Volume Title: Annals of Economic and Social Measurement, Volume 5, number 4

Volume Author/Editor: Sanford V. Berg, editor

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm76-4

Publication Date: October 1976

Chapter Title: Quantal Choice Analysis: A Survey

Chapter Author: Daniel L. McFadden

Chapter URL: http://www.nber.org/chapters/c10488

Chapter pages in book: (p. 363 - 390)

Annals of Economic and Social Measurement, 5/4, 1976

QUANTAL CHOICE ANALYSIS: A SURVEY*

BY DANIEL MCFADDEN

This article surveys quantal choice analysis, focusing on derived selection probabilities, revealed choice models, and the Luce Model. Statistical analysis of selection probabilities is examined from the standpoint of functional forms, methods for estimation and inference, and extensions of the statistical choice problem. Particular attention is given to multivariate choice, separability, and independence, and unsolved problems are discussed. The survey concludes with some economic applications of these choice models.

I. INTRODUCTION

In traditional economic analysis of consumer decisions, for example, the demanded quantities of meat and potatoes, acts and outcomes are treated as real variables. In contrast, decision problems in other social and biological sciences often lead to acts or outcomes indexed by finite or countable sets; we term these qualitative or *quantal response* problems. Examples are migration in social demography, voting behaviour in political science, and mortality in bioassay.

In recent years, economists have recognized that many important economic decisions involve quantal response. Examples are choice of occupation, family size, labor force participation, ownership and brand of consumer durables, household and workplace location, and shopping trip mode and destination. Investigation of these problems has led economists to rediscover or translate models and methods from psychology and statistics, and to elaborate and extend some results. I will attempt in this survey to summarize these developments and point out problems for exploration.

The prototypical study of quantal choice considers individual subjects placed in one of N possible experimental choice settings. Each choice setting requires one of J possible responses from the subject. Choice setting n is characterized in general by a vector c_n of measured characteristics of the subject and vectors $x_{1n}, \ldots, x_{Jn}$ of measured attributes of the alternatives. (We shall often deal with a vector, denoted z_{jn}, of numerical functions of the measured data x_{jn}, c_n, and with $z^n = (z_{1n}, \ldots, z_{Jn})$.) In R_n repetitions of the presentation of the choice setting n to

* Research was supported primarily by National Science Foundation grant GS-35890X, and was also supported in part by NSF grants SOC72-05551A02 and SOC75-22657, and by NSF grants GI-43740 and APR74-20392, Research Applied to National Needs Program, to the University of California, Berkeley. This paper was distributed as background material for the joint session of the Mathematical Social Science Board Workshop on the Theory and Measurement of Economic Choice and the National Science Foundation/National Bureau of Economic Research Conference on Decision Rules and Uncertainty, June 19–22, 1976, at the University of California, Berkeley. I wish to express my blanket appreciation to the long list of persons who have enlightened me by providing materials and comments. I apologize for the sins of commission and omission in this survey.

a subject, S_{jn} responses of alternative j are observed. (R_n may be one.) These data can be summarized in a $J \times N$ contingency table:

Responses		Choice setting 1	... n	... N
	1	S_{11}	... S_{1n}	... S_{1N}
	⋮	⋮	⋮	⋮
Alternatives	j	S_{j1}	... S_{jn}	... S_{Jn}
	⋮	⋮	⋮	⋮
	J	S_{J1}	... S_{Jn}	... S_{JN}
Explanatory variables		z^1	... z^n	... z^N
Repetitions		R_1	... R_n	... R_N

Figure 1

The response of a subject in choice setting n can be depicted as a drawing from a multinomial distribution with probabilities $(P_{1n}, \ldots, P_{Jn})$; we term these the *selection probabilities*, and note that they are "observable" in the sense that they are estimated consistently by the observed relative frequencies $(S_{1n}/R_n, \ldots, S_{Jn}/R_n)$ as the number of repetitions approaches infinity.

We can now pose two major questions:

1. How are the selection probabilities related to the explanatory variables and to underlying theories of individual choice behavior?

2. Given a structure for the selection probabilities, how can the observations be utilized for estimation and tests of hypotheses?

I shall draw on the early literature on quantal choice to illustrate a consistent answer to these two questions, and then organize my survey with the view that the ensuing literature is directed to improving and generalizing this solution.

Consider the problem of binary choice ($J=2$). In the search for laws governing psychophysical phenomena, Thurstone (1927a,b) suggested a random utility model of choice in which individuals draw at random a member of a set of numerical scale functions, and then select the alternative whose attributes maximize the value of the chosen scale function. Thurstone considered in particular the case in which scale values for different stimuli (attributes of alternatives) z_{jn} are independently identically normally distributed about mean values $v(z_{jn})$, and established that the binary selection probabilities satisfy

$$(1) \qquad P_{1n} = \Phi\left(\frac{v(z_{1n}) - v(z_{2n})}{\sigma}\right)$$

where Φ is the standard normal cumulative distribution function, and σ is a parameter. Suppose the argument in (1) can be written as an unknown linear combination of numerical functions of the data.

$$\frac{v(z_{1n})-v(z_{2n})}{\sigma}=\beta' z^n \equiv \sum_{k=1}^{K} \beta_k z_k^n. \tag{2}$$

Then equation (1) becomes the familar *binary probit model* of statistics,

$$P_{1n}=\Phi(\beta' z^n). \tag{3}$$

Thurstone's construction is appealing to an economist because the assumption that a single subject will draw independent utility functions in repeated choice settings and then proceed to maximize them is formally equivalent to a model in which the experimenter draws individuals randomly from a population with differing, but fixed, utility functions, and offers each a single choice; the latter model is consistent with the classical postulates of economic rationality.

A statistical procedure for estimating β was suggested by R. A. Fisher (1935) in the context of testing the hypothesis that $\beta=0$ for the case of two choice settings ($N=2$) and a single explanatory variable; in a generalization due to Berkson (1944), this takes the form of applying ordinary least squares to the equation

$$\Phi^{-1}(S_{1n}/R_n)w_n=\beta' z^n w_n+\varepsilon_n, \tag{4}$$

where Φ^{-1} is the inverse cumulative normal function and w_n are weights satisfying

$$w_n=\phi(\Phi^{-1}(S_{1n}/R_n))\sqrt{R_n^3}/S_{1n}S_{2n}, \tag{5}$$

with ϕ the standard normal density.[1] It is not difficult to show that as the number of repetitions approaches infinity, the usual rank and limit conditions assumed in least squares analysis are sufficient to guarantee that the estimator of β is consistent and asymptotically efficient. Thus, the Thurstone and Berkson arguments taken together provide a practical econometric model for binary choice that is consistent with the classical postulates of consumer behavior. The remainder of this survey will be devoted to generalizations of this solution.

II. Selection Probabilities and Theories of Choice

1. *Derived Selection Probabilities*

Starting from a specified theory of choice, one may derive properties of the selection probabilities. Further, this can be done abstractly or parametrically. Thurstone's law of comparative judgement discussed above is an example of a parametric analysis starting from a theory of choice. Generalizing this approach, let $U(z_{jn})$ denote a random utility associated with alternative j, and let $V(z_{jn})$

[1] The argument in section III.2 justifies this form. The weights w_n which correct for heteroscedasticity may be omitted without affecting consistency.

denote the mean utility associated with this alternative. Write $U(z_{jn}) = V(z_{jn}) + \varepsilon(z_{jn})$. Assume respondents draw a random utility function U, and then select the alternative which maximizes utility. This random utility model then determines the selection probabilities from the condition

$$
\begin{aligned}
P_{jn} &= \operatorname{Prob}[U(z_{jn}) > U(z_{kn}) \quad \text{for } k = 1, \ldots, J \quad \text{and } k \neq j] \\
&= \operatorname{Prob}[\varepsilon(z_{kn}) - \varepsilon(z_{jn}) < V(z_{jn}) - V(z_{kn}) \quad \text{for } k = 1, \ldots, J \quad \text{and } k \neq j].
\end{aligned} \tag{6}
$$

Specifying a parametric joint distribution for the stochastic terms in (6) leads to parametric functional forms for selection probabilities. With judicious choice of the joint distribution to reflect expected or hypothesized variations in tastes and perceptions in the population of respondents, this approach has the potential of yielding flexible and realistic selection probability functions. Unfortunately, the problem of computing selection probabilities generated in this way is usually formidable, particularly for multinomial response.

In a series of seminal papers, Richard Quandt (1966, 1968, 1969) considered the selection probabilities generated by a population of consumers with log-linear utility functions with random coefficients. A generalization of the Quandt model due to Domencich and McFadden (1975) illustrates the conceptual flexibility and computational drawbacks of this approach. Suppose the random utility model has the form

$$U(x_{jn}, c_n) = \alpha' Z(x_{jn}, c_n) + \varepsilon(x_{jn}, c_n), \tag{7}$$

where α is a $K \times 1$ vector of parameters which vary across the population, $z_{jn} = Z(x_{jn}, c_n)$ is a $K \times 1$ vector of numerical transformations of the data, and $\varepsilon(x_{jn}, c_n)$ is a random component in utility.

This model permits a wide range of structures on the selection probabilities. If the taste parameters α_k have degenerate distributions and the $\varepsilon(x_{jn}, c_n)$ are independent, then we can have the Luce model described below. On the other hand, if the α_k have much larger variances than the $\varepsilon(x_{jn}, c_n)$, then choice between alternatives which differ substantially in their attributes will be governed by the α_k distribution, while choice between objects of similar aspects will be governed by the $\varepsilon(x_{jn}, c_n)$. Taking a classical illustration, suppose subjects are offered a choice of a bicycle or a pony. Aspects such as tastes for animate or mechanical objects, reflected in the α_k parameters, effectively determine the binary selection probability for subjects drawn randomly from the population. Suppose the choice is now expanded to include a second bicycle which differs from the first only in color and trim. The model conforms to intuition in yielding the result that subjects will continue to choose between the pony and the bicycles in roughly the same proportions (and, hence, the odds that the first bicycle is chosen over the pony falls). Since the two bicycles have similar attributes and the $\varepsilon(x_{jn}, c_n)$ effects are small, subjects who definitely prefer the pony to the first bicycle will almost certainly also prefer the pony to the second bicycle. The choice between bicycles will be influenced more strongly by the $\varepsilon(x_{jn}, c_n)$ terms since the dominant effects of tastes over aspects tend to cancel out.

We consider the form of the selection probabilities generated by this model. Write $\varepsilon_{jn} = \varepsilon(x_{jn}, c_n)$, and assume that the random variables $\alpha_1, \ldots, \alpha_K$, $\varepsilon_{1n}, \ldots, \varepsilon_{jn}$ $(n = 1, \ldots, N)$ are independent normal with $E\alpha_k = \beta_k$, $E\varepsilon_{jn} = 0$, $E(\alpha_k - \beta_k)^2 = \sigma_k^2$, $E\varepsilon_{jn}^2 = \sigma_0^2$. Let $D = \mathrm{diag}\,(\sigma_1^2, \ldots, \sigma_K^2)$ and define the $K \times J$ matrix $Z^n = (z_{1n}, \ldots, z_{Jn})$. Then $(U(x_{1n}, c_n), \ldots, U(x_{Jn}, c_n))$ is multivariate normal with mean vector $\beta' Z^n$ and covariance matrix $\Omega = \sigma_0^2 I + Z^{n\prime} D Z^n$. With a suitable linear transformation, equation (6) for the selection probabilities can be expressed as an iterated integral of standard normal densities. For the case of binary choice, a modification of the probit formula results,

$$(8) \qquad P_{1n} = \Phi(\beta' Z^n t / \sqrt{t'\Omega t}),$$

where $t' = (1, -1)$. For $J > 2$, the selection probabilities cannot be expressed in a convenient closed form, and for $J > 4$ the computational task of evaluating the selection probabilities makes the model impractical for use in standard iterative statistical algorithms. Current efforts to break this impasse are discussed in the section on statistical methods.

