Discussion Paper

Discussion Paper Nr. 94-12

Some Recent Develpoments in Microeconometrics - A Survey

Heinz König / Michael Lechner

ZEW

Zentrum für Europäische Wirtschaftsforschung GmbH

Industrial Economics and International Management Series

Discussion Paper Nr. 94-12

Some Recent Develpoments in Microeconometrics - A Survey

Heinz König / Michael Lechner

Some Recent Developments in Microeconometrics - A Survey

Heinz König and Michael Lechner[1]

April 21, 1994

Universität Mannheim,
Institut für Volkswirtschaftslehre und Statistik
Postfach 103462
D-68131 Mannheim, Germany
email: Lechner@mailrum.uni-mannheim.de

Zentrum für Europäische
Wirtschaftsforschung (ZEW)
Postfach 103443
D-68043 Mannheim, Germany

[1] Financial support by the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

Abstract

This paper summarizes some recent developments in microeconometrics with respect to methods for estimation and inference in non-linear models based on cross-section and panel data. In particular we discuss recent progress in estimation with conditional moment restrictions, simulation methods, semiparametric methods, as well as specification tests. We use the binary cross-section and panel probit model to illustrate the application of some of the theoretical results.

Zusammenfassung

Diese Arbeit faßt einige neuere Entwicklungen auf dem Gebiet der Mikroökonometrie zusammen. Methoden für die Schätzungen und die Inferenz in nicht-linearen Modellen auf Querschnitts- als auch Paneldaten stehen dabei im Vordergrund. Es werden Schätzmethoden basierend auf bedingten Momentenrestriktionen, Simulationsmethoden, semiparametrische Methoden sowie Spezifikationstests diskutiert. Anhand des binären Probitmodells für Querschnitts- wie auch für Paneldaten veranschaulichen wir einige der theoretischen Erörterungen.

Résumé

Cet article présente un tour d'horizon de développements récents en microéconométrie, en mettant l'accent sur l'estimation et l'inférence pour les modèles non linéaires estimés sur coupe transversale ou sur données de panel. Nous discutons en particulier les progrès récents liés aux méthodes de moments conditonnels, aux méthodes de simulation, aux méthodes semiparamétriques, ainsi que les tests de spécification correspondants. Nous fournissons des illustrations basées sur l'estimation de modèles probit sur coupe transversale et sur données de panel.

1. Introduction

For most of the period since the foundation of the Econometric Society econometrics have been dominated by macroeconomic model building and time series analysis. Except for some fields like agricultural economics (HANAU, 1928), demand analysis (STONE, 1954, BALESTRA and NERLOVE, 1966), human capital theory (MINCER, 1958, 1974), investment activity (MEYER and KÜH, 1957) and production function analysis (DOUGLAS, 1948), microeconomics played a minor role in applied research. Since the 70's, however, microeconometrics became increasingly important both in methodological and applied research. The availability of large data sets for households, firms etc. on the one side, as well as the generally felled need for the microeconomic foundation of macroeconomics on the other side are the main reasons for this development, the extent of which has been made possible by the dramatic decrease computational costs. While econometric textbooks in the 50's and 60's contained at best the linear model for cross-section analysis, there exists now a broad literature dealing with microeconometric estimation and inference problems. This growing interest in the field of microeconometrics may be seen by the fact that the Handbook of Econometrics in its three volumes contains eight paper dealing at length with topics covering the range from labour econometrics, demand analysis, and firm behaviour to typical microeconometric inference problems. And the next edition of the Handbook of Econometrics will show, as did the recent edition of the Handbook of Statistics (volume: *econometrics*), that the speed of the appearance of new developments has not yet decreased.

Technically speaking, in typical microeconometric work datasets with characteristics of many heterogeneous individual economic agents, such as households or firms, are analysed with econometric methods using causal relations among certain random variables of interest. These causal relations are directly or indirectly obtained from behavioural assumptions for the agents. Generally agents are supposed to maximise certain quantities, such as utility or profit, subject to economic and institutional constraints. The resulting model provides restrictions on the class of admissible data generating processes which describe the population. Common restrictions are the nullity of some conditional moments or parametric forms for certain conditional distributions, which could then be used to analyse the available random sample from this population for example by the Generalised Method of Moments (GMM) or by Maximum Likelihood (ML). Furthermore overidentifying restrictions could be used to check the compatibility of the chosen model and the sample. The goal of the analysis is to describe aspects of this behaviour of the agents and to analyse the outcome of potential behavioural changes, such as the reaction of labour supply to changes in the tax-benefit systems, firms' decisions to invest or to innovate, or household decisions to buy durable goods.

The restrictions used for estimation and inference may be complicated for various reasons: First of all, in order to allow for sufficiently realistic models, flexible functional forms may have been chosen for the utility or profit functions. Secondly, the possible choices may be discrete by nature, for example when analysing the use of different modes of transportation, or for sampling reasons, when for instance firms are asked in so-called business surveys to forecast changes in the business climate. Thirdly, complicated restrictions may have to be taken into account, such as the non-convex budget sets generated by nearly all 'real' tax-benefit systems. Furthermore the available sample may contain variables, which are censored, truncated, clustered, qualitative, selectively sampled, or measured with error.

The econometric methods developed, some of which will be described in the following, take specifically account of these problems by exploiting the typical advantage of today-micro-econometrics: the availability of a disaggregated dataset containing a large number of economic units, which can be treated as the result of independent draws from the population of interest. These datasets contain much information about the agents and may even be repeated for the same or different agents over time. Although the assumption of random draws may seem inappropriate in particular for certain company datasets, it is at the heart of nearly all microeconometric estimation procedures, and so we will stick to it in this survey, too.

However, the desired generality of the microeconomic models used for the empirical analysis has always been limited by at least three interrelated factors: The quality of the datasets, the required econometric methods, and the price and availability of the software and computation time necessary for a sophisticated analysis. In the last decade a significant relaxation of these constraints could be observed:

(i) More datasets containing more information became available. In particular, datasets with a panel character, i.e. where the same agent is sampled over more periods, offered new possibilities for identification and estimation of microeconometric models (see HAMERMESH, 1990, for the field of labour economics). In this context it is argued (see e.g. PUDNEY, 1989) that compared to the traditional macroeconometric modelling the microeconometric approach has important advantages insofar that the behaviour of the individual agent is easier to understand while taking into account individual heterogeneity (compared to the representative agent approach of macroeconomics), and that the researcher has more insight into the process by which the data are collected. A warning, however, seems to be appropriate already here. Microeconomic data (if they are not already aggregated by the sampling procedure, e.g. family income instead of individual income, price expectations for commodity groups) may exhibit a larger noise to signal ratio which might be reduced by aggregation. Self-selection or

panel mortality play an important role and the assumption of independence of individual behaviour might be questioned (see GRILICHES, 1986). We will come back to these issues, restricting policy conclusions of microeconometric studies, later.

(ii) The cost of computation time has decreased drastically due to the widespread use of inexpensive, but powerful personal computers. This does not only foster the use for more general microeconomic models which are estimated with more complex and robust econometric methods, but also drastically decreases the costs of handling large survey data-sets at all stages from data collection to the final analysis. The latter obviously has positive external effects on (i).

(iii) Furthermore, many new econometric methods become part of standardised software packages, and new software was developed which allowed easy programming of methods exactly tailored to the specific problems of the analysis. The development of econometric methods, which will be the main focus of this survey, is certainly related to the other developments, since the prospects of potential applications of more sophisticated, and more burdensome, methods increased. Clearly the process from inventing new estimation methods to their routine use in empirical analysis is subject to considerable time lags. But even these time lags seem to get shorter as the number of microeconometricians increases.

A decade ago the toolbox of the microeconometricians contained estimation methods for many models which account for the basic problems mentioned above. Typically they are developed for cross-section analysis, imply tight distributional assumptions on unobservable variables, and are based more on fully parametric ML estimation. These models, pioneered by TOBIN (1958), are well documented for example in the surveys by AMEMIYA (1981, 1984), MADDALA (1983) and MCFADDEN (1984). All modern econometric textbooks contain some of them (e.g. DAVIDSON and MACKINNON, 1993, RONNING, 1991). Subsequently these methods have been applied extensively to all sorts of microeconometric problems. However, researchers became worried about the impact of distributional assumptions on the results of the analysis, very often difficult to justify. It seems reasonable to argue that the following developments tackle this problem in three different ways: First of all, lots of specification tests have been developed which can be used to check the validity of the chosen specification. Secondly, more flexible and more sophisticated parametric models are used for the econometric analysis. However, ML estimation is too burdensome or even impossible to conduct for many of them, so that some efficiency is sacrificed and GMM methods or simulation methods are proposed for the estimation of these models. Thirdly, more and more semi- and nonparametric models have been proposed to avoid unnecessary overidentifying restrictions and, thus, allowing for more robust estimations.

