David A. Hensher*
Macquarie University, Australia

Achieving Representativeness of the Observable Component of the Indirect Utility Function in Logit Choice Models: An Empirical Revelation

Although probabilistic discrete-choice models such as multinomial logit are widely used, little consideration has been given to the appropriateness of the assumption of a representative utility function. The use of socioeconomic variables as proxies for unobserved attributes directly influencing choice, when a single cross section of data is used, has not been adequately tested. In this paper, experimental design enables us to obtain repeated observations on each individual and to investigate the extent to which taste variation in a population can be explained by socioeconomic differences in the population. We show empirically that individual-specific coefficients result in a better model for product positioning but that a group model is preferable for predicting overall market share, because it produces results as good as the individual-specific models at a substantially lower cost.

I. Introduction

Traditionally the parameter estimates of probabilistic discrete-choice models (PDCMs) have been obtained using single observations on a sample of individuals. Its many advocates (e.g., Domencich and McFadden 1975, pp. 50–52) commonly state that a PDCM predicts behavior at the extensive (discrete) margin of choice, and this behavior is related to the individual level rather than only to an aggregate level.

The justification of this interpretation is as follows. In conventional consumer analysis with a continuum of alternatives, the arguments in the utility expression of the individual are quantities of various commodities. With finely divisible commodities we assume that a change in price will have an observable effect on the individual's demand for the commodity. This effect is achievable if we assume that all individuals in a population have a common behavior (except for pure optimization errors) and that systematic variations in aggregate choice reflect common variations in individual choice at a margin, which typi-

* I thank the Australian Bicentennial Authority for permission to conduct further analysis of a data set collected on their behalf. Parts of this paper draw on earlier work (Hensher and Louviere 1982; Hensher and Truong 1982). I am indebted to Lester Johnson who provided many valuable comments, to staff of the Macquarie University Computer Centre for their patience, to an anonymous referee, and to Albert Madansky.

(*Journal of Business,* 1984, vol. 57, no. 2)

0021-9398/84/5702-0005$01.50

cally assumes small changes. This has been referred to by McFadden (1974) as the intensive margin. When a change in price may either leave the consumption of a commodity unaffected or involve zero consumption in preference to the consumption of another commodity, the conventional assumptions are inappropriate in the identification of individual behavior. In order to maintain the benefits of marginal analysis when the objects of choice are discrete, McFadden replaced the common behavior rule with a set of individual behavior rules. The resulting model of population choice behavior is assumed to be relevant to each individual because each individual's behavior rule is contained within the set of individual behavior rules. By assuming that an element of utility is random (see below), is individual specific, and is linked to the probability of choice via a defined distribution of decision rules, we are able to define a population margin to be used to assess the effect on each individual of a change in an observable attribute. This change occurs at what McFadden referred to as the extensive margin.

Although such a model is theoretically able to satisfy these requirements, empirically this requires repeated observations of the same individuals. If responses are not available, one has to pool the data across randomly selected individuals in a sample. The model then represents the behavior of the sample as a whole. The concept of a discrete *individual* behavior response is not applicable. The probabilistic choice model, then, is one which predicts intensive (continuous) marginal changes at the aggregate level.

The indirect utility function of the PDCM is conventionally assumed to be decomposable into two additively separable components, a deterministic component, V, and a random component, ϵ. This specification, the random utility form, is (Hensher and Johnson 1981)

$$U_{jq} = V(S_q, x_{jq}) + \epsilon_{qj} = V_{jq} + \epsilon_{jq}, \tag{1}$$

where V_{jq} is the utility for the representative individual defined on a vector of attributes of individual q, $\mathbf{S}_q$, and a vector of attributes for each element in the choice set, $\mathbf{x}_{jq}, j = 1, \ldots, J$, and ϵ_{jq} is a stochastic component assumed to follow some distribution function. When a PDCM is estimated with one observation per individual we are imposing a great deal on the notion of a representative utility (RU) expression. What I would like to examine is the extent to which the RU expression satisfies this interpretation and, to the extent that it does not, whether there are any identifiable individual-specific dimensions that might be incorporated in the RU expression to account for individual differences. This will then help to reinforce the usefulness of the RU notion when analysis is restricted to conventional data.

