REFERENCES

1. A. B. Sosslau and G. E. Brokke. Appraisal of Sample Size Base on Phoenix O-D Survey Data. HRB, Bulletin 253, 1960, p. 114-127.
2. K. E. Heanue, L. B. Hamner, and R. M. Hall. Adequacy of Clustered Home-Interview Sampling for Calibrating a Gravity-Model Trip-Distribution Formula. HRB, Highway Research Record 88, 1965, pp. 116-136.
3. P. S. Parsonson and P. D. Cribbins. Estimating Urban Trip Production and Attraction. Journal of the Highway Division, Proc., ASCE, Nov. 1967, pp. 115-134.
4. V. G. Stover, J. D. Benson, and L. J. Ringer. Accuracy of Travel-Pattern Estimated From the Home-Interview Survey. Texas Transportation Institute, Texas A&M Univ., College Station, Res. Rept. 167-7, Aug. 1973.
5. J. D. Benson, D. F. Pearson, and V. G. Stover. Accuracy of Travel-Pattern Estimates From the Home-Interview Survey. Texas Transportation Institute, Texas A&M Univ., College Station, Res. Rept. 167-8, March 1974.
6. V. G. Stover, J. D. Benson, and J. Buechler. Sensitivity Evaluation of the Traffic Assignment Process. TRB, Transportation Research Record 610, 1976, pp. 50-52.
7. D. F. Pearson, V. G. Stover, and J. D. Benson. A Procedure for the Estimation of Trip-Length Frequency Distributions. Texas Transportation Institute, Texas A&M Univ., College Station, Res. Rept. 17-1, May 1974.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

An Application of Diagnostic Tests for the Independence From Irrelevant Alternatives Property of the Multinomial Logit Model

Daniel McFadden and Kenneth Train, University of California, Berkeley
William B. Tye, Charles River Associates, Inc., Cambridge, Massachusetts

Statistical tests are proposed to diagnose the validity of the independence from (of) irrelevant alternatives property of the multinomial logit model. Application of the tests is illustrated by the use of actual travel data representing urban modal choice in the San Francisco area. The property as it applies to travel demand forecasting is discussed, and the common misconception that the property holds for market shares in heterogeneous populations is shown by examples to be incorrect. The relation of the property to the basic assumptions of the model is described, and it is shown that the validity of the property in disaggregate modeling is an empirical issue that depends on the model specification and data in a particular application. A series of diagnostic tests for the property are developed and applied to actual travel data.

The most widely used functional form for choice probabilities in disaggregated transportation-demand analysis is the multinomial logit (MNL) model,

$$P(i|C) = \exp V(x^i, s) \Big/ \sum_{j \in C} \exp V(x^j, s) \qquad (1)$$

where

C = finite choice set,
P(i|C) = choice probability for alternative $i \in C$,
x^i = vector of the observed characteristics of alternative i, and
s = vector of the observed characteristics of the decision maker and the choice environment.

The scale function $V(x^i, s)$ may be interpreted as the representative utility of alternative i and is normally assumed to be linear in the parameters. The MNL model has significant advantages over the available alternatives in terms of flexibility and computational efficiency and permits a simple behavioral interpretation of the parameters of the scale function.

The MNL model also has the property that the ratio of the probabilities of choosing any two alternatives

$$P(i|C)/P(k|C) = \exp V(x^i, s)/\exp V(x^k, s) \qquad (2)$$

is independent of the attributes or the availability of a third alternative (j), which is termed the independence from (of) irrelevant alternatives (IIA) property. This property greatly reduces the complexity of estimation and forecasting and in this respect is quite useful. However, it imposes restrictions on the structure of choice probabilities and cross elasticities; these restrictions may be invalid in some applications. Hence, tests of the validity of the IIA property should be made whenever a violation of the assumption is suspected.

This paper analyzes the IIA property and discusses several diagnostic tests. Complete descriptions of the tests and thorough instructions for construction of the test statistics can be found in a National Cooperative Highway Research Program (NCHRP) report (1) and McFadden, Tye, and Train (2).

INDEPENDENCE FROM IRRELEVANT ALTERNATIVES PROPERTY OF THE MULTINOMIAL LOGIT MODEL

In applications of the MNL model to individual modal choice, the IIA property (Equation 2) requires that if two modes are available and a new mode is introduced, the ratio of the probabilities of the two preexisting modes will be unchanged regardless of the probability of choice for the new mode. For example, if the new mode will be chosen with a probability of 0.10 and each preexisting mode had a 0.50 probability before the introduction of the new mode, the probability of each of the preexisting modes will be 0.45 after the new mode is introduced, thus preserving the one-to-one ratio of probabilities of the preexisting modes.

The IIA property also greatly facilitates the forecasting problems associated with new modal-choice predictions. If 100 persons have the same observed characteristics of alternatives, the same observed characteristics of the decision maker, and the same choice set, i.e., they have the same $V(x^i, s)$'s, the demand for a new mode can be calculated by adding another term to the denominator of Equation 1 and recomputing all choice probabilities. The new probabilities can then be multiplied by 100 to estimate the demand for each mode. If the old modes formerly shared the market equally and the probability of the new mode is 0.10 for each individual, the predicted modal demands will be 45, 45, and 10.