The object of this survey is a selective overview of these developments and their practical applications. We try to stick to an essentially non-technical presentation and will use examples to clarify some points of the more theoretical discussion. For the respective complete sets of assumptions for which the results are valid and the associated proofs, the reader is referred to the original papers. In order to keep to this non-technical approach we will not discuss two very technical, but important areas which are of interest for proofing properties of parametric and semiparametric estimators, that is the concept of stochastic equicontinuity (ANDREWS, 1994, NEWEY, 1991, PÖTSCHER and PRUCHA, 1994) and that of semiparametric efficiency bounds (CHAMBERLAIN, 1986, 1992, COSSLETT, 1987, NEWEY, 1990b, NEWEY and POWELL, 1993, THOMPSON, 1993).

A lot of different econometric models have been developed in the literature and applied to microeconometric analysis. Discussing their particular advantages and problems is far beyond our scope and space limits, and is not necessary anyway because many excellent surveys are available, some of which will be mentioned in the following. The most prominent example is the linear model which still has its merits for many microeconometric applications in particular for panel data (RAJ and BALTAGI, 1992). Other examples are models for count data (GURMU and TRIVEDI, 1992, POHLMEIER, 1994), duration models (FLORENS, 1990, LANCASTER, 1990), simultaneous non-linear models (BLUNDELL and SMITH, 1993), LISREL, LISCOMP and similar models (ARMINGER and MÜLLER, 1990, MUTHÉN, 1987), dynamic non-linear models for panel data (HECKMAN, 1981, LECHNER, 1993a, b), models for errors in variables (MARIANO and BROWN, 1993), discrete choice models (MADDALA, 1983, MCFADDEN, 1984), many variants of limited dependent variable models (AMEMIYA, 1981, 1984, MADDALA, 1983, 1987), dynamic discrete choice models (ECKSTEIN and WOLPIN, 1989a, HOTZ and MILLER, 1993, MANSKI, 1991, 1993a, RUST, 1991, 1994), and disequilibrium models (MADDALA, 1986, LAROQUE and SALINIÉ, 1994), among others. From a methodological point of view we completely ignore Baysian analysis (FLORENS and MOUCHART, 1993), and the problem of optimal prediction in non-linear models (MANSKI and THOMPSON, 1989, MADDALA, 1993) not because we consider these topics as unimportant, but because it is beyond the amount of material our survey can carry. For this same reason we will also completely ignore the important topics of optimal sample design and choice based sampling (COSSLETT, 1993, IMBENS, 1992) and pseudo panel data (MOFITT, 1993, VERBEEK and NIJMAN, 1993).

The paper is organised as follows. Section two discusses the general notation of a binary choice model. Estimation problems and approaches are discussed in section three. The fourth section presents an overview about recent development in diagnostic testing, followed by a

brief discussion of some issues of non- and semiparametric estimations in section five. Finally, in section six we conclude with some considerations about topics for future research.

2 An Example: The Binary Choice Model

In section 2 we discuss some econometric methods for estimation and inference in cross-sections and panel data which have been proposed recently. In order to clarify the basic intuition underlying the various procedures we use the binary choice model as an example. Consider the following model:

$$\begin{aligned} y_i^* &= g(x_i, \beta) + u_i, \\ y_i &= 1(y_i^* > 0), \qquad i = 1, \ldots, N. \end{aligned} \tag{1}$$

The $T \times 1$ and $T \times K$ dimensional matrices $(y_i, x_i) = z_i$ denote the 'i'-th observed realisation from N random draws in the joint distribution of the random variables $(Y, X) = Z$. The unobserved vectors y_i^* and u_i serve merely as a device to rationalise restrictions on certain conditional and joint distributions of Z. All procedures which are discussed in the following are based on the 'analogy'-principle (MANSKI, 1988a), which means that we specify some characteristics of the distribution of Z, such as conditional moments or conditional densities, and use the appropriate sample analogs, such as the sample mean, to obtain consistent estimates of the parameters of interest θ, or other quantities of interest, which characterise parts of the distribution of Z. In general θ contains the K × 1 dimensional parameter vector β and parameters of the covariance matrix of the errors Σ. $\theta^0, \beta^0, \Sigma^0$ denote the parameter values by which the true distribution of Z is characterised. . The positive integer 'T' may denote a fixed time dimension in the case of a panel data model, but can also be interpreted as a number of choices in the discrete choice problem for example. y_{ti} equals one if the expression in the 't'-th component of the indicator function $1(\cdot)$ is true and zero otherwise. By choosing other functions as $1(\cdot)$ for the transformation of the latent dependent variable to its observed counterpart, other popular models, such as ordered choice models or tobit-type models can be generated. From the point of view of the loss of information by observing y_i instead of y_i^*, the binary choice model is the worst case in this class of latent models, because only the sign of y_i^* is observed. The implications are at least twofold: First, this poses difficult identification problems for θ and other quantities of interest. These problems are extensively discussed in MANSKI (1988b) and reviewed in HOROWITZ (1993a). One important conclusion is that conditional mean independence, e.g. $E(U|X = x; \theta^0) = 0$, has no identifying

power whatsoever. Second, all methods which 'work' for this model, could also be applied to the other models in that class when y_i carries more information about y_i^*.

3 Estimation Methods for Parametric Models

In this section we use the term 'parametric' in the sense that the researcher knows or pretends to know the complete conditional distribution of $Y|X$, denoted by $F_{Y|X}$, or certain features of it like conditional moments, up to a finite set of parameters θ. Furthermore, a crucial condition for the applicability of all the methods discussed in this section will be that the respective objects ($F_{Y|X}$ or the moments) will be smoothly differentiable w.r.t. the parameters. Hence it will be the typical feature of the semi- and nonparametric models, discussed in section 4, that they are characterised either by an infinite set of (nuisance) parameter and/or non-differentiability.

3.1 Maximum Likelihood

Assume that the researcher knows the conditional distribution of $Y|X\left(F_{Y|X}\right)$. Given the usual regularity conditions (see for example GOURIEROUX and MONFORT, 1989) hold, maximum likelihood estimation (ML) gives consistent, asymptotically efficient and $\sqrt{N}$-normal estimates for θ^0. Typically the specification of $F_{Y|X}$ is done on the level of the latent model by assuming a parametric form for $g(X,\beta)$ and $F_{U|X}$, such as linearity, e.g. $g(X,\beta)=X\beta$, multivariate normality of $F_{U|X}$ and independence of U and X. For this T=1 is the well-known cross-section probit model which can easily be estimated with modern software packages. The other popular model, the logit model, results from assuming a logistic distribution for $F_{U|X}$ instead. Both models are very difficult to distinguish empirically.