To test this requires repeated observation of each individual so that an indirect utility expression can be obtained and compared with the indirect utility expression of the representative individual. By examining any deviations and relating them to standard socioeconomic variables that "describe" individual differences, we might be able to recommend incorporation of specific socioeconomic variables in the RU expression, acting as partial indicators of taste variation. The merit of this approach over other known methods that assume that the parameters of the RU expression are drawn from a specified statistical distribution (e.g., Hausman and Wise 1978; Daganzo 1979; Johnson and Hensher 1982) is that no distributional assumptions are necessary for V_{jq} (see Beggs, Cardell, and Hausman 1981). A distribution assumption is still required for the stochastic term, however. This enables us to estimate individual-specific PDCMs on a small number of repeated observations, obtained from a controlled experimental design. Fischer and Nagin (1981) were the first authors to estimate individual-specific, (pooled) discrete-choice models as an appropriate basis of identifying the magnitude of taste variations in the sampled population. They estimated binary probit models on a small sample of 20 individuals, each faced with 60 pairs of parking alternatives defined on two attributes (price, walking distance). The current study independently pursued a similar approach but with a substantially larger sample of 351 individuals, each faced with 16 combinations of nine alternatives on six attributes. Multinomial logit was used, given the computationally intractable nature of multinomial probit when the number of alternatives is large (McFadden 1981, p. 225). Recently Beggs et al. (1981) have adopted a similar approach to this paper; however, whereas they propose a procedure for using preference data in a choice context, I draw on an empirical study that generated responses directly from a choice design.[1]

The plan of the paper is as follows. In the next section I outline the

1. Since completing the first draft of this paper I have become aware of a study (Beggs, Cardell, and Hausman 1981) in which a logit model is estimated on controlled experimental design data for each of 193 individuals who rank ordered 16 car designs each defined on 16 attributes. No details are supplied of how the design was derived. That study differs from the current paper in that an ordered logit model is applied, estimated using all ranked preference data. The model is

$$P_i = \prod_{i=1}^{I} \left[\exp V_i \Big/ \sum_{j=1}^{i} \exp V_j \right] \quad \begin{array}{l} j = 1, \ldots, J \text{ alternatives} \\ i = 1, \ldots, \text{subset of alternatives, } 1 \leq J. \end{array}$$

While the idea of using all the information from a preference ranking is not new (see, e.g., Punj and Staelin 1978), there is concern that the information declines in relevance by a geometric progression as the preference ordering decreases, such that very little information (but increased computational expense) is obtained beyond rank 5 (see Louviere 1981). This is, however, an argument not for rejecting the ordered logit model but for limiting the rankings modeled as the "choice alternative." In contrast, I apply the conventional first-preference ranking logit; but unlike Beggs et al., who impose a fixed preference set on an individual, in my design I vary the preference sets and seek a first-preference rank only. Fischer and Nagin (1981, p. 281) used a paired comparison procedure; however it is not clear what procedure (e.g., fractional factorial design?) was used to create 60 binary choice sets out of 20 alternatives.

procedure for investigating the RU assumption, including the data base. In Section III, I present the parameter estimates obtained from a logit model of the pooled sample, which assumes that tastes are identical in the population; and the results of the individual coefficient models is on 351 individuals. Various measures of central tendency are used to aid interpretation. In Section IV, I compare the indirect utility of the actual individual with the representative indirect utility of the representative individual.[2] The difference is then regressed against available socioeconomic variables in the search for inclusion in a global model of any proxies for taste variation. The concluding section summarizes the main findings.

II. Procedure for Investigating the Representative (Indirect) Utility Assumption

The overall method is outlined in figure 1. It is used to identify the adequacy of assuming identical parameters in the population and the ability of socioeconomic variables to account for the unobserved attributes of the alternatives in the choice set and hence unexplained taste variation.

The starting point is the development of a data base that permits unique estimation of the parameters for each sampled individual. Market data do not usually provide an appropriate base unless the commodity category under study is associated with sufficient variation in individual response during a period in which we can reliably control for period, aging, and temporal effects. Such data sets are a rarity outside of proprietary research (i.e., consumer panels). The alternative is a controlled experimental design that can generate repeated observations on each individual at a point in time.

