Transpn Res., Vol. 12, pp. 167–174.
© Pergamon Press Ltd., 1978. Printed in Great Britain

0041-1647/78/0601-0167/$02.00/0

A VALIDATION TEST OF A DISAGGREGATE MODE CHOICE MODEL†

KENNETH TRAIN
Cambridge Systematics, Inc./West, Berkeley, CA 94704, U.S.A.

(*Received 7 March 1977; in revised form 5 October 1977*)

Abstract—A model of work trip mode choice was developed on a sample of workers taken before Bay Area Rapid Transit (BART) opened for service. Validation tests of the model were performed on a sample of workers taken after BART service began. Two validation methods were used: (1) the actual mode shares in the post-BART sample were compared to the mode shares predicted by the models estimated on the pre-BART sample, and (2) the parameters of models estimated on the post-BART sample were compared with the parameters of the models estimated pre-BART. Three possible reasons were explored for the differences in actual and predicted shares and in the pre- and post-BART model parameters: (1) failure of the independence from irrelevant alternatives (IIA) property of the multinomial logit model, (2) non-genericity in the attributes of one of the alternatives, and (3) incorrect data for walk times. It was found that non-genericity and incorrect data contributed substantially to the mispredictions, while failure of the IIA property contributed less. The present study concerns only one model and one transportation environment. The results of this test, however, can be viewed along with the results of other validation studies to obtain a sense of the predictive ability of disaggregate mode choice models.

1. INTRODUCTION

In 1973, the Bay Area Rapid Transit (BART) system opened for service in the San Francisco Bay Area. The introduction of this new transit mode offers an exceptional opportunity for testing the validity of disaggregate travel demand models. Models developed before BART was built can be used to predict behavior after BART opened; predicted behavior can then be compared with actual behavior for an indication of how well the models actually represent behavior.

This paper is concerned with evaluating one type of disaggregate travel demand model: mode choice models for work trips. Models were developed on a sample of workers taken before BART service was available. The models were evaluated in two ways on a sample of workers taken after BART opened. First, actual mode shares in the post-BART sample were compared with the mode shares which the pre-BART models predicted. Second, the parameters of models estimated on the post-BART sample were compared with the parameters of the pre-BART model. Sections 2–4 present and evaluate a particular model which was estimated on the pre-BART sample. Specifically, Section 2 presents and discusses this pre-BART model; Section 3 test the model in the two ways described above; and Section 4 analyzes several possible reasons for the differences between predicted and actual shares and between pre- and post-BART model parameters.

This paper is an abbreviated version of Train (1976c). The original paper contains more detailed discussions of the forecasting ability of models, compares the forecasting ability of models of various levels of complexity, and points out the difficulties in obtaining "reasonable" forecasts from the models.

†This research was supported in part by National Science Foundation grant APR74-20392, Research Applied to National Needs Program, National Science Foundation, to the University of California, Berkeley.

2. A PRE-BART MODEL

The pre-BART model upon which evaluation tests are performed in the following section is a multinomial logit (MNL) model of individual choice probabilities. The model expresses the probability that a person with certain observed socioeconomic characteristics and facing a choice among several alternatives will choose a particular alternative. The function is expressed as:

$$P_n(i/C_n) = e^{\beta' z(x_n^i, s_n)} / \sum_{j \in C_n} e^{\beta' z(x_n^j, s_n)} \qquad (1)$$

where C_n is the set of alternatives among which person n may choose; $P_n(i/C_n)$ is the probability that person n will choose alternative $i \in C_n$; x_n^i is a vector of observed characteristics of alternative i for person n; s_n is a vector of observed characteristics of person n; z is a vector-valued function of x and s; and β is a vector of parameters to be estimated.

