Advances in Random Utility Models

Horowitz, Joel and Keane, Michael and Bolduc, Denis and Divakar, Suresh and Geweke, John and Gonul, Fusun and Hajivassiliou, Vassilis and Koppelman, Frank and Matzkin, Rosa and Rossi, Peter and Ruud, Paul

1994

Advances in Random Utility Models
Report of the Workshop on Advances in Random Utility Models
Duke Invitational Symposium on Choice Modeling Behavior

JOEL L. HOROWITZ
Workshop Co-Chairman, University of Iowa

in collaboration with workshop participants

DENIS BOLDUC
Université Laval

SURESH DIVAKAR
State University of New York at Buffalo

JOHN GEWEKE
University of Minnesota

FÜSUN GÖNÜL
Carnegie Mellon University

VASSILIS HAJIVASSILIOU
Yale University

FRANK S. KOPPELMAN
Northwestern University

MICHAEL KEANE
Workshop Co-Chairman, University of Minnesota

ROSA MATZKIN
Northwestern University

PETER ROSSI
University of Chicago

PAUL RUUD
University of California, Berkeley

Abstract

In recent years, major advances have taken place in three areas of random utility modeling: (1) semiparametric estimation, (2) computational methods for multinomial probit models, and (3) computational methods for Bayesian estimation. This paper summarizes these developments and discusses their implications for practice.

1. Introduction

Random utility models aim at modeling the choices of individuals among discrete sets of alternatives. It is assumed in these models that the preferences of an individual among the available alternatives can be described by a utility function. The individual chooses the alternative with the highest utility. The utility of an alternative depends on attributes of the alternative and individual that the analyst observes (e.g., price of a product if utility is indirect, age of an individual) and attributes that the analyst does not observe (e.g., whether the individual is buying for a special occasion). Observed attributes are represented in the utility function by explanatory variables. Unobserved ones are represented as random variables, thereby giving rise to the term "random utility model." See Ben-Akiva and Lerman (1985), Domencich and McFadden (1975), Hensher and Johnson (1981) and McFadden (1974, 1981) for discussions of the theory of random utility modeling. Heckman (1981a, 1981b) discusses the use of random utility models for the analysis of panel data.

Utility is a random variable in a random utility model, so one cannot use such a model to predict an individual's choice with certainty. Instead, a random utility model gives the probability with which each alternative is chosen. Consider an individual who chooses among J alternatives. Let the utility of alternative j be

$$U_j = \beta'X_j + \varepsilon_j,$$

where X_j is a column vector of observed attributes of alternative j and the individual, β is a conformable vector of constant parameters, and ε_j is a random variable that accounts for the effects on preferences of unobserved attributes of the alternative and individual. Then the probability that the individual chooses alternative i is

$$P(i|X) = P(\beta'X_i + \varepsilon_i > \beta'X_j + \varepsilon_j \text{ for all } j = 1, \ldots, J; j \neq i), \qquad (1.1)$$

where $X = (X_1', \ldots, X_J')'$.

Except in special cases, $P(i|X)$ is a complicated function of X that has no closed analytic form and may be very difficult to compute. An important special case in which $P(i|X)$ has a simple analytic form is when the components of $\varepsilon \equiv$

$(\varepsilon_1, \ldots, \varepsilon_j)'$ are independently and identically distributed (*iid*) across alternatives with the Type 1 extreme value distribution. $P(i|X)$ is then given by the multinomial logit (MNL) model:

$$P(i|X) = [\exp(\beta'X_i)]/\sum_{j=1}^{J} [\exp(\beta'X_j)].$$

The MNL model is widely used in applications because of its analytic and computational tractability. However, the assumption that the components of ε are *iid* is highly restrictive. Among other things, it excludes the possibility that tastes vary among individuals randomly (i.e., for reasons that are not observed by the analyst) or that different alternatives have similar unobserved attributes. It also precludes modeling repeated choices by the same individuals over time (panel data) unless one makes the restrictive and frequently incredible assumption that choices by the same individual on different occasions are independent conditional on the model's explanatory variables.