The average log-likelihood function of a random sample for the cross-section probit model, denoted by $L_N(\cdot)$, is given by:

$$L_N(\theta)=\frac{1}{N}\sum_{i=1}^{N} y_i \ln\Phi(x_i\theta)+(1-y_i)\ln\left[1-\Phi(x_i\theta)\right]. \qquad (2)$$

$\Phi(a)$ denotes the cumulative distribution function of the univariate standard normal distribution evaluated at a. The ML-estimate $\hat{\theta}_N$ of the scaled coefficients $\theta^\circ = \beta^\circ/\sigma_1^\circ$, where σ_1° denotes the square root of the first element of Σ, can be computed by standard iterative procedures without any problems. Now consider the situation that a panel data set of T waves is available. Assume that the errors are independent of all regressors and jointly normally distributed. The average log-likelihood function is given by:

$$L_N(\theta) = \frac{1}{N}\sum_{i=1}^{N} \ln \Phi^{(T)}\big(a_i(\theta), b(\theta)\big). \tag{3}$$

$\Phi^{(T)}$ denotes the cumulative distribution function of the multivariate normal distribution with mean zero and covariance matrix $b(\theta)$ evaluated at $a_i(\theta)$. Let $\theta = (\theta_1, \theta_2)$ with $\theta_1 = \beta/\sigma_1$ and θ_2 contain the scaled variance σ_1/σ_t and the correlations appearing in Σ. A typical element of the T-dimensional vector $a(\theta)$ is given by $a_{it} = x_{it}\theta_1(2y_{it}-1)$, a typical element of the $(T \times T)$-dimensional matrix $b(\theta)$ is $b_{ts} = \sigma_{ts}/\sigma_1(2y_{it}-1)(2y_{is}-1)$ (AMEMIYA, 1986). The $T-1$ additional waves lead to a gain in efficiency for the estimation of the scaled coefficients β/σ_1 and, in sufficiently regular cases, to the identification of the relative variances σ_1/σ_t and the correlation matrix of the error terms. However, note that there are $T(T+1)/2-1$ additional parameters to be estimated and that a T-dimensional integration over the multivariate normal density has to be performed. This complicates the estimation considerably, in the sense that the number of iterations necessary for convergence increases. Local extremas may be a problem and the estimates of the correlations or their transformations may tend towards the boundary of the parameter space by approaching +1 or -1, and thus lead to ill-conditioned estimates. Since analytical closed form formulas for the integration are not available, they have to be computed numerically, which is infeasible for $T > 4$, and very burdensome for $T > 2$.

As a way to reduce the number of parameters and the dimension of integration, the following factor-analytic decomposition of the error terms has become popular:

$$\begin{aligned} &u_{it} = \delta_i c_i + \varepsilon_{it} \quad ; \quad c_i \sim N(0,1), \quad \varepsilon_{it} \sim N(0,\sigma_t^2); \\ &E\, c_i \varepsilon_{it} = 0, \quad E\, \varepsilon_{it}\varepsilon_{si} = 0, \quad t \neq s. \end{aligned} \tag{4}$$

This implies that $P(Y_t|Y_s, X, C) = P(Y_t|X, C), t \neq s$, and so the log-likelihood function simplifies to a one dimensional integral over products of the cumulative distribution function

of univariate standard normal distributions. Denote the respective density function by $\phi(c)$ and E_c the expectation operator w.r.t. to the marginal distribution of c, then:

$$L_N(\theta) = \frac{1}{N}\sum_{i=1}^{N} \ln E_c\left\{\prod_{t=1}^{T} y_{it}\Phi\left(\frac{x_{it}\beta + \delta_t c}{\sigma_t}\right) + (1-y_{it})\left[1-\Phi(\cdot)\right]\right\}$$

$$= \frac{1}{N}\sum_{i=1}^{N} \ln \int_{-\infty}^{+\infty}\prod_{t=1}^{T}\left\{y_{it}\Phi\left(\frac{x_{it}\beta + \delta_t c}{\sigma_t}\right) + (1-y_{it})\left[1-\Phi(\cdot)\right]\right\}\phi(c)\,dc, \tag{5}$$

and appropriately normalised coefficients can be estimated. The algorithm of BUTLER and MOFITT (1982), based on Gaussian quadrature, provides an efficient tool for the unidimensional numerical integration necessary (HSIAO, 1986, 1992). This approach serves as a flexible method for the analysis of models with a small time dimension, but for larger T the restrictions implied for Σ may, if not true, still lead to inconsistent estimates. For example, if $T > 4$, u_{it} can only be stationary for the special case of $\delta_t = \delta, \forall t$, which is the pure random or equicorrelation case and rules out for example AR and MA error processes (AMEMIYA, 1986, HECKMAN, 1981). The following methods will, by sacrificing efficiency, overcome this problem in different ways.

3.2 Conditional Moment Restrictions

In applied work there are many reasons why it may be desirable to avoid the complete specification of the distribution of the dependent variables given the independent variables. The most prominent reasons are that the researcher has no information, or no confidence in arbitrary assumptions, on higher conditional moments describing the full distribution. Additional parameters describing the complete conditional distribution may increase the complexity of the model and can lead to problems related to identification, computation and small sample properties. If the researcher is prepared to specify functions which depend on the variables in the population and a finite number of parameters, have mean zero, identify the parameters of interest, are smoothly differentiable with respect to parameters, and satisfy some additional regularity conditions, then the generalised method of moments (GMM) can be used for the estimation. This method introduced by HANSEN (1982) is in full accordance to the analogy principle. It is based on estimating the parameters by minimising quadratic forms of functions which are sample analogs of the respective population moments. The major

distinction to the semi- and nonparametric approaches discussed later will be the requirement of smoothness, differentiability and the finiteness of the number of parameters on which the moments depend. We chose this distinction for the ease of presentation. In econometrics a specific form of moment restrictions play a prominent role, namely the conditional moment restrictions. Recent important insights in the properties of estimators based on conditional moment restrictions have been obtained by CHAMBERLAIN (1987) and NEWEY (1990a). The excellent survey by NEWEY (1993) summarises these results and elaborates on them. The following exposition borrows heavily from this source.

The $T \times 1$ dimensional function $M(Z, \theta)$ depends on the $K \times 1$ dimensional parameter vector θ, and satisfies the following conditional moment restrictions for the true parameter value θ^0.

$$E\left[M(Z, \theta^0) | X = x_i\right] = 0 \tag{6}$$

For identification purposes it necessary that there is no other value of θ in the parametric space Θ (which is part of the R^K) that also fulfils this restriction. The conditional moment restriction implies that at the true value of the paramameters -and only at the true value- all functions of the conditioning variables (X) are uncorrelated with the moments $M(Z,\theta^0)$. Hence the following unconditional moment restrictions can be formed:

$$E\, A(X) M(Z, \theta^0) = 0. \tag{7}$$

Note that the $T \times 1$ dimensional moment condition in (6) has now be transformed to the p x 1 dimensional moment condition in (7). p denotes the row dimension of the $p \times T$ dimensional matrix $A(X)$, and must be at least as large as K to identify the parameters. If p > K there are overidentifying restrictions. It has become common in this literature to call $A(X)$ the *instrument matrix*. Note however that this may cause some confusion related to the standard IV-terminology used in linear regression. Here we are not instrumenting a particular endogenous regressor, but exploiting (6) to form unconditional moments to be used in estimation. But that there is no contradiction with linear IV-estimation. In the case of a linear model, $M(Z,\theta^0)$ would denote the residual, Z would include the dependent variable Y and the endogenous and exogenous regressors. The conditioning variables X would contain only the exogenous regressors plus additional variables to be used as instruments for the endogenous regressors. Those would be used to form the instrument matrix $A(x)$.

The sample analogs for the expectation appearing in (7) are given by the following arithmetic means over the N independent observations:

$$g_N(\theta) = \frac{1}{N}\sum_{i=1}^{N} A(x_i)M(z_i,\theta) \tag{8}$$

The idea of GMM estimation is that for an large sample size (8) should approach zero for the true value of the parameters. Hence the value of the parameters which sets these empirical moments to zero is a *natural* estimate ($\hat{\theta}$) of the true value. If there are overidentifying restrictions, this may not be possible and the following quadratic form is minimised instead:

$$\hat{\theta}_N = \arg\min_{\theta\in\Theta} g_N(\theta)' W g_N(\theta). \tag{9}$$

W denotes any choice of a *p x p* dimensional positive definite weighting matrix. Under suitable regularity conditions on g_N and W, $\hat{\theta}$ is $\sqrt{N}$-consistent and asymptotically normal.

Since in this approach the amount of prior information used, e.g. the conditional moment restrictions, may be much lower than for example in ML estimation, it is particularly important to exploit the available information fully to obtain asymptotically efficient estimators given this information. The tools which can be used to achieve this are the optimal choices of the instruments *A(x)* and of the weighting matrix *W*. As shown by Hansen (1982) in a more general setting, the optimal choice of W is $\left\{E\left[A(X)M(Z,\theta^0)M(Z,\theta^0)'A(X)'\right]\right\}^{-1}$ or any consistent estimator of that expression. CHAMBERLAIN (1987) and NEWEY (1990a) derived the optimal choice of *A(x)*. Let

$$D(x_i) = E\left[\partial M(Z,\theta_0)/\partial\theta' | X = x_i\right], \qquad \Omega(x_i) = E\left[M(Z,\theta^0)M(Z,\theta^0)' | X = x_i\right].$$

The optimal choice for $A(x_i)$ denoted by $A^*(x_i)$ equals:

$$A^*(x_i) = C\,D(x_i)'\Omega(x_i)^{-1}, \tag{10}$$

where *C* is any non-singular $K \times K$ matrix. Note that the column-dimension of $A^*(x_i)$ equals *K*, so that the choice of *W* is irrelevant. $D(x_i)$ and $\Omega(x_i)$ may be substituted by consistent estimates. Note that in the linear regression model $D(x_i)$ will just denote the regressors, and

$\Omega(x_i)$ corrects for conditional heteroscedasticity and correlations, so that this approach results simply in a feasible GLS estimator.