In the MNL model the ratio of the probabilities of choosing any two alternatives is independent of the availability or attributes of other alternatives. This property is called the independence from irrelevant alternatives (IIA) property and can be demonstrated as follows. Consider the ratio of the probability of person n choosing alternative i to that of choosing alternative k, given that set C_n of alternatives is available. It can be seen from (1) that:

$$\frac{P_n(i/C_n)}{P_n(k/C_n)} = \frac{e^{\beta' z(x_n^i, s_n)}}{e^{\beta' z(x_n^k, s_n)}}.$$

This ratio is constant for any C_n which contains i and k (including, of course, the set containing only i and k)

and any attributes of alternatives (except i and k) in C_n. The IIA property greatly facilitates estimation and forecasting, particularly in the situation of a new alternative being introduced. Estimation of the choice model (that is, estimation of β) can be performed on the alternatives available before the new alternative is introduced and forecasting can proceed with an expanded choice set consisting of the old alternatives plus the new one. Therefore, forecasting the demand for a new alternative can be accomplished before the alternative is actually introduced. The IIA property has the disadvantage, however, in that it imposes restrictions on the structure of choice probabilities. In applications in which the ratio of true probabilities is not independent of the availability or attributes of other alternatives, the MNL model is inappropriate.

The IIA property is exploited in the present estimation and forecasting. MNL models were calibrated on a sample of workers living in the San Francisco Bay Area before BART was introduced. The alternative modes which were considered available for the work trip were: auto alone, carpool, bus with walk access to bus, and bus with auto access to bus. Forecasting was performed on a sample of people taken after BART was introduced, with the choice set expanded to include the alternatives of BART with walk access, BART with auto access and BART with bus access. The model which was estimated on the pre-BART sample and is subject to evaluation tests in Section 3 is given in Table 1. The first column lists the elements of $z(x_n^i, s_n)$ in eqn (1); the second and third columns list the estimates and t-statistics, respectively, of the elements of β in eqn (1). Estimation was performed by the maximum likelihood method described in McFadden (1973).

Most of the variables are self-explanatory and their coefficients readily interpretable. For instance, the coefficient of walk time being negative indicates that when time spent walking for a particular mode increases, the probability of that mode being chosen decreases, all other things held constant. Since the ratio of the walk time coefficient to the cost divided by wage coefficient is 2.43, the estimated value of time is 243% of wage. Some of the variables, however, require explanation. The variable "cost divided by wage" was included rather than simply "cost" so that people's values of time could vary with their wage rates. The variable is expressed in units of time (since cents divided by cents per minute is equal to minutes) and can be considered to represent the number of minutes which the person must work in order to earn the cost of his transportation. Rather than expressing cost in terms of time equivalents, the time variables could have been expressed in money equivalents by multiplying the times by the person's wage. Train and McFadden (1976) explores the implications of these two approaches in terms of utility maximization theory.

The headway of a bus line is the number of minutes between bus arrivals at a particular stop for that bus line. Initial wait time is often calculated as half of the headway for the first bus. With this approximation, the value of initial wait time is twice the value of initial headways. The family income variables can be understood with reference to Fig. 1. Family income was segmented so that the effect of an incremental dollar of income on the probability of choosing auto alone could be different at different income levels. For family income up to \$7500, the positive sign of the first income variable indicates that an extra dollar of income increases $\beta' z(x_n^i, s_n)$ of eqn (1) (for i = auto alone) and hence the probability of choosing auto alone. This is represented by the first line segment in Fig. 1 having a positive slope. Negative coefficients for the second and third income variables indicate that for family income between \$7500 and \$15,000, an extra dollar decreases $\beta' z(x_n^i, s_n)$ for i = auto alone and hence the probability of choosing auto alone. No fourth income variable was included, indicating that, for incomes above \$15,000, an incremental dollar does not affect the probability of choosing auto alone. Note, however, that none of the income terms have t-statistics exceeding one, and so the hypothesis cannot be rejected that there is no relation between income and the probability of choosing auto.† The income terms were included not because the relation in Fig. 1 is considered "true", but because their omission changed the other parameters considerably.