These problems can be overcome by assuming that ε is multivariate normal with a covariance matrix that may depend on X (as happens if there is random taste variation) or have nonzero off-diagonal elements (as happens if different alternatives have similar unobserved attributes). When ε is multivariate normal, $P(i|X)$ is called a multinomial probit (MNP) model. MNP is a highly flexible model whose assumptions are much less restrictive than those of MNL. However, there is no closed analytic form for $P(i|X)$ in a MNP model; $P(i|X)$ can be expressed only as an integral of the multivariate normal density. If the number of alternatives is large, $P(i|X)$ is a high-dimensional multivariate integral that cannot be evaluated accurately using simple analytic approximations such as the Clark approximation or standard numerical methods such as Gaussian quadrature. Similar computational problems occur if the data consist of a multiwave panel in which there is unobserved heterogeneity or other forms of serial dependence in the random component of utility. As a result, there have been few applications of MNP models. As is discussed below, recent research has largely overcome the difficulty of evaluating MNP choice probabilities.

Even more flexibility in modeling can be achieved by completely avoiding any parametric assumptions about the distribution of ε. Instead, one assumes only that this distribution satisfies certain regularity conditions. Models of this type are called "semiparametric" to reflect the fact that they involve an unknown finite-dimensional parameter (the coefficients of the explanatory variables X) and an unknown function (the distribution of ε). Semiparametric models are discussed in more detail in Section 2.1 of this paper.

In applications, the values of the utility-function parameters β and any unknown parameters of the distribution of ε (e.g., the covariance matrix in a MNP model) must be estimated from data consisting of observations of the choices made by

individuals and the corresponding values of the explanatory variables X. When the distribution of ε is specified parametrically, estimation is usually carried out by the method of maximum likelihood. Statistical inference is then based on the asymptotic theory of maximum likelihood estimators and test statistics. See, for example, Amemiya (1985) or Ben-Akiva and Lerman (1985) for descriptions of this theory. An alternative approach is Bayesian estimation and inference. Compared to classical (e.g., maximum likelihood) estimation, Bayesian estimation has the advantages of avoiding reliance on asymptotic approximations and providing a natural way of incorporating prior information or beliefs about parameter values into the estimation process. In addition, the posterior distribution of the taste coefficients (nonintercept components of β) in random coefficients models can be obtained much more easily than in classical models. A disadvantage of the Bayesian approach is that its results may be quite sensitive to the choice of a prior distribution of the parameters, and there is often no compelling reason for preferring one prior to others. Bayesian estimation can present severe computational difficulties when conventional numerical techniques are used, owing to the high-dimensional multivariate integrals that usually must be evaluated to obtain posterior moments. These difficulties have been largely overcome in recent research, as will be discussed below.

2. Summary of recent advances

In recent years, major advances have taken place in three areas of random utility modeling:

1. Semiparametric estimation,
2. Computational methods for multinomial probit models, and
3. Computational methods for Bayesian estimation.

These developments are summarized in the remainder of this section.

2.1. Semiparametric estimation

Methods for semiparametric estimation of binary response models are now ready for use in applications. There is software available for implementing some methods. Others do not have readily available software but are quite easy to program and implement. There have been applications indicating that semiparametric methods can yield results that are different in important ways from the results obtained using conventional parametric models (Bult, 1993; Das, 1991; Gabler, Laisney, and Lechner, 1993; Goodwin, 1993; Horowitz, 1993a). However, because semiparametric methods for binary response models are relatively new, experience in using them in applications is limited. It would be very useful for other

researchers to undertake applications of these methods and to compare the results thus obtained with those obtained with standard parametric models.

