Transportation Center and Department of Civil Engineering, Northwestern University, Evanston, Ill. 1984.
3. S. R. Lerman and M. E. Ben-Akiva. "Disaggregate Behavioral Model of Automobile Ownership." In *Transportation Research Record 569*, TRB, National Research Council, Washington, D.C., 1976, pp. 34–55.
4. C. A. Lave, S. Mehring, and R. Kuzmyak. "Price Elasticites of Intercity Passenger Travel." *Proc., Transportation Research Forum*, 1977.
5. G. S. Cohen, N. S. Erlbaum, and D. T. Hartgen. "Intercity Rail Travel Models." In *Transportation Research Record 673*, TRB, National Research Council, Washington, D.C., 1978, pp. 21–25.
6. S. F. DiRenzo and L. P. Rossi. "Diversion Model for Estimating High-Speed Rail Use." In *Highway Research Record 369*, HRB, National Research Council, Washington, D.C., 1971, pp. 15–25.
7. A. Stachurski and R. Rice. "Development of Direct Intercity Rail Passenger Demand Models." *Proc., World Conference on Transportation Research*, London, England, April 1980.
8. G. Kraft. *Demand for Intercity Passenger Travel in the Washington-Boston Corridor*. NTIS: PB166–884. Systems Analysis and Research Corporation, Boston, Mass., 1963.
9. R. E. Quandt and W. J. Baumol. The Abstract Mode Model: Theory and Measurement. *Journal of Regional Science*, Vol. 6, No. 2, 1966, pp. 13–26.
10. S. E. Eriksen. *Demand Model for U.S. Domestic Air Passenger Markets*. Flight Transportation Report R78–2. Massachusetts Institute of Technology, Cambridge, June 1978.
11. P. L. Watson. "Comparison of the Model Structure and Predictive Power of Aggregate and Disaggregate Models of Intercity Mode Choice." In *Transportation Research Record 527*, TRB, National Research Council, Washington, D.C., 1974, pp. 59–65.
12. G. R. Leake and J. R. Underwood. Comparison of Intercity Bi-Modal Split Models. *Transportation Planning and Technology*, Vol. 5, 1978, pp. 55–69.
13. *1977 National Personal Transportation Study, User Guide for Public Use Tapes*. FHWA, U.S. Department of Transportation, 1980.
14. *National Travel Survey—Travel During 1977*. Bureau of the Census, U.S. Department of Commerce, Oct. 1979.

Publication of this paper sponsored by Committee on Intercity Passenger Guided Transportation.

Constraints on Individual Travel Behavior in a Brazilian City

JOFFRE D. SWAIT, JR., AND MOSHE BEN-AKIVA

In this paper the statistical and predictive performance of two disaggregate choice models that incorporate probabilistic choice set formation are compared with a standard logit specification. The empirical work is conducted with work mode choice data from a Brazilian city. For the type of travel demand analyzed it is found that, although statistically inferior to the probabilistic choice set specifications, the standard logit specification, allied with market segmentation, is a robust formulation in both statistical and predictive terms. Recommendations for future research work in probabilistic choice set modeling are presented.

The principal issue addressed by this paper is the appropriateness of choice theory, as it is now interpreted, for modeling travel demand. In a highly constrained environment, such as can be found in low-income areas, observed choice may well be the result of the elimination of alternatives through active constraints, as opposed to the exercise of a choice prerogative by the decision maker.

The effect of constraints on travel behavior is particularly important for analyses in developing nations. Swait et al. *(1)* present an extensive discussion of a disaggregate travel demand model system for a medium-sized Brazilian city. Because of its unique nature, many substantive conceptual and modeling issues have arisen during the course of the study. These issues highlight fundamental differences between developed and developing countries in terms of travel demand. These practical experiences and conceptual concerns have led to the investigation and formulation of a number of probabilistic choice set formation models and to empirical testing of these to investigate their performance with respect to choice models with fixed choice sets.

The overall methodology and the alternative models that incorporate probabilistic choice sets are described in Ben-Akiva and Swait *(2)* and in Swait and Ben-Akiva *(3)*. In this paper two of these models are implemented with data for work mode choice from Maceió, Brazil, and their statistical fit and forecasts are compared with those of a standard logit model.

HYBRID APPROACH TO MODELING CHOICE SET GENERATION

The approach used in this work is based on the following two-stage choice process: first, constraints (of a personal, household, and social nature) act on the individual to define his choice set; second, the individual exercises choice according to some decision rule.

From the perspective of an analyst who normally does not know either the specific alternatives that constitute an individual's choice set or the exact decision rule used to make a choice, the two-step choice paradigm leads to the following probability of observing alternative j being chosen by individual n *(4)*:

$$P_n(j) = \sum_{C \in G_n} P_n(j|C)\, P_n(C) \tag{1}$$

Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

where

M = the universal choice set, made up of all possible alternatives available for the choice context and population in question;

M_n = the set of all deterministically feasible alternatives for individual n ($M_n \subseteq M$);

G_n = the set of all nonempty subsets of M_n; and

C = an element of G_n ($C \subseteq M_n$).

Expression 1 reflects a three-part model of the choice process:

1. A probabilistic choice model, $P_n(j|C)$, conditioned on the choice set being $C \in G_n$, which by definition yields choice probabilities of zero for $j \notin C$;
2. A deterministic choice set generation model that determines the subset M_n from the set M; and
3. A probabilistic choice set generation model, $P_n(C)$, expressing the probability that set $C \subseteq M_n$ is the individual's actual choice set.

This reflects the assumption that the analyst may be willing to impose certain constraints deterministically because of a high level of assurance about their effect (e.g., no automobile driver mode for individuals without a driver's license) but unwilling to do the same for other constraints (e.g., acceptable walk access distances at the origin and destination of a specific trip).

A high degree of computational complexity is implied by Expression 1. If $|C|$ denotes the number of elements in set C, then $|G_n|$ is equal to $(2^{|M_n|} - 1)$, of which $(2^{|M_n|-1})$ choice sets actually contain any given alternative $j \in M_n$. To illustrate how the number of possible choice sets can quickly become overwhelming, if M_n has 3 alternatives, then 4 terms must be summed; with 10 alternatives, the number of possibilities has increased to 512. These sizes are applicable for model estimation; for prediction, when choice probabilities must be evaluated for all the alternatives in M_n, there are $(2^{|M_n|} - 1)$ possible choice sets (e.g., if $|M_n| = 3$ then $|G_n| = 7$).

