CHAPTER 3

VALIDATION OF DISAGGREGATE TRAVEL DEMAND MODELS: SOME TESTS

Introduction

In the previous chapters, models were developed on a sample of workers interviewed before BART service was available. These are called the pre-BART models. After BART opened for service the same sample of people were interviewed again.[1] These events, the introduction of new transit service and a sample of behavior interviews before and after the opening of service, offer an exceptional opportunity for testing the forecasting 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 chapter reports the results of that comparison.

The validation of the pre-BART models on the post-BART sample is done in two ways. First, actual modal shares in the post-BART sample are compared with the modal shares that the pre-BART models predict. Second, the parameters of models estimated on the post-BART sample are compared with the parameters of the pre-BART model. This validation task is conducted using four different pre-BART models. These are the Models 8, 9, 11, and 12. Model 8 is chosen because it does not incorporate the endogenous variable "cars per driver," Model 9 for historical reasons that become clear later, Model 11 because it is the most complex and the best predictor with pre-BART data and is the focus of the validation tests, and Model 12 because it is the popular naive model. The chapter concludes with a discussion of reasons for the prediction errors of the post-BART behavior with the pre-BART model.

[1]Because of attrition in the pre-BART sample, the post-BART analysis was carried out on an augmented sample enriched with new respondents.

Validation Tests of Pre-BART Models

Before presenting the results it is good to review briefly why the production of forecasts from pre-BART models--estimated using four alternatives--to the post-BART situation, where the number of alternatives is from five to seven, is technically a fairly straightforward operation. The pre-BART models upon which evaluation tests are performed in this section are multinomial logit (MNL) models of individual choice probabilities. The models express the probability that a person with certain observed socioeconomic characteristics and facing a choice among several alternatives each of which exhibits certain measured attributes will choose a particular alternative. The function is expressed as:

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

where C_n is the set of alternatives among which person n may choose; $P_n\ (i \mid C_n)$ 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 of 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:

$$\frac{P_n(i \mid C_n)}{P_n(k \mid C_n)} = \frac{e^{\beta' z(x_n^i, s_n)} \Big/ \sum_{j \in C_n} e^{\beta' z(x_n^j, s_n)}}{e^{\beta' z(x_n^k, s_n)} \Big/ \sum_{j \in C_n} e^{\beta' z(x_n^j, s_n)}}$$

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

This ratio is constant for any C_n that 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, of imposing 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 assumed and 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 that 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. Without the IIA property, forecasting demand under the expanded choice set would not be possible, or at least it would be much more difficult.

The specification above of alternative-specific effects for BART alternatives was based on subjective judgments by project personnel on the similarities of unobserved attributes of alternatives. This subjectivity is a weak link in the forecasting process, and points out the drawback of new-mode forecasting with a non-generic model. Much more satisfactory for forecasting purposes would be generic models in which the assessment of alternatives requires only variables that can be calculated from observable mode attributes and socioeconomic characteristics, and their coefficients. However, we found the explanatory power of generic models using only time and cost/wage variables to be low, and the coefficient estimates to be implausible. For policy purposes, it thus seemed to be the lesser of two evils to include alternative-specific variables in the calibration, and attempt the forecasting of new mode dummy effects using the judgment of experts, consumer panels, or more elaborate market research methods. The validation tests are, in effect, tests of the joint hypothesis that the alternative-specific effects have been assigned correctly for the new modes and that the model specifies new mode demand correctly.

Forecasting Ability of the Final Pre-BART Model (Model 11)

In evaluating the pre-BART model in the post-BART situation 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 11 for predicting post-BART shares, a value of 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 can simply be calculated. For the socioeconomic variables and alternative-specific dummies, some assumptions must be made. For instance, in the pre-BART model, the variable "number of persons in household who can drive (3)" takes the described value for the bus-with-auto-access alternative and zero for other pre-BART alternatives. The question arises whether the variable should take the value of zero for all the BART alternatives, or should it take the described value for, say, the BART-with-auto-access alternative and zero in the other BART alternatives. The former approach is equivalent to considering all the BART alternatives to be similar to the bus-with-walk-access alternative; the latter is equivalent to considering BART-with-auto-access to be similar to bus-with-auto-access and BART-with-walk-and bus-access to be similar to bus-with-walk-access.