The model of Table 1 was developed after extensive testing of specifications. Some of the tests are described in Train (1976a) and Train and McFadden (1976). Variables with coefficients which are not significantly different from zero were included since omitting them changed the values of other estimated coefficients. By classical statistics, including these variables does not produce bias. The values of time and headways, particularly auto on-vehicle time, are higher than expected. These higher values result from allowing auto and transit times to have different coefficients. When auto and transit times are constrained to have the same value, the estimated values of time and headways are lower. The hypothesis that the coefficients of auto and transit on-vehicle time are equal can be rejected (at the 0.05 confidence level), indicating that the coefficients should not be constrained. The value of auto on-vehicle time is estimated to be higher than that of transit on-vehicle

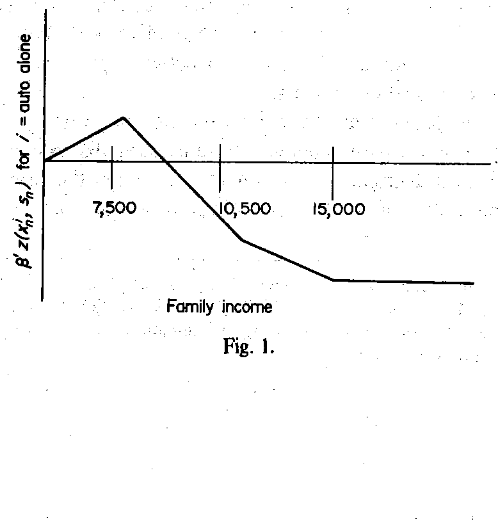

Fig. 1.

†Rather than include these income terms, an alternative approach is to segment the sample into income groups and estimate a separate model for each group. This procedure was not followed since the sample was too small to allow for reasonable estimation on subsamples.

Table 1. Work trip mode choice model, estimated pre-BART. (Mode 1—Auto Alone; Mode 2—Bus, Walk Access; Mode 3—Bus, Auto Access; Mode 4—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 (1-4)	-.0284	4.31
Auto on-vehicle time, in minutes (1,3,4)	-.0644	5.65
Transit on-vehicle time, in minutes (2,3)	-.0259	2.94
Walk time, in minutes (2,3)	-.0689	5.28
Transfer wait time, in minutes (2,3)	-.0538	2.30
Number of transfers (2,3)	-.105	0.776
Headway of first bus, in minutes (2,3)	-.0318	3.18
Family income with ceiling of $7,500, in $ per year (1)	.00000454	0.0511
Family income minus $7,500 with floor of $0 and ceiling of $3,000, in $ per year (1)	-.0000572	0.430
Family income minus $10,500 with floor of $0 and ceiling of $5,000, in $ per year (1)	-.0000543	0.907
Number of persons in household who can drive (1)	1.02	4.81
Number of persons in household who can drive (3)	.990	3.29
Number of persons in household who can drive (4)	.872	4.25
Dummy if person is head of household (1)	.627	3.37
Employment density at work location (1)	-.00160	2.27
Home location in or near CBD (2=in CBD, 1=near CBD, 0 otherwise) (1)	-.502	4.18
Autos per driver with a ceiling of one (1)	5.00	9.65
Autos per driver with a ceiling of one (3)	2.33	2.74
Autos per driver with a ceiling of one (4)	2.38	5.28
Auto alone alternative dummy (1)	-5.26	5.93
Bus with auto access dummy (3)	-5.49	5.33
Carpool alternative dummy (4)	-3.84	6.36

Likelihood ratio index around zero	.4426
Likelihood ratio index around market shares	.2961
Log likelihood at zero	-1069.0
Log likelihood at market shares	-846.4
Log likelihood at convergence	-595.8
Percent correctly predicted	67.83

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

Auto on-vehicle time	227	(3.20)
Transit on-vehicle time	91	(2.43)
Walk time	243	(3.10)
Transfer wait time	190	(2.01)

Value of initial headways as a percent of wage: 112 (2.49)