Most choice contexts of interest are, unfortunately, characterized by many, rather than few, alternatives. A possible approach to reducing the dimensionality of the choice set generation problem is to place a priori restrictions of the members of G_n. That is, modeling the choice situation at hand requires only a subset of the $(2^{|M_n|} - 1)$ possible sets. One useful restriction is the captivity model, in which an individual is assumed either to be captive to a single alternative or to be free to choose from among the full set of deterministically available alternatives. Assuming that the choice model has the logit form, the following logit captivity model is obtained:

$$P_n(i) = [\delta_i/(1 + \sum_{j \in M_n} \delta_j)] + [1/(1 + \sum_{j \in M_n} \delta_j)]$$

$$\exp(V_{in}) / \sum_{l \in M_n} \exp(V_{ln}) \qquad (2)$$

where V_{in} is the systematic utility of the ith alternative, and δ is a vector of nonnegative parameters that represents the odds of the individual being captive to each specific alternative. The first term on the right side of Expression 2 represents the probability that the individual is captive to alternative i, in which case the probability of i being chosen is obviously one. The second term has two parts: the one involving the δ vector represents the probability that the choice is to be from the full choice set M_n, and the other is a logit model of the probability of choosing i given that the choice set is M_n. The reader is referred to Ben-Akiva and Swait *(2)* for a more detailed development of this model.

This logit captivity model was derived by McFadden (unpublished memorandum of September 30, 1976), Ben-Akiva *(5)*, and Gaudry and Dagenais *(6)*, the first two motivated by the probabilistic captivity concept and the last, who refer to this model as "dogit," by the desire to circumvent the independence of irrelevant alternatives (IIA) property of the logit model.

A probabilistic choice set formation model with no a priori restrictions on G_n will also be used here. One specific model, called the independent availability logit model, assumes that the probability of availability of an alternative is independent of the availability or lack thereof of any other alternative. This strong assumption is necessary to achieve a manageable model specification. The mathematical formulation of this model is

$$P_n(C) = \prod_{i \in C} \gamma_i \prod_{j \in M_n - C} (1 - \gamma_j) / [1 - \prod_{l \in M_n} (1 - \gamma_l)],$$

$$C \in G_n \qquad (3)$$

$$P_n(j|C) = \exp(V_{jn}) / \sum_{i \in C} \exp(V_{in}), \quad j \in C \qquad (4)$$

where γ_i is the probability that alternative $i \in M_n$ is available and other quantities are as previously defined. The notation $M_n - C$ denotes the set of the alternatives in M_n less the alternatives in C. Expressions 3 and 4 can be substituted into Expression 1 to obtain the unconditional probability of choice of an alternative.

In Expression 3 the first term in the numerator represents the probability of availability of all of the alternatives in C, and the second the probability of unavailability of all the alternatives in M_n not in C. The denominator is a normalization factor to exclude the event of all alternatives being unavailable.

In the two models, the representation of constraints is done in a simple manner, either by the captivity restriction on possible choice sets or by the simplifying assumption of independent availability. In addition, in both specifications the aggregate impact of these constraints is represented by a single parameter per alternative (i.e., δ_i and γ_i, $i \in M_n$, in the captivity and independent availability models, respectively). Swait and Ben-Akiva *(3)* present an example of a logit captivity model in which this latter restriction is relaxed.

The calibration results of standard logit, logit captivity, and independent availability logit models of work mode choice for Maceió, Brazil, are presented in the following section. The various models are compared on the basis of statistical performance. Following this, the three models are used to produce forecasts in a variety of policy scenarios. These forecasts are compared and their implications are discussed.

ESTIMATION RESULTS

Choice Context

The city of Maceió and its travel patterns have been extensively described in Swait et al. *(1)* and in Geltner and Barros *(7)*. The particular choice dimension to be investigated is that of home-based work mode choice for full-time workers. Because of the widespread habit of returning home for lunch and important policy implications of this type of behavior, the unit of observation is the modal choice pattern for a working day. An investigation of the observed modal choice patterns of Maceió workers, captured in a 1977 household survey, reveals that fewer than 5 percent of the workers chose travel patterns that involved more than one mode. Hence, the universe of alternatives (i.e., the set *M*) is reduced (for modeling purposes) to

- Bus,
- Taxi,
- Automobile driver, and
- Automobile passenger.

Thus "modal alternative" actually refers to the use of that mode by the worker for all home-based work trips taken that day.

The following deterministic constraints were applied to the alternatives:

- The automobile driver alternative is available only to individuals from automobile-owning households who are 18 years or older (no information was available on driver's license) and
- If the one-way network travel time for the mode is greater than 2 hr, it is unavailable.

Thus bus, taxi, and automobile passenger are ubiquitous; automobile driver is limited to those eligible for a driver's license and whose households own a vehicle. The travel time limitation is a further imposition.

To provide a basis for comparison, a standard logit model is first estimated with a random sample of 1,477 workers. Next, market segmentation is used as a first attempt to account for the impact of constraints and taste variations on choice. Following that, the estimation results of the logit captivity and independent availability logit models are presented in turn.

Standard Logit Model

Table 1 gives the estimation results for the standard logit model of home-based full-time worker mode choice for the full data set and three income market segments. The 19 parameter models include time, cost, income, family size, automobile availability, and role-related variables, which, with one exception, show high levels of significance and correct signs. Though no extensive efforts were expended to obtain an improved specification, it is believed that the pooled model as it now stands represents a reasonable standard for comparison.

Inclusion of variables such as automobile availability, income, and family size in the utility functions of alternatives can be interpreted as an ad hoc model of alternative availability in much the same way that size variables are used to correct for aggregation of alternatives in logit models of destination choice *(8)*.

To maintain uniformity during model comparisons, this same specification has been used for the choice model throughout the study of Maceió; exceptions have been opened only in the case of unidentifiability.

Market segmentation is a useful technique for accounting for taste variations in a population, but it can also be used to bring out the impact of constraints on choice. The market segmentation used is based on household income; for Maceió monthly household electrical energy consumption is used as a proxy measure of income [see Swait et al. *(1)* for more discussion of this measure]. The three income groups are *(a)* less than 80 KwH/month (low), *(b)* 80 to 130 KwH/month (medium), and *(c)* greater than 130 KwH/month (high). Because Maceió is located in an economically depressed area of Brazil, it is to be expected that income should play a significant role in determining mode choice. For the three income segment logit models in Table 1 the hypothesis of parameter equality across the segments is rejected with a very high level of significance (more than 99 percent for a chi-squared statistic of 71.2, compared with a critical value of 63.7 with 40 degrees of freedom). The apparent parameter differences appear to be concentrated mainly in the socioeconomic attributes, such as income (the significance of which is quite diminished in the income segment models, which indicates that the segmentation has reduced within-group variation with respect to this variable), household size, and automobile availability. The travel impedance parameters are not very different across segments.

Although the market segmentation results are encouraging, it is impossible to attribute any part of the improvement to a better choice model specification because of accounting for taste variations, or to improved "modeling" of constraints on choice with ad hoc availability variables.

Logit Captivity Model

The logit captivity model represents a choice context in which the decision maker either is captive to one alternative or is free to choose from the full set of available alternatives. Table 2 gives the estimation results for this specification; the choice model parameters (i.e., those for the logit model) are directly comparable with the parameters in Table 1. Note that the model in Table 2 maintains the hypothesis of no captivity to the automobile driver mode for workers who have this alternative. Although this restriction appears to be plausible for the city of Maceió, it is important to realize that this restriction is not arbitrary: it is the result of the parameter being driven to zero during optimization of the log-likelihood function for the Maceió sample. This type of parameter restriction will be seen in each of the choice set models presented in this paper.