An example of a choice setting in which the IIA property is inappropriate is the classic blue automobile versus red automobile case. Assume that the bus mode and the blue automobile each capture 50 percent of a given travel market as shown in the first column of the table below.

Mode	Modal Choice (%) True and MNL (binary choice)	Predicted MNL (3 modes)	True (3 modes)
Bus	50	33.3	50
Blue automobile	50	33.3	25
Red automobile	0	33.3	25
Total	100	100	100

Assume then that a new automobile mode is introduced with exactly the same service attributes as the blue automobile mode except that the automobile is painted a different color, e.g., red (patrons are assumed to be indifferent to color). Assume also that the red automobile is leased for this trip only, to remove questions of automobile ownership and competing demands for the automobile. The true modal shares will now be 50, 25, and 25 percent, for bus, blue automobile, and red automobile respectively; i.e., no bus users will switch to the new mode and automobile users will split evenly between the two automobile modes. However, the MNL model will forecast that each of the three modes captures one-third of the market, as shown by the second column of the table above, because the IIA property requires that the ratio of the bus share to the blue automobile share be unaffected by the introduction of the red automobile. In this example, the ratio is 1.0: When the red automobile is introduced, the ratio of the blue automobile share to the red automobile share is 1.0 (because patrons are assumed indifferent to color). The only shares that allow both the ratio of bus share to blue automobile share and the ratio of blue automobile share to red automobile share to equal one are one-third shares for each mode. Thus, the MNL model predicts shares of 33, 33, and 33 percent when the actual shares are 50, 25, and 25 percent for bus, blue automobile, and red automobile, respectively.

If the problem were confined to this simple example, it would be trivial. The new automobile mode is clearly irrelevant and should not be introduced as a mode. However, this extreme case points to a gray area, where the demand forecast for a new mode could be seriously compromised by incorrectly applying the IIA property.

In the MNL model, the IIA is a property of individual probabilities and market shares in homogeneous populations, but not a property of market shares in heterogeneous populations. Much unwarranted criticism of the MNL model has been based on the erroneous application of the IIA property to market shares in heterogeneous populations. It should be emphasized that the MNL model does not predict that the ratio of market shares in a heterogeneous population will be invariant with the introduction of a new alternative.

To take a specific example, the MNL model does not in general predict that, if a new mode is introduced to a population composed of different market segments that have different observed socioeconomic characteristics and level-of-service attributes [different $V(x^i, s)$'s for the individuals], the percentage of automobile drivers who will use the new mode is equal to the percentage of transit users who will shift.

This principle may be illustrated by an example. Table 1 presents the case of a population composed of two market segments of 100 persons each. Each segment is composed of homogeneous individuals; i.e., each person in the segment assigns the same representative utility to each alternative. Assume that the choice environment of observed attributes is identical for all persons within each segment and that it differs significantly between the two market segments. Segment 1 is automobile oriented, splitting 90 to 10 in favor of the automobile, and segment 2 is transit oriented, splitting 90 to 10 in favor of transit.

A new mode—dial-a-bus—is introduced. The MNL model predicts that it will capture 5 percent of segment 1 and 15 percent of segment 2. As indicated in Table 1, the ratio of the automobile market share to the bus market share is preserved within each homogeneous market segment. However, the ratio of the bus modal share to the automobile modal share is not constant after the new bus mode is introduced, but decreases from 1.0 to 0.91 for the entire population (86 ÷ 94 = 0.91).

Although the percentage diversions from the bus and the automobile to the dial-a-bus are the same within each homogeneous market segment (e.g., in segment 1,

Table 1. Effect of IIA property on a forecast of behavior in a population of heterogeneous market segments.

Mode	Modal Share: MNL (binary choice) Market Segment 1	Market Segment 2	Total Market	Predicted MNL (3 modes) Market Segment 1	Market Segment 2	Total Market
Bus	10	90	100	9.5	76.5	86.0
Automobile driver	90	10	100	85.5	8.5	94.0
Dial-a-bus	0	0	0	5.0	15.0	20.0

5 percent of both bus and automobile patrons switch), the predicted diversions from the automobile and the bus are not the same for the population as a whole. Of the 100 total bus patrons in the binary-choice situation, 14 percent (100 - 86) were predicted to switch to the dial-a-bus, but only 6 percent of the total automobile users were predicted to switch to the dial-a-bus.

The IIA property is obviously a key assumption of the MNL model. Previous studies of it have tended to discuss its reasonableness or unreasonableness on logical grounds. This paper argues that the issues raised by the property are essentially empirical. The convenience of the IIA property in estimating and forecasting makes it extremely attractive to use when it is valid. But the undesirable consequences of assuming the IIA property when it is invalid are reason for caution in applying the MNL model without assurances of the reasonableness of the IIA property.

This dilemma is addressed here by the development of statistical tests that can identify whether or not the IIA property is reasonable in the particular circumstances. These tests are comparable to the standard statistics that are routinely calculated as part of regression programs to identify whether or not the assumptions of the least-squares model are reasonable.