2. *Revealed Choice Models*

An alternative to deriving selection probabilities from a parametric choice model is to postulate a particular structure for selection probabilities and then attempt to verify the consistency of these structures with the random utility model or other choice theory. In its most abstract form, this process is analogous to the theory of revealed preference in conventional consumer demand analysis; we seek necessary and sufficient conditions on selection probabilities to ensure their consistency with the random utility model. This problem was first investigated by Block and Marshak (1960), who obtained conditions for universes containing three or four objects of choice. Their result has been generalized to arbitrary universes of choice by McFadden and Richter (1970, 1971); a re-examination of the three-object case provides the key to the generalization.

Suppose, in a universe of objects $\{1, 2, 3\}$, subjects are offered one of the binary choices $\{1, 2\}$, $\{2, 3\}$, or $\{3, 1\}$. Let P_{12} denote the selection probability for 1 from the choice set $\{1, 2\}$. The random utility model postulates maximization of some preference ranking of 1, 2, 3. Then, there are six decision rules generated by the permutations of 1, 2, 3. The problem is to determine whether there exists a distribution $(q_1, \ldots, q_6)$ of decision rules in the population yielding the observed selection probabilities, i.e., whether the following system of linear inequalities has a nonnegative solution $(q_1, \ldots, q_6)$ with $q_1 + \ldots + q_6 = 1$, where an element of the coefficient matrix is one if the associated decision rule is successful in producing the observed response in the associated choice setting, and is zero otherwise.

$$
(9)\qquad \begin{pmatrix} p_{12} \\ p_{23} \\ p_{31} \\ p_{21} \\ p_{32} \\ p_{13} \end{pmatrix} \geq \begin{array}{c} \begin{array}{cccccc} 123 & 231 & 312 & 132 & 321 & 213 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \end{pmatrix}
$$

If a solution exists, then the left-hand-side vector lies in the convex polytope spanned by the columns of the coefficient matrix, (9) holds with equality, and $(q_1, \ldots, q_6)$ defines the distribution of decision rules in the population. A necessary and sufficient condition for the existence of a solution is that no hyperplane with a nonnegative integral normal separate the left-hand vector from the convex polytope; i.e.,

$$
(10)\qquad \sum_{ij} r_{ij} p_{ij} \leq \max_k \sum_{ij} r_{ij} a_{ij,k},
$$

where ij indexes the pairs 12, 21, etc., $a_{ij,k}$ is an element in column k of the coefficient matrix, and r_{ij} is any nonnegative integer. This condition generalizes directly to any finite universe of choice, and straightforwardly to non-finite universes. The particular result that is useful in the context of quantal choice is the following:

Suppose the universe of attributes of alternatives is a metric space and each choice setting presents a finite set of alternatives. A necessary and sufficient condition for the existence of a σ-additive probability on the preference orderings of all attribute vectors which yields the selection probabilities in equation (7) is the following axion of revealed stochastic preference: *For each finite sequence $\{j_l n_l\}$ of alternatives and choice settings, $l = 1, \ldots, L$ (with repetitions allowed), the sum $\sum_{l=1}^{L} P_{j_l n_l}$ cannot exceed the maximum number of choices of j_l (in setting n_l) for $l = 1, \ldots, L$ consistent with tne behavior of some preference-maximizing individual.*

This result can be used readily to obtain necessary conditions for consistency with the random utility model; however, verification of sufficiency becomes computationally intractable for a universe of more than a few objects.

Once the question of the existence of a random utility model for given selection probabilities is resolved, the next interesting question is the extent to which the selection probabilities bound the probability distribution on the underlying decision rules. Returning to the example of the three-object universe, we see for instance that $p_{12} \geq q_1$, $p_{23} \geq q_1$, $p_{13} \geq q_1$ imply $q_1 \leq \min(p_{12}, p_{13}, p_{23})$, and that $q_3 + q_4 \leq p_{32}$ and $q_1 = p_{12} - q_3 - q_4$ imply $q_1 \geq p_{12} - p_{32}$. In a recent paper, Robert Hall (1973) has pointed out that these are examples of Tchebycheff systems (Karlin and Studden (1966)), and has extended this theory to develop systems of bounds in the case of choice theories indexed by a single parameter. Bounds for the general consumer choice problem have been investigated by McFadden (1975).

3. *The Luce Model*

A practical approach to the specification of selection probabilities and determination of revealed choice theories is to postulate plausible axioms on selection probabilities and deduce the implications of these axioms for choice behavior. An influential work by Luce (1959) takes this tack, starting from an *independence from irrelevant alternatives* axiom on the selection probabilities. This axiom states, roughly, that the relative odds in a binary choice will remain the same for these alternatives when additional alternatives become available, and implies that selection probabilities can be written in the form

$$P_{jn} = \frac{e^{v(z_{jn})}}{\sum_{k=1}^{J} e^{v(z_{kn})}}, \tag{11}$$

where the $v(z)$ are scale functions of the stimuli. This is often termed the *strict utility* model. We note that when v is linear in parameters, this is the multinomial logit formula encountered in the statistical literature.

The results of Thurstone and Luce were introduced to economics by Marschak (1960) and Block and Marschak (1960), who established that each strict utility model could be derived from some random utility model. The Block–Marschak proof is a non-constructive demonstration that a joint distribution on the $\varepsilon(z_{jn})$ variables exists such that (6) and (11) both hold. A simpler constructive demonstration due to A. Marley is reported by Luce and Suppes (1965); following this line, one can establish (McFadden (1973)) that a *necessary* and sufficient condition for the random utility model (11) with the $\varepsilon(z_{jn})$ independently identically distributed to yield the strict utility model (6) is that $\varepsilon(z_{jn})$ be Weibull distributed.[2]

The multinomial logit model combined with the results relating it to the random utility model is particularly appealing in applications because of computational advantages and consistency with a theory of sampling from a population of classical utility-maximizing consumers. The primary drawback of this formulation is that it ignores some of the structural aspects of choice which make the independence from irrelevant alternatives property inappropriate in some applications. As pointed out by Debreu (1960), the strict utility model predicts too high a joint probability of selection for two alternatives which are in fact perceived as "similar" rather than "independent" by the subject.

4. *Extensions of the Luce Model*

The questionable validity of the independence from irrelevant alternatives axiom in some applications has led to a search for alternatives which are empirically practical and are compatible with a plausible theory of choice. Examination of the properties of the Luce model suggests conditions which more flexible functional forms must satisfy. The Luce model is a member of a class of

[2] A random variable Y is Weibull distributed if $\text{Prob}[Y \le y] = \exp[-e^{-y}]$.

functional forms with selection probabilities satisfying

$$P_{in} = F(V(z_{in}) - V(z_{1n}), \ldots, V(z_{in}) - V(z_{Jn})), \tag{12}$$

where F is some increasing function of its arguments and the $V(z_{jn})$ are mean scale functions associated with the alternatives. Such models, in which the selection probabilities are functions solely of mean scale values and do not depend on the "orientation" or "similarity" of alternatives in attribute space, are termed *simply scalable*. Tversky (1972) has shown that simple scalability is equivalent to a condition of *order independence*: the selection probability for alternative 1 exceeds that for 2 when both are available *if and only if* the selection probability for 3 when 2 is available and 1 is not exceeds the selection probability for 3 when 1 is available and 2 is not. For example, $p_{12} \geq p_{21}$ if and only if $p_{13} \geq p_{23}$. The bicycle-pony example cited earlier shows that models satisfying order independence yield implausible conclusions when there are strong contrasts in similarity among the alternatives. Suppose alternatives 1 and 2 are bicycles differing in color, and 3 is a pony. Respondents are observed to be indifferent on the average between a pony and a bicycle, and to treat the two bicycles as equivalent. Then, the selection probabilities are $p_{12} = p_{13} = p_{23} = 1/2$ for binary choice and $p_{123} = p_{213} = 1/4$, $p_{312} = 1/2$ for multiple choice, where p_{ijk} is the probability of choosing i from the alternative set $\{i, j, k\}$ and $p_{312} = p_{31} = 1/2$ since the two bicycles are viewed as equivalent. Order independence implies that if $p_{312} > p_{123}$, then $p_{32} > p_{31}$, contradicting the observed selection probabilities. We conclude that useful extensions of the Luce model must avoid the simple scalability (order independence) property. No functional form which evaluates alternatives in terms of mean scale values without accounting for similarities in observed or unobserved attributes can avoid the restrictive implications of simple scalability. In particular, simple probit or transformed logit (McLynn (1973)) versions of (12) share the basic drawback of the Luce model.

The multinormal model in (7) avoids the simple scalability restriction, provided covariances are allowed to vary to reflect differential patterns of similarity. This model appears to have the greatest potential as a flexible choice model, provided computational barriers can be overcome. An alternative approach which may yield practical functional forms for applications is to seek more general axioms characterizing classes of selection probabilities. One such construction is the *elimination by aspects* model of Tversky (1972); the problems and potential of this model in economic applications have been discussed by McFadden (1975b).

5. *Unsolved Problems*

Each of the approaches to specification of selection probabilities discussed above suggests a number of questions for future study. Equation (6) yields a general form for selection probabilities,

$$P_{1n} = \int_{-\infty}^{+\infty} F_1(t, v_2 - v_1 + t, \ldots, v_J - v_1 + t)\, dt \tag{13}$$

where $F(\varepsilon_1, \ldots, \varepsilon_J)$ is the cumulative joint distribution of the $\varepsilon_j = \varepsilon(x_{jn}, c_n)$, F_1 is the derivative of F with respect to its first argument, and $v_j = V(x_{jn}, c_n)$ is the non-stochastic scale function for alternative j. Can joint distribution functions F be found which permit order non-independence and for which the selection probability (13) can be expressed in simple closed form? Suppose, more specifically, that utility has the structure given in (7). Given a distribution for the $\varepsilon(x_{jn}, c_n)$, one can use (6) to derive selection probabilities conditioned on α, which we denote by $p_{jn}(\alpha' Z^n)$. For example, assuming the $\varepsilon(x_{jn}, c_n)$ to be independently identically Weibull distributed would lead to the multinomial logit formula for the $p_{jn}(\alpha' Z^n)$. The selection probabilities would then satisfy

$$P_{jn} = \int_{-\infty}^{+\infty} p_{jn}(\alpha' Z^n) g(\alpha)\, d\alpha, \tag{14}$$

where g is the joint density function of the random parameter vector α. Can a plausible joint density g be found, say in the case that the p_{jn} are multinomial logistic, which allows (11) to be expressed in a simple closed form? Which allows P_{jn} to be computed easily?