First, the standard logit and logit captivity estimated with the full sample are compared. With a chi-squared statistic of 3.8 with 3 degrees of freedom, the hypothesis, at a 90 percent significance level, that the captivity parameters are jointly zero for the pooled sample cannot be rejected. Further, the hypothesis that each parameter is individually zero also cannot be rejected at reasonable significance levels. Thus there appears to be little evidence of captivity for the sample of workers as a whole. This is not, of course, a surprising result: the radical choice set structure (i.e., captivity or complete freedom of choice) of this model is unlikely to be generally applicable to the population.

This lack of significant improvement over the fit of the logit specification and the significant improvement obtained by the income segmentation (Table 1) compared to the pooled sample led to the hypothesis that evidence of captivity could perhaps be uncovered by calibrating logit captivity models by income group.

The income segmentation results for the logit captivity model in Table 2 indeed bring to light significant captivity to the bus mode in the low-income group and to the bus and automobile passenger modes in the medium-income group. There is indicated a small degree of captivity to automobile passenger in the high-income

TABLE 1 MACEIÓ HOME-BASED WORK TOUR MODE CHOICE MODEL—LOGIT SPECIFICATION

Parameters	Estimated Parameters: Low Income	Medium Income	High Income	All
Alternative-specific constants				
Bus	0	0	0	0
Taxi	0.05 (0.0)	-1.30 (-0.5)	0.78 (0.7)	-1.29 (-2.5)
Automobile passenger	-3.47 (-2.0)	-3.30 (-1.8)	-2.17 (-2.3)	-2.88 (-7.5)
Automobile driver	1.14 (0.6)	0.50 (0.3)	0.12 (0.1)	0.02 (0.0)
Total travel time (min/day)	-0.008 (-1.0)	-0.014 (-2.3)	-0.011 (-1.8)	-0.012 (-3.5)
Total travel cost (Cr$ 1977/day divided by ℓn (household income, Kwh/month)	-0.245 (-2.1)	-0.318 (-3.8)	-0.342 (-2.6)	-0.296 (-5.9)
Household income (KwH/month)				
Bus	0	0	0	0
Taxi	0.001 (0.1)	-0.010 (-0.4)	0.001 (0.3)	0.005 (2.7)
Automobile passenger	0.020 (1.0)	0.016 (0.9)	0.003 (1.2)	0.007 (5.1)
Automobile driver	0.008 (0.4)	0.026 (1.5)	0.002 (0.9)	0.006 (4.8)
No. of household members				
Bus	0	0	0	0
Taxi	-0.44 (-2.5)	0.18 (1.7)	-0.32 (-2.9)	-0.11 (-2.1)
Automobile passenger	0.26 (-1.9)	-0.16 (-1.7)	-0.17 (-2.1)	-0.15 (-3.0)
Automobile driver	-0.26 (-1.1)	-0.57 (-4.1)	-0.17 (-2.2)	-0.21 (-3.9)
Automobile availability (cars/workers)				
Bus	0	0	0	0
Taxi	3.94 (2.8)	1.37 (1.4)	2.03 (3.3)	1.81 (4.3)
Automobile passenger	3.62 (2.8)	2.47 (6.2)	3.84 (8.3)	3.10 (11.3)
Automobile driver	2.04 (2.1)	1.89 (3.3)	3.92 (8.2)	2.88 (9.6)
CBD work location and lunch trip home				
Bus	0	0	0	0
Taxi	0.3 (0.3)	0.3 (0.5)	1.3 (2.4)	0.7 (2.1)
Automobile passenger and driver	-0.3 (-0.4)	-0.5 (-1.2)	0.1 (0.3)	-0.2 (-0.7)
Female worker				
Bus, taxi, automobile passenger	0	0	0	0
Automobile driver	-1.7 (-1.8)	-2.0 (-4.2)	-1.6 (-5.0)	-1.7 (-7.0)
Professional worker and lunch trip home				
Bus	0	0	0	0
Taxi	-0.5 (-0.4)	0.4 (0.5)	1.4 (2.8)	0.8 (2.3)
Automobile passenger and driver	0.9 (1.2)	0.6 (1.4)	1.5 (3.6)	1.0 (4.2)
Summary Statistics				
Log-likelihood at zero	-596.5	-532.8	-649.8	-1779.0
Log-likelihood at convergence	-129.0	-237.3	-319.0	-720.9
Rho-squared	0.7837	0.5546	0.5090	0.5948
Adjusted rho-squared[a]	0.7519	0.5190	0.4798	0.5841
Sample Description				
Choosing				
Bus	467	301	163	931
Taxi	12	18	25	55
Automobile passenger	14	39	67	120
Automobile driver	35	88	248	371
Total observations	528	446	503	1,477

Note: Asymptotic *t*-statistics in parentheses for the hypothesis the parameter is zero.
[a]See Expression 5 for the definition of this measure.

group, but this is not statistically significant because of the large variance of the respective captivity parameter. It is clear that the income segment captivity models are a statistically significant improvement over the pooled logit model of Table 1 and the pooled logit captivity model of Table 2.

The income segment logit captivity models are also jointly a statistically significant improvement over the income segment logit models of Table 1. The hypothesis that the captivity parameters are all jointly zero is tested by using a chi-squared statistic of 24.8, which can be compared with a critical value of 23.2 at the 99 percent significance level with a conservative 10 degrees of freedom. Therefore this hypothesis is rejected; the data indicate that in addition to the taste variations that are captured by the income segmentation in Table 1, there is a variation in the choice set structure of individuals that must be accounted for in the choice model specifications. Bear in mind, however, that the major source of improvement stems not from choice set modeling but from income segmentation.

Also note that the logit captivity models provide statistically better fit across all three income segments than do their standard logit counterparts of Table 1. It is also interesting to note some of the significant changes that have occurred in certain individual parameters of the logit model utilities.

Consider, for example, the coefficients of the travel time and cost variables. Those in the logit captivity models are uniformly larger than the corresponding parameters in Table 1. Conceptually, the removal of captives from consideration in the calibration of the choice model removes their diluting effect on its parameters; only the true choosers affect the choice model parameters. Indeed, all of the travel impedance and socioeconomic parameters grow in magnitude, some of them quite drastically (e.g., automobile availability in the low-income group).