SOURCES OF VIOLATION OF THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES PROPERTY

In developing statistical tests to determine the validity of the IIA property it is useful to consider the basic assumptions of the MNL model and the ways in which they might be violated. The utility for the individual of the ith alternative is assumed to be a function of the observed characteristics of that alternative, the observed characteristics of the decision maker and choice environment, and an unobserved component that represents the effects of omitted random taste variations, choice attributes, and socioeconomic variables.

$$U_i = U_i(x^i, s, \mu_i) \qquad (3)$$

where

- U_i = utility of the ith alternative,
- x^i = vector of observed characteristics of the ith alternative, $(x_{i1}, \ldots, x_{iN})$,
- s = vector of observed characteristics of the decision maker, and
- μ_i = vector of unobserved characteristics of the decision maker and the ith alternative.

Without loss of generality, U_i can be separated into two parts: $V(x^i, s)$, a function of the observed data, and ϵ_i, a random component that is not observed; i.e.,

$$U_i = V(x^i, s) + \epsilon_i \qquad (4)$$

The nonstochastic term is called the representative utility and is specified to be linear in the parameters:

$$V(x^i, s) = \beta Z(x^i, s) \qquad (5)$$

where Z = vector-valued function of x^i and s and β = vector of the parameters.

Assume that alternative i is chosen if, and only if, it has greater utility than any other alternative; i.e., if $U_i > U_k$ for all $k \neq i$. Because the ϵ_i are random variables, the event $U_i > U_k$ for all $k \neq i$ is also random. The probability that the ith alternative is chosen is given by

$$P(i|C) = P(U_i > U_k) \quad (k \neq i) \qquad (6)$$

and, from Equation 4,

$$P(i|C) = P[\epsilon_k - \epsilon_i < V(x^i, s) - V(x^k, s)] \quad (k \neq i) \qquad (7)$$

To determine the probability that U_i satisfies Equation 7, we must know the probability distribution of ϵ_i. Assume that ϵ_i has a reciprocal exponential (Weibull) distribution, distributed identically and independently across all alternatives; i.e.,

$$P(\epsilon_i \leq t) = \exp(-e^{-t}) \qquad (8)$$

Given this assumption, it is possible to derive the MNL model (Equation 1) and the IIA property (Equation 2) (3, 4, 5).

Any significant violation of the assumptions of the MNL model will usually cause the IIA property to fail to be valid. Generally, the violations may be traced to the MNL assumption that the unobserved-utility component is independent across alternatives and independent of the observed attributes (1, 2).

Because the unobserved terms are defined simply as the difference between the true utility and the representative utility, the independence or nonindependence of the ϵ_i's depends on the specifications of the representative utility. In a given choice situation, two different specifications of representative utility will result in two different sets of ϵ_i's. One set of ϵ_i's might be independent, while the other might not. Thus, the IIA property might be valid for one specification of representative utility and not for another, even though both specifications relate to the same choice situation. This means that the IIA property is or is not valid for a particular specification of representative utility in a logit model of a particular choice situation, not for the choice situation itself. Consequently, it is meaningless to say, for example, that the IIA property is or is not valid for a traveler's choice of mode. It is only possible to state that the IIA property is or is not valid for a particular specification of the representative utility of the various modes.

Intuitively, the IIA property plays a role in the MNL model that is analogous to the assumption of independent-error terms in least-squares regression. The IIA property implies that the factors omitted from the analysis (the ϵ_i's) are independent random variables.

APPLICATION OF DIAGNOSTIC TESTS FOR THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES PROPERTY

Suppose a set of qualitative choice data is hypothesized to satisfy a particular specification of an MNL form. If the hypothesis is valid, the data and fitted models should have internal consistency properties; these can form the basis for diagnostic tests of the IIA property.

Model Specification

Table 2 presents an MNL model of the choice of mode for the work trip. The estimation was performed by the maximum likelihood method described by McFadden (5) on a sample of 641 workers in the San Francisco-Oakland Bay Area.

The model considers seven alternative modes: two automobile modes (automobile alone and car pool), two bus modes (one with walk access to bus and the other with automobile access), and three Bay Area Rapid Transit (BART) modes (with walk, bus, and automobile access). Most of the independent variables are self-explanatory, and their coefficients are readily interpreted. For example, the negative coefficient of on-vehicle time indicates that, when time spent in the vehicle for a particular mode increases, the probability of that mode being chosen decreases, all else being constant. Since the ratio of the on-vehicle time coefficient to the cost divided by wage coefficient is 0.43, the estimated value of on-vehicle time is 43 percent of the wage. Three income variables are included to allow for a nonlinear relation between income and the representative utility of the automobile-alone alternative. [These variables can be understood most readily by reference to Train (8).]