The direct specification of selection probability functions and tests of the consistency of these functions with plausible choice theories has received little attention other than the elimination by aspects model of Tversky. Problems to be solved include development of (1) a practical parametric version of the Tversky model, (2) generalizations and simplifications of the Tversky model, for example the elimination by strategy model discussed by McFadden (1975b), and (3) links between properties of taste variations at the level of the individual and properties of the selection probability functions for the population.

The empirical consistency of choice theories also deserves examination. Although we have emphasized the compatibility of the random preference model with the classical economic model of fixed preference maximization, the analyses may be on sounder behavioral ground in postulating random preferences *within* each subject. It would be desirable to go further and test the validity of the random preference model of individual choice using psychological and economic data; the McFadden–Richter axiom provides a framework in which such tests could be carried out. One possible outcome of such an analysis would be the conclusion that behavior of a population of economic consumers can be described satisfactorily by a probability mixture of individuals, each of whom satisfies a random preference model. An interesting question would then be the "components of variance" within and between individuals and the problem of "spurious contagion" in predicting individual behavior.

Finally, the work of McFadden, Richter, and Hall on characterization of the underlying probability measures induced by selection probabilities suggests several questions: Is it possible in principle to identify a unique underlying measure from specified classes of experiments yielding selection probabilities? Can useful computational procedures be developed to provide Tchebycheff bounds?

III. Statistical Analysis of Selection Probabilities

1. *Functional Forms*

Theories of individual choice can be used to deduce a structure on selection probabilities, or to motivate the inclusion or exclusion of explanatory variables.

On the other hand, parametric specifications of the selection probabilities such as the multinomial logit model can be used (misused) in empirical analysis without reference to choice theoretic foundations. In this sense, parametric quantal response models are analogous to the linear statistical models used throughout econometrics, and all the usual econometric issues appear: consistency, efficiency, and robustness of estimators, errors in variables, multivariate systems, components of variance structures, structural identification, and simultaneous equations estimation. Unsurprisingly, most large sample theory for linear statistical models carries over to the quantal response case, although differences in detail often require argument. Small sample theory is in a much less satisfactory state, and is hampered by the lack of closed form estimators with satisfactory statistical properties.

The statistical literature divides into non-parametric tests of association in contingency tables, and parametric estimation and hypothesis testing in models of the probit or logit type. The first of these topics is reviewed by Lewis (1962), Goodman and Kruskal (1959), and Mosteller (1968). Parametric models have been much more influential in economic applications, and will receive the most attention here. An excellent general survey of results and handbook covering this area is C. R. Cox (1970); other surveys by Maxwell (1961) and Finney (1971) emphasize applications in psychology and bioassay, and a survey by McFadden (1973a) concentrates on multinomial response. Recent books by Bishop, Fineberg, and Holland (1975) and Haberman (1974) emphasize contingency table analysis, but also treat probit and logit models.

The most extensive statistical investigation of quantal response models has been made for binary choice. It was quickly recognized that the binary probit model is a special case of the class of functional forms

$$P_{1n} = F(\beta'(z_{1n} - z_{2n})), \tag{15}$$

where β is a $K \times 1$ vector of unknown parameters, the z_{jn} are $K \times 1$ vectors of numerical functions of the data, and F is any cumulative distribution function on the real line. It is worth noting some of the cases covered by the general linear-in-parameters specification $\beta'(z_{1n} - z_{2n})$. The components of z_{jn} may be simple or complex (non-linear) numerical functions of the data (x_{jn}, c_n); in particular, they may be terms in a "Hamel basis" for the space of all "smooth" scale functions of (x_{jn}, c_n) so that $\beta' z_{in}$ is a uniform approximation to an arbitary smooth function $V(x_{jn}, c_n)$. The interpretation of $\beta' z_{in}$ as an approximation is discussed further in McFadden (1975). In a case with $z_{jn_k} = x_{jn_k}$ (where k denotes the component) for $j = 1, 2$, aspect k of the alternatives is valued *generically*. In a case with $z_{1n_k} = x_{1n_k}, z_{2n_k} = 0, z_{1n_{k'}} = 0, z_{2n_{k'}} = x_{2n_k}$, the expression $\beta'(z_{1n} - z_{2n})$ contains the term $\beta_k x_{1n_k} - \beta_{k'} x_{2n_k}$, and an "alternative-specific" valuation of aspect k occurs. If in the last case, we had taken $z_{1n_k} = c_{n_k}, z_{2n_k} = 0$, $z_{1n_{k'}} = 0$, $z_{2n_{k'}} = c_{n_k}$, then $\beta'(z_{1n} - z_{2n})$ contains the term $(\beta_k - \beta_{k'})c_{n_k}$. Then a normalization, say $\beta_k + \beta_{k'} = 0$, is required. Models with this structure are termed *tolerance models*.

In addition to the probit model, commonly used transformations in (15) are the *logit model* (Berkson, 1944),

(16) $$P_{1n} = 1/(1+e^{-\beta'(z_{1n}-z_{2n})}),$$

the arctan model derived from the Cauchy distribution (Urban, 1910),

(17) $$P_{1n} = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\beta'(z_{1n}-z_{2n}),$$

and the linear model derived from the uniform distribution,

(18) $$P_{1n} = \begin{cases} 1 & (\beta'(z_{1n}-z_{2n}) \geq 1) \\ \beta'(z_{1n}-z_{2n}) & (0 < \beta'(z_{1n}-z_{2n}) < 1) \\ 0 & (\beta'(z_{1n}-z_{2n}) \leq 0) \end{cases}$$

2. *Methods for Estimation and Inference*

The parameter vector β in any of these models can be estimated from individual observations (i.e., $R_n = 1$) by maximum likelihood methods, and from repetitions (i.e., $R_n > 1$) by maximum likelihood, minimum chi-square, or modified minimum chi-square; iterative procedures are normally required to obtain solutions. An alternative method first exploited by Berkson (1944) for the logit model is to make a Taylor's expansion of the inverse transformation $F^{-1}(P_{1n}) = \beta'(z_{1n}-z_{2n})$ of (15) about the observed relative frequency S_{1n}/R_n, obtaining

(19) $$y_n \equiv F^{-1}(S_{1n}/R_n) = \beta'(z_{1n}-z_{2n}) + \eta_n$$

where

(20) $$\eta_n = \frac{(S_{1n}/R_n) - P_{1n}}{f(\xi_n)},$$

f is the frequency function corresponding to F and ξ_n is a point between $\beta' x_{1n}$ and $F^{-1}(S_{1n}/R_n)$. When the number of repetitions R_n becomes large, the second term in (19) is approximately normally distributed with mean zero and asymptotic variance $P_{1n}(1-P_{1n})/R_n f(\beta'(z_{1n}-z_{2n}))^2$. Weighted least squares applied to (19) then provides a consistent asymptotically efficient (as $R_n \to +\infty$) estimator of β under mild conditions on F and the explanatory variables. A good discussion of this method is given by Cox (1970).

When repetitions of choice settings are observed, the Berkson procedure has great computation advantages over the maximum likelihood estimator; Monte Carlo experiments (Berkson, 1953, 1955) suggest that it may also be statistically more satisfactory in small samples. However, many economic applications do not provide exact repetitions, making it necessary to group data in order to apply these methods. Monte Carlo experiments with grouping find generally that the Berkson procedure yields lower variances, larger biases, and comparable mean square errors when compared with maximum likelihood estimation (Domencich and McFadden, 1975, p. 112). This suggests that application of the Berkson method to grouped data, combined with a correction for the bias due to grouping, may be superior to other estimators. However, it should be noted that for more than three or four independent variables, it is often impossible to group data satisfactorily.

The linear probability model can be fitted to individual observations by ordinary least squares; the primary disadvantage of this model is that the estimator is extremely sensitive to specification error in the response curve (Domencich and McFadden, 1975; Nerlove and Press, 1973). Ladd (1966) has shown that the usual discriminant model estimator is formally equivalent to the ordinary least squares estimator of the linear probability model. Further, the relative odds of correct binary classification are given by the logit formula. Nevertheless, the discriminant model is based on a statistical structure that is incompatible with most choice theories, and the discriminant estimator is not in general a consistent estimator of the parameters in the logit model when the selections are generated by the latter model (McFadden, 1976b).

Charles Manski (1975) has recently suggested a class of *maximum score methods* which provide consistent estimators of the (normalized) parameters in the quantal response model (15) without specific distributional assumptions of F. These techniques are based on maximization of the (weighted) number of predictive "successes" of the model, where a success in choice situation n is defined as observation of a response j such that $\beta' z_{jn} > \beta' z_{in}$ for $i \neq j$. Maximum score methods require iterative computational procedures, and cannot in general be efficient, but have the asset of robustness with respect to the transformation F.

The statistical analysis of multinomial response models has developed rapidly in the past decade, with considerable impetus from economists. The multinomial logit model was first developed for a special case of tolerance analysis by Gurland, Lee, and Dahm (1960), and more generally by Bloch and Watson (1967), Bock (1968), Rassam *et al.* (1971), McFadden (1973a, 1975a), Nerlove and Press (1973), Stopher (1969), and Theil (1969, 1970).[3]

Theil (1969) has provided a generalization of Berkson's method for the multinomial logit model in which the parameters are estimated by application of multivariate least squares with linear restrictions across equations. When repetitions of choice settings are observed, or the data can be satisfactorily grouped, this procedure has considerable computational merit. Maximum likelihood estimation is most commonly used when data cannot be conveniently grouped. The log likelihood function for the multinomial logit model

$$L = \sum_{n=1}^{N} \sum_{j=1}^{J} S_{nj} \log P_{jn}, \tag{21}$$

[3] These developments seem to have occured in at least three independent streams. The statistical literature (Bloch, Bock) stems from the Gurland, Lee and Dahm paper, and concentrates on problems in bioassay. The work of Theil (1969), and unpublished works for the same period by John Cragg, stem from the Warner (1962) analysis of binary travel decisions, which also influenced Stopher and others working on problems of travel demand. The approach of McFadden (1968, 1973) grows out of the choice theory analyses of Luce (1959) and Marschak (1960), and was stimulated by a problem of Phoebe Cottingham (1966) on the analysis of highway route selection decisions. The author developed a maximum likelihood program in 1965 to perform multinomial logit estimation, and became aware of the parallel developments of Cragg and Theil during a visit to the University of Chicago during 1966–67.

with

$$P_{jn} = e^{\beta' z_{jn}} \Big/ \sum_{i=1}^{J} e^{\beta' z_{in}}, \tag{22}$$

is strictly concave, with a Hessian matrix which is independent of the observed responses. Consequently, efficient computer programs can be written to produce the estimates.[4] Details of the maximum likelihood method applied to multinomial logit, including its statistical properties, are discussed in McFadden (1973). The log likelihood function (21) assumes a sample randomly drawn from the population without reference to the observed choices. Let $p_i(z)$ denote the selection probability, $g(z)$ denote the prior distribution of the explanatory variables, $q_i(z)$ denote the posterior distribution of the explanatory variables conditioned on choice of i, and P_i denote the population mean selection probability. Then, the joint distribution of z and i is $p_i(z)g(z) = q_i(z)P_i$, and the log likelihood function for observations obtained by random sampling of z is (except for combinatorial factors which do not enter optimization)

$$L = \sum_{n=1}^{N} \sum_{i=1}^{J} S_{in} \log (p_i(z)g(z)). \tag{23}$$

Since $g(z)$ does not depend on the selection probability parameters under the choice theory specification, this equation is equivalent to (21).