TABLE 2 MACEIÓ HOME-BASED WORK TOUR MODE CHOICE MODEL—LOGIT CAPTIVITY SPECIFICATION

	Estimated Parameters			
	Low Income	Medium Income	High Income	Pooled
Choice Model Parameter				
Alternative-specific constants				
Bus	0	0	0	0
Taxi	2.14 (0.6)	0.37 (0.1)	1.35 (1.0)	-1.13 (-2.0)
Automobile passenger	-2.16 (-0.8)	-5.72 (-1.3)	-2.32 (-2.0)	-3.02 (-6.3)
Automobile driver	2.79 (0.7)	2.15 (0.6)	0.17 (0.2)	-0.07 (-0.2)
Total travel time (min/day)	-0.012 (-0.4)	-0.041 (-2.4)	-0.013 (-1.8)	-0.014 (-3.4)
Total travel cost (Cr$ 1977/day) divided by ℓn (household income, KwH/month)	-0.543 (-1.4)	-0.826 (-2.8)	-0.451 (-2.8)	-0.356 (-4.7)
Household income (KwH/month)				
Bus	0		0	0
Taxi	0.112 (0.9)	-0.014 (-0.4)	0.001 (0.3)	0.006 (2.7)
Automobile passenger	0.034 (1.0)	0.029 (0.7)	0.002 (0.8)	0.007 (4.2)
Automobile driver	0.023 (0.6)	0.041 (1.3)	0.001 (0.4)	0.006 (3.9)
No. of household members				
Bus	0	0	0	0
Taxi	-3.12 (-1.2)	0.24 (1.7)	-0.36 (-2.8)	-0.11 (-1.9)
Automobile passenger	-0.71 (-1.6)	-0.36 (-1.4)	-0.25 (-2.0)	-0.17 (-2.7)
Automobile driver	-0.58 (-1.1)	-1.06 (-2.9)	-0.22 (-2.2)	-0.22 (-3.6)
Automobile availability (cars/workers)				
Bus	0	0	0	0
Taxi	11.67 (1.7)	2.58 (1.7)	2.11 (1.5)	1.87 (3.1)
Automobile passenger	10.79 (2.1)	4.61 (3.1)	5.16 (5.0)	3.63 (8.1)
Automobile driver	10.43 (2.2)	2.83 (2.0)	5.57 (5.3)	3.44 (7.6)
CBD work location and lunch trip home				
Bus	0	0	0	0
Taxi	-0.2 (-0.1)	1.8 (1.8)	1.7 (2.5)	0.8 (2.2)
Automobile passenger and driver	-0.3 (-0.2)	-0.4 (-0.5)	0.2 (0.4)	-0.1 (-0.5)
Female worker				
Bus, taxi, automobile passenger	0	0	0	0
Automobile driver	-4.4 (-3.0)	-3.2 (-3.1)	-1.9 (-4.3)	-1.7 (-6.6)
Professional worker and lunch trip home				
Bus	0	0	0	0
Taxi	NI[a]	0.2 (0.2)	1.8 (2.9)	1.0 (2.4)
Automobile passenger and driver	NI	1.2 (1.5)	1.9 (2.9)	1.1 (3.8)
Captivity Parameters				
Bus	0.167 (2.1)	0.099 (2.0)	0.011 (1.1)	0.013 (1.2)
Taxi	0.016 (2.4)	0.010 (1.3)	0.007 (1.1)	0.004 (0.9)
Automobile passenger	0.007 (1.0)	0.058 (2.7)	0.044 (1.4)	0.008 (0.8)
Automobile driver	0	0	0	0
Summary Statistics				
Log-likelihood at $\beta = 0, \delta = 0$	-596.5	-532.8	-649.8	-1779.0
Log-likelihood at convergence	-124.9	-234.5	-313.5	-719.0
Rho-squared	0.7907	0.5599	0.5175	0.5958
Adjusted rho-squared	0.7571	0.5186	0.4837	0.5835

Note: Asymptotic *t*-statistics in parentheses.
[a]NI = not identifiable or not included.

These large changes have occurred in the presence of what is not a large captivity effect. After all, the most significant degree of captivity is that to bus in the low-income segment where there is an estimated probability of 0.14 of captivity to that mode. To arrive at this figure, the first term on the right of Expression 2 was used so that the likelihood of captivity to bus is 0.167/1.90 ≈ 0.14.

Independent Availability Logit Model

Table 3 gives the estimation results for the independent availability logit model. First, models for the full data set will be compared. Unlike the captivity model, the independent availability model provides a significantly better fit to the pooled sample than does the standard logit model: the hypothesis that the availability parameters are all jointly one (indicating deterministic availability of all alterntives in M_n for all individuals) is rejected at the 95 percent level. This improvement is explained by the independent availability model's complete representation of the choice set structure as opposed to the extreme assumption underlying the captivity model of the previous section.

Again, as in the case of the logit captivity models, a general increase in the magnitude of the choice model parameters is noted. In the independent availability model, this increase is attributable not only to consideration of captivity but also to consideration of all of the trade-off situations that each decision maker can possibly face. For example, if an individual has available bus (B), taxi (T), and automobile passenger (AP), not only is there a probability that his choice of bus is from the set (B,T,AP), but there is now a probability that the choice is from (B,T) and (B,AP).

The improvement in fit provided by the pooled independent availability logit model, albeit statistically significant, is certainly not dramatic (the chi-squared statistic is 11.8 with 4 degrees of freedom, compared with a critical value of 9.5 at a 95 percent signficance level). Once again, this has led to segmentation of the sample of workers along the income dimension and estimation of separate models for each (Table 3). The hypothesis that all of the parameters are equal across income segments can be rejected at the