The model given in Table 2 seems particularly appropriate for testing violations of the IIA property. Because some of the alternative modes are similar, unobserved attributes of each mode may be correlated across modes. For example, the comfort on on-vehicle travel is similar for bus with walk access and bus with automobile access, and yet no comfort variable is included in the model. Failure of the IIA property could also result from the attributes of the alternatives not being exogenous. If the choice of how many automobiles to own is related to the work-trip modal choice, then the automobiles per driver variables are endogenous, and if the choice of where to live is related to the work-trip modal choice, then the cost and time variables are also endogenous.

Table 2. Work-trip modal-choice model.

Independent Variable	Estimated Coefficient	t-Statistic
Cost divided by posttax wage, ¢ ÷ ¢/min	-0.0380	6.83
On-vehicle time, min	-0.0162	1.91
Walk time,[a] min	-0.1006	4.25
Transfer-wait time,[a] min	-0.0122	0.923
Headway of first bus,[a] min	-0.0341	3.51
Automobiles per driver (ceiling of one)[b]	2.38	6.16
Automobiles per driver (ceiling of one)[c]	1.48	1.92
Dummy if person is head of household[b]	0.494	2.62
Number of persons in household who can drive[b]	0.5242	4.18
Number of persons in household who can drive[c]	0.7567	3.82
Family income (ceiling of $7500),[b] $/year	-0.000 308	2.18
Family income minus $7500 (floor of $0 and ceiling of $3000),[b] $/year	0.000 139	1.05
Family income minus $10 500 (floor of $0 and ceiling of $5000),[b] $/year	-0.000 096 6	1.78
Automobile-alone dummy[d]	-1.84	1.74
Bus-with-automobile-access dummy[e]	-5.38	5.69
BART-with-walk-access dummy[f]	1.94	3.18
BART-with-bus-access dummy[g]	-0.159	0.285
BART-with-automobile-access dummy[h]	-4.06	4.38
Car-pool dummy[i]	-2.39	5.28

Notes: Likelihood ratio index = 0.4119; log likelihood at zero = -982.6; log likelihood at convergence = -577.9; degrees of freedom = 2460; percent correctly predicted = 64.27; values of time saved as a percentage of wage: on-vehicle time = 43, walk time = 265, and transfer-wait time = 32 respectively.
All cost and time variables calculated for round trip. Dependent variable is alternative choice (1 for chosen alternative, 0 otherwise).
Sample size = 641.

[a] Variable is 0 for automobile-alone and car-pool alternatives and takes dummy value for other alternatives.
[b] Variable takes dummy value for automobile-alone alternative and is 0 otherwise.
[c] Variable takes dummy value for bus-with-automobile-access and BART-with-automobile-access alternatives and is 0 otherwise.
[d] Variable is 1 for automobile-alone alternative and 0 otherwise.
[e] Variable is 1 for bus-with-automobile-access alternative and 0 otherwise.
[f] Variable is 1 for BART-with-walk-access alternative and 0 otherwise.
[g] Variable is 1 for BART-with-bus-access alternative and 0 otherwise.
[h] Variable is 1 for BART-with-automobile-access alternative and 0 otherwise.
[i] Variable is 1 for car-pool alternative and 0 otherwise.

Universal Logit Method

The universal logit (UL) model is a more general model than is the MNL model; it takes advantage of the fact that every choice model with positive probabilities can be written in apparent MNL form, except that the scale function of alternative i will depend on attributes of other alternatives.

To use the universal logit method, a model is specified that includes all the variables in Table 2 plus some variables that are defined so that the attributes of one alternative are allowed to enter the representative utility of another. The hypothesis that the coefficients of all of the extra variables are zero is tested. If the hypothesis of zero coefficients is rejected, then the joint hypothesis of the MNL form and the specification in Table 2 is rejected.

The more general model includes the variables given in Table 2 and the following others:

1. Cost divided by posttax wage of automobile alone, with bus-with-walk-access and BART-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0;
2. Cost divided by posttax wage of bus with walk access with automobile-alone and BART-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0;
3. Cost divided by posttax wage of BART with walk access wage, with automobile-alone and bus-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0;
4. Total weighted time (sum of on-vehicle time, 2.5 times walk time, 1.25 times transfer-wait time, and 1.25 times first headway) of automobile alone, with bus-with-walk-access and BART-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0;
5. Total weighted time of bus with walk access, with automobile-alone and BART-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0; and
6. Total weighted time of BART-with-walk-access, with automobile-alone and bus-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0.

The null hypothesis that the interaction effects are zero (i.e., the MNL model and the IIA are true) can be tested by using the likelihood ratio; when the MNL model and the UL model are fitted by maximum likelihood estimation, the likelihood-ratio statistic

$$\chi^2 = 2(\text{log likelihood of UL model - log likelihood of MNL model}) \quad (9)$$

is asymptotically distributed chi square (χ^2) with degrees of freedom (df) equal to the number of parameter restrictions imposed by the null hypothesis. [This has been discussed in the NCHRP report (1) (Vol. 2, pp. C-172-175)].

The log likelihood at convergence for the more general model is -567.6. The log likelihood at convergence for the model in Table 2 is -577.9. Therefore, the test statistic (from Equation 9) is 20.6. The critical (0.05-level) value of χ^2 with six df is 12.6. The joint hypothesis that the MNL form and the specification of Table 2 are correct is rejected.