An alternative sampling procedure, based on observed choices, has been investigated by Manski and Lerman (1976). This approach is of great importance for applications, where respondents are most readily observed at the "place of choice" (e.g., on-board surveys of commuters, surveys of convicted criminals), or where enriched sampling of a low-probability outcome is desired (e.g., welfare recidivists, migrants). With this sampling procedure, the log likelihood function is (excluding combinatorial factors):

$$\begin{aligned} L &= \sum_{n=1}^{N} \sum_{i=1}^{J} S_{in} \log q_i(z) \\ &= \sum_{n=1}^{N} \sum_{i=1}^{J} S_{in} \log (p_i(z)g(z)/P_i), \end{aligned} \tag{24}$$

with

$$P_i = \int p_i(z)g(z)\, dz. \tag{25}$$

[4] The following routines are currently available: (1) XLOGIT, a program written by McFadden, Varian, and Wills for the CDC 6400 and adapted to the IBM 360 and Burroughs computers, (2) ULOGIT, a program written by the U.S. Bureau of Standards for IBM 360, (3) a multinomial logit program written by Manski for the IBM 360, and used at M.I.T. and elsewhere, (4) QUAIL, a program written by Berkman, Brownstone, McFadden, Murano, and Wills for the CDC 6600, and (5) a multivariate multinomial logit program written by Nerlove for the RAND Corporation. There are a number of commercial programs also available. The Manski program is the most readily adapted and used for small and intermediate estimation tasks. ULOGIT is available at installations which have the U.S.D.O.T. Urban Transportation Planning Package. QUAIL is a flexible program which permits data transformation and storage, and retrieval of intermediate results, in addition to basic estimation commands.

Since P_i depends on the parameters of the selection probabilities, this likelihood function is distinct from (21), and maximization of (21) using a choice-based sample will, in general, produce inconsistent estimates. Lerman and Manski establish a procedure for weighting the observations S_{in} in (21) so that maximization of this function yields consistent estimates.

The lack of explicit forms for multinomial selection probabilities except in the logit case has hindered progress in developing alternative statistical procedures. Several investigations have employed models of the form

$$P_{jn} = \frac{F(\beta' z_{jn})}{\sum_{i=1}^{J} F(\beta' z_{in})}, \tag{23}$$

where F is a positive increasing real valued function. This becomes the multinomial logit model when $F(b) = e^b$. Alternatives include taking F to be a standard cumulative normal function. This is sometimes misleadingly termed the multinomial probit model; this term might more appropriately be reserved for the model of equation (7) when the distribution of the stochastic components of utility is multivariate normal. Models of the form (23) share with the multinomial logit model the sometimes undesirable property of independence from irrelevant alternatives, and can be interpreted as versions of the strict utility model with the nonlinear-in-parameters scale functions $v(z_{in}) = \log F(\beta' z_{in})$. There appears to be little gain in functional generality in this specification, and the computational cost entailed in the loss of the concave programming structure of the multivariate logit model may be considerable.

Estimation of alternatives to the multinomial logit model by the maximum likelihood method raises a computational problem of evaluating the selection probabilities in an iterative procedure. Analysis of methods has concentrated on the multinomial normal model (7). Let $\Omega = (\omega_{ij}) = \sigma_0^2 I + Z'DZ$ denote the covariance matrix of the deviations of utility levels from the mean scale values $V_j = \beta' z_{jn}$. Define

$$\lambda_{2j} = \omega_{2j} + \omega_{11} - \omega_{21} - \omega_{j1} \tag{24}$$

and recursively, for $i = 3, \ldots, J$,

$$\lambda_{ik} = \omega_{ik} + \omega_{11} - \omega_{i1} - \omega_{k1} - \sum_{j=2}^{i-1} \lambda_{ji}\lambda_{jk}/\lambda_{jj} \qquad (k = 2, \ldots, J). \tag{25}$$

Then Domencich and McFadden (1975, p. 67) show that the selection probability for the first alternative can be written

$$\begin{aligned} P_{1n} = {} & \frac{1}{\sqrt{\lambda_{22} \ldots \lambda_{JJ}}} \int_{t_2=-\infty}^{V_1 - V_2} \phi\left(\frac{t_2}{\sqrt{\lambda_{22}}}\right) \int_{t_3=-\infty}^{V_1 - V_3 - (\lambda_{23}/\lambda_{22})t_2} \phi\left(\frac{t_3}{\sqrt{\lambda_{33}}}\right) \cdots \\ & \cdots \int_{t_J=-\infty}^{V_1 - V_J - \sum_{j=2}^{J-1} (\lambda_{jJ}/\lambda_{jj}) t_j} \phi\left(\frac{t_J}{\sqrt{\lambda_{JJ}}}\right) dt_J \ldots dt_2, \end{aligned} \tag{26}$$

where ϕ is the standard univariate normal density. Evaluation of this formula by numerical methods is elementary for $J=3$, but is costly for $J=4$ and 5 and prohibitive for $J>5$. The use of advanced numerical methods (Dutt, 1976; Hausman and Wise, 1976) promise improvements in computation time, but do not appear to make this approach practical for large J.

Several indirect methods for calculating multinomial selection probabilities are being investigated. A promising Monte Carlo method developed by Manski substitutes calculated deviates in (7), and then estimates probabilities from the relative frequencies of calculated utility maxima. A second approach, suggested by Talvitie, is to approximate the ε error terms in (7) with Weibull errors, yielding the formula

$$P_{1n} \doteq \int \frac{e^{k\alpha' z_{in}}}{\sum_{i=1}^{J} e^{k\alpha' z_{in}}} x(\alpha)\, d\alpha, \tag{27}$$

where $x(\alpha)$ is the density of the random taste parameters α and k is a large positive parameter. If the dimension of the taste parameter vector is low, then direct numerical evaluation of this formula is practical. Alternatively, approximations can be obtained by making a series expansion of the logistic formula in (26) such that the expectation with respect to α can be calculated term by term. This has been done by Talvitie using a Taylor's series expansion in powers of α, by Spence using an expansion in terms of moment generating functions, and by McFadden using Fourier series expansions. Results to date suggest that these expansions converge slowly, particularly for extreme probabilities, making them impractical in interative statistical procedures. The question of the existence of expansions with rapid convergence deserves further exploration.

In maximum likelihood estimation of multinomial quantal response models, it is convenient to define residuals and measures of goodness of fit analogous to those in linear statistical models. The likelihood ratio statistic provides a convenient basis for defining an index of "proportion of variance explained,"

$$\rho^2 = 1 - L_1/L_2 \tag{28}$$

where L_1 is the log likelihood function evaluated at the maximum likelihood estimate for the model under consideration, and L_2 is the log likelihood evaluated at the maximum likelihood estimate for a model in which only pure alternative effects appear. The properties of this index are discussed by McFadden (1973). A more conventional measure of goodness of fit, corresponding to the multiple correlation coefficient, can be based on the weighted sum of squared deviations of observations from predicted values,

$$\chi^2 = \sum_{n=1}^{N} \sum_{j=1}^{J} R_n D_{jn}^2, \tag{29}$$

with

$$D_{jn} = (S_{jn} - R_n P_{jn})/\sqrt{R_n P_{jn}}. \tag{30}$$

The statistic χ^2 is asymptotically chi-squared with $N(J-1)-K$ degrees of freedom when the $R_n \to +\infty$; but is less stable in small samples when valued at the maximum likelihood estimator than is the likelihood ratio. In samples without repetitions ($R_n = 1$), this measure is poorly behaved, and its use is not recommended. The residuals defined in (30) can be used in tests for the importance of excluded variables and for the multinomial logit specification. A transformation of the residuals suggested by Cox and Snell (1968) for the binary case can be generalized to provide residuals which are approximately uncorrelated, with mean zero and unit variance, when the model specification is correct. The details of these constructions are given in McFadden (1973).

3. *Extensions of the Statistical Choice Problem*

An area of multiple quantal response where models other than the multinomial logit form have been estimated successfully is the study of binary transition probabilities in a Markov chain. Data on waiting times, for example, can be described by a negative binomial distribution, with the binomial probability given by a binary probit model or similar form. Papers by Heckman and Willis (1974) and Boskin and Nold (1974) apply such models to the estimation of conception probabilities and welfare recidivism. Multinomial transition probabilities can also be estimated using the logit model; a typical formulation is

$$P_{ij} = e^{\beta_i' z_j} \Big/ \sum_{k=1}^{J} e^{\beta_i' z_k}, \tag{31}$$

where P_{ij} is the transition probability from i to j and $(z_1, \ldots, z_J)$ describe the attributes of the alternative states.

The model in (31) assumes a homogeneous population; the estimation problem is complicated considerably under the assumption of heterogeneity. This problem, which is a generalization of the classical "mover-stayer" problem in sociology and demography, has been analyzed by Hall (1973), Heckman and Willis (1974), and Spilerman (1972).

4. *Multivariate Choice, Simultaneity, and Independence*

We have thus far viewed the quantal choice experiment as the observation of one of J possible alternatives along a single dimension of choice; for example, the choice among several colleges. However, the alternatives $j = 1, \ldots, J$ can also be taken to index the joint responses on several quantal choices. Suppose, for example, an individual is considering J_1 travel modes and J_2 possible destinations. Then there are $J_1 \cdot J_2 = J$ possible joint responses, now more conveniently given the double index ik, where i is the mode and k is the destination, rather than a single running index j. We see from this example that the problem of multivariate multinomial quantal response can always be treated formally as one of univariate multinomial response. However, the multivariate structure introduces a series of important new questions and problems. The first is that interest in applications is often on the *conditional* or *marginal* selection probabilities for a particular quantal response variable. For example, one may wish to forecast mode choice conditioned on destination, or to estimate the marginal destination choice selection probabilities for all mode choices. Concentration on conditional and marginal selection probabilities may also make it possible to avoid treating a prohibitive number of total alternatives.

One approach to the problem of formulating simultaneous choice models is to begin with the random utility model, and utilize separability properties of the underlying utility functions to derive statistical properties of the selection probabilities. This tack has been followed by Domencich and McFadden (1975) and by Ben-Akiva (1972). To illustrate the structural features of this method, consider an example in which the utility associated with alternative ij is

$$U(x_i, y_i, z_{ij}) = \alpha' z_{ij} + \beta' x_i + \gamma' y_i + \varepsilon_{ij}, \tag{32}$$

where z_{ij}, x_i, and y_j are explanatory variables, ε_{ij} is a random effect, and we now suppress the choice setting index n. The most interesting and challenging case occurs when ε_{ij} has a components of variance structure; however, we initially assume the ε_{ij} are independently Weibull distributed, yielding the multinomial logit model

$$P_{ij} = \frac{\exp[\alpha' z_{ij} + \beta' x_i + \gamma' y_j]}{\sum\sum_{kl} \exp[\alpha' z_{kl} + \beta' x_k + \gamma' y_l]}. \tag{33}$$

The marginal probabilities are then

$$P_i = \sum_j P_{ij} = \frac{\sum_j \exp[\alpha' z_{ij} + \beta' x_i + \gamma' y_j]}{\sum\sum_{kl} \exp[\alpha' z_{kl} + \beta' x_k + \gamma' y_l]}, \tag{34}$$

and the conditional probabilities satisfy

$$P_{j|i} \equiv P_{ij}/P_i = \frac{\exp[\alpha' z_{ij} + \gamma' y_j]}{\sum_k \exp[\alpha' z_{ik} + \gamma' y_k]}. \tag{35}$$

The conditional probabilities from the multivariate multinomial logit model are again multinomial logit; this is a consequence of the independence from irrelevant alternatives property and is *not* true for general selection probabilities. The condition for the two responses to be statistically independent (i.e., $P_{ij} = P_i P_j$) is that $\alpha' z_{ij}$ be zero.