TABLE 3 MACEIÓ HOME-BASED WORK TOUR MODE CHOICE MODEL—INDEPENDENT AVAILABILITY LOGIT

	Estimated Parameters			
	Low Income	Medium Income	High Income	Pooled
Choice Model Parameters				
Alternative-specific constants				
Bus	0	0	0	0
Taxi	1.35 (0.6)	-1.28 (-0.5)	1.75 (1.1)	-0.44 (-0.3)
Automobile passenger	-1.80 (-0.7)	-3.41 (-1.9)	-2.28 (-1.9)	-2.91 (-6.6)
Automobile driver	1.75 (0.8)	1.20 (0.5)	0.46 (0.4)	0.82 (1.1)
Total travel time (min/day)	0.005 (0.4)	-0.017 (-2.5)	-0.015 (-2.1)	-0.015 (-3.5)
Total travel cost (Cr$ 1977/day) divided by ℓn(household income, KwH/month)	-0.238 (-1.2)	-0.336 (-3.8)	-0.450 (-2.9)	-0.325 (-5.2)
Household income (KwH/month)				
Bus	0	0	0	0
Taxi	0.022 (0.6)	-0.010 (-0.4)	0.001 (0.4)	0.005 (2.3)
Automobile passenger	0.037 (1.3)	0.017 (0.9)	0.003 (1.0)	0.007 (4.3)
Automobile driver	0.018 (0.9)	0.040 (1.5)	0.001 (0.2)	0.007 (3.5)
No. of household members				
Bus	0	0	0	0
Taxi	-1.35 (-1.6)	0.18 (1.6)	-0.34 (-2.4)	-0.11 (-1.8)
Automobile passenger	-0.57 (-2.0)	-0.17 (-1.7)	-0.17 (-1.8)	-0.15 (-2.9)
Automobile driver	-0.36 (-1.4)	-0.83 (-2.9)	-0.23 (-1.9)	-0.31 (-3.3)
Automobile availability (cars/workers)				
Bus	0	0	0	0
Taxi	4.70 (2.6)	1.31 (1.3)	2.19 (2.2)	2.06 (3.4)
Automobile passenger	4.55 (2.6)	2.39 (5.6)	5.40 (5.6)	3.59 (8.2)
Automobile driver	2.20 (2.0)	2.03 (2.4)	7.25 (5.0)	5.11 (5.5)
CBD work location and lunch trip home				
Bus	0	0	0	0
Taxi	0.0 (0.0)	0.3 (0.6)	1.5 (2.3)	0.7 (2.0)
Automobile passenger and driver	-0.6 (-0.6)	-0.7 (-1.6)	-0.1 (-0.2)	-0.3 (-1.2)
Female worker				
Bus, taxi, automobile passenger	0	0	0	0
Automobile driver	-2.1 (-1.9)	-2.9 (-3.3)	-2.7 (-4.1)	-3.1 (-5.1)
Professional worker and lunch trip home				
Bus	0	0	0	0
Taxi	-0.7 (-0.4)	0.4 (0.6)	1.8 (2.7)	0.9 (2.1)
Automobile passenger and driver	1.3 (1.3)	0.7 (1.6)	1.8 (3.5)	1.1 (3.8)
Availability Parameters				
Bus	0.98 (129.8)	1.00	1.00	1.00
Taxi	1.00	1.00	0.69 (2.4)	0.50 (1.0)
Automobile passenger	0.36 (1.4)	1.00	0.81 (7.3)	0.88 (7.9)
Automobile driver	1.00	0.87 (12.1)	0.87 (24.5)	0.83 (28.6)
Summary Statistics				
Log-likelihood at β=0, γ=1	-596.5	-532.8	-649.8	-1779.0
Log-likelihood at convergence	-126.3	-263.3	-311.7	-715.0
Rho-squared	0.7882	0.5565	0.5203	0.5981
Adjusted rho-squared	0.7531	0.5190	0.4865	0.5857

Note: Asymptotic *t*-statistics in parentheses.

95 percent significance level, so it has been a definite statistical improvement to segment the sample.

When interpreting the availability parameters in Table 3, it should be kept in mind that deterministic alternative availability rules have been applied to construct choice sets for the estimation sample. For example, the low-income segment independent availability model has estimated a probability of availability of 1.0 for the automobile driver mode; however, as was stated before, only individuals from automobile-owning households who are 18 years or older actually have the automobile driver alternative deterministically available. Thus the correct interpretation of this specific parameter is that the best fit to the observed modal choices in the low-income segment is achieved when a probability of 1.0 is assigned to the availability of the automobile driver mode, given that it is deterministically available to the decision maker. Similarly, in the high-income group, the probability of availability of the automobile driver mode is estimated to be about 0.87 for those who have the alternative in their set M_n. This value contrasts with the probability of availability of 1.0 assigned to this alternative in the standard logit model.

Comparison of Alternative Probabilistic Choice Set Models

In this subsection the independent availability logit income segment models will be compared with the logit captivity models. Although it is possible to perform a formal statistical test (recall, however, that the logit captivity specification is not nested within the independent availability model) the two specifications can be compared using a corrected likelihood ratio based on the Akaike information criterion (AIC). This latter measure, defined as the log-likelihood at convergence minus the number of parameters in the model, was first proposed by Akaike *(9)* and is discussed in Amemiya *(10)*. It can be used to compare nonnested hypotheses; the model with the larger value of AIC is preferred.

Alternately, use can be made of an adjusted likelihood ratio index ($\bar{\rho}^2$) based on the AIC and defined as

$$\bar{\rho}^2 = 1 - [L(\hat{\beta}) - K]/L(0) \quad (5)$$

where

$L(\hat{\beta})$ = the log-likelihood of the sample at the maximum likelihood estimates $\hat{\beta}$ of the parameters,

$L(0)$ = the log likelihood of the sample assuming equal probability of choice for all alternatives, and

K = the number of parameters in β.

Following an analysis identical to that of Horowitz *(11)*, it can be shown that if the $\bar{\rho}^2$ of two nonnested models differ by 0.002 or more for a sample of 1,147 observations and a four-alternative choice context, then almost certainly the model with the lower $\bar{\rho}^2$ is incorrect.

The three market segment logit captivity models have a joint $\bar{\rho}^2$ of 0.586 compared with 0.586 for the income segment independent availability logit models. In the aggregate there appears to be no difference between the two probabilistic choice set models.

The following table gives the $\bar{\rho}^2$-values for the individual choice set formation models by income segment, pooled income segments, and the pooled models.

Model	$\bar{\rho}^2$ *Low*	*Medium*	*High*	*All*	*Pooled*
Logit	0.752	0.519	0.480	0.583	0.584
Logit captivity	0.757	0.519	0.484	0.586	0.584
Independent availability logit	0.753	0.519	0.487	0.586	0.586

For the sample sizes in each segment, a difference in $\bar{\rho}^2$ of 0.002 is still significant. Hence the logit captivity model performs better than the independent availability logit model in the low-income group; the reverse is true in the high-income group; and in the middle-income group the choice between the two models is indifferent.

This result highlights an important practical conclusion. It indicates that the restrictions imposed on the probabilistic choice set generation process cannot be arbitrary; instead, they must reflect the population in question and the source of the constraints on it. Hence, in the present context, it would appear reasonable to adopt the logit captivity model for both the low- and medium-income groups (for the latter group, the decision is arbitrary) and the independent availability specification for the most unconstrained group, the high-income segment of the workers.

A last comparison between the two types of probabilistic choice set models is given in Table 4, in which the predicted choice set probabilities according to the logit captivity and independent availability logit models are given (both for the pooled data set and for the three income segments). The table has two parts, the first of which corresponds to a decision maker with all four modal alternatives in M_n and the second of which corresponds to an individual without the automobile driver alternative.

Although many useful inferences can be drawn from the table, one of the most interesting comes from the first part for the independent availability model for the high-income segment. This group is naturally the one that displays the higher rate of automobile ownership and is therefore the one in which workers will most often have the automobile driver alternative allocated to them by the choice set construction rules. Yet, for these individuals, there is predicted a less than 50 percent chance that they will actually be selecting from the full choice set that includes automobile driver, as opposed to the usual assumption of 100 percent in a standard choice model. Medium-income workers who have the automobile driver alternative available, on the other hand, are predicted to have a probability of 87 percent of choosing from the full choice set of four alternatives. A third observation can be made concerning low-income workers who have bus, taxi, and automobile passenger available. The choice set construction rules adopted allowed automobile passenger to all workers; there is only a 35 percent chance, however, that a low-income worker with this three-alternative M_n actually chooses from M_n. It is nearly twice as likely that he will choose between bus and taxi instead.