The signs of the coefficients of the extra variables are consistent with the hypothesis that the value of automobile on-vehicle time is higher than that of transit on-vehicle time. Variables 5 and 6 entered with negative signs (the latter with a t-statistic of 3.0), and the coefficient of variable 4 was estimated to be positive. Train (6) found the value of automobile on-vehicle time to

Table 3. Tests based on conditional choice.

Statistic	All Except BART Modes	All Except Bus Modes	All Except Bus and BART-With-Automobile-Access Modes	All Except BART-With-Walk-Access Mode	All Except Car-Pool Mode
	Alternatives Included in Subset				
Log likelihood at convergence for subsample choosing an alternative within subset of alternatives	-452.6	-400.1	-452.3	-557.8	-230.7
Log likelihood with coefficients restricted to values given in Table 2	-454.6	-403.3	-455.0	-557.9	-247.3
test	4.0	6.4	5.4	0.2	33.2
df	16	17	17	18	18
Critical (0.05 level) value of χ^2 with appropriate df	26.3	27.6	27.6	28.9	28.9

Table 4. Tests of association.

	Alternative											
	Automobile Alone			Car Pool			Bus With Walk Access			Bus With Automobile Access		
	No. of Residuals		Avg Probability for Cell	No. of Residuals		Avg Probability for Cell	No. of Residuals		Avg Probability for Cell	No. of Residuals		Avg Probability for Cell
Cell	Positive	Negative		Positive	Negative		Positive	Negative		Positive	Negative	
1	17	5	0.93	6	16	0.060	14	3	0.66	1	16	0.090
2	21	1	0.89	9	13	0.055	10	7	0.52	0	17	0.053
3	17	5	0.87	11	11	0.050	7	10	0.43	2	15	0.045
4	20	2	0.85	8	14	0.046	7	10	0.37	2	15	0.038
5	16	6	0.82	6	16	0.042	4	13	0.31	2	15	0.034
6	17	5	0.81	6	16	0.038	7	10	0.26	0	17	0.030
7	20	2	0.80	4	18	0.034	1	16	0.22	0	17	0.028
8	19	3	0.78	4	18	0.030	3	14	0.19	1	16	0.025
9	17	5	0.77	5	17	0.028	2	15	0.17	0	17	0.022
10	18	4	0.75	6	16	0.025	2	15	0.15	0	17	0.020
11	14	8	0.74	4	18	0.021	3	14	0.13	1	16	0.018
12	19	2	0.72	2	19	0.018	1	16	0.11	0	17	0.016
13	14	7	0.69	5	16	0.015	3	14	0.10	0	17	0.014
14	18	3	0.67	6	15	0.010	2	15	0.085	0	17	0.013
15	13	8	0.65	2	19	0.004	0	17	0.072	0	17	0.012
16	14	7	0.63	5	16	0.17	2	15	0.059	0	17	0.011
17	12	9	0.62	6	15	0.164	0	17	0.051	0	17	0.009
18	12	9	0.60	7	14	0.158	0	17	0.044	0	17	0.008
19	14	7	0.57	1	20	0.154	1	16	0.038	0	17	0.008
20	11	10	0.53	1	20	0.147	0	17	0.034	0	17	0.007
21	11	10	0.49	4	17	0.141	1	16	0.028	0	17	0.006
22	7	14	0.47	4	17	0.135	0	17	0.022	0	17	0.005
23	10	11	0.42	6	15	0.129	0	17	0.019	0	17	0.005
24	11	10	0.38	3	18	0.122	0	17	0.015	0	17	0.004
25	3	18	0.35	1	20	0.116	0	17	0.013	0	17	0.003
26	5	16	0.32	5	16	0.110	0	17	0.010	0	17	0.003
27	5	19	0.27	4	17	0.104	0	17	0.008	0	17	0.002
28	5	19	0.20	2	19	0.085	0	16	0.006	0	16	0.001
29	4	17	0.13	0	21	0.080	0	16	0.004	0	16	0.001
30	3	18	0.05	4	17	0.054	0	16	0.001	0	16	0.000

Table 5. Tests of association (continued).