The latter approach was chosen for forecasting purposes. That is, in creating the socioeconomic variables for the BART alternatives, 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. The alternative-specific dummy variables were created analogously: the bus-with-auto-access alternative dummy takes the value of one not only in the bus-with-auto-access alternative but also in the BART-with-auto-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 15 presents the prediction success table for predictions based upon the model of Table 11. The table requires explanation. The $ij^{\underline{th}}$ element of the 7×7 central matrix (where i denotes the row and j denotes the column) is the probability of person n choosing mode j, summed over all persons who actually chose mode i:

$$ij^{\underline{th}} \text{ element} = \sum_{n \in S_i} P_n(j \mid C_n) \quad ,$$

where S_i is the set of persons in the sample who actually chose mode i. For instance, the element in the first row, second column (21.11) is the sum over all persons who chose auto alone of the probability of choosing bus with walk access.

The most interesting property of this matrix is that its row and column sums are immediately interpretable. Summing across a particular row gives the number of people who actually chose that mode:

$$\text{row total} = \sum_{j \in C_n} \left[\sum_{n \in S_i} P_n(j \mid C_n) \right] = \sum_{n \in S_i} \left[\sum_{j \in C_n} P_n(j \mid C_n) \right] = \sum_{n \in S_i} 1 = N_i \quad ,$$

where N_i is the number of persons in set S_i. For example, the row total for the auto-alone row is 378 , meaning that 378 persons actually chose auto-alone. Summing down a particular column gives the sum over all persons in the sample of the probability of taking that mode:

$$\text{column total}_j = \sum_{i \in C_n} \left[\sum_{n \in S_i} P_n(j \mid C_n) \right] = \sum_{\text{all } n} P_n(j \mid C_n) \quad .$$

TABLE 15 Prediction Success Table (based on Model 11)

Actual Alternatives	Predicted Alternatives								
	(1) Auto Alone	(2) Bus/Walk	(3) Bus/Auto	(4) BART/Walk	(5) BART/Bus	(6) BART/Auto	(7) Carpool	Row Total	Actual Share (%)
(1) Auto Alone	246.5	21.11	5.984	17.63	1.188	10.91	74.71	378.0	59.53
(2) Bus/Walk	10.97	33.20	2.799	6.689	1.104	.8186	12.43	68.00	10.71
(3) Bus/Auto	1.108	2.367	.5929	1.571	.0057	1.041	2.314	9.000	1.417
(4) BART/Walk	.4660	.2832	.0758	1.307	.1829	.6247	1.060	4.000	0.6299
(5) BART/Bus	.7687	1.526	.0963	1.287	1.040	.1813	1.101	6.000	0.9449
(6) BART/Auto	7.350	2.302	1.365	7.709	.5612	6.201	7.512	33.00	5.197
(7) Carpool	70.65	11.39	3.071	11.62	1.149	5.228	33.89	137.0	21.57
Column Total	337.8	72.17	13.98	47.81	5.230	25.01	133.0	635.0	
Predicted Share (%)	53.19	11.37	2.202	7.529	0.8236	3.938	20.95		
Percent Correct	72.97	45.99	4.240	2.735	19.88	24.79	25.48		

Total Percent Correct: 50.82

Root Mean Squared Error: 9.53

This is the best prediction of the total number of people to choose the particular mode. Dividing the row totals and column totals by the number of people in the sample gives, respectively, the actual and predicted mode shares. A comparison of actual and predicted shares indicates that the pre-BART model:

- underpredicts use of auto-alone;
- overpredicts use of both the bus alternatives;
- greatly overpredicts the use of BART-with-walk-access;
- underpredicts the use of the other two BART alternatives;
- underpredicts the use of carpool.

Summing the columns and rows of Table 15 over the five transit modes gives an actual transit share of 18.9% and a predicted share of 25.9%. That is, the predicted transit share is 37% larger than the actual transit share.