One could estimate the parameter vectors α and γ from (35) employing conditional data on j given i. Considering the case that j is binary, and assuming repetitions, we can write (35) in the form

$$\log \frac{S_{1i}}{R_i - S_{1i}} = \gamma'(y_1 - y_2) + \alpha'(z_{i1} - z_{i2}) + \eta, \tag{36}$$

where R_i is the number of responses where i is chosen and S_{1i} is the number of these for which $j = 1$ is chosen, and

$$\eta = (S_{1i} - R_i P_{1|i}) / R_i P_{1|i}(1 - P_{1|i}). \tag{37}$$

One can easily verify that η and $(z_{i1}-z_{i2})$ are uncorrelated, so that ordinary least squares applied to (36) provides unbiased estimates of the coefficients. This conclusion implies that in applications such as estimation of traveler mode choice conditioned on auto ownership or household monthly probability of conception conditioned on contraceptive technique, no bias is introduced by including the value of the conditioning quantal response as an explanatory variable. Consider the particular case with $z_{11}=1$ and $z_{12}=z_{21}=z_{22}=0$. Then the model can be written

$$\log\frac{S_{1i}}{R_i-S_{1i}}=\gamma'(y_1-y_2)+\alpha\delta_i+\eta, \tag{38}$$

where δ_i is a dummy variable which is one if $i=1$ and zero otherwise. Then, δ_i can be interpreted as the outcome of choice in the dimension i, and (38) can be interpreted as an equation for the conditional choice j which depends via the shift effect $\alpha\delta_i$ on the choice of i. We then have the remarkable conclusion that the presence of an "endogeneous" quantal response variable on the right-hand-side of (38) does *not* disturb the consistency of the ordinary single equation maximum likelihood estimator. However, it should be emphasized that this conclusion depends critically on the assumption that the random components in utility ε_{ij} are independent. If, to the contrary, these errors contain unobserved components which tend to increase the probability that $j=1$ when $i=1$, the classical statistical problem of "spurious contagion," also known as the "mover-stayer" problem, arises. Intuition suggests that in a model such as (36) where the possibility of a misspecification exists which would lead to a correlation of $(z_{i1}-z_{i2})$ and η, an instrumental variables procedure should produce more satisfactory estimates. In an application, McFadden (1974) has used as an instrumental variable a "reduced form" fitted logit probability for δ_i in (38). However, no argument is given for the statistical desirability of this procedure.

A comprehensive treatment of simultaneous equations with quantal response has recently been made by Heckman (1975). Methods for direct analysis of multivariate responses in the multinomial logit framework have been developed by Nerlove and Press (1973), who develop an analysis of variance framework for multivariate multinomial logit analysis. Their structure is an extension of the example above, and is closely related to the log linear models treated by Bishop *et al.* (1975) and Haberman (1974). A second approach by Amemiya (1974) begins with a specification of the selection probabilities that is related to earlier statistical work on tolerance analysis. In the case of two binary responses, a binomial probit model is defined.

$$\begin{aligned}
&P_{11}=N(\alpha'x,\beta'y,\Omega)\\
&P_{10}=N(\alpha'x,+\infty,\Omega)-N(\alpha'x,\beta'y,\Omega)\\
&P_{01}=N(+\infty,\beta'y,\Omega)-N(\alpha'x,\beta'y,\Omega)\\
&P_{00}=1+N(\alpha'x,\beta'y,\Omega)-N(+\infty,\beta'y,\Omega)-N(\alpha'x,+\infty,\Omega),
\end{aligned} \tag{39}$$

where $\alpha'x$ and $\beta'y$ are linear-in-parameters functions of the respective binary alternatives, Ω is a 2×2 covariance matrix, and N is the bivariate cumulative normal distribution. There does not appear to be a random utility model which will produce this specification of the selection probabilities. However, the model itself is of considerable empirical appeal, and will be applicable in many cases where a choice theory rationalization is irrelevant.

5. *Unsolved Problems*

We conclude this section by listing some problems of statistical analysis for future study. First, there is need for further work on estimation methods in simple quantal response models, and on the small sample properties of the estimators. It would be particularly interesting to pursue distribution-free methods such as Manski's maximum score estimators. Since maximum likelihood estimation will probably continue to be the most useful general purpose estimation procedure, work should be done on correcting these estimators to improve their statistical properties in samples of moderate size. Data grouping methods for use of Berkson's methods should be investigated. There is need for further development of computer programs to carry out estimation, particularly when large numbers of observations are involved or forms other than multinomial logit are adopted.

There is a pressing need for development of practical models for selection probabilities which do not have the independence of irrelevant alternatives property. The issues arising in such constructions may also be relevant to construction of components of variance models where the mover-stayer problem can be analyzed. It would be particularly useful to achieve a computational breakthrough on the multinomial normal model.

The subject of multivariate and structural estimation of quantal response requires further work in almost every aspect; in particular, there has been little work on estimators for structural models.

IV. Economic Applications

Economic applications employing binary quantal choice models have been most common in studies of travel demand modal split, migration, and the demand for consumer durables. A selection of papers are Allouche (1972), Amemiya and Boskin (1972), Fisher (1967), Korbel (1966), Lave (1970), Lee (1963), Lisco (1967), McGillivray (1970), Moses *et al.* (1967), Reichman and Stopher (1971), Stopher (1969), Stopher and Lisco (1970), Talvitie (1972), Thomas and Thompson (1971), Uhler (1968), Walker (1968), Warner (1962), and Zellner (1965).

The most extensive use of multinomial response models has been in travel demand analysis; the reader is referred to the following contributions: Ben-Akiva (1972), Brand (1972), Lave (1970), McFadden (1973a), Domencich and McFadden (1975), McFadden and Reid (1974), McLynn (1971), Stopher and Lisco (1970), Talvitie (1972, 1973), Watson and Westin (1973), Watson (1972), and Wigner (1973). Of particular interest are the issues of choice structure and independence discussed in ben-Akiva (1972) and Domencich and McFadden (1975), and the problems of forecasting in aggregate models discussed by McFadden and Reid (1974), Talvitie (1973), and Watson and Westin (1973).

Other areas of application are analysis of college-going behavior (Miller and Radner, 1970, 1974, and Kohn, Manski, and Mundel, 1976); occupational choice (Boskin, 1972) and housing location (Friedman, 1974, Pollakowski, 1975, and Lerman, 1975). Areas in which quantal choice models have been used on only a limited basis, and in which substantial potential exists for useful exploitation of these methods are fertility behavior, migration, economic determinants of voting behavior, ownership and brand of consumer durables, and labor market behavior, including job search and occupational choice. The techniques for quantal response analysis surveyed in this paper may prove useful for these problems, and for related problems in other social sciences. On the other hand, the topics suggested here for further investigation and the requirements of new applications may lead to the development of practical quantal response models which provide a more accurate description of individual choice among qualitative alternatives.

University of California, Berkeley

References

Adler, T. J. and M. Ben-Akiva (1975), "A Joint Frequency, Destination, and Mode Choice Model for Shopping Trips," paper forthcoming in a Transportation Research Board *Record*.

Aitchison, J. and S. Silvey (1957), "The Generalization of Probit Analysis to the case of Multiple Responses," *Biometrika*, pp. 131-140.

Aitchson, J. and J. Bennett (1970), "Polychotomous Quantal Response by Maximum Indicant," *Biometrika*, 57, pp. 253–262.

Allouche, J. (1972), "Approach to Probability Distribution of Value of Walking Time and Pedestrian Circulation Models," Highway Research Board *Record*, 392, pp. 121–133.

Amemiya, T. and M. Boskin (1974), "Regression Analysis When the Dependent Variable is Truncated Lognormal, with an Application to the Determinants of the Duration of Welfare Dependency," *International Economic Review*, 15, pp. 485–496.

Amemiya, T. (1973), "Regression Analysis When the Dependent Variable is Truncated Normal," *Econometrica*, 41, pp. 997–1016.

Amemiya, T. (1974), "Multivariate Regression and Simultaneous Equations Models When the Dependent Variables are Truncated Normal," *Econometrica*, 42, pp. 999–1012.

Amemiya, T. (1974), "Bivariate Probit Analysis: Minimum Chi-Square Methods," *Journal of the American Statistical Association*, 69, pp. 940–944.

Amemiya, T. (1974), "The Maximum Likelihood Estimator vs. The Minimum Chi-Square Estimator in the General Qualitative Response Model," Institute for Mathematical Studies in the Social Sciences, Technical Report No. 136, Stanford University.

Amemiya, T. (1974), "The Equivalence of the Nonlinear Weighted Least Squares Method and the Method of Scoring in the General Qualitative Response Model," Institute for Mathematical Studies in the Social Sciences, Technical Report No. 137, Stanford University.

Amemiya, T. (1975), "Qualitative Response Models," *Annals of Economic and Social Measurement*, 4, 363–372.

Amemiya, T. and F. Nold (1975), "A Modified Logit Model," *Review of Economics and Statistics*, 57, 255–57.

Anscombe, E. J. (1956), "On Estimating Binomial Response Relations," *Biometrika*, 43, 461–464.

Antle, C., L. Klimko, and W. Harkness (1970), "Confidence Intervals for the Parameters of the Logistic Distribution," *Biometrika*, 57, pp. 397–402.

Ashford, J. R. and R. R. Sowden (1970), "Multivariate Probit Analysis," *Biometrics*, 26, pp. 535–546.

Ashton, W. (1972), *The Logit Transformation*, Hafner, New York.

Atherton, T. J. and M. Ben-Akiva (1976), "Transferability and Updating of Disaggregate Travel Demand Models," paper prepared for presentation at the 55th annual meeting of the transportation Research Board, Washington, DC, January 1976.

Atkinson, A. (1972), "A Test of the Linear Logistic and Bradley–Terry Models," *Biometrika*, Vol. 59, pp. 37–42.