Experimental design procedures for choice models have been considered recently (e.g., Louviere 1981); however, only four known studies have developed and applied choice designs in the framework of a PDCM (Batsell 1980; Batsell and Lodish 1981; Fischer and Nagin 1981; and Hensher and Louviere 1983). Furthermore, only Batsell and Fischer and Nagin are known to have estimated a unique model for each individual. A number of studies have aggregated the responses by an individual associated with each alternative and estimated, for the total sample, a share model (e.g., Louviere and Woodworth 1982). The

2. As outlined later in the text, each individual is faced with a nine-alternative choice set (but varying contents) on 16 occasions. For an individual-specific model, to obtain the indirect utility expression for each alternative we treat each observation on each alternative as unique, and talk of the indirect utility of the actual individual. Studies such as Beggs et al. do not have a varying choice or preference set and thus lead to possible confounding of the selection of choice sets with selection of an alternative given a choice set.

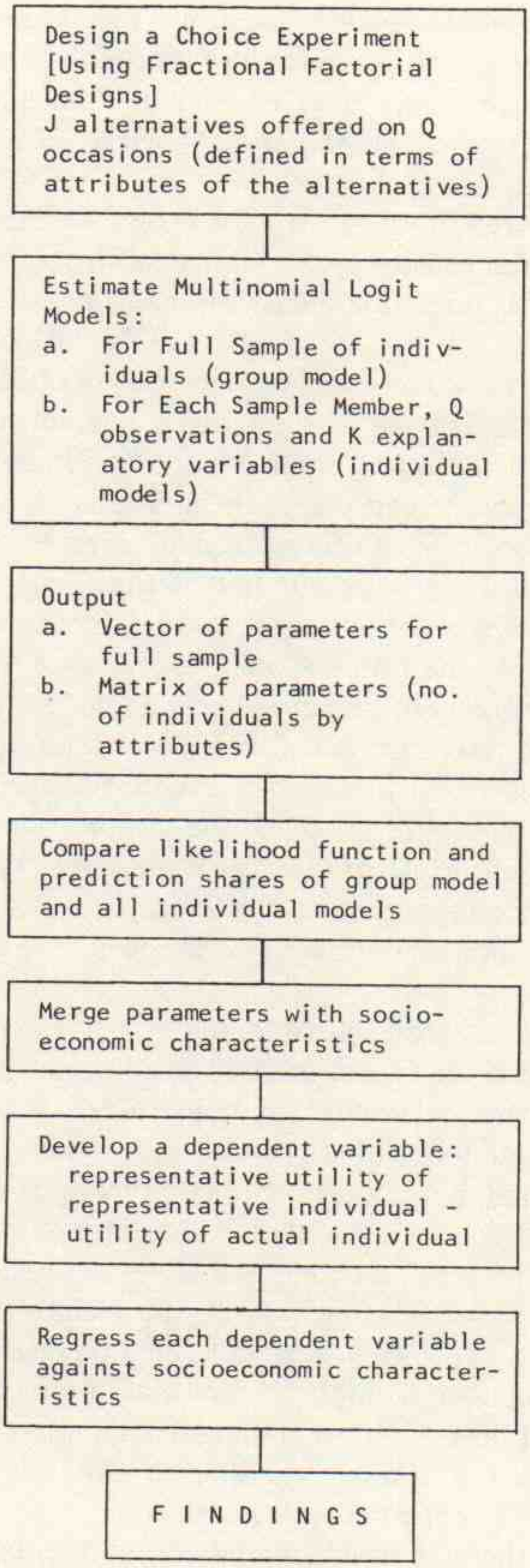

FIG. 1.—Procedure for investigating the representativeness of the indirect utility function "representative" component.

majority of experimental designs used in a PDCM context are preference designs (using ranks), which are manipulated through various ingenious procedures (e.g., data explosion: Punj and Staelin [1978]; ordered logit: Beggs et al. [1982]) and applied to the sample as a whole or each individual.

An important strength of an explicit choice design is that it can be designed so that the selection of a choice set from the universal finite set of choice sets is unconfounded with the selection of an alternative from a given choice set. In McFadden's (1974) classic derivation of the logit model he invoked an axiom (irrelevance of alternative set effect) to assume the nonexistence of such confoundment. Clearly if we can design an empirical setting to eliminate confoundment we are one step ahead. A preference design does not have this characteristic; hence one has to be especially careful in transposition to a choice context. Srinivasan (1980) also points out that if rank-order preference data are initially collected and then converted to the equivalent $n(n - 1)/2$ paired comparisons, the property of independence from irrelevant alternative in the basic logit model is likely to be violated.