The percent correct for an alternative is the element in the diagonal for the particular column of the 7 × 7 matrix divided by the column total. For instance, the percent correct for bus-with-auto-access is (.5929/13.98) = 4.24 . To interpret these percents correctly, it is useful to compare them to the percents correct that would be obtained from "chance." Any model that assigns the same probability of choosing an alternative to all persons in the sample would obtain a percent correct for each alternative equal to the actual share for that alternative. This result is shown as follows. By definition, the percent correct is:

$$(2) \qquad \frac{\sum_{n \in S_i} P_n(i \mid C_n)}{\sum_{\text{all } n} P_n(i \mid C_n)} \quad .$$

$P_n (i \mid C_n)$ is the same for all persons in the sample (hence the model is called one of "chance"), and so the probability can be denoted as P_i , independent of n . Substituting into (2):

$$\frac{\sum_{n \in S_i} P_i}{\sum_{\text{all } n} P_i} = \frac{N_i P_i}{N P_i} = \frac{N_i}{N} \quad ,$$

where N is the total number of people in the sample.

Comparing the percents correct in Table 15 with those that would obtain from "chance" gives an indication of how well the model is predicting. For each alternative, the percent correct in Table 15 is higher than that which "chance" would produce. The percent correct for bus-with-auto-access is 4.24, which is about three times better than "chance" (which would obtain 1.417 percent correct). Therefore, even though the percent correct is small for this alternative, the percent correct is better compared to "chance" for this alternative than for the carpool alternative.

The total percent correct is the sum of the elements in the diagonal of the 7×7 matrix divided by the total number of people in the sample. This statistic can be interpreted by comparing its value for a particular forecasting model with the total percent correct obtained by a model containing only alternative specific dummies and based on knowledge of the post-BART shares. With the "dummy-only" model, each person is assigned a probability of taking a given mode equal to the aggregate share for that mode. Thus, the total percent correct for the "dummy-only" model is equal to

$$100 \times \sum \left(\frac{N_i}{N} \right)^2 \quad .$$

With the actual shares of Table 15, the "dummy-only" model obtains a total percent correct of 41.54. The total percent correct for the model of Table 11 is 50.82.

The comparison between the "dummy-only" model described above and any forecasting model is, however, somewhat misleading. The "dummy-only" model is based upon knowledge of the post-BART actual shares, while the forecasting model is not. A more revealing comparison might be between the total percent correct obtained by a "dummy-only" model based only on knowledge of pre-BART shares and that obtained by the forecasting model. Calculation of the total percent correct from this type of "dummy-only" model is made with the assumption that the dummy variables for the BART-with-walk and bus-access alternatives are the same as that for the bus-with-walk-access and that the dummy variable for BART-with-auto-access equals that of bus-with-auto-access. The total percent correct for this "dummy-only" model, given the actual pre- and post-BART shares, is 29.21.

The root mean square error is a statistic that allows quick (but crude) comparison of the predictive ability of different models. The statistic is defined as

$$\mathrm{RMSE} = \sqrt{\sum_i (q_i - r_i)^2}$$

where q_i and r_i are the predicted and actual shares, respectively, for alternative i.

The major discrepancy between observed and predicted shares occurred in the BART-with-walk access mode, which was the observed choice of only four persons in the sample. There is some tendency for the model to overpredict all transit-with-walk access alternatives; this is due to the network calculations of transit walk time, which were predicted on a denser transit system than currently prevails, particularly in suburban areas. However, this effect can account for only a minor part of the overprediction of BART-with- walk-access. A second possible source of the difficulty is the assumption that BART-with-walk-access has average unobserved characteristics similar to bus-with-walk-access. In fact, bus access distances are typically shorter than for BART, and bus stops are located in residence neighborhoods rather than at commercial/transportation centers as are BART stations. The predicted BART/walk patronage that fails to materialize comes primarily from urbanized areas. This suggests that special characteristics of short urban trips may be important, or that the tastes of urban residents differ from those of suburban residents. The problem of BART unreliability may loom larger for short urban trips than for long suburban ones, depressing its use by the population segment who would be expected to employ walk access. For many urban trips, BART-with-walk-access is dominated in time and cost by bus-with-walk-access. It may be the case that discrimination among modes is much stronger when dominance obtains relative to when it does not than the linear "trade-off" utility function assumed in the multinomial logit model permits.

Finally, there may be substantial taste variation in individual attitudes toward walking, particularly with respect to the maximum distance perceived as a reasonable walk. By forcing a common importance weight on walk time, the logit model may overpredict the willingness of individuals to extend walk times.

If the BART-with-walk-access alternative is dropped, and forecasts are made for the remaining six alternatives conditioned on the BART/walk mode not being chosen, then the predictive success of the pre-BART model is quite good. Table 15A is the conditional six-mode prediction success table for Model 11.

