PROCEEDINGS —
Eighteenth Annual Meeting

Theme:
"Transportation in Transition"

October 17-18-19, 1977
Colony Square Hotel
Atlanta, Georgia

Volume XVIII • Number 1 1977

T R F

TRANSPORTATION RESEARCH FORUM

A Summary of the Results of a Validation Test of Disaggregate Travel Demand Models

*by Kenneth Train**

I. 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 on a sample of workers taken after BART opened by comparing actual mode shares in the post-BART sample with the mode shares which the pre-BART models predicted.

Sections II-IV present and evaluate a particular model which was estimated on the pre-BART sample. Specifically, section II presents and discusses this pre-BART model; section III tests the model in the way described above; and section IV 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 various models of different specifications, and points out the difficulties in obtaining "reasonable" forecasts from the models.

II. 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 socio-economic 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:

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

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 $i \epsilon C_n$; 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. The IIA property greatly facilitates estimation and forecasting, particularly in the situation of a new alternative being 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. Without the IIA property, forecasting demand under the expanded choice set would not be possible.

**University of California, Berkeley.*

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.

The model which was estimated on the pre-BART sample and is subject to evaluation tests in section III is given in Table 1. The first column lists the elements of $z(x_n^i,s_n)$ in Equation (1); the second and third columns list the estimates and t-statistics, respectively, of the elements of β in Equation (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 percent of wage. Further explanation of the model is contained in Train (1976c).

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).

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 .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 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.

III. EVALUATING THE PRE-BART MODEL

The 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 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 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

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	.4426
Log likelihood at zero	—1069.0
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

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

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 equation (1), and N is the sample size.

A comparison of the 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 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.

Section IV explores possible reasons for the mispredictions of Table 2. The possibilities which are discussed are:

(A) Failure of the IIA property;
(B) Non-genericity in the attributes of BART with walk access;
(C) Incorrect data for walk times in the post-BART sample.

IV. REASONS FOR MISPREDICTIONS

A. Failure of IIA.

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

Since neither the Maximum nor Log-sum models entails the property 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 better, but not substantially better, than those of the MNL model. 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 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 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 to 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 of these four attributes were performed and no significant (at the .05 confidence level) non-genericities were found. These results are detailed in Train (1976b). These tests 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 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 dummy is zero (by normalization). If no non-genericity of the first type exists, then the estimated coefficient of BART with walk access alternative dummy is expected to be close to zero. A post-BART model with the same specification as the model 1 was estimated (Train (1976c), Table 3). For this model, 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. Non-genericity of the first

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

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 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 under-predicted and the bus alternatives over-predicted. The possibility that bad data for walk times, especially in the bus alternatives, 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 years. 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, 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.

The unrealistically low walk times could explain the mispredictions of Table 5 (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 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

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

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.

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, is tentative. Hand-calculated data on the attributes of transit are being prepared, and the validation tests will be repeated on these data. In particular, the issue of incorrect walk time data will be addressed directly. This further analysis will hopefully provide more information as to the causes of the mispredictions.

REFERENCES

Charles River Associates (1976), "The Independence of Irrelevant Alternatives Property of the Multinomial Logit Model," National Cooperative Highway Research Program, 8-13.

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

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 Tests," *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 D. McFadden (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.