However, finding these consistent estimates of the optimal instruments may be formidable task in a complex model, as will be exemplified for the relatively simple panel probit case. To circumvent these problems NEWEY (1990a, 1993) suggests the use of nonparametric methods, such as nearest neighbour estimation and series approximations instead, and derives the conditions necessary for these methods to result in consistent and asymptotically efficient estimates.

Some aspects of the estimation with conditional moment restrictions will be clarified in the cases of cross-section and panel probit models. In a single cross-section the conditional expectation of the observed variable, as defined in section 2.1, is given by the probability that it equals one, and so the following conditional moment can be used for estimation:

$$M(Z,\theta) = Y - \Phi(X\theta) \quad ; \quad \theta = \beta/\sigma. \tag{11}$$

In order to find the asymptotically optimal matrix of instruments, the conditional expectations of derivative *w.r.t.* θ and the conditional variance, $\Omega(x)$, have to be evaluated at the true value:

$$\begin{aligned} E\left[\left.\frac{\partial M(Z,\theta^0)}{\partial\theta}\right| X = x_i\right] &= -\phi(x_i\theta^0)x_i' \\ E\left[M(Z,\theta^0)M(Z,\theta^0)\middle| X = x_i\right] &= \Phi(x_i\theta^0)\left[1-\Phi(x_i\theta^0)\right]. \end{aligned} \tag{12}$$

Since there are no overidentifying restrictions, the optimal GMM-estimator, $\hat{\theta}$ fulfils:

$$g_N(\hat{\theta}) = 0 = \frac{1}{N}\sum_{i=1}^{N}\frac{\left[y_i - \Phi(x_i\hat{\theta})\right]\phi(x_i\hat{\theta})}{\Phi(x_i\hat{\theta})\left[1-\Phi(x_i\hat{\theta})\right]}x_i'. \tag{13}$$

Note that $g_N(\cdot)$ equals the first derivative of the log-likelihood function, hence the optimal GMM-estimator is identical to the ML-estimator.

The analysis of panel data provides the more interesting case. The estimator which is proposed in the following is based on assumptions about the marginal distribution of U_t, the joint distribution of U is otherwise unrestricted, except that it is supposed to fulfil certain regularity conditions which ensure the existence of second moments. The marginal distribution could be derived from some multivariate distribution for U, but this is not necessary.

Suppose that each U_t is normally distributed with mean zero and variance σ_t^2. Furthermore, it is assumed to be independently distributed from all elements of X. The $T \times 1$-dimensional moment function is given by:

$$\begin{aligned} M(Z,\theta) &= \left[m_1(Z_1,\theta),\ldots,m_t(Z_t,\theta),\ldots,m_T(Z_T,\theta)\right]' \\ m_t(Z_t,\theta) &= Y_t - \Phi\left(\frac{X_t\beta}{\sigma_t}\right). \end{aligned} \tag{14}$$

The use of these types of moments as a basis of GMM estimation has first been advocated by AVERY et al. (1983). It can be shown that the ML-estimators, under the assumption of independent errors (AVERY et al., 1983), and the sequential random effect probit estimator suggested by CHAMBERLAIN (1980, 1984), belong to this class of estimators and differ only in the choice of the instrument matrices $A(X)$ and W, and by using one-step vs. two-step estimation procedures (BREITUNG and LECHNER, 1994). Note that embedding the unrestricted ML-estimator in this framework would require specification of the probability of the complete sequence $(Y_1,\ldots,Y_T)$ and appropriately defined indicator variables for the occurrence of a particular sequence of events. However, as for the ML estimation this would require T-dimensional integration and estimation of Σ.

In order to obtain the optimal GMM-estimator based on (10) the following notation is introduced. Let $\theta=(\theta_1,\theta_2)'$, $\theta_1=\beta/\sigma_1$, $\theta_2=(\theta_{22},\ldots,\theta_{2T})'$, $\theta_{2t}=\sigma_1/\sigma_t$ and $\rho_{ts}=\sigma_{ts}/(\sigma_1\sigma_2)$, where σ_{ts} denotes $E\,U_tU_s$. Denote the $T \times K$-matrix of first derivatives of M as M_θ, then:

$$M_\theta(Z,\theta) = \left[m_{\theta 1}(Z_1,\theta)', \ldots, m_{\theta T}(Z_T,\theta)' \right]'$$

$$E\left[m_{\theta t}(Z_t,\theta) \middle| X = x_i \right] = \left[-\phi(x_{it}\theta_{2t}\theta_1)\theta_{2t}x_{ih}', 0, \ldots, -\phi(x_{it}\theta_{2t}\theta_1)x_{it}\theta_1, \ldots, 0 \right].$$

Let ω_{tsi} a typical element of $\Omega(x_i)$. Furthermore, assume that (U_t, U_s) are jointly normally distributed with correlation coefficient ρ_{ts}. For notational convenience let $\Phi_{ti} := \Phi(X_t\theta_1\theta_{2t})$, $\Phi_{ti} := \Phi(x_{it}\theta_1\theta_{2t})$ and $\Phi^{(2)}_{tsi} := \Phi^{(2)}(x_{it}\theta_1\theta_{2t}, x_{si}\theta_1\theta_{2s}, \rho_{ts})$. $\Phi^{(2)}(\cdot)$ denotes the cumulative distribution function of the bivariate standardised normal distribution:

$$\omega_{tsi} = \left[E(Y_t - \Phi_t)(Y_s - \Phi_s) \middle| X = x_i \right] = \begin{cases} \Phi_{ti}(1 - \Phi_{ti}) & \text{if } t = s, \\ \Phi^{(2)}_{tsi} - \Phi_{ti}\Phi_{si} & \text{if } t \neq s. \end{cases}$$

Note that $\omega_{ts}(x)$ has the same sign as ρ_{ts} and that $\omega_{ts} = 0$ if $\rho_{ts} = 0$, so that the GMM-estimator collapses to the ML-estimator for uncorrelated errors, but generally the optimal GMM-estimator is less efficient than the ML-estimator. However, the estimation of the optimal GMM-estimator is still difficult, because it depends on the unknown correlation coefficients of Σ. Although the unknown coefficients could be substituted by consistent estimates without affecting the asymptotic distribution of the estimates, obtaining them would require $(T-1)T/2$ bivariate probits, which can be cumbersome for large T. An alternative, suggested by NEWEY (1993), is to use nonparametric methods, such as nearest neighbour or series estimation to obtain consistent estimates of $\Omega(x_i)$. The idea behind nearest neighbour estimation (NN) is very simple: If there is only a finite number of (J) of configurations for X, each containing a large number of observations (N_j), then averaging the squared residuals within each sub population having the same x^j e.g. $\frac{1}{N_j}\sum_{i=1}^{N_j}\left[y_i^j - \Phi\left(x^j\hat{\theta}\right) \right]\left[y_i^j - \Phi\left(x^j\hat{\theta}\right) \right]$ would give a consistent estimate for $\Omega(x^j = x_i)$, where $\hat{\theta}$ is a consistent, but inefficient estimate. However, in most finite sample there are only a few observations having the same values for all explanatory variable x_i. NN weights the residuals according to their similarity to x_i. Under regularity conditions (NEWEY, 1993) this will give consistent estimates of $\Omega(x_i)$ for each individual without the need for estimating ρ_{ts}. Another alternative is to sacrifice more efficiency and to use sub optimal choices for $A(x_i)$, such as the regressors only. Some of the

possible choices have been investigated by AVERY et al. (1983) and BREITUNG and LECHNER (1994).

The panel data example showed the trade-off between asymptotic efficiency, robustness to arbitrary assumptions and computational convenience which can be controlled by using specific types of GMM-estimators, and which makes GMM estimation such a useful tool for applied microeconometric work.