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	429
Bus with walk access	134
Bus with auto access	30
Carpool	178
Total sample size	771

time. This result was explored in Train (1976a) and the explanation was given that, while autos are more comfortable than transit, the difficulty of driving an auto during rush hour congestion makes auto time more onerous than transit time. While the result seems to be contrary to popular belief about the disutility of transit travel, this belief is perhaps based upon a consideration of all transit time, including walk and wait time, rather than simply on-vehicle time. Furthermore, the result relates only to the value of a marginal unit of on-vehicle time. Many of the attributes of transit use which are considered onerous, such as lack of comfort and the possibility of crime, do not vary substantially with length of time spent on-vehicle and are captured by the alternative specific dummy variables rather than the on-vehicle time coefficient.

3. EVALUATING THE PRE-BART MODEL

The first evaluation method is to compare predicted with actual mode shares in the post-BART sample. In order to use the model of Table 1 for predicting post-BART shares, a value for each independent variable in the model must be created for each BART alternative: BART with walk access, BART with bus access, and BART with auto access. For the transportation system variables, such as on-vehicle and walk times, the BART attributes were simply calculated. The socioeconomic and alternative-specific variables for the BART alternatives were created as follows: the value for the BART with auto access alternative was set equal to the value for the bus with auto access alternative, and the values for the other two BART alternatives were set equal to the value for the bus with walk access alternative.

In predicting post-BART demand, the auto alone alternative was considered unavailable to a person if no autos were available to his household. Any of the transit alternatives was considered unavailable to a person if going to work by that alternative entailed more than three transfers either to or from work, a total weighted travel time of more than four hours either to or from work, or other excessive attributes.

Table 2 presents the actual and predicted shares, with predictions based on the model of Table 1. The actual share for a particular alternative is the share of people in the post-BART sample who actually chose the alternative. The predicted share is the share of the post-BART sample which the model of Table 1 predicts will choose the alternative. This predicted share is defined as:

$$S_i = \frac{1}{N} \sum_n P_n(i/C_n)$$

where $P_n(i/C_n)$ is expressed as eqn (1), and N is the sample size.

Table 2. Actual and predicted shares, with predictions based on model of Table 1

	Actual Share (%)	Predicted Share (%)
(1) Auto alone	59.53	53.19
(2) Bus/walk	10.71	11.37
(3) Bus/auto	1.42	2.20
(4) BART/walk	0.63	7.53
(5) BART/bus	0.94	0.82
(6) BART/auto	5.20	3.94
(7) Carpool	21.57	20.95

Root Sum of Squared Error: 9.53 (n = 635)

A comparison of the actual and predicted shares indicates that the pre-BART model: (i) underpredicts use of auto alone; (ii) overpredicts use of both the bus alternatives; (iii) greatly overpredicts the use of BART with walk access; (iv) underpreducts the use of the other two BART alternatives; (v) underpredicts the use of carpool. Summing the columns and rows of Table 2 over the five transit modes gives an actual transit share of 18.9% and predicted share of 25.9%. That is the predicted transit share is 37% larger than the actual transit share.

The second method for evaluating the pre-BART model of Table 1 is to estimate a model with the same specification on the post-BART sample. If the estimates and specification of Table 1 are accurate, then the estimates obtained on the post-BART sample should be similar. Comparison of the pre- and post-BART estimates not only provides a test of the accuracy of the pre-BART mode; the comparison can also give indications as to the problems in the pre-BART model which give rise to the discrepancies between predictive and actual post-BART shares.