Horowitz (1993b) reviews semiparametric estimators for binary response models. An outline of the most important estimators will now be given. A binary response model can be written in the form

$$Y = \begin{cases} 1 & \text{if } \beta'X + v > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where X is a vector of explanatory variables, β is a vector of parameters whose values are unknown and must be estimated from data, and v is an unobserved random variable. In an important special class of binary response models, called single-index models, the probability distribution of v is either independent of X or depends on X only through the index $\beta'X$. When this assumption is satisfied, the probability that $Y = 1$ is

$$P(Y = 1|X) = F(\beta'X), \tag{2.1}$$

where F is a function that is unknown in semiparametric models. Semiparametric estimators of β in (2.1) when X has both discrete and continuous components have been described by Han (1987), Ichimura (1993), Klein and Spady (1993), and Sherman (1993). Some of these estimators are conceptually similar to the parametric estimator that would be obtained if F were known. For example, if F were known, β could be estimated by maximizing the log-likelihood function

$$\log L(b) = \sum_{i=1}^{n} \{Y_i \log[F(b'X_i] + (1 - Y_i)\log[1 - F(b'X_i)]\},$$

where $\{Y_i, X_i: i = 1, \ldots, n\}$ is a random sample of observations of (Y, X). Apart from technical details that are unimportant in applications, the estimator of Klein and Spady (1993) is obtained by replacing the unknown function F with a nonparametric estimate. The estimator of Ichimura amounts to nonlinear least-squares estimation with the unknown F replaced with a nonparametric estimate.

If all of the components of X are continuous, β can be estimated by the density-weighted average derivative method of Powell, Stock, and Stoker (1989). This method has the considerable advantage of not requiring nonlinear optimization. This makes it easy to program and fast to compute.

Given an estimate b of β, $P(Y = 1|X)$ in a single-index model can be estimated by carrying out the nonparametric mean regression of Y on $b'X$. See Härdle (1990) for a discussion of nonparametric regression techniques.

In models with random taste variation, $P(Y = 1|X)$ does not have the single-index form (2.1), so the estimation methods just described cannot be used. The parameter β in semiparametric models with random taste variation can be esti-

mated using either the maximum score method (Manski, 1975, 1985) or the smoothed maximum score method (Horowitz, 1992). Maximum score estimation consists, essentially, of choosing the estimator b of β to maximize the number of correct predictions of Y when one predicts that $Y = 1$ if $b'X \geq 0$ and $Y = 0$ otherwise. In smoothed maximum score estimation, the predictor of Y is a continuous function of $b'X$, so there can be predicted values between 0 and 1. The estimator b is chosen to optimize a measure of the distance between the predicted and observed values of Y. Smoothing speeds the rate of convergence of the estimator and greatly simplifies its asymptotic distribution.

The maximum score and smoothed maximum score estimators do not require the analyst to know the form of any random taste variation. Software for implementing the maximum score estimator is available in the LIMDEP package. A GAUSS program for implementing the smoothed maximum score estimator is available by e-mail at the Internet address JHorowit@scout-po.biz.uiowa.edu.

Semiparametric methods for multinomial response models are not as well developed as methods for binary response. The maximum score method can be extended to certain types of multinomial response models (Manski, 1975), but there have been no applications of this method. Progress in semiparametric estimation is being made very rapidly, however, and semiparametric methods for multinomial response models are likely to be available for applied use soon.

There has also been research on completely nonparametric estimation of binary response models. In nonparametric estimation, the term $\beta'X$ on the right-hand side of (2.1) is replaced by an unknown function of X. Thus, both this function and the distribution of ε are unknown in nonparametric estimation. See Matzkin (1992) for details.

2.2. Computational methods for classical estimation of MNP models

There has been much recent progress in the development of simulation methods for estimating the parameters and computing the choice probabilities of MNP models. Using the best simulation methods, the time required to compute choice probabilities and other quantities required for estimation is roughly linear in the number of alternatives in the choice set. The required computations are very fast on a modern microcomputer, so the advent of efficient simulation methods has essentially removed the number of alternatives as a factor in computing MNP choice probabilities.