Another potential in GMM estimation is that this framework easily allows to integrate information from outside the basic sample. The outside information can be in the form of a second sample of a similar type or of information on the population as a whole, which may stem from a caucus, for example. For the theoretical considerations the reader is referred to the work of Arrelano and MEGHIR (1992) and IMBENS and LANCASTER (1991). First applications which are contained in ARRELANO and MEGHIR (1992), GEIB et al. (1992) and LAISNEY and LECHNER (1993) show the potential of the additional gains in efficiency and identification.

3.3 Simulation Methods

Recently econometric methods based on simulation techniques have become increasingly popular. This has been induced by the dramatic decrease of the cost of computing power. Parallel to this there are significant advances in statistical methods dealing with simulations. Simulation methods play an important role in problems when the analytical derivation of a certain expression of interest is not possible and when fast and sufficiently exact numerical methods, such as Gaussian quadrature for the evaluation of certain integrals are not available. Fields where simulations have been applied for a long time are for example Monte Carlo studies (see DAVIDSON and MACKINNON, Ch. 21, 1993) and Bootstrap methods (JOENG and MADDALA, 1993, VINOD, 1993).

Recent efforts have been made on integrating simulation methods in 'standard' theory for estimation and inference in parametric models. Besides classical estimation problems with independent observations on which the following exposition focuses, PESARAN and PESARAN (1993) show how these techniques can be fruitfully applied to the computation of the Cox-statistic in complicated classical inference problems. Furthermore, the special issue of the *Journal of Applied Econometrics* (BROWN et al., 1993) contains applications to Baysian estimation (ANDREWS et al., 1993, GEWEKE, 1993, KLEIBERGEN and van DIJK, 1993) and time series models (SMITH, 1993, SHEPHARD, 1993).

An important innovation for the use of simulations in classical estimation problems is the paper by LERMAN and MANSKI (1981). They suggest an approximation of the choice probabilities conditional on regressors in a multinomial probit model by drawing in the distribution of the error terms for each individual and by averaging the simulated 'observed' outcomes over simulations. These averages are used instead of the actual probabilities to form the likelihood function. The performance of this estimator in terms of draws necessary to achieve a performance similar to ML was disappointing, since it violates at least the first three of the following conditions which later turned out to be very important for the construction of efficient simulation estimators: the simulated probabilities should be smooth functions of the parameters of the model, the same draws should be used in different iterations, the simulated quantities should be unbiased, the draws should be independent for each individual, and the additional variance included by the simulations should be small. Under some additional regularity conditions these requirements are enough to obtain estimators which are consistent and $\sqrt{N}$-asymptotically normal for a fixed number of simulations. These insights are basically due to a seminal paper by MCFADDEN (1989) which provides a rigorous treatment of simulation estimations using GMM-type estimators with particular respect to the multinomial probit model. Another seminal paper by PAKES and POLLARD (1989) provides useful conditions to prove consistency and asymptotic normality in cases when the simulations may not be smooth functions of the parameters. They apply their methods also to a similar GMM-type estimation of the multinomial probit model. For the case of smooth simulators GOURIEROUX and MONFORT (1991) derived the asymptotic properties of simulated ML (SML), simulated GMM (SGMM), and simulated Pseudo-ML (SPML, see also GOURIEROUX and MONFORT, 1993c) estimators. HAJIVASSILOU and MCFADDEN (1990) suggested using the scores of the log-likelihood function directly as objects for simulation. Those methods are termed as method of simulated scores (MSS). The comprehensive surveys by HAJIVASSILOU (1993) and KEANE (1993, 1994) discuss in great detail the properties of the different methods which can be used for estimation. Furthermore, they discuss the various possibilities of constructing the objects for simulations and obtaining unbiased simulators with a small variance. The latter is particularly important for SML of non-linear models because the consistency and asymptotic normality depend on the condition that $\sqrt{N}/H \to 0$, where H denotes the number of independent draws for each individual. Reducing the variance of the simulation decreases the number of draws necessary to be close to the asymptotic distribution. Suggestions by STERN (1992), GOURIEROUX and MONFORT (1993a) and in particular by BÖRSCH-SUPAN and HAJIVASSILIOU (1993), GEWEKE (1991) and KEANE (1993) appeared to be very successful in achieving this goal at least for discrete choice models.

The 'indirect inference' approach suggested by GOURIEROUX et al. (1993b) exploits the potential of simulation techniques in another, ingenious way. They propose to estimate a possibly incorrect but simple model, and to use simulation methods to establish the relation of the parameters from the incorrect model and the parameters of the correct model. This relation can be used to obtain consistent and asymptotically normal estimates. The advantage is that it is not necessary to specify any conditional moments of the true model, which are perhaps very complicated. It is sufficient that the true models can be simulated. For the asymptotic results obtained and the numerical convenience in the application, it is important that the necessary simulations of the true model depend smoothly on the parameters, which is not the case for example in discrete choice models. Further research is needed to find ways to smooth that approach sufficiently. Its enormous potential will be worth the effort necessary.

4 Diagnostic Testing of Parametric Models

The assumptions on $F_{Y|X}$, perhaps with the exception of the first moment, are generally not derived from some structural economic model, but imposed to achieve a computationally convenient estimator. However, the consistency and efficiency of the estimated coefficients and standard errors generally depend on the validity of these assumptions. Although there are some exceptions for particular models and data generating processes (RUUD, 1983, 1986), the incurred biases when these assumptions are incorrect could be substantial (GABLER et al., 1993, MANSKI and THOMPSON, 1986, HOROWITZ, 1993a, for the binary choice model, or ARABMAZAR and SCHMIDT, 1982 for the tobit model). One possibility to avoid these biases is to estimate more general models in the beginning, which could accommodate the suspected features in a parametric way, for example by assuming particular forms of heteroskedasticity (DAVIDSON and MACKINNON, 1984) or using more general distribution functions nesting the normal (GABLER et al., 1993) or the logistic (LECHNER, 1991, POIRIER, 1980, THOMAS, 1993) distribution. The drawback is that one cannot be sure that the chosen specification is still general enough. Furthermore, the computations may be very cumbersome and the potential efficiency loss substantial. This remark applies as well to the estimation of semi- and non-parametric models which avoid many of the assumptions of the parametric models and will be discussed in the next section.

The alternative is to estimate the simple parametric models by ML and subject the results to extensive specification tests. This is most convenient when these procedures do not require the estimation of a more complicated alternative model. In the following we will discuss some of these methods, termed *diagnostics* by ENGLE (1984). Furthermore, some developments

will be presented for the more complicated situation, when a researcher has to choose one of two competing and possibly non-nested models, which are equally plausible from an economic point of view.

The seminal papers by NEWEY (1985a, b) and TAUCHEN (1985) provide a unified framework for the distribution theory of specification tests when a parametric model is tested against a parametric alternative. They observe that many test procedures are based on a criterion function indexed by a finite number of parameters which has zero expectation provided the hypothesis under test is true. If it is false the criterion function should be chosen so that its expectation with respect to the true data generating process is large. By the analogy principle the appropriate sample analogs, e.g. the means of the respective sample functions and estimates of their variances, are used to form quadratic forms which, given regularity conditions, have a central χ^2-distribution under the null hypothesis when N tends to infinity. Within this framework two important tasks remain: The choice of a criterion function satisfying the zero expectation requirement and regularity conditions which has a lot of power in the 'desired' direction. The 'desired' direction will typically be an alternative model which, if true, is particularly harmful for the interpretation of the results of the performed estimations. The second task is to find efficient estimators for the covariance matrix of the sample moments of the criterion to guarantee that the small sample distribution of the test statistic is close to the asymptotic distribution for a reasonable sample size.

Discussing all procedures which have been suggested in the literature to test various aspects of specific models is beyond our capabilities. Instead we classify the various procedures according to the way they define the respective alternatives for which particular power is desired. The first group consists of procedures specifying a full parametric model as an alternative, the second one specifies only particular moments of possible alternatives, and the third group leaves the alternative unspecified.