In earlier work (Hensher and Louviere 1982) a choice experiment was developed for use in forecasting attendance at a planned international exposition as part of the 1988 celebrations marking 200 years of European settlement in Australia. The experiment consists of a set of choice sets. To accommodate variation in both attributes and choice sets, two interlocking fractional factorial designs were developed, as follows:[3]

a) Five activity categories were treated as a 2^5 factorial design (present, absent),[4] and a one-fourth fraction (eight treatment combinations; see Hahn and Shapiro [1966]) was constructed in which some main effects are confounded with two-way and higher interactions. This enables eight expo types consisting of combinations of activity categories. This design generates the choice alternatives.

b) The eight expo types are then treated as factors in a second design to generate choice sets. Using cost of entry and participation as another consideration, a 16-treatment combination (1/16) fractional factorial design was selected from the 2^8 factorial of the eight expo types at two levels of cost (low, high) such that all main effects are independent of two-way interactions (Hahn and Shapiro 1966). Each of the 16 treatment combinations represents a choice set.

By examining the way in which choices among the expo types vary as a function of both activities and costs, we can develop logit models

3. The explanation of the choice experiment design is drawn from Hensher and Louviere (1983). Further detail is given therein.

4. Categories were national and/or cultural exhibits and displays; industrial and/or technological exhibits and displays; merchandised foods and beverages of different national or ethnic origin; shows and spectacles; rides, amusements, and games.

either for each individual or for the entire sample. The respondent was required to indicate which of the eight expos she or he would most likely attend or to indicate that she or he would probably not attend. The final design is given in the Appendix. The sample of 351 individuals each has 16 observations.[5] Individual-specific and group logit models were estimated for nine alternatives. The explanatory variables are activity-specific dummy variables and cost. The group model, estimated on 16 observations from 351 individuals, combines intraperson and interperson variations in behavior. Errors in choice at the individual level are due to confusion, indifference, and inattention and are assumed to be randomly distributed within and between individuals. In pooling the repeated choices of the sampled individuals there are also errors due to taste variation within the population. For such a model, intraperson and interperson variations are indistinguishable in their effect on the observed distribution of choice (McFadden 1981). Both are assumed additive and identically distributed.

The estimated parameters from each individual model were used to calculate the indirect utility associated with each alternative (V_{jq}, $j = 1, \ldots, 9; q = 1, \ldots, 351$). For each individual 16 indirect utilities are obtained for each of nine alternatives. I treat each alternative on each observation as unique and obtain a utility of alternative j for the intraperson observation for the actual individual (*UAI*). For the model estimated on the full sample, we can calculate the "representative utility of alternative j for the representative individual" (*RURI*). Since the difference $RURI - UAI$ is attributable to the differences in the estimated parameters, we might be able to "explain" part of it by socioeconomic variation across the population. Socioeconomic variables are frequently included as proxies for unobserved attributes of the alternatives and other possible sources of unexplained taste variation (e.g., Train and McFadden 1978). *RURI-UAI* can be regressed against the socioeconomic variables. The remaining sections apply this procedure.

III. Estimation Results for Group and Individual Coefficient Models

The group logit model with fixed parameters is given in table 1 and the summary of the results of 351 separate logit models in tables 2 and 3. Each individual logit model has 11 degrees of freedom. A likelihood ratio test can be used to compare the group model with identical coefficients to the individual-specific coefficients model.[6] Twice the

5. The 351 logit models required 14 hours of central processing unit time on a UNIVAC 1106.

6. The log likelihood associated with the individual models is the sum of all the log likelihoods at convergence.

TABLE 1 Logit Model Results—Group Model with Common Taste Parameters, 351 Individuals, 5,616 Observations

Variable*	Estimated Coefficient	*t*-statistic	Normalized Coefficient Estimates
Cost (1–8)	−.2765	−64.6	−1.0
Culture dummy (1, 5, 7, 8)	.9949	25.2	3.6
Industry dummy (2, 6, 7, 8)	1.1880	31.8	4.3
National food dummy (3, 4, 7, 8)	.6327	16.5	2.29
Likelihood ratio index		.25	
Log likelihood at convergence		−7938.5	

* A variable takes the described value in the alternatives listed in parentheses and zero otherwise.

difference of the log likelihoods is 4,129.6, which under the null hypothesis of equal coefficients across individuals is distributed as χ^2 with 1,256 degrees of freedom. The null hypothesis of equal coefficients can be confidently rejected. The individual coefficient model fits the data much better than the model specification based on constant taste parameters in the population.