Roughly speaking, a simulator evaluates choice probabilities by repeatedly sampling the probability distribution of ε, determining the utility-maximizing choice corresponding to each sample, and computing the proportion of draws on which each alternative is chosen. Although this idea is simple, several difficult problems must be solved to construct a practical, accurate simulator. For example, to use a simulator in connection with standard numerical techniques for maximizing a log-likelihood function, the simulated choice probabilities must be dif-

ferentiable functions of the utility-function parameters and bounded away from 0 and 1. In addition, to achieve computational efficiency, the simulator must produce accurate results with a small number of draws from the distribution of ε.

Hajivassiliou (1993a) reviews all known simulation methods for classical estimation of MNP models. Here we outline one easy-to-use method, the smoothed simulated maximum likelihood (SSML) method based on the GHK simulator. "GHK" stands for John Geweke, Vassislis Hajivassiliou, and Michael Keane, who independently developed the GHK simulator. A detailed description of the SSML/GHK method is given by Börsch-Supan and Hajivassiliou (1993). Software for using this method to estimate MNP models with cross-sectional or panel data is available by anonymous FTP at the Internet site "econ.yale.edu" in directory "pub/vassilis/simulation." Comparisons of the performances of the SSML/GHK and other simulation estimators are given by Geweke, Keane, and Runkle (1992), Börsch-Supan and Hajivassiliou (1993), Hajivassiliou (1993b), Hajivassiliou, McFadden, and Ruud (1992), and Keane (1994). These comparisons have found that the SSML/GHK method is very accurate except when the random components of utility include autoregressive terms with first-order correlation coefficients that are near one. Keane (1994) describes a different simulation method that is more accurate than SSML/GHK in the presence of highly correlated autoregressive terms. Other simulation approaches to MNP estimation are described by Lerman and Manski (1981), McFadden (1989), Pakes and Pollard (1989), Hajivassiliou and McFadden (1992), and Stern (1992).

To outline the SSML/GHK method, consider a sample consisting of observations on a cross-section of n individuals. Let $P_i(\theta)$ denote the MNP probability of the alternative that individual i is observed to choose, where θ denotes the parameters of the model. The log-likelihood function is

$$\log L(\theta) = \sum_{i=1}^{n} \log P_i(\theta). \tag{2.2}$$

In general, $P_i(\theta)$ depends on θ through a multivariate integral that is difficult or impossible to evaluate using conventional numerical methods such as quadrature. The idea of the SSML/GHK method (as well as several other simulation estimators) is to estimate the value of $P_i(\theta)$ by Monte Carlo simulation and replace $P_i(\theta)$ in (2.2) with the estimate. The GHK simulator is obtained by writing $P_i(\theta)$ as the mean value of a certain function of independent random variables that have truncated standard normal distributions. The function is a product of standard normal probabilities, and its value is easy and fast to compute. A simulation estimate of its mean can be obtained by sampling variables independently from the appropriate univariate truncated normal distributions, evaluating the function at the sampled values of the variables, and averaging the results obtained from several samples. Draws from the univariate truncated normal distribution can be generated easily by sampling from the uniform distribution and applying the inverse truncated normal distribution function to the result.

This simulation method produces an estimate of $P_i(\theta)$ that is a continuous, differentiable function of θ. Moreover, accurate results can be obtained using a relatively small number of repetitions of the sampling process. Börsch-Supan and Hajivassiliou (1993) and Geweke, Keane, and Runkle (1992) obtain very accurate results with only twenty to thirty repetitions in models with three to seven alternatives.

2.3. Computational methods for Bayesian estimation

Bayesian estimation of MNP models also presents computational difficulties that are intractable with conventional numerical techniques. This is because evaluating Bayesian posterior moments requires computing complicated multivariate integrals whose integrands depend on the MNP choice probabilities. Recently, however, McCulloch and Rossi (1992) (hereinafter referred to as MR) and Albert and Chib (1993) have developed computationally efficient simulation methods for evaluating Bayesian posterior moments. The MR approach is outlined below. Software for implementing this approach is available via anonymous FTP at the Internet site gsbper.uchicago.edu. The program is in pub/rossi/mnp, and the McCulloch and Rossi (1992) paper is in pub/rossi/papers/mnp.