Bard, Y. (1974), *Nonlinear Parameter Estimation*, Academic Press, New York.
Bartlett, M. S. (1935), "Contingency Table Interactions," *Supplement to the Journal of the Royal Statistical Society*, 2, 248–252.
Bates, T. (1973), "An Econometric Analysis of Lending to Black Buisnessmen," *Review of Economics and Statistics*, Vol. 55, pp. 272–283.
Beals, R., D. H. Krantz, and A. Tversky (1968), "Foundations of Multidimensional Scaling," *Psychological Review*, 75, 127–142.
Ben-Akiva, M. (1973), "Structure of Passenger Travel Demand Models," Ph.D. dissertation, Massachusetts Institute of Technology, Department of Civil Engineering, Cambridge, MA, unpublished.
Ben-Akiva, M. (1974), "Structure of Passenger Travel Demand Models," Transportation Research Board *Record*, 526.
Ben-Akiva, M. (1974), "Multi-Dimensional Choice Models: Alternative Structures of Travel Demand Models," Transportation Research Board *Special Report*, 149.
Ben-Akiva, M. and S. Lerman (1974), "Some Estimation Results of a Simultaneous Model of Auto Ownership and Mode Choice to Work," *Transportation*, 3, pp. 357–376.
Ben-Akiva, M. and M. G. Richards (1975), "A Disaggregate Multi-modal Model for Work Trips in the Netherlands," paper prepared for presentation at the 54th annual meeting of the Transportation Research Board, Washington, DC, January, 1975. Forthcoming in Transportation Research Board *Record*.
Bennett, B. M. (1964), "A Non-Parametric Test for Randomness in a Sequence of Multinomial Trials," *Biometrika*, 20, 182–190.
Berkson, J. (1955), "Estimate of the Integrated Normal Curve by Minimum Normit Chi-Square with Particular Reference to Bio-Assay," *Journal of the American Statistical Association*, 50, pp. 529–549.
Berkson, J. (1951), "Why I Prefer Logits to Probits," *Biometrika*, 7, 327–339.
Berkson, J. (1944), "Application of the Logistic Function to Bio-Assay," *Journal of the American Statistical Association*, 39, pp. 357–365.
Berkson, J. (1953), "A Statistically Precise and Relatively Simple Method of Estimating the Bio-Assay with Quantal Response, Based on the Logistic Function," *Journal of the American Statistical Association*, 48, pp. 565–599.
Berkson, J. (1955), "Maximum Likelihood and Minimum Chi-Square Estimations of the Logistic Function," *Journal of the American Statistical Association*, 50, pp. 130–161.
Bishop, Y., S. Fienberg, and P. Holland (1975), *Discrete Multivariate Analysis*, MIT Press.
Bliss, C. I. (1934), "The Method of Probits," *Science*, 79, 38–39.
Bloch, D. and G. Watson (1967), "A Bayesian Study of the Multinomial Distribution," *Annals of Mathematical Statistics*, 38, pp. 1423–1435.
Block, H. and J. Marschak (1960), "Random Orderings and Stochastic Theories of Response," in I. Olkin, ed., *Contributions to Probability and Statistics*, Stanford University Press.
Blood, D. M. and C. B. Baker (1958), "Some Problems in Linear Discrimination," *Journal of Farm Economics*, 40, 674–683.
Bock, R. D. (1968), "Estimating Multinomial Response Relations," in *Contributions to Statistics and Probability: Essays in Memory of S. N. Roy*, University of North Carolina Press, Chapel Hill.
Bock, R. D. and L. V. Jones (1968), *The Measurement and Prediction of Judgement and Choice*, San Francisco, Holden-Day.
Boskin, M. and F. Nold (1974), "A Markov Model of Turnover in Aid to Families with Dependent Children," Technical Report No. 125, Institute of Mathematical Studies in the Social Sciences, Stanford University.
Boskin, M. (1974), "A Conditional Logit Model of Occupational Choice," *Journal of Political Economy*, 82, 389–398.
Brand, D. (1973), "Travel Demand Forecasting: Some Foundations and a Review," Highway Research Board *Special Report* 143, *Urban Travel Demand Forecasting*, Washington, DC.
Brand, D. (1976), "Travel Behavior Research Approaches," Transportation Research Board *Record*, forthcoming.
Brand, R. J., D. E. Pinnock, and K. L. Jackson (1973), "Large Sample Confidence Bands for the Logistic Response Curve and Its Inverse," *The American Statistician*, 27, 157–159.
Cacoullos, T. (1973), *Discriminant Analysis and Applications*, ed., Academic Press, New York.
Chambers, E. A. and D. R. Cox (1967), "Discrimination between Alternative Binary Response Models," *Biometrika*, 54, 573–578.
Coleman, J. S. (1968), "The Mathematical Study of Change," in H. Blalock and A. Blalock, eds., *Methodology in Social Research*.
Cook, P. (1975), "The Correctional Carrot: Better Jobs for Parolees," *Policy Analysis*, 1, 11–54.

Corbin, R. and A. A. J. Marley (1974), "Random Utility Models with Equality: An Apparent, but not Actual, Generalization of Random Utility Models," *Journal of Mathematical Psychology*, 3, 174–293.

Cottingham, P. (1966), "Measurement of Non-User Benefits," Center for Planning and Development Research, University of California, Berkeley, unpublished.

Cox, D. (1958), "The Regression Analysis of Binary Sequences," *Journal of the Royal Statistical Society*, Series B., pp. 215–242.

Cox, D. (1966), "Some Procedures with the Logistic Qualitative Response Curve," in F. David, ed., *Research Papers in Statistics*, Wiley, New York.

Cox, D. and E. Snell (1968), "A General Definition of Residuals," *Journal of the Royal Statistical Society*, Series B., 30, pp. 248–265.

Cox, D. (1970), *Analysis of Binary Data*, Methuen, London.

Cox, D. and E. Snell (1971), "On Test Statistics Calculated from Residuals," *Biometrika*, 58, pp. 589–594.

Cox, D. R. (1972), "The Analysis of Multivariate Binary Data," *Applied Statistics*, 21, 113–120.

Cragg, J. and R. Uhler (1970), "The Demand for Automobiles," *Canadian Journal of Economics*, 3, pp. 386–406.

Debreu, G. (1960), "Review of R. Luce, Individual Choice Behavior," *American Economic Review*, 186–188, Vol. 50.

Domencich, T. and D. McFadden (1975), *Urban Travel Demand: A Behavioral Analysis*, North-Holland Publishing Company, Amsterdam.

Dutt, J. (1976), "Numerical Aspects of Multivariate Normal Probabilities in Econometric Models," *Annals of Economic and Social Measurement*, this issue.

Efron, B. (1975), "The Efficiency of Logistic Regression Compared to Normal Discriminant Analysis," *Journal of the American Statistical Association*, 70, 892–898.

Eisenbeis, R. A. and R. B. Avery (1972), *Discriminant Analysis and Classification Procedures: Theory and Applications*, D. C. Heath, Lexington, MA.

Elashoff, J. D., R. M. Elashoff, and G. E. Goldman (1967), "On The Choice of Variables in Classification Problems with Dichotomous Variables," *Biometrika*, 54, pp. 668–670.

Ergun, G. (1971), "Development of a Downtown Parking Model," Highway Research Board *Record*, 369, pp. 118–134.

Fan, K. (1956), On Systems of Linear Inequalities, in H. Kuhn and A. W. Tucker, eds., *Linear Inequalities and Related Systems*, *Annals of Mathematical Studies*, 38, Princeton University Press.

Feldstein, M. S. (1966), "A Binary Variable Multiple Regression Method of Analyzing Factors Affecting Perinatal Mortality and Other Outcomes of Pregnancy."

Ferber, R. (1954), "Factors Influencing Durable Goods Purchases," in L. H. Clark, ed., *Consumer Behavior*, Vol. 2, New York University Press.

Finney, D. (1971), *Probit Analysis*, 3rd edition, Cambridge University Press.

Fisher, R. A. (1936), "The Use of Multiple Measurements in Taronomic Problems," *Annals of Eugenics*, 7, 179–188.

Fisher, R. A. (1935), "The Logic of Induction Inference (with discussion)," *Journal of the Royal Statistical Society*, 98, 39–54. Reprinted without discussion in R. A. Fisher, *Contributions to Mathematical Statistics*, Wiley, New York, 1950.

Fisher, J. (1967), "An Analysis of Consumer Goods Expenditures in 1957," *Review of Economics and Statistics*, Vol. 49, pp. 64–71.

Freedman, D. and R. Purves (1969), "Bayes Method for Bookies," *Annals of Mathematical Statistics*, Vol. 40, pp. 1117–1186.

Friedman, J. (1974), "Housing Location and Supply of Local Public Services," Ph.D. dissertation, Department of Economics, University of California, Berkeley, unpublished.

Friedman, P. (1973), "Suggestions for the Analysis of Qualitative Dependent Variables," *Public Finance Quarterly*, 1, 345–355.

Gart, J. and J. Zweifel (1967), "On the Bias of Various Estimators of the Logit and Its Variance," *Biometrika*, 54, pp. 181–187.

Gilbert, E. (1968), "On Discrimination Using Qualitative Variables, *Journal of the American Statistical Association*, Vol. 63, pp. 1399–1412.

Goldberger, A. S. (1964), *Economic Theory*, Wiley, New York.

Goldberger, A. S. (1971), "Econometrics and Psychometrics: A Survey of Communalities," *Psychometrika*, pp. 83–107, Vol. 36.

Goldberger, A. S. (1973), "Correlations between Binary Outcomes and Probabilistic Predictions," *Journal of the American Statistical Association*, 68, p. 84.

Goldfeld, S. M. and R. E. Quandt (1972), *Nonlinear Methods in Econometrics*, Amsterdam, North-Holland Publishing Company.

Goldstein, M. and M. Rabinowitz (1975), "Selection of Variates for the Two-Group Multinomial Classification Problem," *Journal of the American Statistical Associations*, 70, 776–781.

Goodman, L. A. (1965), "On the Multivariate Analysis of Three Dichotomous Variables," *American Journal of Sociology*, 71, 290–301.

Goodman, L. A. (1970), "The Multivariate Analysis of Qualitative Data: Interactions Among Multiple Classifications," *Journal of the American Statistical Association*, 65, 226–256.

Goodman, L. A. (1971), "The Analysis of Multidimensional Contingency Tables: Stepwise Procedures and Direct Estimation Methods for Building Models for Multiple Classifications," *Technometrics*, 13, 33–61.

Goodman, L. A. (1972), "A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables," *American Sociological Review*, 37, 28–46.

Goodman, L. A. (1972), "A General Model for the Analysis of Surveys," *American Journal of Sociology*, 77, 1035–1086.

Goodman, L. A. (1972), "A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables," *American Sociological Review*, 37, 28–46.

Goodman, L. A. (1973), "Causal Analysis of Panel Study Data and Other Kinds of Survey Data," *American Journal of Sociology*, 78, pp. 1135–1191.

Goodman, L. A. (1974), "The Relationship between the Modified and the More Usual Multiple Regression Approach to the Analysis of Dichotomous Variables," Technical Report No. 9, Department of Statistics, University of Chicago.

Goodman, L. A. and W. H. Kruskal (1954), "Measures of Association for Cross Classifications," *Journal of the American Statistical Association*, 49, 732–764.

Goodman, L. A. and W. H. Kruskal (1959), "Measures of Association for Cross Classification II, Further Discussion and References," *Journal of the American Statistical Association*, 54, 123–163.

Griliches, A. (1973), "Errors in Variables and Other Unobservables," Department of Economics, Harvard University, unpublished.