4.1 Completely Specified Model as Alternative

When a full parametric alternative is specified two basic cases have to be distinguished: nested and non-nested models. In the case of nested models the validity of the model under test implies implicit or explicit restrictions on the parameters of the alternative model. In the case of ML estimation, for example, the classical trinity of Lagrange Multiplier (LM) or Score tests, Likelihood ratio (LR) tests, and Wald tests can be applied (ENGLE, 1984) to check whether these restrictions are violated. There is an excellent introduction to this topic in

chapters 11-13 of the textbook by DAVIDSON and MACKINNON (1993). For other $\sqrt{N}$-consistent and asymptotic normal estimators generalised versions of the LM, Score and Wald tests (GOURIEROUX and MONFORT, 1989) can be applied. As a specification test the LM has certain advantages: Since it is based on checking whether the scores of the alternative model are significantly different from zero when evaluated at the estimates of the restricted model, an estimation of the alternative model is not necessary. This allows the choice of alternatives which are too complicated to yield sensible parameter estimates, but which can easily be evaluated under the null hypothesis. An example for this is the use of the Pearson-family of distributions as an alternative for the normal distribution (BERA et al., 1984). Furthermore, in cases when the likelihood function is misspecified but the resulting estimates are still consistent, generalised versions of the test still have the same asymptotic distribution under the null (WHITE, 1982) as opposed to the LR test. Compared to the Wald test which uses the estimates under the alternative, it is invariant (if its covariance matrix is not based on the empirical mean of the hessian) to different representations of the same null hypothesis, reparametrisations or one-to-one transformations of the parameter space (DAGENAIS and DUFOUR, 1991, LAFONTAINE and WHITE, 1986). Furthermore, there is evidence that the power properties of Wald-tests are badly approximated by the asymptotic local power function when the alternative is far away from the null (NELSON and SAVIN, 1990). The computation of the LM test statistic is simplified by regression based methods (ENGLE, 1984, DAVIDSON and MACKINNON, 1984) or the use of generalised or simulated residuals (CHESHER and IRISH, 1987, GOURIEROUX et al., 1987a, b). The LM statistic depends on a consistent estimate of the covariance matrix of the score under the null, which is based on the empirical mean of the hessian (H), the outer product of gradients (OPG), the expectation of the hessian conditional on explanatory variables (I), or a combination of OPG and either the inverse of H or I (WHITE, 1982). There is accumulating evidence that H, OPG and the combination of H and OPG should be avoided in applications, whenever possible. When H is used alone or in combination with OPG the LM test is not invariant to reparametrisations (DAGENAIS and DUFOUR, 1991). This means that the statistic may have different values for different parametrisations of exactly the same model, which is undesirable. Furthermore, H may not have full rank when evaluated under the null. The OPG version of the test is very popular because there is no need to compute second derivatives of the likelihood function, which may be complicated depending on the alternative model. However, there is accumulating evidence from Monte Carlo studies that these versions have excessive size under the null in small samples, e.g. they reject the null too often when it is correct (CHESHER and SPADY, 1991, DAVIDSON and MACKINNON, 1984, LECHNER, 1991).

Since these classical test procedures have a long history, many important aspects of them are now well understood, they are available for many different models and they become more common in empirical applications (BLUNDELL, 1987, BLUNDELL et al., 1993, LAISNEY et al., 1991, LEE and MADDALA, 1985). However, this approach also has drawbacks, for example there are important alternatives in which the estimated model cannot be nested. Furthermore, because of the specification of a particular parametric model, the tests may have little power in other directions which are of equal importance, in particular since these procedures are used as diagnostics for the estimated model.

So far we considered the case where the estimated model is parametrically nested in the alternative model used for the construction of the diagnostic. However, there may be interesting alternatives which do not have this property. Since the tests against non-nested alternatives (TNNA) have a very similar structure as the probabilistic tests for selection of possibly non-nested models (SNNM), they will be discussed jointly. Note that nevertheless the decision framework is quite different: A TNNA is used to subject the estimated model to a specification test against a non-nested alternative. The estimated model can be rejected or not rejected. A SNNM test is used to allow the data to decide which of two plausible models is the most 'correct' one. The classical TNNA for non-linear models is the one proposed by COX (1961, 1962). It is based on the difference of the maximum value of the log-likelihood function of different models and compares this difference with its expected value evaluated with the maintained conditional distribution of the estimated model. In the case of nested models this reduces to the conventional LR statistic. PESARAN (1987) investigates the power characteristics of this test and obtains some results similar to the classical tests for specific types of non-nested alternatives, which have certain overlaps. Using the fact that two linear models with different regressors can easily be nested in a general model, ZABEL (1993) investigates small sample properties of the Cox test compared to conventional tests, such as the Wald test (DAVIDSON and MACKINNON, 1981). In his results it is stated that the Cox test has superior power properties for local alternatives for most data generating processes considered. However, it seems that the sometimes difficult calculation of the expectations of the likelihood function of the alternative model with respect to the maintained distribution has prevented a widespread application of this test. The use of simulation techniques to evaluate this quantity as suggested by PESARAN and PESARAN (1993) may be a useful tool to overcome this difficulty. SMITH (1992) generalises the Cox test by suggesting a variant which is based on GMM estimation, and so overcomes the limitation that a full parametric density has to be specified to compute the test.

VUONG (1989) suggested likelihood ratio tests for model selection by testing whether two competing models are equally distant from the true model against the alternative that one model is significantly closer to the true one. A natural measure for this distance is the Kullback-Leibler contrast (see GOURIEROUX and MONFORT, 1989, for reference) which measures the distance between a given distribution and the true distribution. The proposed statistic has a clear interpretation in probabilistic terms and is not too difficult to compute. It has been applied for example by LAISNEY et al. (1991) and LECHNER (1991) for the case of testing Probit vs. Logit models. An approach which is similar has been suggested by POLLACK and WALES (1991). They compare adjusted likelihood ratio tests of each of the models under test with a model nesting both of them, without requiring an estimation in that general model. However, their approach does not allow the outcome that the distance of both models from the true model is not significantly different.

The theory of specification tests based on the encompassing principle (MIZON and RICHARD, 1986, HENDRY and RICHARD, 1989) provides a unifying framework for various tests against non-nested and nested alternatives. The test is based on the idea of checking whether salient features of one model can also be found in other models. More formally, the tests are based on the expectation of a statistic which has been computed in an alternative model, while the expectation of the statistic is computed w.r.t. the maintained model. It seams to be the major problem of this approach that the computation of the respective expectation may become very cumbersome in non-linear models.

4.2 Incompletely Specified Model as Alternative

Now, we turn to procedures which do not require the specification of a full parametric alternative model. Instead, they are based either on some population moments implied by the maintained model, but not used for its estimation, or some moments of an alternative model which parametrically nest moments of the estimated model and which are zero when the estimated model is the true one.

An approach taken by SMITH (1989) is to specify the alternative density as the product of the null density and an infinite series of orthogonal polynomials. This is a very general approach since in principle all smooth alternative densities can be generated in this way. In order to use LM test principles for a finite number of parameters, the expansion is truncated at some chosen degree and the significance of the remaining expansion terms is tested. Although this approach allows the construction of general specifications tests at least in the case of ML

estimation, its considerable complexity seems to have prevented, at the best of our knowledge, any empirical application of this procedure.

NEWEY (1985a) and TAUCHEN (1985) suggested a general principle for specification tests in models with explanatory variables based on functions which have zero expectation conditional on explanatory variables under the null model, and non-zero expectation in desired directions of departure from the model. By the law of iterated expectations unconditional expectations of products of these functions with functions of the explanatory variable have the same property and, given usual regularity conditions, can be used to form appropriate diagnostics. These tests have been termed conditional moments (CM) tests and nest all the procedures discussed so far. This framework unifies many different approaches and so allows us to obtain results concerning the properties of these tests, such as their power properties (BIERENS, 1990), for various estimation procedures, and to construct additional tests. The excellent surveys of PAGAN and VELLA (1989) and PAGAN and PAK (1993) give an account of the potential of this framework. WOOLDRIDGE (1991) discusses CM tests in models estimated by Pseudo ML (PML). The most important restriction of this approach is that the regularity conditions require a certain degree of smoothness and a finiteness of the parameter vector of the conditional moments under the alternative, so that a distribution theory for testing against many nonparametric and semiparametric alternatives is not provided. However, a very recent paper by WHANG and ANDREWS (1993) provides considerable generalisations in that direction.