	Alternative								
	BART With Walk Access			BART With Automobile Access			BART With Bus Access		
	No. of Residuals		Avg Probability for Cell	No. of Residuals		Avg Probability for Cell	No. of Residuals		Avg Probability for Cell
Cell	Positive	Negative		Positive	Negative		Positive	Negative	
1	0	12	0.057	6	6	0.390	1	4	0.206
	0	12	0.037	4	8	0.289	1	4	0.148
2	1	11	0.029	2	10	0.222	1	4	0.119
3	0	12	0.025	5	7	0.195	1	4	0.103
4	1	11	0.021	3	9	0.177	1	4	0.088
5	0	12	0.019	1	11	0.160	0	5	0.080
6	0	12	0.017	2	10	0.138	0	5	0.067
7	1	11	0.016	2	10	0.124	1	4	0.060
8	0	12	0.014	1	11	0.113	0	4	0.052
9	0	12	0.012	0	12	0.103	0	4	0.044
10	1	11	0.011	2	10	0.093	0	4	0.034
11	0	12	0.010	1	11	0.086	0	4	0.031
12	0	12	0.008	1	11	0.077	0	4	0.027
13	0	12	0.007	2	10	0.069	0	4	0.023
14	0	12	0.007	1	11	0.065	0	4	0.018
15	0	12	0.006	0	12	0.060	0	4	0.016
16	0	12	0.005	0	12	0.055	0	4	0.014
17	0	12	0.005	0	12	0.050	0	4	0.014
18	0	11	0.005	0	11	0.046	0	4	0.012
19	0	11	0.004	0	11	0.042	0	4	0.010
20	0	11	0.004	0	11	0.038	0	4	0.008
21	0	11	0.003	0	11	0.034	0	4	0.007
22	0	11	0.003	0	11	0.030	0	4	0.006
23	0	11	0.002	0	11	0.028	0	4	0.005
24	0	11	0.002	0	11	0.025	0	4	0.005
25	0	11	0.002	0	11	0.021	0	4	0.004
26	0	11	0.001	0	11	0.018	0	4	0.003
27	0	11	0.001	0	11	0.015	0	4	0.002
28	0	11	0.001	0	11	0.010	0	4	0.001
29	0	11	0.000	0	11	0.004	0	4	0.000
30				0	11				

be higher than that of bus on-vehicle time and gave the explanation that, while automobiles are more comfortable than transit, the difficulty of driving an automobile during rush-hour congestion makes automobile on-vehicle time more onerous than transit on-vehicle time. The model of Table 2 requires that automobile and bus times be valued equally; this constraint may contribute to the failure of the model of Table 2 in the test against the more general model.

Tests Based on Conditional Choice

If two dependent modes are included in the calibration sample, a different set of model coefficients will be generated than those generated from a model in which one of the the dependent modes is eliminated; i.e., violation of the IIA property will cause the maximum-likelihood parameter estimates to be biased. If the IIA property is valid, however, the coefficients estimated from the full choice set will coincide with the coefficients for a smaller choice set. An obvious test of the validity of the IIA property is whether or not the coefficients estimated from a reduced choice set are statistically different from those estimated from the full choice set.

In applying this test, the estimation is performed on the subsample of individuals who chose an alternative in the subset of alternatives to be tested for dependence. The coefficients of representative utility are estimated on the subsample and the log likelihood at convergence is calculated; the log likelihood is also calculated on the subsample with the coefficients restricted to the values given in Table 2. By using the likelihood-ratio test statistic analogous to that applied to the UL model test (Equation 9), the hypothesis that the coefficients estimated on the subsample are the same as those given in Table 2 is tested. The results of the tests for various subsets of alternatives are given in Table 3. The subsets chosen for testing were those that seemed most probable to cause rejection of the hypothesis of equal coefficients. For example, models similar to that of Table 2 estimated on a sample taken before BART was providing service greatly overpredict the use of BART with walk access; hence, the subset consisting of all alternatives except BART with walk access seemed particularly relevant for testing models based on conditional choice.

The hypothesis that the coefficients estimated on the subsample are the same as those of Table 2 (the hypothesis that the property IIA is valid) is accepted for each subset of alternatives except the subset that includes all alternatives except car pool. The failure of this test for this subset is probably the result of measurement errors in the observed attributes of the car-pool alternative. The exact attributes of the car-pool mode depend on such factors as the number of persons in the car pool, each person's home and work locations, and the allocations of costs among car-pool members. Because these variables cannot be determined for persons who do not choose car pool, crude estimates were used in calculating car-pool attributes.

Residuals Tests

Violations of the IIA property will cause systematic errors in the predicted choice probabilities. The difference between the observed choices and the predicted choice frequencies (the residuals) will therefore depend on whether the IIA property is valid or not (1).

To illustrate the way in which the residuals may be used to test the validity of the IIA property, suppose that an MNL model is estimated. Then the residuals

$$D_{jn} = (S_{jn} - R_n P_{jn})/(R_n P_{jn})^{1/2} \qquad (10)$$

can be defined, where

$n = 1, \ldots, N$ is a sample,
$P_{jn} = P(j \mid C, X_n, s_n)$ for $j \epsilon C$ is the estimated choice probability,
R_n = number of repetitions (possibly one) of sample point n, and
S_{jn} = number of choices j.

To avoid statistical dependence in the above residuals, it is sometimes more convenient to work with the transformed residuals,

$$Y_{jn} = D_{jn} - D_{ln}(P_{jn})^{1/2}[1 - (P_{ln})^{1/2}]/(1 - P_{ln}) \qquad (11)$$

where $l \epsilon C$ is a fixed alternative and $j \neq 1$. Under the hypothesis that the estimated model is correct, the residuals D_{jn} have, asymptotically, zero mean, unit variance, and covariance, $ED_{1n}D_{jn} = -(P_{1n}P_{jn})^{1/2}$. The residuals Y_{jn} are asymptotically independent and have zero mean and unit variance. Further discussion of these residuals and their properties has been given by McFadden (5).