Normalizing the coefficients (in tables 1–3) shows that the relative magnitudes between variables within each model are quite similar, even though the absolute magnitudes of the two models differ markedly, by 50% or more.[7] The ratio of each pair of coefficients within the group and individual models (tables 2, 3) show a remarkable degree of stability. Considerable dispersion exists among the estimated individual coefficients on cost, with a relative dispersion of 3.4. The other variables are much less dispersed.

A further comparison can be made between the ability of the individual models and the group model to predict the choice shares. We also include the results of a model using the average of the individual coefficients. Our analysis requires

1. shares obtained for each individual and aggregated across the sample;[8]
2. shares obtained using the average individual coefficients, with and without socioeconomic variables;
3. the group model with and without socioeconomic variables.

7. Although it may initially seem appealing to compare the magnitudes of the pooled sample coefficients with those of the means of the individual coefficients, this is hazardous because the disturbances in the pooled sample model should have higher variance (there is more unexplained) than the disturbances in the individual-based models. Since the logit program normalizes the scale of the disturbances to equal one in all cases, the group coefficients should be smaller in absolute value than the means of the individual ones (as shown in tables 1 and 2). This point was made by an anonymous referee.

8. This first analysis cannot be undertaken on observations not included in model estimation unless some rules are available for mapping other members of the population onto the preference surface of a member of the sample.

TABLE 2 Logit Model Results: Individual Models Coefficients

	Mean	Normalized Coefficient Estimates	Standard Error	Variance	Relative Dispersion	Maximum	Minimum
Cost	−.447	−1.0	.0177	2.29	3.38	−.27	−2.52
Culture dummy	1.744	3.9	.077	2.066	.82	8.06	−1.72
Industry dummy	1.609	3.60	.083	2.39	.96	7.85	−4.66
National food dummy	0.845	1.89	.071	1.78	1.6	7.36	−3.74
Log likelihood at convergence				−5873.683			

TABLE 3 Ratio of Within-Model Coefficients of Pairs of Attributes

	Group Model	Individual Model
Cost/culture	.28	.26
Cost/industry	.23	.28
Cost/national	.44	.53
Culture/industry	.84	1.08
Culture/national	1.57	2.06
Industry/national	1.88	1.90

TABLE 4 **Predicted Choice Shares under Alternative Modeling Assumptions**

		Share Predicted From:				
			Average (Of Individuals) Model		Group Model†	
Alternative	Actual Share	Individual Models*	No Socios	With Socios	No Socios	With Socios
A	3.25	5.3	4.6	4.7	5.5	4.60
B	3.65	8.5	7.9	8.1	8.2	7.11
C	3.87	3.5	2.2	2.2	4.4	3.50
D	3.76	5.2	4.3	4.5	5.2	4.14
E	6.21	6.2	5.4	5.6	5.9	4.88
F	9.36	5.6	4.2	4.3	6.7	5.92
G	30.79	23.2	24.5	24.6	24.6	25.70
H	21.30	26.1	27.6	27.9	25.7	26.56
W	17.50	16.4	19.3	18.1	13.8	17.59

* Exponentiating the level of indirect utility associated with each of the 16 observations on each alternative by each individual (i.e., exp [*UAI*]), calculating the probability of selecting each alternative in each of 16 occasions, and then averaging.

† Twice the difference in the log likelihood between the two (nested) group models equals 120, a χ^2-statistic with 4 degrees of freedom indicates that the socioeconomic variables do have a relevant role.

The socioeconomic variables are selected from the set found significant in the next section. These are entered as alternative specific variables in the utility expression of alternative *W*. The results are given in table 4.

A χ^2 test of the predicted share associated with the individual models approach and the group model (with and without socioeconomic variables) leads me to conclude that there is no significant difference between the prediction distributions.[9] Beggs et al. (1981) did not conduct a prediction test but relied solely on the likelihood ratio test as the criterion for model fit. Their conclusion in support of the individual model cannot be accepted on a prediction test. If the interest is in identification of relevant variables and their valuation, but not prediction, the individual model might be preferred; however, for prediction its merits seem questionable on my evidence. Since a group model is easier to develop, if prediction of behavior is the sole criterion, a group model is preferred on my evidence.

IV. The Relationship between Unexplained Interperson Taste Variation and Socioeconomic Dimensions

Having established one case for separate taste parameters, my next task is to determine whether any socioeconomic variables of the indi-

9. $\chi^2\ .05_{(8)} = 15.51 > 3.3855$ for individual vs. group (no socioeconomics) models; $\chi^2\ .05_{(8)} = 15.51 > 3.3128$ for individual vs. group (with socioeconomics) models.