Another pattern of note in Table 4 is the decrease in probability of captivity to the bus mode with increases in income, as predicted by the logit captivity specifications by income group. Such a result, although intuitively plausible and in conformance with the constraint-based view of choice set formation, also indicates that some parameterized version of the captivity model, in which captivity is expressed as a function of independent variables (among them income), might result in statistically better models.

It has been shown that probabilistic captivity and independent availability choice set models, combined with market segmentation, result in statistically superior models compared with the standard logit model. This result holds in spite of apparent weaknesses in the choice set models (i.e., the strong assumption of independence of alternative availability, or the extreme scenario of captivity or full choice).

TABLE 4 PREDICTED CHOICE SET PROBABILITIES

	Logit Captivity				Independent Availability Logit			
Choice Set	Pooled	Low	Medium	High	Pooled	Low	Medium	High
Available Alternatives: B, T, AP, AD								
B	0.013	0.141	0.085	0.011	0.010	0	0	0.008
T	0.004	0.013	0.009	0.006	0	0	0	0
AP	0.008	0.006	0.050	0.041	0	0	0	0
AD	0	0	0	0	0	0	0	0
Full	0.976	0.840	0.857	0.942	0.366	0.351	0.868	0.485
All others	0	0	0	0	0.624	0.649	0.132	0.507
Available Alternatives: B, T, AP								
B	0.013	0.141	0.085	0.011	0.062	0	0	0.059
T	0.004	0.013	0.009	0.006	0	0.010	0	0
AP	0.008	0.006	0.050	0.041	0	0	0	0
Full	0.976	0.840	0.857	0.942	0.441	0.351	1.000	0.559
All others	0	0	0	0	0.497	0.639	0	0.382

Note: B = bus, T = taxi, AP = automobile passenger, and AD = automobile driver.

On the other hand, the additional difficulty of calibrating a choice model with probabilistic choice sets can be significant compared with estimating the parameters of a standard choice model. The loss of certain convenient properties of the log-likelihood function of the sample creates serious obstacles for the analyst and jeopardizes the practical usefulness of probabilistic choice set models in general.

Thus it is necessary to go beyond measures of statistical significance to evaluate the practical significance of probabilistic choice set modeling. To do so, the predictions produced by the different models are compared in the following section.

MODEL PREDICTION RESULTS

The predictions of the income segment logit models (Table 1) are compared with the probabilistic choice set specifications that are statistically best for each income group (i.e., the logit captivity models of Table 2 for the low- and medium-income groups and the independent availability logit specification of Table 3 for the high-income group). The two sets of models are used to predict changes in modal shares due to

1. Uniform changes (two levels, low and high) across the population for travel time;
2. Implementation of a specific policy alternative; and
3. Shifts in the distributions of a socioeconomic variable, specifically income.

Uniform Changes in Travel Time and Cost

Two levels of change in travel time are implemented herein, 10 percent (low) and 100 percent (high). The reason for this two-level test is that the benefits of choice set modeling may be nonlinear and appear only under conditions of extreme change in these variables.

Table 5 gives the predicted changes under a 10 percent travel time increase for the income segment logit and the income segment choice set specifications, respectively. The results are presented both by income segment and over the entire sample of workers.

If the low-income predictions are studied, it can be noted that the changes in ridership from the logit captivity formulation are less than or equal to the corresponding prediction from the standard logit model for the group; this is to be expected, of course, given that the predicted degree of captivity to bus is highest in this segment of the workers, for which this mode is also the most frequently chosen. The estimation sample of 528 workers in the low-income segment has an observed frequency of choice of bus of 88 percent, so it is understandable that the dampening effect mentioned is present. The predictions are average changes: error bounds have not been provided for these measures because the differences are small between models (with few exceptions) and it is reasonable to assume that none of the differences are statistically significant at any reasonable level of significance. Even in the case of the taxi mode, for which there is a 100 percent difference in the predictions of the two models, it is unlikely that they are statistically different because this mode is the least well explained by any of the models presented.

For the medium-income segment of the Maceió workers, the opposite result has been found: namely, the choice set specification in general states that the medium-income workers are more sensitive to the 10 percent travel time increase than predicted by the standard logit specification. This segment, like the low-income group, has a high incidence of choice of bus (67 percent), but the choice set specification predicts a smaller degree of captivity in this group compared with the low-income segment. At the same time, the travel time coefficients in the income segment logit and logit captivity specifications differ by a factor of almost 3. However, it again appears that the predicted differences are not significant.

The high-income segment is also predicted to be more sensitive to the 10 percent travel time increase by its independent availability model than by the standard logit specification.

In aggregate, the data in Table 5 show a tendency of the standard logit specification to underpredict the effect of travel time increases on the worker population compared with the choice set model. The source of this disparity between the models is the medium- and high-income groups, which the choice set models predict to be more sensitive to the change than does the logit formulation. Because of the aggregation, the overall changes in demand predicted to occur by each set of models are even more uniform than if viewed by income segment, as has just been done.

TABLE 5 PREDICTED IMPACT (% change in demand) OF 10 PERCENT TRAVEL TIME INCREASE

	Predicted Response in Mode							
	Bus		Taxi		Automobile Passenger		Automobile Driver	
Change in Mode	L	PCS	L	PCS	L	PCS	L	PCS
Low income (<80 KwH/month)								
Bus	-0.3	-0.2	3.3	1.7	4.0	3.4	1.4	0.8
Taxi	0.1	0	-3.3	-1.7	0	0	0.3	0.3
Automobile	0.1	0.1	0.8	0.8	-1.3	-1.4	-0.6	0
Medium income (80-130 KwH/month)								
Bus	-1.8	-2.6	6.7	12.8	6.6	7.3	1.9	2.8
Taxi	0.3	0.6	-6.7	-12.3	0.3	0.1	0.2	0.2
Automobile	0.5	0.6	0.6	0.6	-1.8	-1.5	-0.9	-1.3
High income (>130 KwH/month)								
Bus	-2.3	-2.3	3.2	4.4	1.9	2.7	0.6	0.4
Taxi	0.4	0.6	-6.0	-6.4	0.4	0.6	0.2	0.1
Automobile	0.6	0.6	1.2	1.2	0.7	-0.3	-0.8	-0.4
Overall								
Bus	-1.1	-0.3	4.2	6.6	3.7	4.3	1.0	1.0
Taxi	0.2	0.3	-5.6	-7.3	0.3	0.4	0.2	0.2
Automobile	0.3	0.3	0.9	0.9	-0.3	-0.8	-0.8	-0.6

Note: L = logit model and PCS = probabilistic choice set model.