A CM test which has appeared as a useful device for testing models against general misspecification is the information matrix test (*IM*) introduced by WHITE (1982). The IM test is based on checking the validity of the fundamental information matrix identity of ML estimation, namely that the conditional expectation of the outer product of the gradients equals the conditional expectation of minus the hessian. The only specified feature of the alternative model is that this identity does not hold. Although the basic version of the test is confined to models estimated by ML, LECHNER (1992) showed that certain other estimators imply similar conditions which could be used to form diagnostics in the same spirit as the IM test. Recent Monte Carlo studies (CHESHER and SPADY, 1991, LECHNER, 1991, ORME, 1990) showed that accurate estimation of the covariance matrix of the indicators is crucial for getting small sample distributions which close to the asymptotic distribution used for inference. OPG versions which avoid the need to compute third derivatives of the log-likelihood function (CHESHER, 1984, LANCASTER, 1984), and to some extent also White's original version, overreject drastically under the null. ORME (1988, 1990) suggested to use the asymptotic efficient covariance estimator of the test indicators by exploiting the whole

structure of the model under the null. This version of the test has a distribution under the null which is close to the asymptotic one even for moderate samples, at least for the probit model, but is more cumbersome to compute.

The diagnostics suggested by ANDREWS (1988a, b) and KLEIN (1993) are based on the predictions of the estimated model. KLEIN'S (1993) approach for single index models estimated by maximum likelihood is based on comparing the predictions of the model, e.g. the expectation conditional on a specific interval of the index with the observed sample analogue or with a prediction of a semiparametric model, evaluated at the parameters of the null model. The regions necessary to define the test are solely based on the single index. KLEIN (1993) derives the asymptotic distribution under the null and local power properties for these statistics and gives conditions on the type of semiparametric model allowed. The latter is based on the type of partially nonparametric density estimation techniques proposed in KLEIN and SPADY (1993). This approach should allow the identification of regions in which the estimated model has a good or bad predictive performance, which is important for policy analysis. Building on previous work by HECKMAN (1984), ANDREWS (1988a, b) generalises the classical Pearson test to general types of econometric models. This approach is more general than the previous one in the sense that it allows any $\sqrt{N}$-consistent and asymptotically normal estimator and more general types of cell building necessary to define the test. Furthermore, it is not restricted to single index models. The test is based on partitioning the space spanned by the endogenous and exogenous variables in disjunct cells. The cells may depend on estimated coefficients. The test is computed by comparing the conditionally expected number of realisations with the observed number of realisations in each cell. ANDREWS (1988a) derives the asymptotic distribution under the null and local alternatives. The big advantage of this test is the flexibility in forming the cells. This flexibility can be used to obtain power in directions which cannot be addressed with other tests. Applications of this test can be found for example in BLUNDELL et al. (1993), LAISNEY et al. (1991) and HOROWITZ and NEUMANN (1989).

5 Estimation in Semi- and Nonparametric Models

The interest in the estimation of semiparametric and nonparametric models (see the recent special issue of the *Journal of Econometrics*, HÄRDLE and MANSKI, 1993) stems from the intention to diminish the influence of stochastic assumptions, which are often not derived from the economic model, on the estimation results. In terms of the example of the binary choice model discussed so far, different economic models, e.g. utility functions, lead to

different forms of the function capturing the influence of the regressions, $g(x_i,\beta)$. However, the choice of assumptions for the error term, except that it has mean zero conditional on regressors, could very rarely be motivated by some behavioral hypothesis. Normality and independence of the regressors, which in turn implies conditional homoscedasticity, lead to the conventional probit model. They are chosen merely for computational convenience.

Typical semiparametric models use a parametric specification of the function $g(\cdot)$ and relax some assumptions on the error term. In the last decade different sets of assumptions, and hence different estimators, have been proposed for cross-section data. An excellent survey by HOROWITZ (1993a) shows the relationship between different plausible (non-nested) sets of assumptions and their importance with respect to identification and asymptotic properties of the estimators designed for binary choice model. For any details the reader is referred to that source. Although it seems that the low informational content of the binary choice model posed a challenge that led to the development of numerous different estimators, these types of estimators have also been proposed for various types of cross-section tobit and selection models (NEWEY et al., 1990, POWELL, 1984, 1986, 1989, POWELL et al., 1989, ROBINSON, 1988, MOON, 1989). Some extensions to panel data models appeared in the recent literature. MANSKI (1987) suggested an estimator for the binary choice model which is based on the principles of maximum score estimation (MANSKI, 1975, 1985). The seminonparametric estimator for the binary choice model (GABLER et al., 1993) has been extended to random effects panel data by LAISNEY et al. (1992), and HONORÉ (1992, 1993) proposed estimators for fixed effect tobit models which are similar to those suggested by POWELL (1984, 1986) for cross sections.

Nevertheless there is a price to pay for the robustness offered by semiparametric methods. Many semiparametric estimators share some of the following drawbacks: Their computation is difficult because of objective functions which may not be differentiable or exhibit many local extremes. Recently there is progress toward reducing the impact of that problem (see GOFFE et al., 1994, PINSKE, 1993 and VEALL, 1990), it is still a major obstacle in applications. Furthermore, some estimators exhibit lower convergence rates than $\sqrt{N}$, have non-normal distributions and may require bootstrapping in order to obtain the distrubution of the coefficient estimates. Furthermore, the ability of the model to be used for policy simulations may be reduced due to the unspecified error distribution. But it should be noted that this is not true for all these estimators. The optimal choice of a particular estimator for a specific application depends very much on the type of results desired.

Recently the use of semiparametric estimators in applications increased. Examples are labor supply studies (GABLER et al., 1993, GERFIN, 1993, NEWEY et al., 1990) and transport choice problems (HOROWITZ, 1993b) and innovation activity (LAISNEY et al., 1992) among others.

Whereas in semiparametric estimation it is still the aim to identify and estimate 'parameters', the nonparametric approach is based on the estimation of the functional $g(x,\cdot)$ itself. Different ways to do this are discussed for example in DELGADO and ROBINSON (1992) or HÄRDLE (1990). These methods may be applied either to identify and estimate non-linear relationships between two, or more variables, or as an intermediate step in multistep parametric or semiparametric estimation problems. The former is demonstrated for example by BERTSCHEK and ENTORF (1993) regarding the relation of firm size to innovative activity or by HÄRDLE et al. (1991) for the estimation of Engel-curves. Examples for the latter are the semiparametric estimators of KLEIN and SPADY (1993) and HOROWITZ (1992), the estimation of optimal instruments in CM-estimation (NEWEY, 1990, 1993), as already discussed in section 3, or the non-parametric estimation of expectations in a dynamic discrete choice problem (AHN and MANSKI, 1993, MANSKI, 1991, 1993a).

6 Topics for Future Research

In the previous sections we sketched some recent developments of econometric methods which provided useful tools for applied microeconometric work. However, they will certainly not represent the end of a phase of rapid developments in that area. In this section we suggest some fields for future research.

The use of simulation methods for ML or GMM estimation of structural models will be more common in applications. Many different ways to perform these simulations have been suggested so far, but it is still difficult to decide which particular simulator is the best in terms of computational convienience, efficiency, consistency and good small sample performance given a specific model and a dataset. More comparisons of the different methods based on Monte Carlo studies and applications will be useful. Similar issues related to small sample performance arise in the estimation with conditional moment restrictions. It may, or may not, be worth the computational inconvience in order to obtain gains in asymptotic efficiency by using the asymptotically optimal instruments, and (or) base the estimation on more complicated second or higher order conditional moments. These small sample considerations are also important in semiparametrics. In particular the questions arise whether the data at hand would really support such flexible models or whether the gain of confidence in the

consistency of the estimates is accompanied by a large small sample mean square error which renders the results worthless for any policy analysis.

With respect to specification testing, many ways have been suggested to test parametrics models estimated by ML or GMM, but little is known in cases when simulation estimation or semiparametric estimation is performed.

Artificial neural networks appeared to be useful tools for prediction purposes in financial econometrics, because they can capture highly nonlinear relations between different variables of interest. Since these non-linear relations between variables, and the ability to predict outcomes either in the time or individual dimension are also important in micro-econometrics, these methods have a potential application in this field too. However, in order to use them we need to understand the statistical properties of the "black boxes" employed. The papers in the book by WHITE (1992) show that neural networks can be analysed like semi- and nonparametric, or flexible parametric methods so that the appropriate distribution theories could be applied. However, there is still a long way to go in order to use them in applied microeconometrics.

References

Ahn, H. and Manski, C.F. (1993): "Distribution Theory for the Analysis of Binary Choice ander Uncertainty with Non-parametric Estimation of Expectations", *Journal of Econometrics*, 56, 291-321.