TABLE 5 Difference between *RURI* and *UAI*

Alternative	Mean	Standard Error	Variance	Kurtosis	Skewness	Minimum	Maximum
A	1.519	.055	16.755	10.65	1.72	24.0	40.0
B	1.886	.058	19.165	8.56	1.75	22.0	40.0
C	2.147	.056	17.864	8.50	1.66	18.0	40.0
D	2.147	.056	17.864	8.50	1.66	18.0	40.0
E	1.519	.055	16.765	10.65	1.72	−24.0	40.0
F	1.886	.058	19.165	8.56	1.75	−22.0	40.0
G	.720	.054	16.37	8.64	1.31	−25.0	38.0
H	.720	.054	16.37	8.64	1.31	−25.0	38.0

vidual or the household can act as proxies to explain part of the unexplained interperson differences. A number of bases for identifying possible socioeconomic explanations can be considered. The most informative involves regressing the representative utility of the representative individual minus the utility of the actual individual (i.e., *RURI* − *UAI*) (table 5) against a set of socioeconomic regressands.[10]

The results are given in table 6. The intercorrelation matrix (available on request) shows negligible pairwise correlation between all variables. Overall the socioeconomic descriptors are poor in explaining the difference between *RURI* and *UAI*, $R^2 = .05$, even though most are statistically significant. This, together with the predicted share result, supports the findings of the consumer marketing literature (e.g., Green 1977) and of Fischer and Nagin (1981) that incremental variance explained by demographic and socioeconomic variables is usually quite modest. The prediction share findings in the previous section, however, do not support the recent conclusion of Beggs et al. (1982) that population variation modeled by socioeconomic variables is not as good as a specification that allows for separate taste parameters across individuals.

This conclusion may be of the most concern to the business economist and others interested in product design (where specific attribute effects are critical) and product marketing, where average preferences of the market (leading to majority fallacy) is not a useful concept. Persons interested in aggregate equilibrium behavior of markets may be less concerned. Indeed, my evidence that the individual model fails to improve prediction but performs better in parameter (and hence variable) identification suggests its role in product design; the group model cannot be rejected for predicting aggregate market behavior.

10. Analysis of the relationship between the socioeconomic variables and *RURI* and *UAI*, respectively, were performed but added no additional further insight. The results are available on request.

TABLE 6 **Relationship between *RURI-UAI* and Socioeconomic Variables of the Sample, Ordinary Least Squares Regression**

Socioeconomic Variables	Dependent Variables (*RURI-UAI*)*j**								Mean of Socioeconomic Variables
	j = *A*	B	C	D	E	F	G	H	
Number of persons 0–14 years old	1.22	1.64	1.44	1.44	1.22	1.64	2.36	2.36	.923
Number of persons 15–24 years old	4.78	5.35	4.46	4.46	4.78	5.35	5.36	5.36	.749
Number of persons 25–34 years old	1.94	2.69	2.69	2.69	1.94	2.69	0.55	0.55	.539
Number of persons 35–49 years old	6.64	5.29	5.17	5.17	6.64	5.29	3.07	3.07	.718
Number of persons 50–60 years old	4.45	3.14	2.95	2.95	4.45	3.14	1.44	1.44	.385
Number of persons over 60 years old	.32	1.68	1.45	1.45	.32	1.68	1.38	1.38	.239
Number of professional persons	3.42	4.38	6.04	6.04	3.42	4.38	3.79	3.79	.456
Household gross income/100	7.82	7.74	8.37	8.37	7.82	7.74	8.52	8.52	20.878
Number of motor vehicles	.32	2.48	2.43	2.43	.32	2.48	4.13	4.13	1.490
Number employed in household	2.87	2.13	1.71	1.71	2.87	2.13	3.65	3.65	1.447
Age of individual interviewed	.52	4.07	3.70	3.70	.52	4.07	3.32	3.32	44.540
Sex of individual interviewed	1.62	4.04	1.48	1.48	1.62	4.04	3.28	3.28	.328
R^2	.05	.05	.05	.05	.05	05	.06	.06	. . .

* Figures in matrix are *t*-statistics except for last (R^2) row. Parameter estimates are avilable on request from the author.