TABLE 6 PREDICTED IMPACT (% change in demand) OF 100 PERCENT TRAVEL TIME INCREASE

	Predicted Response in Mode							
	Bus		Taxi		Automobile Passenger		Automobile Driver	
Change in Mode	L	PCS	L	PCS	L	PCS	L	PCS
Low income (<80 KwH/month)								
Bus	-3.7	-2.5	40.8	13.3	54.7	54.4	12.3	6.4
Taxi	0.6	0.2	-30.0	-12.5	1.3	1.4	1.7	0.8
Automobile	0.7	0.5	3.3	2.5	-11.3	-10.9	-5.7	-2.5
Medium income (80-130 KwH/month)								
Bus	-20.7	-36.5	84.9	219.0	88.2	144.8	14.2	14.5
Taxi	2.4	3.3	-49.7	-63.1	2.0	1.3	0	1.1
Automobile	4.6	5.1	4.5	5.6	-12.0	4.8	-11.4	-20.2
High income (>130 KwH/month)								
Bus	-21.8	-24.1	34.3	49.2	22.3	28.6	4.9	3.2
Taxi	3.4	4.8	-45.4	-48.4	3.3	4.0	1.4	0.7
Automobile	6.4	5.6	14.3	9.6	14.3	2.7	-9.6	-5.3
Overall								
Bus	-12.4	-17.3	52.4	96.7	47.6	69.7	7.8	6.2
Taxi	1.7	2.0	-43.5	-45.4	2.6	2.8	1.3	0.8
Automobile	3.0	2.9	8.9	6.7	2.6	1.7	-9.7	-8.6

Note: L = logit model and PCS = probabilistic choice set model.

Next, the differences in model predictions under a high (100 percent) uniform perturbation in travel time are considered. Table 6 gives the model predictions due to a large change in travel time for the income segment logit and choice set specifications. Note that the income segment logit specification predicts that a uniform doubling of automobile travel time results in more than an 11 percent decrease in demand for both the automobile passenger and the automobile driver modes. The corresponding prediction for the choice set models, however, shows a 20 percent loss of demand in the automobile driver mode and about a 5 percent increase in demand for the automobile passenger mode. Thus the standard logit specification for the medium-income group suggests that a 100 percent increase in automobile travel time causes a shift away from the mode entirely; the captivity specification, however, suggests that there will instead be a shift within the automobile mode via the mechanism of increased ridesharing.

Careful study of Table 6 does bring to light one interesting pattern of differences between the two sets of models. Note that in the aggregate prediction results the income segment logit specification generally predicts a smaller response to a doubling of bus travel time than predicted by the choice set specifications; conversely, a 100 percent increase in automobile travel time is said to result in greater changes than predicted by the choice set models. Further study indicates that these aggregate-level differences between the two models stem from identical patterns in the income groups, though both of the effects mentioned are not necessarily present in each segment. What is observed here is perhaps the result of a twofold effect:

1. The choice set specifications predict a greater response to a change in bus travel time because of the increased sensitivity that these models display to travel impedance compared with the standard logit specifications (compare the travel time coefficients of Table 1 to those of Tables 2 and 3, noting that the former are uniformly less in absolute value than the latter) and
2. The choice set models predict a smaller impact of changing automobile travel time because of their fuller consideration of alternative availability (i.e., an individual's captivity to the automobile passenger mode makes him insensitive to changes in the mode's travel time).

Swait *(12)* reports prediction tests analogous to these two model systems but involving the travel cost variable. The inferences to be drawn from those results are identical to the ones drawn here for travel time.

The results presented thus far are not supportive of any strong superiority of the probabilistic choice set specifications to the standard logit formulation for the choice dimension being examined. Certain differences of note between the predictions of the two model systems have been pointed out, but they may not be worth the extra effort necessary to estimate probabilistic choice set models. On the other hand, neither is the uniform change scenario reflected in the previous predictions necessarily a realistic one for application of these models. This has led to testing for differences in predictions when the two model systems are applied in the context of evaluation of a specific policy scenario.

Evaluation of a Specific Policy

The policy scenario to be used in this subsection is inspired by an actual policy evaluation previously reported by Geltner and Swait *(13)* for Maceió. The specific policy considered envisions extensive traffic engineering improvements in the central business district (CBD) of the city, including the implementation of "bus only" streets and improved loading and unloading spaces and procedures and prohibition of parking of private automobiles in certain areas of the CBD. The impact of such changes is assumed to affect trips to and through the CBD in the following manner:

1. Bus trips—decrease of 10 min per leg of the trip due to improved flow of traffic,
2. Automobile trips—increase of 5 min per leg due to increased walking distances in the CBD, and
3. Taxi trips—no effect.

Table 7 gives the predicted average impacts of implementing this policy. In the first part of the table the predictions of the logit models are given, and in the second part those of the choice set models are given. The income segment logit specification understates the impact of the policy on the bus and taxi modes for the medium- and high-income groups and conversely overstates

TABLE 7 PREDICTED IMPACT (% change in demand with respect to base case) OF CBD IMPROVEMENT POLICY ALTERNATIVES

	Predicted Response in Mode			
	Bus	Taxi	Automobile Passenger	Automobile Driver
Income segment logit specification				
Low income	0.8	-5.0	-10.0	-4.3
Medium income	5.4	-15.6	-16.9	-8.0
High income	8.3	-4.4	-6.9	-3.2
Overall	3.6	-8.4	-10.5	-4.4
Income segment probabilistic choice set specifications				
Low income	0.5	-1.7	-12.0	0.0
Medium income	7.2	-30.2	-12.3	-12.8
High income	8.5	-10.0	-8.7	-2.3
Overall	4.0	-14.7	-10.3	-4.6

the impacts on the low-income group compared with the corresponding choice set model predictions. This result can be explained by the sensitivity of the medium- and high-income groups to travel time in the choice set models being greater than the alternative availability effect; the opposite holds in the low-income group. For the private modes there is no such clear-cut pattern. In either case, however, it is unclear that any of the observed differences in predictions between model systems is actually statistically significant.

This result is not unexpected given the homogeneity of predictions presented previously for uniform changes in travel time. It has been hoped that by targeting a specific group of the workers' population, namely those working in the CBD or traveling through it to reach their workplace, significant differences between the model systems could be detected. It is possible, however, that differences would indeed be found if the impacts on only CBD workers or those traveling through that part of the city were examined.

Shifts in a Socioeconomic Characteristic

In the previous two subsections differences in predictions between the two model systems under consideration have been evaluated in contexts that could best be labeled short range. In both cases, although certain trends are apparent, it remains unclear if one of the model systems is undoubtedly superior to the other. The purpose of this section is to evaluate the differences when the simulated scenario corresponds to long-range shifts in the composition of the worker population in Maceió. Specifically, two different shifts in income distribution will be simulated.

Table 8 gives the observed worker household income distribution and the postulated shift in that distribution. This hypothesis represents a significant worsening of income distribution compared with the observed case. The shift in income distribution is simulated by assigning a weight to each observation corresponding to the ratio of the postulated to the observed frequency for its income group (e.g., 16.9/15.1 for the lowest income category). Note that the actual income value of an observation is not changed, merely the weight given to the predictions for the observation. This methodology assumes that all other characteristics remain constant within the sample (e.g., there are no accompanying shifts in the conditional automobile ownership distribution).