TABLE 15A Prediction Success Table for Pre-BART Model and Post-BART Data

Actual Alternatives	Predicted Alternatives					
	(1) Auto-Alone	(2) Bus/Walk	(3) Bus/Auto	(4) BART/Bus	(5) BART/Auto	(6) Carpool
(1)Auto-Alone	255.1	22.21	6.632	1.513	13.72	79.07
(2)Bus/Walk	11.56	36.43	2.988	1.679	1.421	13.92
(3)Bus/Auto	1.249	2.811	.687	.0066	1.625	2.622
(4)BART/Bus	.858	1.934	.120	.1391	.258	1.440
(5)BART/Auto	8.898	3.149	1.756	.695	8.828	9.674
(6)Carpool	74.68	12.43	3.305	1.357	7.497	37.73
Column Total:	352.4	78.97	15.22	6.642	35.35	144.4
Predicted Share (%) (standard error)	55.8 (11.4)	12.5 (3.4)	2.4 (1.4)	1.0 (0.5)	5.3 (2.4)	22.9 (10.7)
Row Total	378	68	9	6	33	137
Observed Share	59.9	10.8	1.4	0.95	5.2	21.7
Percent Correct	72.4	46.1	4.5	21.0	26.5	26.1
Success Index	1.30	3.69	1.88	21.0	5.0	1.14
Predicted Share less Observed Share	-4.1	1.7	1.0	0.05	0.1	1.2

Overall Percent Correct: 53.9 (42.0 by chance)

Overall Success Index: 1.28

Forecasting Ability of Other Pre-BART Models (Models 8, 9, and 12)

The pre-BART models (8, 9, and 12) were specified and estimated without reference to the post-BART sample, where the model of Table 11 was developed after the post-BART model estimation had begun. A strict test of the pre-BART models would, therefore, be based upon the "best" truly pre-BART model rather than Model 11.

The model that attained the highest value of the log likelihood function is Model 9. This model is different from Model 11 in that (1) on-vehicle time is included as a single variable rather than as a different variable for auto and transit on-vehicle time; (2) single "autos per driver" and "number of drivers" variables were included, entering the auto-alone and bus-with-auto-access alternatives, rather than the three "autos per driver" and three "number of drivers" variables in Model ll; (3) two initial headway variables were included, one for headways less than eight minutes and another for headways exceeding eight minutes.[1]

Not only Model 11 but also Model 9 is more complex than most of the disaggregate travel demand models that are being used for planning by planning agencies. An important question is whether these less complex models entail considerably worse predictions. Two less complex pre-BART models were chosen to examine this issue. The first is Model 8, which is a simplification of Model 9 in

[1]Model 11 was chosen as the focus of validation because it was a later step in the on-going process of model development work even though it was "discovered" after some experience using post-BART data. Model 11 is a natural outgrowth of the specification tests conducted pre-BART. For instance, the hypothesis that auto and transit on-vehicle times have the same coefficient, implied by Model 9, can be rejected at .05 level of significance. This test was performed pre-BART, Models 8 and l0. Similarly, all the other changes from Model 9 to Model 11 follow analogous tests of hypotheses. Thus, Model 11 is a justifiable improvement over earlier models.

In retrospect, the real reason we did not "discover" Model 11 earlier is that, because of its endogenous nature, we were unhappy with including "autos per driver" as a variable and also unhappy with the non-genericity of the on-vehicle time, even though we knew that these variables had an important bearing on forecasting accuracy. Thus, *a priori* theoretical considerations prevented us from fully exploring "obvious" model specifications. Discussions in Part III, Chapter 1, are relevant here.

The other changes, three separate coefficients for "drivers" and only one headway coefficient, did not improve forecasting accuracy. Examination of Model 11 readily shows that the three coefficients for "drivers" are almost equal and could be entered as one variable. The resolution adopted for headway is tentative and subject to change once more detailed data are available (See Part III, Chapter 5).

The reader is asked to study carefully the pre-BART model development work reported in Part II, Chapter 2, and to learn from our experience.

that it does not include the variables "head of household," "employment density at work zone," "home location in or near CBD," or the endogenous variable "autos per driver." The second is Model 12. This "naive" model includes only the variables "cost divided by wage," "on-vehicle time," "excess time," and the alternative-specific dummies. Excess time was defined to be the sum of walk and transfer times, and one-half of first headways.