Table 3 presents a model estimated on the post-BART sample with the same specification as the model of Table 1. The estimates are fairly similar. The differences between the pre- and post-BART estimates which seem most relevant to the forecasting errors of Table 2 are (a) the value of walk time is much higher post-BART than pre-BART, and (b) the BART with walk access dummy is significantly less than zero. Since only the transit modes entail walk time, the different estimates for the value of walk time pre- and post-BART could be related to the overprediction of transit. The significantly negative estimate of the BART with walk access dummy could be related to the large overprediction of the BART with walk access alternative. In forecasting BART usage, the BART with walk access alternative was considered to have a value of zero for its dummy, which is the value for the bus with walk access dummy in the pre-BART model (see the discussion above). This procedure is equivalent to assuming that the effect of unincluded variables on demand for BART with walk access is the same as that for bus with walk access. The significantly negative estimate for the BART with walk access dummy indicates that this assumption is not valid.

The following Section explores possible reasons for the mispredictions of Table 2.

4. REASONS FOR MISPREDICTIONS

(a) *Failure of IIA.*

If the five transit alternatives are not actually independent, then the MNL model would be expected to

Table 3. Work trip mode choice model, estimated post-BART, with non-generic auto and transit on-vehicle time. (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)	-.0266	3.92
Auto on-vehicle time, in minutes (1,3,6,7)	-.0473	3.48
Transit on-vehicle time, in minutes (2-5)	-.0197	2.03
Walk time, in minutes (2-6)	-.0900	3.36
Transfer wait time, in minutes (2-6)	-.0438	1.81
Number of transfers (2-6)	-.120	0.856
Headway of first transit carrier, in minutes	-.0290	2.60
Family income with ceiling of $7,500, in $ per year (1)	-.000289	1.78
Family income minus $7,500 with floor of $0 and ceiling of $3,000, in $ per year (1)	.0000522	0.364
Family income minus $10,500 with floor of $0 and ceiling of $5,000, in $ per year (1)	-.0000419	0.738
Number of persons in household who can drive (1)	1.48	5.26
Number of persons in household who can drive (3,6)	1.65	5.16
Number of persons in household who can drive (7)	1.28	4.85
Dummy if person is head of household (1)	.668	3.19
Employment density at work location (1)	-.00164	3.45
Home location in or near CBD (2=in CBD, 1=near CBD, 0 otherwise) (1)	.1546	0.835
Autos per driver with a ceiling of one (1)	4.79	3.70
Autos per driver with a ceiling of one (3,6)	3.63	4.81
Autos per driver with a ceiling of one (7)	3.26	3.19
Autos alone alternative dummy (1)	-4.18	2.82
Bus with auto access dummy (3)	-8.24	6.67
BART with walk access dummy (4)	-2.28	3.36
BART with bus access dummy (5)	-.473	0.708
BART with auto access dummy (6)	-7.30	5.93
Carpool alternative dummy (7)	-5.31	5.56

Likelihood ratio index	.4599
Log likelihood at zero	-964.4
Log likelihood at convergence	-520.9
Percent correctly predicted	67.24

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

Auto on-vehicle time	178 (2.53)
Transit on-vehicle time	74 (1.84)
Walk time	338 (2.46)
Transfer wait time	165 (1.65)

Value of initial headways as a percent of wage: 109 (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

overpredict transit use (see Charles River Associates (1976)). Since transit use was indeed overpredicted, it is possible that failure of IIA is the cause. To explore this possibility, two non-MNL models were estimated on the pre-BART sample and used for forecasting post-BART behavior. Neither of these models entails the IIA property. The two non-MNL models are called the Maximum model and the Log-sum model† Both of the models assume a two step procedure for a person deciding which mode to choose: first, a choice among auto alone, transit and carpool is made; second, if transit is chosen in the first step, then a choice is made among the transit alternatives (bus with walk access, bus with auto access, etc.). In both the Maximum and Log-sum models, the first choice is specified to be an MNL model of choice among auto alone, transit, and carpool, and the second choice is specified to be an MNL model of transit mode choice. The models differ in how the attributes of transit in the first choice are calculated. In the Maximum model, the transit attributes faced by a person in the first choice are considered to be the attributes of the transit mode which the person has the highest probability of choosing in the second choice. In the Log-sum model the transit attributes in the first choice are calculated as a function of the attributes of all the transit modes. The function is:

$$x_n{}^t = -\log \sum_{i \in T_n} e^{-x_n{}^i}$$

where $x_n{}^t$ is the calculated transit attributes in the first choice, $x_n{}^i$ is the attribute of transit mode i; and T_n is the set of all transit modes available to person n.