The MR approach consists of using the Gibbs sampler to generate draws from the posterior distribution of the MNP model, thereby avoiding the need to compute MNP choice probabilities. Posterior moments are estimated by the corresponding moments of the resulting sample. The Gibbs sampler is a method for generating draws from a multivariate distribution by successively drawing from suitable conditional distributions. For example, suppose we want to generate draws from the bivariate density $f(x_1, x_2)$ but are able to sample only the conditional densities $f_1(x_1|x_2)$ and $f_2(x_2|x_1)$. The Gibbs sampler in this case consists of first drawing an X_1 from f_1 conditional on an initial value of X_2, then drawing an X_2 from f_2 conditional on the sampled X_1, then sampling a new X_1 from f_1 conditional on the sampled X_2, etc. It can be shown that the resulting sequence of (X_1, X_2) pairs converges in distribution to the desired bivariate distribution, $f(x_1, x_2)$.

The MR approach specifies the prior distribution of the MNP utility-function parameters β to be multivariate normal and the prior distribution of the inverse of the covariance matrix of $\varepsilon_j - \varepsilon_i$ to be Wishart independent of β. The Gibbs-sampler simulation procedure consists of repeatedly sampling:

1. Each utility difference $U_j - U_i$ conditional on the remaining differences, β, the covariance matrix of $\varepsilon_j - \varepsilon_i$, and the observed choice.
2. β conditional on the utility differences and the covariance matrix.
3. The covariance matrix conditional on the utility differences and β.

The resulting draws converge in distribution to draws from the Bayesian posterior distribution of the MNP parameters. This approach can be extended to handle

models with random coefficients and models for panel data. See MR for details. The computational requirements for implementing the MR procedure are similar to those of the SSML/GHK method for classical maximum likelihood estimation.

3. Other issues in multinomial probit modeling

The development of efficient simulation procedures has eliminated the computation of MNP choice probabilities as a barrier to the use of MNP models. Nonetheless, MNP modeling remains a delicate matter that requires careful judgment by the modeler. Some of the reasons for this will now be outlined.

a. Because MNP choice probabilities depend only on the signs of utility differences, the parameters of a MNP model are identified only up to certain location and scale normalizations. In particular, the covariance parameters of a fully parameterized MNP model are not all identified. See Bunch (1991) and Dansie (1985) for discussions of identification in MNP models.

b. The use of MNP models can lead to severe proliferation of random effects and parameters. A MNP model with K explanatory variables other than alternative-specific constants may have K random effects associated with random taste variation as well as additive effects that are independent of the explanatory variables. If the estimation data are cross-sectional, there are M alternatives, and the additive effects are independent of the taste effects, there are $[K(K + 1) + M(M - 1)]/2 - 1$ distinct covariance matrix elements that must either be estimated from data or set by assumption. Thus, for example, a MNP model with ten alternatives and five explanatory variables has fifty-nine covariance parameters. Carrying out a nonlinear optimization over such a large number of variables is rarely easy. Moreover, in MNP estimation, the log-likelihood function may have ridges or relatively flat regions in which the second derivative matrix of the log-likelihood function is nearly singular. In addition, the MNP log-likelihood function can be multimodal. These characteristics add to the usual difficulties of nonlinear numerical optimization. Moreover, the parameters of a large MNP covariance matrix may be imprecisely estimated, and this imprecision may be transmitted to estimates of utility function parameters, choice probabilities, and market shares.

c. Proliferation of probit covariance parameters can be avoided by imposing prior restrictions on the covariance matrix. One way of doing this is by imposing a factor-analytic structure on ε. This is analogous to selecting the tree structure in a nested logit model. Care is needed in choosing the structure, however, since different structures can yield substantially different predictions of market shares (Forinash and Koppelman, 1993). Particular attention should be given to the possibility of random taste variation, which has been found in a variety of applications of probit models (Fischer and Nagin, 1981; Gönül and Srinivasan, 1993; Hausman and Wise, 1978; Horowitz, 1993a; McCulloch and Rossi, 1992). Random taste variation causes ε to be heteroskedastic and, therefore, has effects on choice

probabilities that are different from those of correlated components of ε that are independent of the explanatory variables.