Grizzle, J. (1962), "Asymptotic Power of Tests of Linear Hypotheses Using the Probit and Logit Transformations," *Journal of the American Statistical Association*, Vol. 57, pp. 877–894.

Grizzle, J. (1971), "Multivariate Logit Analysis," *Biometrics*, 27, pp. 1057–1062.

Grizzle, J., C. Starmer, and G. Koch (1969), "Analysis of Categorial Data by Linear Models," *Biometrics*, 25, pp. 498–504.

Gumbel, D. J. (1961), "Bivariate Logistic Distributions," *Journal of the American Statistical Association*, 56, pp. 335–349.

Gupta, S., A. Qureishi, and B. Shah (1967), "Best Linear Unbiased Estimators of the Parameters of the Logistic Distribution Using Order Statistics, *Technometrics*, 9, pp. 43–56.

Gurland, J., I. Lee, and P. Dahm (1960), "Polychotomous Quantal Response in Biological Assay," *Biometrics*, 16, pp. 382–398.

Haberman, Shelby (1974), *The Analysis of Frequency Data*, University of Chicago Press.

Haldane, J. (1955), "The Estimation and Significance of the Logarithm of a Ratio of Frequencies," *Annals of Human Genetics*, 20, pp. 309–311.

Hall, R. (1973), "On the Statistical Theory of Unobserved Components," Department of Economics, Massachusetts Institute of Technology, unpublished.

Hall, R. (1974), "Annotated Bibliography on Inhomogeneous Markoff Models and Applications to Social Mobility," for NSF-NBER Conference on Decision Rules and Uncertainty, March 22–23, 1974, University of California, Berkeley, unpublished.

Harter, J. and A. Moore (1967), "Maximum Likelihood Estimation, from Censored Samples, of the Parameters of a Logistic Distribution," *Journal of the American Statistical Association*, 62, 675–683.

Hausman, J. and D. A. Wise (1976), "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences, unpublished.

Hausman, J. (1975), "A Generalization of Probit Analysis to Polytomous Choice," Department of Economics, Massachusetts Institute of Technology, unpublished.

Heath, D. C. and Sudderth, W. D. (1972), "On a Theorem of de Finetti, Oddsmaking, and Game Theory, *Annals of Mathematical Statistics*, 43, pp. 2072–2078.

Heckman, J. and R. Willis (1974), "Estimation of a Stochastic Model of Reproduction: An Econometric Approach," in N. Terleckyj, ed., *Household Production and Consumption*, National Bureau of Economic Research, New York.

Heckman, J. (1975), "Simultaneous Equations Models with Both Continuous and Discrete Endogeneous Variables with and without Structural Shift in the Equations," University of Chicago, unpublished.

Hodges, J. L., Jr. (1950), "Discriminatory Analysis, I. Survey of Discriminatory Analysis," USAF School of Aviation Medicine Report No. 1, Randolf Field, Texas.

Hulett, J. R. (1973), "On the Use of Regression Analysis with a Qualitative Dependent Variable," *Public Finance Quarterly*, 1, 339–344.

Hymans, S. H. and D. L. Rubinfeld (1973), "Specification and Estimation of a Linear Scoring Model," University of Michigan, unpublished.

Jacoby, S. L., J. S. Kowalik, and J. T. Pizzo (1972), *Iterative Methods of Nonlinear Optimization Problems*, Prentice-Hall, Englewood Cliffs, NJ.

Johansson, Methods K. (1973), "A Generalized Logistic Function with an Application to the Effects of Advertising," *Journal of the American Statistical Association*, 68, 824–827.

Johnson, N. L. (1949), "Systems of Frequency Curves Generated by Methods of Translation," *Biometrics*, Vol. 5, pp. 149–176.

Johnson, N. L. and S. Kotz (1970), *Continuous Univariate Distributions*, Vol. 1 and 2, Houghton Mifflin, Boston.

Johnson, T. (1972), "Qualitative and Limited Dependent Variables in Economic Relationships," *Econometrica*, 40, 455–462.

Jorgenson, D. W. (1961), "Multiple Regression Analysis of a Poisson Process," *Journal of the American Statistical Association*, 56, 235–245.

Juster, F. T. (1966), "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design." *Journal of the American Statistical Association*, 61, 658–696.

Karlin, S. and W. Studden (1966), *Tchebycheff Systems: With Applications in Analysis and Statistics*, Wiley, New York.

Kastenbaum, M. A. (1969), "A Review of Contingency Tables," in R. C. Bose, *et al.*, eds., *Essays in Probability and Statistics*, University of North Carolina Press, Chapel Hill.

Kohn, M., C. Manski, and D. Mundel (1976), "An Empirical Investigation of Factors which Influence College Going Behavior," *Annals of Economic and Social Measurement*, this issue.

Koppelman, F. (1975), "Travel Prediction with Disaggregate Choice Models," Ph.D. dissertation, Massachusetts Institute of Technology, Department of Civil Engineering, unpublished.

Koppelman, F. (1976), "Methodology for Analysis of Errors in Prediction with Disaggregate Choice Models," paper presented at the 55th annual meeting of the Transportation Research Board, Washington, DC, January 1976.

Koppelman, F. (1976), "Rules for Aggregate Travel Prediction with Disaggregate Travel Choice Models," paper presented at the 55th annual meeting of the Transportation Research Washington, DC, January 1976.

Korbel, J. (1966), "Labor Force Entry and Attachment of Young People," *Journal of the American Statistical Association*, Vol. 61, pp. 117–127.

Krantz, D. H. (1974), "Measurement Theory and Qualitative Laws in Psychophysics," in D. H. Krantz, *et al.*, eds., *Measurement, Psychophysics, and Neutral Information Processing*, Vol. II, Freman, San Francisco.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971), *Foundations of Measurement*, Vol. 1, *Additive and Polynomial Representations*, Academic Press, New York.

Krzanowski, W. J. (1975), "Discrimination and Classification Using Both Binary and Continuous Variables," *Journal of the American Statistical Association*, 70, 782–790.

Lachenbruck, P. A. (1975), *Discriminant Analysis*, Hafner, New York.

Lachenbruck, P. A., C. Sneeringer, and L. T. Revo (1973), "Robustness of the Linear and Quadratic Discriminant Function to Certain Types of Non-normality," *Communications in Statistics*, 1, 39–56.

Ladd, B. (1966), "Linear Probability Functions and Discriminant Functions," *Econometrica*, Vol. 34, pp. 873–885.

Lave, C. (1970), "The Demand for Urban Mass Transit," *Review of Economics and Statistics*, Vol. 52, pp. 320–323.

Lee, T. (1963), "Demand for Housing: A Cross-Section Analysis," *Review of Economics and Statistics*, Vol. 45, pp. 190–196.

Lerman, S. R. (1975), "A Disaggregate Behavioral Model of Urban Mobility Decisions," Ph.D. dissertation, Massachusetts Institute of Technology, Department of Civil Engineering, unpublished.

Lerman, S. R. and M. Ben-Akiva (1975), "A Disaggregate Behavioral Model of Automobile Ownership," paper prepared for presentation at the 54th annual meeting of the transportation Board, Washington, DC, January 1975. Forthcoming in a Transport Research Board *Record.*

Lerman, S. R. (1976), "Location, Housing, Auto Ownership, and Mode to Work: A Joint Choice Model," paper presented at the 55th annual meeting of the Transportation Research Board, Washington, DC, January 1976.

Lewis, B. N. (1962), "On the Analysis of Interaction in Multidimensional Contingency Tables," *Journal of the Royal Statistical Society*, A125, pp. 88–117.

Light, R. J. and B. H. Margolin (1971), "An Analysis of Variance for Categorical Data," *Journal of the American Statistical Association*, 66, 534–544.

Little, R. E. (1968), "A Note on Estimation for Quantal Response Data," *Biometrika*, 55, 578–579.

Lisco, T. (1967), "The Value of Commuters' Travel Time: A Study in Urban Transportation," Ph.D. dissertation, Department of Economics, University of Chicago, unpublished.

Luce, R. D. and P. Suppes (1965), "Preference, Utility, and Subjective Probability," in R. Luce, R. Bush, and E. Galanter, eds., *Handbook of Mathematical Psychology*, III, Wiley, New York.

Luce, R. D. (1959), *Individual Choice Behavior*, Wiley, New York.

Maddala, G. S. and F. D. Nelson (1974), "Analysis of Qualitative Variables," National Bureau of Economic Research Working Paper No. 70, Cambridge, MA.

Magidson, J. (1975), "Qualitative Variables and Simultaneous Equation Econometric Models," Center for Mathematical Studies in Economics and Management Science, Discussion Paper No. 142, Northwestern University.

Manski, C. F. (1975), "Maximum Score Estimation of the Stochastic Utility Model of Choice," *Journal of Econometrics*, 3, 205–228.

Manski, C. F. (1976), "The Structure of Random Utility Models," forthcoming in *Theory and Decision.*

Manski, C. F. and S. R. Lerman (1976), "The Estimation of Choice Probabilities from Choice Based Samples," forthcoming, *Econometrica.*

Mantel, N. (1966), "Models for Complex Contingency Tables and Polychotomous Response Curves," *Biometrics*, 22, 83–110.

Marley, A. (1968), "Some Probabilistic Models of Simple Choice and Ranking," *Journal of Mathematical Psychology*, Vol. 5, pp. 311–332.

Marschak, J. (1960), "Binary Choice Constraints on Random Utility Indicators," in K. Arrow, ed., *Stanford Symposium on Mathematical Methods in the Social Sciences*, Stanford University Press.

Maxwell, A. E. (1961), *Analyzing Qualitative Data*, Metheun, London.

McFadden, D. (1975), "The Revealed Preferences of a Government Bureaucracy: Theory," *Bell Journal of Economics*, Vol. 6, pp. 401–416.

McFadden, D. (1976), "The Revealed Preferences of a Government Bureaucracy: Evidence," *Bell Journal of Economics*, Vol. 7.

McFadden, D. and C. Manski (1976), "Alternative Estimators and Sample Designs for Discrete Choice Analysis," Department of Economics, Universtiy of California, Berkeley, unpublished.

McFadden, D. (1975), "Tchebycheff Bounds for the Space of Agent Characteristics," *Journal of Mathematical Economics*, Vol. 2, pp. 225–242.

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

McFadden, D. (1974), "The Measurement of Urban Travel Demand," *Journal of Public Economics*, Vol. 3, pp. 303–328.

McFadden, D. (1975), "Economic Applications of Psychological Choice Models," Working Paper No. 7519, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley, unpublished.

McFadden, D. and F. Reid (1974), "Aggregate Travel Demand Forecasting from Disaggregated Behavioral Models," Transportation Research Board *Record*, 534, pp. 24–37.

McFadden, D. and M. K. Richter (1970), "Revealed Stochastic Preference," University of California, Department of Economics, unpublished.

McFadden, D. and M. K. Richter (1971), "On the Extension of a Set Function to a Probability on the Boolean Algebra Generated by a Family of Events, with Applications," Working Paper No. 14, Mathematical Social Science Board Workshop on the Theory of Markets under Uncertainty, Department of Economics, University of California, Berkeley, unpublished.

McGillivray, R. (1970), "Demand and Choice Models of Modal Split," *Journal of Transport Economics and Policy*, 4, pp. 192–207.