The two models are expressed symbolically as follows. The first choice is an MNL model of choice among auto alone, transit and carpool:

$$P_n(i|A_n) = e^{\beta' z(x_n{}^i, s_n)} / \sum_{j \in A_n} e^{\beta' z(x_n{}^j, s_n)}$$

where A_n is the subset of the set {auto alone, transit, carpool} which is available to person n. The second choice is an MNL model of choice among transit modes:

$$P_n(l|T_n) = e^{\alpha' z(x_n{}^l, s_n)} / \sum_{k \in T_n} e^{\alpha' z(x_n{}^k, s_n)}$$

†See McFadden (1974) concerning the Maximum model and Templeton and Talbitie (1976, eqn (2)) concerning the Log-sum model.

where T_n is the set of transit modes available to person n.

The difference between the two models is in the calculation of the variable $x_n{}^t$ where t denotes "transit" in the set A_n. For the Maximum model:

$$x_n{}^t = \sum_{i \in T_n} x_n{}^i s_n{}^i$$

where

$$s_n{}^i = \begin{cases} 1 \text{ if } \alpha' z(x_n{}^i, s_n) \geq \alpha' z(x_n{}^j, s_n) \quad \forall j \in T_n \\ 0 \text{ otherwise.} \end{cases}$$

For the Log-sum model:

$$x_n{}^t = -\log \sum_{i \in T_n} e^{-x_n{}^i}.$$

Since neither of these models entails the property of IIA, each should predict better than the MNL model if failure of IIA is the reason for the mispredictions of the MNL model.

Table 4 presents the predicted shares for the choice among auto alone, transit, and carpool for both the Maximum and the Log-sum models. The predictions are somewhat better than those of the MNL model, but transit use is still being greatly overpredicted. The Maximum and Log-sum models overpredict transit use by 35 and 22% respectively, whereas the MNL model overpredicts transit by 37%. Since the non-MNL models greatly overpredict transit use, it seems that failure of IIA is not a primary cause of the overprediction of transit by the MNL model. It is possible, however, that failure of IIA contributes somewhat to the overprediction.

(b) *Non-genericity of attributes of* BART *with walk access*

If BART with walk access exhibits some important attributes which none of the pre-BART modes exhibits, then the value of these attributes cannot be estimated with pre-BART data. Similarly, if some attributes of BART with walk access (such as walk time to BART) are valued differently than similar attributes of bus, then the value of the BART attributes cannot be estimated with pre-BART data. The overprediction by the pre-BART model of BART with walk access might result from the existence of either of these two types of non-genericity. If non-genericity exists for the BART with walk access

Table 4. Predictions based on non-logit models

	Actual Share	Predicted Share Based on Maximum Model	Predicted Share Based on Log-sum Model
Auto alone	59.15	53.44	54.34
Transit	19.56	26.36	23.89
Carpool	21.28	20.21	21.76
Root Sum of Squared Error		8.94	6.49
(n = 639)			

alternative, then it should appear in models estimated on the post-BART sample. To determine if significant non-genericities exist, tests were performed on post-BART models.

Several tests on the post-BART models attempted to determine whether any attribute which is similar for bus and BART (such as on-vehicle time) is valued differently for the two modes. Since BART trains are generally more comfortable than buses, the value of on-vehicle time is perhaps lower for BART than bus. Similarly, since waiting for BART trains is generally done indoors, perhaps the value of initial headways and transfer wait time are lower for BART than bus. Walk time to BART is perhaps considered more onerous than walks to bus since many BART stations are surrounded by parking facilities which are less pleasant to walk through than walking on sidewalks. Tests on these four attributes were performed and no significant (at the 0.05 confidence level) non-genericities were found. These results are detailed in Train (1976b). These tests indicate, therefore, that different valuations for similar attributes of bus and BART do not explain the large overprediction for BART with walk access.