Tables 4 and 5 present tests of association of the residuals and estimated probabilities of the model in Table 2. For each alternative, a contingency table is constructed as described by McFadden, Tye, and Train (2): The estimated probabilities for the alternative are ranked and classified into 30 cells, with each cell containing approximately the same number of cases, and the numbers of positive and negative residuals associated with the probabilities in a cell are counted. (The number of positive and negative residuals summed over all cells for a particular alternative is different for different alternatives because the number of persons in the sample who have a given alternative available varies among alternatives.)

If the MNL form and the specification of Table 2 are accurate, then the number of positive residuals is expected to be higher for low-numbered cells than for high-numbered cells (a positive residual is generated if the alternative was actually chosen). This pattern emerges for each alternative.

The goodness-of-fit test

$$\chi^2 = \sum_{m=1}^{M} (S_m - N_m \bar{P}_{jm})^2 / N_m \bar{P}_{jn} \qquad (12)$$

where

m = index of cell,
M = total number of cells,
S_m = number of positive residuals in cell,
N_m = total number of observations in cell,
$\bar{P}_{jm}$ = average probability for alternative j in cell m, and
$\bar{P}_{jn}$ = average probability of alternative j for total sample,

has an asymptotic distribution bounded by χ^2 distributions with M - 1 and M - K - 1 df, where K is the number of estimated parameters. These test statistics are not independent across alternatives.

The test statistic for each alternative is given below.

Alternative	Test Statistic
Automobile alone	17.51
Bus with walk access	14.14
Bus with automobile access	15.75
Car pool	38.63
BART with automobile access	13.81
BART with walk access	15.83
BART with bus access	5.32

Since there are 30 cells and 19 parameters, the test statistic has an asymptotic distribution, under the hypothesis that the MNL form and the specification of Table 2 are correct, bounded by χ^2 distributions with 29 and 10 df. The critical (0.05-level) value of χ^2 with 29 df is 42.56; that with 10 df is 18.31. The values of the test statistic for all alternatives except car pool are below the lower of the two bounding critical values, and therefore the hypothesis is accepted for those alternatives. For the car-pool alternative, the test statistic falls between the two bounding critical values: The test is therefore inconclusive. As in the failure of the test based on conditional choice, measurement errors in the car-pool attributes are probably the reason that the car-pool alternative cannot pass the test of association unambiguously.

Other Tests

Other tests using the properties of residuals are the means test and the variance test. Other tests that may be used are the saturated model test, which was found not to be powerful, and tests using two data sets. Tests using two data sets were found to be particularly powerful in identifying violations of the IIA property (2). For example, a before-and-after data set involving the introduction of a new mode offers a particularly powerful test of the independence of the mode. Both likelihood-ratio and residuals tests can be used. Another alternative that deserves consideration is to test the MNL model against the multinomial probit model with an explicit structure of dependence of unobserved attributes, which is practical if the number of choice alternatives is four or less (7).

Modifications of the Modal Choice Model to Correct for Violations of the Irrelevance of Independent Alternatives Property

The model given in Table 2 failed two of the tests of the IIA property. First, it failed the universal-logit test against a more general model with six extra cross-alternative variables. The probable reason for this failure is that the model constrains the value of automobile and transit time to be equal. Second, it failed the test of equality of coefficients across choice sets when the car-pool alternative was eliminated. The probable reason for this failure is that the car-pool data were poor.

A new model of work-trip modal choice has been given by Train (8). This model is more general than the model given in Table 2 in that, among some other generalizations, automobile and transit on-vehicle times are allowed to have different coefficients and socioeconomic variables are allowed to enter the car-pool alternative.

This more general model passed both of the diagnostic tests that the MNL model (Table 2) failed:

1. The universal logit test—the log likelihood of Train's model is -519.9. The log likelihood of the more general model (which includes the six cross-alternative variables) is -515.5. Therefore, the test statistic is 8.8. The critical (0.05-level) value of χ^2 with six df is 12.6. The model passes the test.
2. The test of equality of coefficients across choice sets—the log likelihood of Train's model with the car-pool alternative removed is -191.0. The log likelihood of the model with the car-pool alternative removed and the parameters restricted to those obtained with all alternatives included is -199.4. The test statistic, therefore, is 16.8. The critical (0.05-level) value of χ^2 with 23 degrees of freedom is about 35. The hypothesis of equal parameters is accepted, and the model passes the test.

These results illustrate that the passing or failing of the diagnostic tests depends on the specification of the model for a particular choice situation, not on the choice situation itself; i.e., the IIA property is or is not valid for a particular model, not for a particular choice situation. These results indicate the way in which the diagnostic tests can be used to find problems in the specification of the model. The diagnostic tests applied to the model in Table 2 indicated that there were problems in the on-vehicle-time variable and the car-pool alternative; these problems were corrected in Train's model.

ACKNOWLEDGMENT

The development of the concepts and statistical tests of this research was supported by the National Cooperative Highway Research Program. The application of the model calibration and of the tests was supported by the National Science Foundation Research Applied to National Needs Program. The opinions and conclusions expressed or implied are those of the authors. They are not necessarily those of the Transportation Research Board, the National Academy of Sciences, the Federal Highway Administration, the American Association of State Highway and Transportation Officials, the individual states participating in the National Cooperative Highway Research Program, or of the National Science Foundation, the University of California, or Charles River Associates, Inc.