Table 16 presents the summaries of the prediction success table of these two simple models, the truly pre-BART model, Model 9, and the final pre-BART model; the actual shares are also shown. The figures in Table 16 indicate that the simpler models predict less well.

We cannot interpret this result to mean that the complex models are necessarily better-suited for forecasting travel demand, because in using the models for prediction, the future values of all the independent variables in the model must be predicted. Model 11 does not require any more data preparation than Model 9, but they require substantially more than Model 8 and in particular Model 12, which needs only the travel times, costs, and wage rate. For Model 11, these variables plus many socioeconomic variables such as "number of drivers" and "autos per driver" must be forecast. The inaccuracies in forecasting these socioeconomic variables must be coupled with the mispredictions of mode choice given the socioeconomic variables to obtain a true measure of the forecasting ability of the model. When these considerations are taken into account, the predictive power of Model 12 may not be that much worse than the values in Table 16 indicate.

It can be argued that for short-term forecasting the socioeconomic variables can be taken as given and hence the complex models are better. In particular, the more complex models should remove from the estimated coefficients of transportation policy variables any "spurious" effects due to the correlation of these variables with omitted socioeconomic variables, and thus improve policy forecasts. Again, for short-term forecasting incremental forecasts are often employed. These require only the coefficient of the variable being affected. The comparison of coefficients of system variables for Models 11 and 12 show they are about equal when excluding the cost and auto on-vehicle time coefficients. There is nothing to prevent one developing a naive model having non-generic on-vehicle time; this change in specification may or may not produce a cost coefficient more nearly to that of Model 11.

TABLE 16 Predictions Based on Logit Models with Simple Specifications

	(1) Actual Share	(2) Predicted Share based on Model 10 (Final Model)	(3) Predicted Share based on Model 8 *	(4) Predicted Share based on Model 7 *	Predicted Share based on Model 11 ("naive" model)
Auto Alone	59.53	53.19	50.74	47.30	44.68
Bus/Walk	10.71	11.37	13.24	13.77	14.08
Bus/Auto	1.417	2.202	2.617	2.934	3.185
BART/Walk	0.630	7.529	7.587	8.335	10.60
BART/Bus	0.945	0.824	1.950	1.844	1.308
BART/Auto	5.197	3.938	3.020	3.133	4.073
Carpool	21.57	20.95	20.85	22.68	22.08
Root Mean Squared Error		9.53	11.82	15.06	18.33
Total Percent Correctly Predicted		50.82	48.37	44.12	42.33
	(n = 635)				

*The model was reestimated without the variable “Length of Residency in Community”

The above paragraph is not meant to convince the reader that the naive models are better; clearly, they are not. Rather, the discussion was presented to arouse the curiosity of the reader; e.g., why does the deletion of many powerful socioeconomic variables not affect the coefficients of the system variables? We will return to this question in Part III.

The discussion turns next to the comparison of pre- and post-BART models.

Comparison of Model Coefficients Developed Using Pre- and Post-BART Data

The second method for evaluating the pre-BART model of Table 11 is to estimate a model with the same specification on the post-BART sample. If the estimates and specification of Table 11 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 that give rise to the discrepancies between predictive and actual post-BART shares.

Table 17 presents a model estimated on the post-BART sample with the same specification as the model of Table 11 (the pre-BART values of the corresponding parameters are given in parentheses). The estimates are fairly similar. The differences between the pre- and post-BART estimates that seem most relevant to the forecasting errors of Table 15 are (1) the value of walk time is much higher post-BART than pre-BART, and (2) the BART-with-walk-access dummy is significantly less than zero. Because only the transit modes entail walk time (data assumption), the different estimates for the value of walk time pre- and post-BART could be related to the over-prediction of transit. The significantly negative estimate of the BART-with-walk-access dummy could be related to the large over-prediction 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 was the value for the bus-with-walk-access dummy in the pre-BART model. 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 BART-with-walk-access indicates that this assumption is not valid. This also suggests that the test for the quality of the pre- and post-BART coefficients must be rejected.[1] This statement must be tempered by the fact that the sample included only four persons who chose BART-with-walk-access and six who chose BART-with-bus-access. The dummy variable coefficients for these alternatives are thus based on very few observations. Even if the coefficients are not transferable from one time period to another the model predictions are encouragingly good. Besides, the reasons for the mispredictions do not rest entirely with the demand models. These possible reasons for mispredictions are discussed next.