V. Conclusion

This empirical enquiry has produced evidence to support the contention that taste variation within a population is poorly explained by the socioeconomic differences within that population. Unlike the previous study by Beggs et al. (1981), which transformed the preference ranking data for use in a choice model, by developing a choice design such as the Fischer-Nagin study (1981) I am able to satisfy the fundamental assumptions of the choice axiom which underlie the basic logit model used to obtain parameter estimates. This avoids the criticism directed by market researchers toward the logit technique (e.g., Srivinasan 1980, p. 49). The Fisher-Nagin approach using binary probit faces the problem of a lack of practical, accurate methods for approximating the choice probabilities when the number of alternatives is large. The logit model embedded in an experimental design is an analytically tractable alternative.

The use of individual-specific coefficients, accounting for part of taste variation, results in a better-fitting model on the likelihood test (tables 1–3) but cannot be distinguished from the group model on a prediction-share test (table 4). Residual unexplained taste variation is not adequately reflected in the inclusion of socioeconomic variables (as proxies) in a group-level model. Since socioeconomic variables in a group model are not by themselves suitable proxies for taste variation accounted for by the specified set of attributes in the individual-specific models, and since the latter set of variables do not explain fully the taste variation of the sampled population, there is a need, on the basis of my evidence, for an improved specification of the set of attributes defining alternatives in the estimation of coefficients in both the individual-specific models and the group model. In design and marketing applications we should use the individual coefficients, seeking out a basis for applying the unique individual-specific models to nonsample members of the population. For prediction of overall market behavior the group model appears preferable simply because it produces results as good as the individual-specific models but at a substantially lower cost. Inclusion of socioeconomic variables contributes marginally to overall predictive performance. The need for continuing research along these lines is clear.

Appendix

The Design Question

We have listed eight (A–H) hypothetical but possible Expos which are being considered. Each Expo (A–H) listed has a different mix of types of activities. We have listed beneath each Expo (A–H) an estimated cost per person to attend. This cost includes a fee to enter the gate, have a meal in a restaurant

	TYPES OF EXPOS FROM WHICH TO CHOOSE								
	EXPO A	EXPO B	EXPO C	EXPO D	EXPO E	EXPO F	EXPO G	EXPO H	If Australia put on an Expo like A, B, etc., I would be most likely to attend (Circle <u>ONE</u> please):
	Cultural Exhibits Only	Industrial/ Technolog-ical Exhibits Only	Shows/ Spectac-les, National Foods/ Drinks	Amuse-ments/ Rides & National Foods/ Drinks	Amuse-ments/ Rides & Shows/ Spectac-les & Cultural Exhibits	Amuse-ments/ Rides & Shows/ Spect-acles & Indust-rial/ Technol-ogical Exhibits	Shows/ Spectac-les & National Foods/ Drinks & Cultural Exhibits & Indust-rial Techno-logical Exhibits	Amuse-ments/ Rides & National Foods/ Drinks & Cultural Exhibits, Industrial/ Techno-logical Exhibits	A,B,C,D,E,F,G,H Probably Wouldn't Go
	Cost per Person	Cost per Person	Cost per Person	Cost per Person	Cost per Person	Cost per Person	Cost per Person	Cost per Person	
CHOICE 1	$4	$6	$10	$2	$8	free	$6	$4	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 2	$6	$10	$2	$14	$14	$20	$12	$4	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 3	$10	$2	$10	$8	$20	$6	$16	$16	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 4	$2	$8	$14	$14	free	$12	$4	$16	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 5	$8	$18	$8	free	$6	$16	$16	$13	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 6	free	$10	free	$20	$12	$4	$4	$10	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 7	$6	$14	$14	$6	$16	$16	$6	$6	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 8	$4	$14	$20	$12	$4	$4	$18	$10	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 9	$16	free	$6	$4	$16	$18	$10	$14	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 10	$18	$6	$4	$16	$4	$6	$10	$14	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 11	$10	$4	$12	$4	$6	$10	$14	$2	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 12	$14	$4	$16	$16	$18	$10	$2	$8	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 13	$14	$20	$4	$6	$10	$2	$14	free	A,B,C,D,E,F,G,G Probably Wouldn't Go
CHOICE 14	$20	$12	$6	$18	$10	$14	$8	$6	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 15	$12	$16	$16	$10	$2	$8	free	$20	A,B,C,D,E,F,G,H Probably Wouldn't Go
CHOICE 16	$16	$16	$18	$10	$14	$14	$20	$12	A,B,C,D,E,F,G,H Probably Wouldn't Go

FIG. 2.—Design of personal interview. Respondent was encouraged to answer for one row at a time (i.e., to treat each row as independent of the others).

representing one of the countries or nations participating, and enjoy four or five of the rides or amusement events. The costs for food and drink and amusement and rides of course apply only to Expos that have these attractions. When foods/drinks and amusements/rides are not part of the Expo (A–H), the cost per person represents just the entrance fee.