REFERENCES

1. The Independence of Irrelevant Alternatives Property of the Multinomial Logit Model. In Disaggregate Travel Demand Models, Charles River Associates, Cambridge, Mass.; NCHRP, Phase 1 Rept., Vols. 1 and 2, 1976.
2. D. McFadden, W. Tye, and K. Train. Diagnostic Tests for the Independence From Irrelevant Alternatives Property of the Multinomial Logit Model. Institute of Transportation Studies, Univ. of California, Berkeley, Working Paper No. 7616, Aug. 1976.
3. A Disaggregated Behavioral Model of Urban Travel Demand. Charles River Associates, Cambridge, Mass., 1972.
4. T. Domencich and D. McFadden. Urban Travel Demand: A Behavioral Analysis. North-Holland, Amsterdam, 1975.
5. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (P. Zarembka, Ed.), Academic, New York. 1973.
6. K. Train. Work-Trip Mode Split Models: An Empirical Exploration of Estimate Sensitivity to Model and Data Specification. Institute of Transportation

Studies, Univ. of California, Berkeley, Working Paper No. 7602, 1976.
7. J. Hausman and D. Wise. A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences. Department of Economics, MIT, Cambridge, Working Paper No. 173, April 1976.
8. K. Train. A Post-BART Model of Mode Choice: Some Specification Tests. Institute of Transportation Studies, Univ. of California, Berkeley, Working Paper No. 7620, 1976.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

Effects of Parking Costs on Urban Transport Modal Choice

David W. Gillen, Department of Economics, University of Alberta, Edmonton

The effects of parking costs on urban modal choice are investigated by using a standard binary-choice model and estimated by using the logit technique. Previous studies have misspecified the form of the parking-cost variable and the model normally estimated. After estimating the traditional and correctly specified models, the claim that parking taxes are an effective substitute for roadway pricing in influencing congestion is only partially supported. Aggregate elasticities for four policy-oriented variables are calculated. The elasticities provide a measure of the bias from misspecification and indicate the most effective policy variable for the reduction of automobile use.

This paper is concerned with estimating the effects of parking costs on individual choice of transportation mode for trips within urban areas. It has three basic objectives:

1. The determination of how to characterize the parking variable and incorporate it into a model of modal choice,
2. The calculation of the elasticity of modal choice with respect to parking costs, and
3. The determination of the way in which changes in one of the characteristics that determine the modal choices of individuals will affect the expected proportion of individuals taking the choice being considered.

{The third objective is implemented by examining the way in which changes in individual characteristics affect the mean of the distribution of population probabilities [cf. Westin (11)]}.

The first section introduces a model of individual choice of transportation alternatives that treats parking as a commodity, the demand for which is derived from the choice of the automobile as the transit mode. The second section describes the data and the implications of this model for the structural forms of the estimating equations. The third section presents the empirical results for an application of this model to data for Toronto. The fourth section presents the derivation of the elasticity of modal choice with respect to instrumental variables and empirical results for aggregate and individual elasticities.

BASIC MODEL

The variable to be explained is the individual's choice of transportation mode (automobile versus public transit). The econometric model used in this paper to represent this binary-choice problem is derived from a choice-theoretic framework, based on a microeconomic behavioral model developed by DeSerpa (2), in which individuals maximize utility in choosing among alternative goods and the times allocated to them, subject to income and time-resource constraints. In this model, the choice of any amount (X_i) of commodity i places only a lower bound on the amount of time (T_i) the individual must use in consuming X_i; a change in relative prices of either goods or times causes the individuals to substitute among goods of various time intersities and, therefore, to implicitly substitute among alternative uses of time. Others (5, 6, 7, 9, 10) have used similar theoretical approaches to demonstrate the relation between the microeconomics of choice behavior and binary-choice economitric models. These models suggest, that modal choice is a function of two categories of variables, transportation-system characteristics that affect the money and time costs of travel and user characteristics that serve as proxies for objective comfort characteristics.

Traditionally, modal-choice studies have simply added the costs of parking to the automobile running costs (7, 10). This procedure implicitly assumes that parking services and automobile use enter into the individual's production function in fixed proportions. It also implies that the decision about where to park is independent of modal choice, so that parking-location decisions are unaffected by variations in time costs.

In this paper, parking is defined as a commodity that is complementary to automobile trips. The individual is assumed to maximize a utility unction [$U(C_i)$], where $C_i = F(X_i, T_i)$, subject to income and time-resource constraints. The explicit specification of the production functions that determine C_i is important for understanding the role of parking use. For transit, the service consumed is generated by the production function

$$C_T = F_T(X_T,T_T) \quad (1)$$

where X_T = transit service purchased and T_T = time spent in using X_T; for automobile use, the service consumed is generated by the function

$$C_A = F_A(X_A,T_A,X_P,T_P) \quad (2)$$