d. Careful specification testing is especially important when developing a MNP model. Virtually any specification error can create the appearance of random taste variation or additive random components of utility that are not iid. In particular misspecifying the systematic component of the utility function tends to have this effect. If specification testing is not carried out, an analyst may reach the conclusion that there is random taste variation or a complicated covariance structure when the real problem is that the utility function is misspecified. Horowitz (1993a), Horowitz and Härdle (1992), and Horowitz and Louviere (1993) discuss some specification tests that may be useful for avoiding this problem.

e. In models for panel data, it is important to distinguish between unobserved heterogeneity and state dependence (e.g., brand-loyalty), which are quite different conceptually and behaviorally but can create similar patterns in observed choices. Heckman (1981a, 1981b) discusses the econometric issues that are involved. Divakar (1993) presents evidence that failing to take proper account of the effects of unobserved heterogeneity can lead to greatly exaggerated estimates of the effects of brand loyalty in brand-choice models.

References

Albert, J.H., and S. Chib. (1993). Bayesian Analysis of Binary and Polychotomous Response Data, *Journal of the American Statistical Association*, 88, 669–686.

Amemiya, T. (1985). *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

Ben-Akiva, M., and S.R. Lerman. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: MIT Press.

Börsch-Supan, A., and V.A. Hajivassiliou. (1993). Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models, *Journal of Econometrics*, 58, 347–368.

Bult, J.R. (1993). Semiparametric versus Parametric Classification Models: An Application to Direct Marketing, *Journal of Marketing Research*, 30, 380–390.

Bunch, D.S. (1991). Estimability in the Multinomial Probit Model, *Transportation Research*, 25B, 1–12.

Dansie, B. (1985). Parameter Estimability in the Multinomial Probit Model, *Transportation Research*, 19B, 526–528.

Das, S. (1991). A Semiparametric Structural Analysis of the Idling of Cement Kilns, *Journal of Econometrics*, 50, 235–256.

Divakar, S. (1993). Measuring Brand Loyalty with Structural State Dependence Models: An Investigation of Some Issues, Ph.D. dissertation, Department of Marketing, University of Iowa, Iowa City.

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

Fischer, G.W., and D. Nagin. (1981). Random Versus Fixed Coefficient Quantal Choice Models. In C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press, 273–304.

Forinash, C.V., and F.S. Koppelman. (1993). Application and Interpretation of Nested Logit Models of Intercity Mode Choice. *Transportation Research Record*, No. 1413, 98–106.

Gabler, S., F. Laisney, and M. Lechner. (1993). Seminonparametric Estimation of Binary-Choice Models with an Application to Labor-Force Participation, *Journal of Business and Economic Statistics*, 11, 61–80.

Geweke, J., M. Keane, and D. Runkle. (1992). Alternative Computational Approaches to Inference in the Multinomial Probit Model, Manuscript, Research Department, Federal Reserve Bank of Minneapolis.

Gönül, F.F., and K. Srinivasan. (1993). Modeling Multiple Sources of Heterogeneity in Multinomial Logit Models: Methodological and Managerial Issues, *Marketing Science*, 12, 213–229.

Goodwin, B.K. (1993). Semiparametric (Distribution-Free) Testing of Speculative Efficiency in a Parimutuel Gambling Market, Manuscript, North Carolina State University.

Han, A.K. (1987). Non-Parametric Analysis of a Generalized Regression Model, *Journal of Econometrics*, 35, 303–316.

Hajivassiliou, V.A. (1993a). Simulation Estimation Methods for Limited Dependent Variable Models. In G.S. Maddala, C.R. Rao, and H.D. Vinod (eds.), *Handbook of Statistics, Vol. 11.* Amsterdam: Elsevier Science Publishers, 519–543.

