people do not require twice the income of one person to attain the same standard of living. Rather, the figure is some function of income and number of persons.

To explore these ideas, models similar to that of Table 24 were estimated with "cost divided by income" replaced with the following function, called "cost divided by income per weighted person":

$$C - \frac{I}{1 + w(N-1)} \quad ,$$

where C is the cost of an auto; I is household income; N is the number of persons in the household, and w is the weight attached to the persons in the household other than the first one.

Table 26 presents the log likelihood of the models estimated with the "cost divided by income per weighted person" term and with various values of w . For w = 1 , the term becomes simply "cost divided by income per person". For w = 0, the term becomes "cost divided by income"; that is, the case of w = 0 is the model of Table 24.

As Table 26 shows, the log likelihood increases as the weight decreases. This indicates that dividing cost by income in the model of Table 24 is preferable to dividing by income per weighted person (for any positive weight).

Two additional, yet related, issues concerning the cost of auto-ownership require discussion. First, the model of Table 24 does not confront the difference between fixed and variable costs of auto-ownership. It is perhaps the case that households consider only fixed costs when deciding how many autos to own and that variable costs only affect trip making decisions. Determining which costs are fixed and which are variable is, however, very difficult. Perhaps the annual depreciation of an auto is the only fixed cost. Depreciation, however, depends to an extent on the number of miles driven each year; hence, even part of the depreciation of an auto can be considered a variable cost. Rather than solve the problem of fixed versus variable costs, the model of Table 24 ignores the issue.

Second, the possibility of a household leasing an auto is not considered in the model of Table 24. Because leasing is perhaps simply a different method of financing auto ownership, households who lease an auto should perhaps be considered to own an auto. Households in the estimation sample were not asked about the number of autos they leased; consequently, the issue of leasing versus buying an auto could not be explored.

TABLE 26 Log Likelihoods of Models Estimated with "Cost Divided by Income per Weighted Person" and Various Weights

Value of Weight	Log likelihood
1	-549.4
.5	-549.0
.25	-548.6
0	-548.4

"Cost divided by income per weighted person" is defined as

$$C \div \frac{I}{1 + w(N-1)}$$

where C is the cost of an auto;
I is household income;
N is the number of persons in the household; and
w is the weight attached to the persons in the household other than the first one.

Differences in parameters for urban and suburban dwellers. The reader will recall that previously we found that the values of time for the work-trip are significantly different for urban dwellers and suburban dwellers. This conclusion indicates that perhaps the parameters of the auto-ownership model are different for urban and suburban dwellers.

A Chow-test was performed of the hypothesis that urban households have the same values of the model parameters as suburban households. A model was estimated in which the two groups of households were allowed to have different parameter values. The log likelihood of this model is -534.4 . The model of Table 24 was estimated under the constraint that the two groups of households have equal parameters. The log likelihood of the model of Table 24 is -548.4 . Using the formula of page 25 the test statistic is 28.0 . The critical (.05 significance level) value of chi-squared with twenty-one degrees of freedom is 32.7 . The hypothesis of equal coefficients is not rejected at the .05 significance level. The hypothesis is rejected only at significance levels exceeding .14 .

Diagnostic tests of the IIA property of marginal probabilities. The model of Table 24 seems particularly likely to violate the IIA property of marginal probabilities. For example, it seems doubtful that the probability of choosing two autos over none remains the same whether or not the possibility of owning one auto exists. If the probability is not constant, then the IIA property does not hold for the marginal probabilities. The question of whether this and other problems are severe enough to reject the IIA property in modeling auto ownership is explored by applying diagnostic tests.

The diagnostic tests applied to the auto ownership model are analogous to those applied to the work-trip mode-choice model described previously. Consequently, the results of the test are presented with little explanatory discussion.

The first test of the model of Table 24 is against a more general model in which attributes of an alternative are allowed to enter the representative utility function of another alternative. The more general model includes the variables of Table 24 plus the following variables:

(1) Cost of one auto divided by household income, taking the described value in the fourth alternative and zero otherwise.

(2) Cost of two autos divided by household income, taking the described value in the first alternative and zero otherwise.

(3) Cost of three autos divided by household income, taking the described value in the third alternative and zero otherwise.

The log likelihood for this more general model is -546.7 . That for the model of Table 24 is -548.4 . Using the formula of page 25, the test-statistic is 3.4 . The critical (.05 significance level) value of chi-squared with three degrees of freedom is 7.81 . The hypothesis that the coefficients of the three extra terms are zero is not rejected at the .05 significance level. The hypothesis is rejected only at significance levels exceeding .33 . Consequently, the hypothesis that the IIA property implicit in Table 24 is correct and is not rejected.

The second test is against the saturated model. The log likelihood of the saturated model attains a value of zero. The log likelihood for the model of Table 24 is -548.4 . Therefore, the test-statistic (using the formula of page 25) is 1096.8. The critical (.05 significance level) value of chi-squared with 1947 degrees of freedom is slightly more than 1947. The hypothesis that the IIA property implicit in the model of Table 24 is correct and is not rejected at the .05 significance level.

The next set of tests are based on the fact that with the IIA property, parameters estimated on restricted choice sets are the same, asymptotically, as the parameters estimated on the full choice set. Estimation is performed in the subsample of individuals who chose an alternative within the subset of alternatives relevant for the test. The hypothesis that the estimates obtained on the subsample are the same as those obtained on the full sample is tested.

The results of the tests for various subsets of alternatives are shown in Table 27. The hypotheses of equal parameters are not rejected at reasonable significance levels for any subset of alternatives. Consequently, the hypothesis is not rejected that the IIA property of marginal probabilities implicit in the model of Table 24 are correct.

TABLE 27 Results of Tests Based on Conditional Choice

Alternatives included in the subset of alternatives	Log likelihood at convergence for subsample choosing an alternative within subset of alternatives	Log likelihood with coefficients restricted to values of Table 25	Test Statistic	Degrees of Freedom	Critical (.05 level) value of chi-squared with appropriate degrees of freedom	Significance level above which the hypothesis is rejected
All alternatives except the alternative of owning no auto	-454.6	-465.6	22.0	14	23.7	.08
All alternatives except the alternative of owning one auto	-169.1	-179.2	21.2	15	25.0	.13
All alternatives except owning two autos	-177.3	-187.4	10.2	15	25.0	.80
All alternatives except owning three or more autos	-384.2	-384.6	.8	15	25.0	.99

Summary of the Model

The estimated joint model of auto-ownership and work-trip mode-choice conforms to expectations concerning its parameters and passed a number of tests of the restrictions implied by the specification of representative utility. The model also passed the tests of the restrictions implicit in the structured logit form: the IIA property of the conditional probabilities and the IIA property of marginal probabilities. Consequently, it seems to be usable in policy analysis of issues affecting auto ownership and/or work-trip mode-choice.