We would like you to examine the eight (A–H) Expos and the costs per person to attend in each choice situation (1–16) and simply circle the one you most probably would attend (A–H) or circle the words "probably wouldn't go" if you probably would not attend any of the eight Expos at the costs listed. For example, suppose you examine choice 1. If you really want an Expo with cultural events and foods/drinks, only Expos G and H have these. G costs about $6.00 per person and H costs about $4.00. G has shows and spectacles which H does not, while H has amusements and rides which G does not. You must make up your mind which you prefer and whether the cost is acceptable. Suppose you find the costs for G and H too high for you to afford, you would circle "probably wouldn't go." However, if you preferred G to H and also thought you could afford it, you would probably circle G. Similarly, you decide whether you would attend each at the cost listed for each of the 16 choices.

References

Batsell, R. R. 1980. Consumer research allocation models at the individual level. *Journal of Consumer Research* 7:78–87.

Batsell, R. R., and Lodish, L. M. 1981. A model and measurement methodology for predicting individual consumer choice. *Journal of Marketing* 18:1–12.

Beggs, S.; Cardell, S.; and Hausman, J. 1981. Assessing the potential demand for electric cars. *Journal of Econometrics* 17:1–19.

Daganzo, C. 1979. *Multinomial Probit*. New York: Academic Press.

Domencich, T., and McFadden, D. 1975. *Urban Travel Demand: A Behavioural Approach*. Amsterdam: North-Holland.

Fischer, G. W., and Nagin, D. 1981. Random versus fixed coefficient quantal choice models. In C. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data: With Econometric Applications*. Cambridge, Mass.: MIT Press.

Green, P. E., 1977. A new approach to market segmentation. *Business Horizons* 20:61–73.

Hahn, G. J., and Shapiro, S. S. 1966. A catalog and computer program for the design and analysis of orthogonal symmetric and asymmetric fractional factorial experiments. General Electric Research and Development Center Technical Report no. 66-C-165. Schenectady, N.Y.: General Electric Research and Development Center.

Hausman, J., and Wise, D. 1978. A conditional probit model for qualitative choice. *Econometrica* 42 (March): 403–26.

Hensher, D. A., and Johnson, L. W. 1982. *Applied Discrete-Choice Modelling*. New York: Wiley.

Hensher, D. A., and Louviere, J. J. 1983. Forercasting consumer demand for a unique cultural event: An approach based upon an integration of probabilistic discrete-choice models and experimental design data. *Journal of Consumer Research* (in press).

Hensher, D. A., and Truong, T. P. 1982. A direct preference approach to non-working time savings valuation using experimental design. Mimeographed.

Johnson, L. W., and Hensher, D. A. 1982. Application of multinomial probit to a two-period data set. *Transportation Research* 16A, nos. 5–6 (September): 457–64.

Louviere, J. J. 1981. On the identification of the functional form of the utility expression and its relationship to discrete choice. Appendix B. In Hensher and Johnson (1982).

Louviere, J. J., and Woodworth, G. 1982. Design and analysis of simulated consumer choice on allocation experiments: An approach based on aggregated data. Working Paper no. 82-7. Ames: University of Iowa, College of Business Administration.

McFadden, D. 1974. A conditional logit model of qualitative choice behaviour. In P. Zarembka et al., *Frontiers of Econometrics*. New York: Academic Press.
McFadden, D. 1981. Econometric models of probabilistic choice. In C. F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Application*. Cambridge, Mass.: MIT Press.
Punj, G. N., and Staelin, R. 1978. The choice process for graduate business schools. *Journal of Marketing Research* 15:588–98.
Srinivasan, V. 1980. Comments on 'On conjoint analysis and quantal choice models.' *Journal of Business* 53 (3): 47–50.
Train, K. E., and McFadden, D. 1978. The goods/leisure trade off and disaggregate work trip mode choice models. *Transportation Research* 12:349–53.