The existence of BART attributes which do not exist for any pre-BART modes can be detected by examining the coefficients of the alternative specific dummy variables in the post-BART models. The coefficients of the dummy variables reflect the "average" or common effect on demand of all the attributes which are not included in the model. For forecasting purposes, it was assumed that the common effect of the unincluded variables of the BART with walk access alternative is the same as that of the bus with walk access alternative. The coefficient of the bus with walk access alternative is zero (by normalization). If no BART attributes exist which are different from all pre-BART attributes, then the estimated coefficient of BART with walk access is expected to be close to zero. As the post-BART model of Table 3 shows, the BART with walk access alternative dummy has an estimated coefficient which is significantly less than zero. This indicates that the unincluded attributes of BART with walk access affect demand for that alternative significantly differently than the unincluded attributes of bus with walk access. This non-genericity contributes to the overprediction by pre-BART models of the BART with walk access mode. If non-genericity exists for one alternative, then the pre-BART model can be used to predict the shares of the other alternatives conditional upon the non-generic alternative *not* being chosen. (The consistency of such conditional prediction is a result of the IIA property.) These predicted shares can be compared with the actual shares to obtain an indication of how well the model predicts in the absence of non-genericity.

Table 5 presents the predicted and actual shares conditional upon BART with walk access not being chosen. The predictional shares are calculated the same as those in Table 2, but the four people who actually chose BART with walk access are removed from the sample and the BART with walk access alternative is removed from each person's choice set. The predicted shares in Table 5 are much closer to the actual shares than those of Table 2. However, the auto alone alternative is still being underpredicted and the bus alternative overpredicted. The possibility that bad data for walk times, especially in the bus alternative, is causing these mispredictions is explored below.

(c) *Incorrect walk time data*

The attributes of the transit alternatives were calculated by computer programs which simulate the Bay Area transit system for particular years. Simulated systems existed (that is, they had been previously constructed for an earlier study) for the years 1965 and 1980. The 1965 system had been constructed to represent the system as it actually existed in 1965. The 1980 system had been constructed to represent the system as transportation planners expected it to exist in 1980. This system included anticipated transit improvements, including the addition of many new bus lines. The simulated systems for the years of interest (1972 for pre-BART and 1975 for post-BART) were constructed as follows. The 1972 pre-BART simulated system was obtained by adjusting the 1965 system to account for the few changes that occurred during the intervening year. Complete information on the system status in 1972 was available at the time of adjustment. The 1975 post-BART attributes were obtained, however, by adjusting the 1980 system. In this adjustment, the extra bus lines that were expected in 1980 but not existing in 1975 were removed,

Table 5. Predictions conditional upon BART with walk access not being chosen

	Actual Share	Predicted Share
Auto Alone	59.90	55.84
Bus/Walk	10.78	12.51
Bus/Auto	1.426	2.411
BART/Bus	0.951	1.053
BART/Auto	5.230	5.286
Carpool	21.71	22.89
Root Sum of Squared Error		4.67
(n = 631)		

but the walk times to transit were not adjusted. The walk times should have been adjusted, since decreasing the number of bus lines increases, on average, the walk times to transit. The ratio of the mean walk time for the bus with walk access alternative in the pre-BART sample to that in the post-BART sample is 1.78. Little change in the bus system has occurred during the years between the pre- and post-BART samples, and it is doubtful that the difference in the means is a result of the sampling procedure. Rather, it seems that the post-BART walk times were calculated to be too short. Calculated walk times being unrealistically low could explain the higher estimated value of walk time in the post-BART model than the pre-BART model. If the walk time variable is biased downward, then its coefficient, and hence its value of time, would be biased upward. The value of time in the pre-BART model is a consistent estimator of the true value, while that of the post-BART model is biased upward.