---

[1]A formal test for the equality of the pre- and post-BART coefficients is postponed until Part III, Chapter 7, where several types of transferability are examined and tested with data from three different cities in the U.S.

TABLE 17 Work-Trip Mode-Choice Model, Estimated Post-BART with Non-Generic Auto and Transit On-Vehicle Time

(Mode l--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 minutes (1-7)	-.0266	(-.0284)	3.92
Auto on-vehicle time, in minutes (1,3,6,7)	-.0473	(-.0644)	3.48
Transit on-vehicle time, in minutes (2-5)	-.0197	(-.0259)	2.03
Walk time, in minutes (2-6)	-.0900	(-.0689)	3.36
Transfer-wait time, in minutes(2-6)	-.0438	(-.0538)	1.81
Number of transfers (2-6)	-.120	(-.105)	0.856
Headway of first transit carrier, in minutes	-.0290	(-.0318)	2.60
Family income with ceiling of $7,500, in $ per year (1)	-.000289	(.00000454)*	1.78
Family income minus $7,500 with floor of $0 and ceiling of $3,000, in $ per year (1)	.0000522	(-.0000572)	0.364
Family income minus $10,500 with floor of $0 and ceiling of $5,000, in $ per year (1)	-.0000419	(-.0000543)	0.738
Number of persons in household who can drive (1)	1.48	(1.02)	5.26

Table 17, continued

Independent Variable	Estimated Coefficient		T-Statistic
Number of persons in household who can drive (3,6)	1.65	(.990)	5.16
Number of persons in household who can drive (7)	1.28	(.872)	4.85
Dummy if person is head of household (1)	.668	(.627)	3.19
Employment density at work location (1)	-.00164	(-.00160)	3.45
Home location in or near CBD (2=in CBD, l=near CBD, 0 otherwise) (1)	.1546	(-.502)**	0.835
Autos per driver with a ceiling of one (1)	4.79	(5.00)	3.70
Autos per driver with a ceiling of one (3,6)	3.63	(2.33)	4.81
Autos per driver with a ceiling of one (7)	3.26	(2.38)	3.19
Autos-alone alternative dummy (1)	-4.18	(-5.26)	2.82
Bus-with-auto-access dummy (3)	-8.24	(-5.49)	6.67
BART-with-walk-access dummy (4)	-2.28	(0.00)[t]	3.36
BART-with-bus-access dummy (5)	-.473	(0.00)[t]	0.708
BART-with-auto-access dummy (6)	-7.30	(-5.49)[t]	5.93
Carpool alternative dummy (7)	-5.31	(-3.84)	5.56

Likelihood ratio index	.4599	(.4426)
Log likelihood at zero	-964.4	(-1069.0)
Log likelihood at convergence	-520.9	(-595.8)
Percent correctly predicted	67.24	(67.83)

Table 17, continued

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

Auto on-vehicle time	178	(2.53)	(227	(3.2))
Transit on-vehicle time	74	(1.84)	(91	(2.4))
Walk time	338	(2.46)	(243	(3.1))
Transfer-wait time	165	(1.65)	(190	(2.0))
Value of initial headways as a percent of wage:	109	(2.13)	(112	(2.5))

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	(429)
Bus-with-walk-access	68	(134)
Bus-with-auto-access	9	(30)
BART-with-walk-access	4	–
BART-with-bus-access	6	–
BART-with-auto-access	33	–
Carpool	137	(178)
Total sample size	635	771

* Has significantly worse t-statistic

** Has significantly better t-statistic

t Used in validation

Reasons for Mispredictions

Three possible reasons were identified as being the culprit for mispredictions. These are (1) failure to satisfy the IIA requirement, (2) non-genericity in the attributes of BART-with-walk-access, and (3) incorrect data for walk times in the post-BART sample. These reasons apply with varying intensity to all the models from the most complex to the "naive" models.