Table 9 gives the predictions for the income shift scenario for each of the model systems. Comparison of the two parts of the table shows little or no difference in the predictions of the standard logit versus choice set specifications.

TABLE 8 OBSERVED AND POSTULATED INCOME DISTRIBUTIONS FOR MACEIÓ WORKERS

Income Category (KwH/month)	Observed Distribution (%)	Scenario Distribution (%)
0-40	15.1	16.9
41-60	9.0	16.9
61-80	12.4	13.5
81-100	12.5	13.5
101-120	13.5	10.2
121-150	12.5	10.2
151-200	10.2	9.3
201-250	5.6	3.4
251-300	2.7	2.7
>300	6.5	3.4
Total	100.0	100.0

Though not presented here, another simulation of a shift in automobile ownership distribution has been carried out with similar results across the two model systems.

CONCLUSIONS

This study was aimed at evaluating the statistical validity of modeling probabilistic choice set formation when the representation of alternative availability is particularly simple (i.e., a single parameter). The estimated models presented here indicate a need to further investigate modeling choice set formulation, particularly in environments such as Maceió, where the traveling pubic is subject to significant constraints of many types that cannot be observed. The choice set formation stage should be of even greater importance in the more discretionary types of behavior, such as mode and destination choice and trip generation for shopping.

Market segmentation, although an indispensible technique to improve the explanatory power of the choice models for a population with taste variations, is too crude a tool to, alone, substitute for explicit models of choice set formation. Allied to the latter, however, market segmentation is of great value. In the empirical work presented here, income segmentation of the sample results in a greater incremental improvement in model fit than is provided by the choice set models that have been tested; nonetheless, it has been demonstrated for this data that choice set modeling provides a statistically significant increase in explanatory power of the work mode choice model system for Maceió.

Another result of the empirical work in Maceió is the confirmation of the important effect of the assumption of choice set struc-

TABLE 9 PREDICTED IMPACT (% change of demand with respect to base case) OF INCOME DISTRIBUTION SHIFT SCENARIO

	Predicted Response in Mode			
	Bus	Taxi	Automobile Passenger	Automobile Driver
Income segment logit specification				
Low income	30.8	29.2	25.3	24.9
Medium income	-9.0	-8.4	-11.2	-10.6
High income	-18.3	-20.7	-26.5	-26.2
Overall	9.3	-6.0	-15.1	-17.7
Income segment probabilistic choice set specifications				
Low income	30.7	26.7	23.3	28.6
Medium income	-9.6	- 8.4	-10.0	- 9.3
High income	-18.3	-21.9	-27.1	-25.9
Overall	9.1	-6.9	-15.3	-16.8

ture on the explanatory power of the full choice model. A strategy that combines market segmentation and appropriate choice set restrictions appears to be most likely to work well, and the logit captivity model appears best for the low-income group, whereas the independent availability logit specification appears to be superior for the high-income group.

This factor may indeed be the reason for the inability of the choice set specifications to present clearly predictions that differ from the standard logit specifications under the various policy scenarios considered. Despite the statistical superiority of the choice set models compared with the standard logit models, it is thought that the homogeneity of the predictions across the two model systems is due in part to limitations of the choice set structure representation inherent in the captivity and independent availability models. Perhaps the assumptions made by each of these choice set models, although somewhat better than the deterministic choice set representation of traditional discrete choice models, are nonetheless inappropriate (even simplistic) for the populations in question. Further, the representation of the impact of constraints via single parameters per alternative is a restrictive and simplistic representation of a complex process. As indicated by Swait and Ben-Akiva *(3)*, the alternative route of parameterization of the availability functions may be more fruitful for further work than is the present approach.

A drawback of the choice set formation models is the greatly increased difficulty of calibrating them. The departure from the standard logit linear-in-parameters formulation can be costly because the convenient property of concavity of the log-likelihood function, which guarantees the uniqueness of the parameters at the point of convergence, is lost. Hence a greater degree of care and sophistication on the part of the analyst is necessary, not to mention specialized estimation software.

ACKNOWLEDGMENTS

The first author wishes to thank the Conselho Nacional de Desenvolvimento Cientifico e Technológico, CNPq, for their support of his graduate studies at the Massachusetts Institute of Technology (MIT). The authors also extend their thanks to the Empresa Brasileira de Planjamento de Transportes, GEIPOT, for the use of the 1977 Maceió Home Interview Survey and to the Department of Civil Engineering of MIT for the computer time without which the empirical work in this paper would have been impossible.

REFERENCES

1. J. Swait, V. Kozel, R. Barros, and M. Ben-Akiva. A Model System of Individual Travel Behavior for a Brazilian City. *Transportation Policy and Decision Making* (forthcoming).
2. M. Ben-Akiva and J. Swait. Choice Models with Simple Probabilistic Choice Set Generation Processes. Submitted for publication to *Transportation Research B*.
3. J. Swait and M. Ben-Akiva. Incorporating Random Constraints in Discrete Choice Models: An Application to Mode Choice in São Paulo, Brazil. Working Paper. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, 1984.
4. C. Manski. The Structure of Random Utility Models. *Theory and Decision*, Vol. 8, 1977, pp. 229–254.
5. M. Ben-Akiva. Choice Models with Simple Choice Set Generating Processes. Working Paper. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, 1977.
6. M. Gaudry and M. Dagenais. The Dogit Model. *Transportation Research*, Vol. 13B, 1979, pp. 105–111.
7. D. Geltner and R. Barros. Travel Behavior and Policy Analysis in a Medium Size Brazilian City. *Transportation Policy and Decision Making* (forthcoming).
8. M. Ben-Akiva, H. Gunn, and L. Silman. "Disaggregate Trip Distribution Models." *Proc., Japan Society of Civil Engineers*, No. 347/IV-1, July 1984, pp. 1–16.
9. H. Akaike. "Information Theory and an Extension of the Maximum Likelihood Principle." In *Second International Symposium on Information Theory* (B. N. Petrov and F. Csaki, eds.), Akademiai Kiado, Budapest, Hungary, 1973, pp. 267–281.
10. T. Amemiya. Selection of Regressors. *International Journal of Economic Literature*, Vol. 19, 1980, pp. 1483–1536.
11. J. Horowitz. Statistical Comparison of Non-Nested Probabilistic Discrete Choice Models. *Transportation Science*, Vol. 17, No. 3, 1983, pp. 319–350.
12. J. Swait. *Probabilistic Choice Set Generation in Transportation Demand Models*. Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge, 1984.
13. D. Geltner and J. Swait. *Illustrative Policy Analysis: Applications of a Disaggregate Mode Choice Model for Work Trips in Maceió, Brazil*. Studies of Urban Travel Behavior and Policy Analysis in Maceió, Vol. 3. Center for Transportation Studies, Massachusetts Institute of Technology, Cambridge, 1981.

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