McGillivray, R. (1970), "Estimating the Linear Probability Function," *Econometrica*, 38, 775–776.

McLynn, J. (1971), "Some Considerations of Modal Split Models for the 1972 Transportation Needs Study," National Bureau of Standards.

McLynn, J. (1973), "A Technical Note on a Class of Fully Competitive Modal Choice Models," DTM Corp., Bethesda, MD, unpublished.

McLynn, J. and K. Goodman (1973), "Mode Choice and the Shirley Highway Experiment," Report No. UMTA-IT-0024-73-1, U.S. Department of Transportation.

Miller, L. and R. Radner (1970), "Demand and Supply in U.S. Higher Education," *American Economic Review*, 60, pp. 326–334.

Miller, L. and R. Radner (1975), *Demand and Supply in U.S. Higher Education*, Carnegie Commission, McGraw-Hill.

Mohan, D. F. (1975), "Alternative Specifications of Qualitative Dependent Variable Models: A Monte Carlo Approach," unpublished Ph.D. dissertation, Rutgers University.

Moore, D. H. (1973), "Evaluation of Five Discrimination Procedures for Binary Variables," *Journal of the American Statistical Association*, 68, 399–404.

Morrison, D. (1972), "Upper Bounds for Correlation between Binary Outcomes and Probabilistic Predictions," *Journal of the American Statistical Association*, Vol. 67, pp. 68–70.

Moses, L., R. Beals, and M. Levy (1967), "Rationality and Migration in Ghana," *Review of Economics and Statistics*, Vol. 49, pp. 480–486.

Mosteller, F. (1968), "Association and Estimation in Contingency Tables," *Journal of the American Statistical Association*, 63, pp. 1–28.

Nerlove, M. and S. J. Press (1973), "Univariate and Multivariate Log-Linear and Logistic Models," RAND Report No. R-1306-EDA/NIH.

Neter, J. and E. S. Maynes (1970), "On the Appropriateness of the Correlation Coefficient with a 0, 1 Dependent Variable," *Journal of the American Statistical Association*, 65, 501–509.

Orcutt, G. H., M. Greenberger, J. Korbel, and A. M. Rivlin (1961), *Microanalysis of Socioeconomic Systems: A Stimulation Study*, Harper and Row, New York.

Planckett, R. L. (1974), *The Analysis of Categorical Data*, Charles Griffin and Co., London.

Poirier, D. J. (1975), "Aggregate Prediction and Goodness-of-Fit in Models with Qualitative Dependent Variables," Institute for Policy Analysis, Working Paper No. 7510, University of Torotonto.

Poirier, D. J. (1975), "A Logistic Interpretation to a Probit Model with a Random Scale Factor," unpublished manuscript, University of Toronto.

Poirier, D. J. (1976), "The Determinants of Home Buying in the New Jersey Graduated Work Incentive Experiment," in H. W. Watts and A. Rees, eds., *Impact of Experimental Payments on Expenditure, Health, and Social Behavior, and Studies on the Quality of the Evidence*, New York, Academic Press.

Pollakowski, H. (1975), "A Conditional Logit Model of Residential Choice," University of Washington, unpublished.

Quandt, R. (1956), "Probabilistic Theory of Consumer Behavior," *Quarterly Journal of Economics*, Vol. 70, pp. 507–536.

Quandt, R. (1968), "Estimation of Modal Splits," *Transportation Research*, 2, pp. 41–50.

Quandt, R. (1970), *The Demand for Travel*, D. C. Heath, London.

Quandt, R. and W. Baumol (1966), "The Demand for Abstract Transport Modes: Theory and Measurement," *Journal of Regional Science*, 6, pp. 13–26.

Quandt, R. and K. Young (1969), "Cross Sectional Travel Demand Models: Estimation and Tests," *Journal of Regional Science*, 9, pp. 201–214.

Quandt, R. (1972), "New Methods for Estimating Switching Regressions," *Journal of the American Statistical Association*, Vol. 67, pp. 306–310.

Quarmby, D. (1967), "Choice of Travel Mode for the Journey to Work: Some Findings," *Journal of Transport Economics and Policy*, Vol. 1, pp. 273–314.

Rao, C. R. (1957), "Theory of the Method of Estimation by Minimum Chi-Square," *Bulletin of the Internation Statistical Institute*, 35, 25–32.

Rassam, P., R. Ellis, and J. Bennett (1974), "The n-Dimensional Logit Model: Development and Application," Highway Research Board *Record*, 369, pp. 135–147.

Reichman, S. and P. Stopher (1971), "Disaggregate Stochastic Models of Travel-Mode Choice," Highway Research Board *Record*, 369, pp. 91–103.

Richards, M. G. and M. Ben-Akiva (1974), "A Simultaneous Destination and Mode Choice Model for Shopping Trips," *Transportation*, 3, pp. 343–356.

Richards, M. G. and M. Ben-Akiva (1975), *A Disaggregate Travel Demand Model*, Saxon House, D.C. Heath, London.

Schultz, T. P. (1075), "The Determinants of Internal Migration in Venezuela: An Application of the Polytomous Logistic Model," paper presented at the third World Congress of the Econometric Society, Toronto, Canada.

Scott, D. (1964), "Measurement Models and Linear Inequalities," *Journal of Mathematical Psychology*, Vol. 1, pp. 233–248.
Smith, V. K. and C. J. Cicchetti (1972), "Regression Analysis with Dichotomous Dependent Variables," paper presented at the Third World Congress of the Econometric Society, Toronto, Canada.
Spilerman, S. (1972), "Extensions of the Mover-Stayer Model," *American Journal of Sociology*, Vol. 78, pp. 599–626.
Stopher, P. (1969), "A Multinomial Extension of the Binary Logit Model for Choice of Mode of Travel," Northwestern University, unpublished.
Stopher, P. (1969), "A Probability Model of Travel Mode Choice for the Work Journey," Highway Research Board *Record*, 283, pp. 57–65.
Stopher, P. and J. Lavender (1972), "Disaggregate Behavioral Travel Demand Models: Empirical Tests of Three Hypotheses," Transport Research Forum *Proceedings*, 13th annual meeting, Denver.
Stopher, P. and T. Lisco (1970), "Modelling Travel Demand: A Disaggregate Behavioral Approach—Issues and Applications," Transportation Research Forum *Proceedings*, pp. 195–214.
Swan, J. H. (1975), "Estimation Procedure for Dichotomous Variables," Department of Sociology, Northwestern University, unpublished.
Talvitie, A. (1972), "Comparison of Probabilistic Modal-Choice Models: Estimation Methods and System Imputs," Highway Research Board *Record*, 392, pp. 111–120.
Talvitie, A. (1973), "Aggregate Travel Demand Analysis with Disaggregate or Aggregate Travel Demand Models," Transportation Research Forum *Proceedings*.
Theil, H. (1971), *Principles of Econometrics*, Wiley, New York.
Theil, H. (1967), *Economics and Information Theory*, North-Holland, Amsterdam.
Theil, H. (1969), "A Multinomial Extension of the Linear Logit Model," *International Economic Review*, 10, 251–259.
Theil, H. (1970), "On the Estimation of Relationships Involving Qualitative Variables," *American Journal of Sociology*, 76, 103–154.
Theil, H. (1972), "Decompositions and Their Determining Factors," in *Statistical Decomposition Analysis*, North-Holland, Amsterdam, Ch. 4.
Theil, H. and R. F. Kosobud (1968), "How Informative are Consumer Buying Intentions Surveys," *Review of Economics and Statistics*, 50, 50–59.
Thomas, T. and G. Thompson (1971), "Value of Time Saved by Trip Purpose," Highway Research Board *Record*, 369, pp. 104–113.
Thurstone, L. (1927), "Psychological Analysis," *American Journal of Psychology*, 38, pp. 368–389.
Thurstone, L. (1927), "A Law of Comparative Judgement," *Psychological Review*, 34, pp. 273–286.
Tinter, G. (1952), *Econometrics*, Wiley, New York.
Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 26, pp. 24–36.
Tocher, K. D. (1949), "A Note on the Analysis of Grouped Probit Data," *Biometrika*, 39, pp. 9–17.
Tversky, A. (1969), "Intransitivity of Preferences," *Psychological Review*, Vol. 76, pp. 31–48.
Tversky, A. and D. H. Krantz (1970), "The Dimensional Representation and the Metric Structure of Similarity Data," *Journal of Mathematical Psychology*, 7, 341–367.
Tversky, A. (1972), "Elimination by Aspects: A Theory of Choice," *Psychological Review*, Vol. 79, pp. 281–299.
Tversky, A. (1972), "Choice by Elimination," *Journal of Mathematical Psychology*, Vol. 9, pp. 341–367.
Tversky, A. and J. Russo (1969), "Substitutability and Similarity in Binary Choices," *Journal of Mathematical Psychology*, Vol. 6, pp. 1–12.
Uhler, R. (1968), "The Demand for Housing: An Inverse Probability Approach," *Review of Economics and Statistics*, Vol. 50, pp. 129–134.
Urban, F. M. (1910), "Die Psychophysischen Massmethoden als Gundlagen Empirischer Messungen (continued)," *Arc. Ges. Psychol.*, 16, pp. 168–227.
Walker, F. (1968), "Determinants of Auto Scrappage," *Review of Economics and Statistics*, Vol. 50, pp. 503–506.
Walker, S. and D. Duncan (1967), "Estimation of the Probability of an Event as a Function of Several Independent Variables," *Biometrika*, Vol. 54, pp. 167–179.
Warner, S. (1962), *Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice*, Northwestern University Press.
Warner, S. L. (1963), "Multivariate Regression of Dummy Variates under Normality Assumptions," *Journal of the American Statistical Association*, 58, 1054–1063.

Watson, P. L. (1974), *The Value of Time: Behavioral Models of Modal Choice*, D. C. Heath, Lexington, MA.

Watson, P. L. and R. B. Westin (1975), "Transferability of Disaggregate Mode Choice Models," *Regional Science and Urban Economics*, 5, 227–249.

Watson, P. and R. Westin (1975), "Reported and Revealed Preferences as Determinants of Mode Choice Behavior," *Journal of Marketing Research*, Vol. 12, pp. 282–289.

Westin, R. (1975), "Statistical Models for Interrelated Discrete and Continuous Choice," University of Toronto, unpublished.

Westin, R. (1974), "Predictions from Binary Choice Models," *Journal of Econometrics*, 2, 1–16.

Westin, R. (1975), "Comment on 'A Modified Logit Model'," University of Toronto, unpublished.

Westin, R. and P. Watson (1975), "Transferability of Disaggregate Mode Choice Models," *Regional Science and Urban Economics*, Vol. 5, pp. 227–250.

Wigner, M. (1973), "Disaggregated Mode Choice Models of Downtown Trips in the Chicago Region," Chicago Area Transportation Study, CATS 341-05.

Zavonia, V. and R. McKelvey (1969), "A Stochastic Model for the Analysis of Legislative Voting Behavior," presented at the annual meeting of the American Political Science Association, New York.

Zellner, A. and T. Lee (1965), "Joint Estimation of Relationships Involving Discrete Random Variables," *Econometrics*, 33, 382–394.

Zinnes, J. (1969), "Scaling," *Annual Review of Psychology*, 20, 447–478.