Failure of independence from irrelevant alternatives

If the five transit alternatives are not actually independent, the MNL model would overpredict transit use (Charles River Associates, 1976). Because 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 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 that 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 to 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:

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

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

$$P_n(\ell \mid T_n) = e^{\alpha' z(x_n^\ell, s_n)} \Big/ \sum_{k \in T_n} e^{\alpha' z(x_n^k, s_n)} \quad ,$$

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:

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

where

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

For the log-sum model:

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

Because 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 18 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 better, if not substantially so, than those of the MNL model. The maximum and log-sum models overpredict transit use by thirty-five percent and twenty-two percent, respectively, whereas the MNL model overpredicts transit by thirty-seven percent.

TABLE 18 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 Mean Squared Error (n = 639)		8.94	6.49

Because the non-MNL models greatly over-predict 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.

Non-genericity of attributes of BART-with-walk-access

If BART-with-walk-access exhibits some important attributes that none of the pre-BART modes do, 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 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 non-genericity of the second type exists, that is, whether any attribute that is similar for bus and BART (such as on-vehicle time) is valued differently for the two modes. Because BART trains are generally more comfortable than buses, the value of on-vehicle time is perhaps lower for BART than bus. Similarly, because 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 walking to bus because many BART stations are surrounded by parking facilities that are less pleasant to walk through than walking on sidewalks. Tests of these four attributes were performed and no significant (at the .05 significance level) non-genericities were found. The results of these tests are detailed in the next section. These test indicate, therefore, that non-genericity of the second type does not explain the large overprediction for BART-with-walk-access.

The existence of non-genericities of the first type (that is, attributes existing for BART that 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 that 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 non-genericity of the first type exists, then the estimated coefficient of BART-with-walk-access is expected to be close to zero. As the post-BART model of Table 17 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 (note, however, the few observations in the sample on which the dummy for BART-with-walk-access is based). Non-genericity of the first type seems indeed to exist and to contribute 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 19 presents the predicted and actual shares conditional upon BART-with-walk-access not being chosen. This table summarizes the results in Table 15A. The predicted shares are calculated the same way as those in Table 15, 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 19 are much closer to the actual shares than those of Table 15. However, the auto-alone alternative is still being unpredicted and the bus alternatives overpredicted. The possibility that bad data for walk times, especially in the bus alternatives, is causing these mispredictions is explored below.

TABLE 19 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 Mean Squared Error (n = 631)		4.67

Incorrect walk time data

The attributes of the transit alternatives were calculated using standard Urban Transportation Planning System (UTPS) type networks and programs that simulate the Bay Area transit system for particular years. These networks existed (that is, they had been previously coded for another study) for the years 1965 and 1976. The 1965 system had been coded to represent the system as it actually existed in 1965. The 1976 network had been scaled down from a 1968 network. The extent of changes that were undertaken in this operation is unknown. It is fair to expect that the 1976 network included planners' anticipations about future transit improvements. The networks for the years of interest, 1972 for pre-BART and 1975 for post-BART were constructed as follows. The 1972 pre-BART network was obtained by adjusting the 1965 system to account for the few changes that occurred during the intervening years. Complete information on the system status in 1972 was available at the time of adjustment. The 1975 post-BART attributes were obtained by adjusting the 1976 network. In these adjustments some bus lines that were expected to exist in 1976 (or perhaps 1980) were deleted and some, mainly transbay, bus lines were reinstated. The values in the walk links were not changed, however.

Three hypotheses come to mind immediately, all of which point toward the existence of too-low walk times. First, the number of bus lines the planners anticipated to exist in 1976 was much higher than actually existed in 1976. Decreasing the number of bus lines increases, in average, the walk time to bus. This effect was not accounted for in the 1976 or 1975 networks because the walk times were not adjusted. Second, there is often a tendency among coders to code the walk times of those who actually use the system rather than the average walk time of the population segment. Third and finally, it is possible that some walk times in the post-BART network were intentionally coded low in order to predict a large BART patronage.

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 coded 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 unrealistically low walk times could also explain the misprediction of Table 19 (that is, 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 alternatives than actually do. BART on-vehicle times were calculated relatively accurately because 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, because walk times for BART-with-auto-access are small. Because 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 alternative. As Table 19 shows, the mispredictions that would be expected from downward biased walk times for buses actually occur.

It seems, therefore, that the mispredictions of the pre-BART model can be traced to two major problems, non-genericity in the BART-with-walk-access alternative and incorrect walk time data. This conclusion, however, like any *ex post* conclusion, is tentative.