The unrealistically low walk times could also explain the mispredictions of Table 5 (i.e. mispredictions which do not result from non-genericity in the BART with walk access alternative). If walk times for the bus are biased downward, more people would be predicted to choose the bus alternative than actually do. BART walk times were calculated relatively accurately since transit planners were fairly sure of the number and placement of BART stations. As a result, the predicted share for the BART with auto access alternative would be expected to be fairly precise. Since walk times for BART with bus access are a combination of walk times to BART and walk times to buses, the predicted share for this alternative would be expected to be too high, though the overprediction would not be expected to be as large as that for the bus alternatives. As Table 5 shows, the mispredictions which would be expected from downward biased walk times for buses actually occur.

5. CONCLUSIONS AND QUALIFICATIONS

The validation test produced several findings concerning the particular model which was used in the test. It was found that the model overpredicted transit usage, and in particular, greatly overpredicted the usage of BART with walk access. Failure of the IIA property was found to contribute only slightly to the overprediction. However, erroneous data for walk times and nongenericity of the BART with walk access mode were found to contribute considerably to the overprediction of transit usage.

It would be desirable to correct these errors and determine how well the corrected model predicts. However, doing so was infeasible for this study. Non-genericity of the BART with walk access mode cannot be corrected with presently available models. Until a method is developed for determining the value of non-generic attributes exhibited by modes for which forecasts are required but which are not available at the time of model estimation, such non-genericity cannot be incorporated into the forecasts. The inaccurate walk time data cannot be corrected without an expenditure of time and money beyond the ability of this research project.

Several issues must be considered before any general conclusions can be drawn concerning the predictive power of disaggregate mode choice models. First, the question of whether the predictions of a model are "good" or not depends upon the cost of misforecasting and the accuracy of alternative methods of forecasting. Consequently, it is not possible to state whether a model predicts well or not in general. Second, the factors which were found to produce the mispredictions are not the only, or even necessarily the most important, sources or error. Only some of the many possible sources of mispredictions were explored. For example, a potential source of error is the problem of determining what modes are available to a person. The rules which the present study used for determining a mode to be available or not were fairly arbitrary. Different rules would have resulted in different predictions. Consequently, the size of the prediction errors depends on the accuracy of the "availability" rules. Last, and most important, is the fact that the present study concerns only one model and one transportation environment. Since the question of how well disaggregate mode choice models predict is an empirical question rather than an analytic one, no definite, general conclusion is possible. Only by examining the performance of the models in a variety of transportation environments will it be possible to have a "feel" for the power of the models. Once a number of validation tests have been performed, it will be possible to have a sense of how well the models predict, under what circumstances they work best and worst, and what problems need to be considered most in building future models.

†Copies of these papers can be obtained from Martin Lacey, 109 McLaughlin Hall, Institute of Transportation Studies, University of California, Berkeley, CA 94720.

REFERENCES

Charles River Associates (1976) The Independence of Irrelevant Alternatives Property of the Multinomial Logit Model. In *Disaggregate Travel Demand Models*, Project 8–13, Phase I Report, Washington, D.C., National Cooperative Highway Research Program.

McFadden D. (1973) Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics* (Ed. by P. Zarembka). Academic Press, New York.

McFadden D. (1974) The measurement of urban travel demand, *J. Public Econ.* **3**, 303–328.

Templeton J. and Talvitie A. (1976) A model system for network access, *Proc. Transportation Research Forum.*

Train K. (1976a) Work trip mode split models: an empirical exploration of estimate sensitivity to model and data specification. *Working Paper No.* 7602, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley.†

Train K. (1976b) A post-BART model of mode choice: some specification test. *Working Paper No.* 7620, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley.†

Train K. (1976c) A validation test of disaggregate travel demand models. *Working Paper No.* 7619, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley.†

Train K. and McFadden D. (1976) The measurement of urban travel demand IIA, *Working Paper No.* 7611, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley.†