———. (1993b). Simulating Normal Rectangle Probabilities and Their Derivatives: The Effects of Vectorization, *International Journal of Supercomputer Applications*, 7, 231–253.

Hajivassiliou, V.A., and D. McFadden. (1992). The Method of Simulated Scores for the Estimation of LDV Models, Manuscript, Cowles Foundation for Research in Economics, Yale University, New Haven, CT.

Hajivassiliou, V.A., D. McFadden, and P. Ruud. (1992). Simulation of Multivariate Normal Orthant Probabilities: Methods and Programs, Cowles Foundation Discussion Paper No. 1021, Yale University, New Haven, CT.

Härdle, W. (1990). *Applied Nonparametric Regression.* New York: Cambridge University Press.

Hausman, J.A., and D.A. Wise. (1978). A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences, *Econometrica*, 46, 403–426.

Hensher, D.A., and L.W. Johnson. (1981). *Applied Discrete-Choice Modeling.* London: Croom-Helm.

Heckman, J.J. (1981a). The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. In C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications.* Cambridge, MA: MIT Press, 179–195.

———. (1981b). Statistical Models for Discrete Panel Data. In C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications.* Cambridge, MA: MIT Press, 114–178.

Horowitz, J.L. (1993a). Semiparametric Estimation of a Work-Trip Mode Choice Model, *Journal of Econometrics*, 58, 49–70.

———. (1993b). Semiparametric and Nonparametric Estimation of Quantal Response Models. In G.S. Maddala, C.R. Rao, and H.D. Vinod (eds.), *Handbook of Statistics, Vol. 11.* Amsterdam: Elsevier Science Publishers, 45–72.

———. (1992). A Smoothed Maximum Score Estimator for the Binary Response Model, *Econometrica*, 60, 505–531.

Horowitz, J.L., and W. Härdle. (1992). Testing a Parametric Model Against a Semiparametric Alternative, Econometric Theory, forthcoming.

Horowitz, J.L., and J.J. Louviere. (1993). Testing Predicted Choices Against Observations in Probabilistic Discrete-Choice Models, *Marketing Science*, 12, 270–279.

Ichimura, H. (1993). Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-Index Models, *Journal of Econometrics*, 58, 71–120.

Keane, M.P. (1994). A Computationally Practical Simulation Estimator for Panel Data, *Econometrica*, 62, 95–116.

Klein, R.L., and R.H. Spady. (1993). An Efficient Semiparametric Estimator for Discrete Choice Models, *Econometrica*, 61, 387–422.

Lerman, S.R., and C.F. Manski. (1981). On the Use of Simulated Frequencies to Approximate Choice Probabilities. In C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press, 305–319.

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

———. (1985). Semiparametric analysis of Discrete Response: Asymptotic Properties of the Maximum Score Estimator, *Journal of Econometrics*, 27, 313–334.

Matzkin, R.L. (1992). Nonparametric and distribution-Free Estimation of the Binary Threshold Crossing and the Binary Choice Models, *Econometrica*, 60, 239–270.

McCulloch, R., and P.E. Rossi. (1992). An Exact Likelihood Analysis of the Multinomial Probit Model, *Journal of Econometrics*, forthcoming.

McFadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration, *Econometrica*, 57, 995–1026.

———. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (ed.), *Frontiers in Econometrics*. New York: Academic Press, 105–142.

———. (1981). Econometric Models of Probabilistic Choice. In C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press, 198–272.

Pakes, A., and D. Pollard. (1989). Simulation and the Asymptotics of Optimization Estimators, *Econometrica*, 57, 1027–1058.

Powell, J.L, J.H. Stock, and T.M. Stoker. (1989). Semiparametric Estimation of Index Coefficients, *Econometrica*, 57, 1403–1430.

Sherman, R.P. (1993). The Limiting Distribution of the Maximum Rank Correlation Estimator, *Econometrica*, 61, 123–138.

Stern, S. (1992). A Method for Smoothing Simulated Moments of Discrete Probabilities in Multinomial Probit Models, *Econometrica*, 60, 943–952.