Amemiya, T. (1981): "Qualitative Response Models: a Survey", *Journal of Economic Literature*, 19, 1483-1536.

Amemiya, T. (1984): "Tobit Models. A Survey", *Journal of Econometrics*, 24, 3-61.

Amemiya, T. (1986): *Advanced Econometrics*, Oxford: Basil Blackwell.

Andrews, D.W.K. (1994): "Asymptotics for Semiparametric Models via Stochastic Equicontinuity", *Econometrica*, 62, 43-72.

Andrews, D.W.K. (1988a): "Chi-square Diagnostic Tests for Econometric Models: Theory", *Econometrica*, 56, 1419-1453.

Andrews, D.W.K. (1988b): "Chi-square Diagnostic Tests for Econometric Models: Introduction and Application", *Journal of Econometrics*, 37, 135-156.

Andrews, D.W.K. (1991): "Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Regression Models", *Econometrica*, 59, 307-345.

Andrews, R.W., Berger, J.O. and Smith, M.H. (1993): "Bayesian Estimation of Manufacturing Effects in a Fuel Economy Model", *Journal of Applied Econometrics*, 8, S5-S18.

Arabmazar, A. and Schmidt, P. (1982): "An Investigation of the Robustness of the Tobit Estimator to Non-Normality", *Econometrica*, 50, 1055-1063.

Arellano, M. and Meghir, C. (1992): "Female Labour Supply and On the Job Search: An Empirical Model Estimated Using Complementary Data Sources", *Review of Economic Studies*, 59, 537-557.

Arminger, G. and Müller, F. (1990): *Lineare Modelle zur Analyse von Paneldaten*, Opladen: Westdeutscher Verlag.

Avery, R., Hansen, L. and Hotz, V. (1983): "Multiperiod Probit Models and Orthogonality Condition Estimation", *International Economic Review*, 24, 21-35.

Balestra, P. and Nerlove, M. (1966): "Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas", Econometrica, 34, 585-612.

Bera, A., Jarque, C. and Lee, C.F. (1984): "Testing the Normality Assumption in Limited Dependent Variable Models", *International Economic Review*, 25, 563-578.

Bertschek, I., and Entorf, H. (1993): "On Nonparametric Estimation of the Schumpeterian Link between Innovation and Firm Size: Evidence from Belgium, France and Germany", *Discussion Paper* 9318, Humboldt-Universität zu Berlin.

Bierens, H.J. (1990): "A Consistent Conditional Moments Test of Functional Form", *Econometrica*, 59, 1443-1458.

Blundell, R.W. (ed., 1987): "Specification Testing in Limited and Discrete Dependent Variable Models", *Journal of Econometrics*, 34, Annals.

Blundell, R.W. and Smith, R.J. (1993): "Simultaneous Microeconometric Models with Censored or Qualitative Dependent Variables", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 5, Amsterdam: North-Holland.

Blundell, R.W., Laisney, F. and Lechner, M. (1993): "Alternative Interpretations of Hours Information in an Econometric Model of Labour Supply", *Empirical Economics*, 18, 393-415.

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

Breitung, J. and Lechner, M. (1994): "GMM-Estimation of Non-linear Models on Panel Data", *mimeo*.

Brown, B.W., Monfort, A. and van Dijk, H.K. (eds., 1993): "Econometric Inference Using Simulation Techniques", Supplement to the *Journal of Applied Econometrics*, 8.

Butler, I.S. and Mofitt, R. (1982): "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model", *Econometrica*, 50, 761-764.

Chamberlain, G. (1980): "Analysis of Covariance with Qualitative Data", *Review of Economic Studies*, 47, 225-238.

Chamberlain, G. (1984): "Panel Data", in Griliches, Z. and Intriligator, M.D. (eds.), *Handbook of Econometrics*, Vol. II, Ch. 22, Amsterdam: North-Holland.

Chamberlain, G. (1986): "Asymptotic Efficiency in Semi-parametric Models with Censoring", *Journal of Econometrics*, 33, 189-218.

Chamberlain, G. (1987): "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions", *Journal of Econometrics*, 34, 305-334.

Chamberlain, G. (1992): "Efficiency Bounds for Semiparametric Regression", *Econometrica*, 60, 567-596.

Chesher, A. (1984): "Testing for Neglected Heterogeneity", *Econometrica*, 52, 865-872.

Chesher, A. and Irish, M. (1987): "Residual Analysis in the Grouped and Censored Normal Linear Model", *Journal of Econometrics*, 34, 33-61.

Chesher, A. and Spady, R. (1991): "Asymptotic Expansions of the Information Matrix Test Statistic", *Econometrica*, 59, 787-815.

Cosslett, S.R. (1987): "Efficiency Bounds for Distribution-free Estimators of the Binary Choice and Censored Regression Models", *Econometrica*, 55, 559-585.

Cosslett, S.R. (1993): "Estimation from Endogenously Stratified Samples", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 1, Amsterdam: North-Holland.

Cox, D.R. (1961): "Tests of Separate Families of Hypotheses", *Proceedings of the 4th Berkley Symposium*, Vol. 1, 105-123, Berkley: University of California Press.

Cox, D.R. (1962): "Further Results on Tests of Separate Families of Hypotheses", *Journal of the Royal Statistical Society Series B*, 24, 406-424.

Dagenais, M.G. and Dufour, J.M. (1991): "Invariance, Nonlinear Models, and Asymptotic Tests", *Econometrica*, 59, 1601-1615.

Davidson, R. and MacKinnon, J.G. (1981): "Several Tests for Model Specification in the Presence of Alternative Hypotheses", *Econometrica*, 49, 781-793.

Davidson, R. and MacKinnon, J.G. (1984): "Convenient Specification Tests for Logit and Probit Models", *Journal of Econometrics*, 25, 241-262.

Davidson, R. and MacKinnon, J.G. (1993): *Estimation and Inference in Econometrics*, Oxford: Oxford University Press.

Delgado, M.A. and Robinson, P.M. (1992): "Nonparametric and Semiparametric Methods for Economic Research", *Journal of Economic Surveys*, 6, 200-249.

Eckstein, Z. and Wolpin, K.I. (1989): "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models - A Survey", *Journal of Human Resources*, 24, 562-598.

Engle, R.F. (1984): "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics", in Griliches, Z. and Intriligator, M.D. (eds.), *Handbook of Econometrics*, Vol. II, Amsterdam: North-Holland.

Florens, J.P. (1990): " Models and Inference for Duration Data: A Survey", in Florens, J.P., Ivaldi, M., Laffont, J.C. and Laisney, F. (eds.), *Microeconometrics: Surveys and Applications*, Oxford: Basil Blackwell, 55-83.

Florens, J.P. and Mouchart, M. (1993): "Baysian Testing and Testing Baysians", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 11, Amsterdam: North-Holland.

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

Gallant, A.R. and Nychka, D.N. (1987): "Semi-Nonparametric Maximum Likelihood Estimation", *Econometrica*, 55, 363-390.

Gerfin, M. (1993): "Parametric and Semiparametric Estimation of the Binary Response Model of Labour Market Participation", *mimeo*.

Geweke, J. (1991): "Efficient Simulation From Multivariate Normal and Student t-distributions Subject to Linear Constraints", in Keramidas, E.M. (ed.), *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, Fairfax: Interface Foundation, 571-578.

Geweke, J. (1993): "Bayesian Treatment of the Independent Student t-linear Model", *Journal of Applied Econometrics*, 8, S19-S40.

Goffe, W.L., Ferrier, G.D. and Rogers, J. (1994): "Global Optimisation of Statistical Functions with Simulated Annealing", *Journal of Econometrics*, 60, 65-99.

Gourieroux, C. and Monfort, A. (1989): *Statistique et Modeles Econometriques*, Paris: Economica.

Gourieroux, C. and Monfort, A. (1991): "Simulation Based Inference in Models with Heterogeneity", *Annales d'Économie et de Statistique*, 20 / 21, 69-107.

Gourieroux, C. and Monfort, A. (1993a): "Simulation-based Inference: A Survey with Special Reference to Panel Data Models", *Journal of Econometrics*, 59, 5-33.

Gourieroux, C. and Monfort, A. (1993b): "Indirect Inference", *Journal of Applied Econometrics*, 8, S85-S118.

Gourieroux, C. and Monfort, A. (1993c): "Pseudo-Likelihood Methods", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 12, Amsterdam: North-Holland.

Gourieroux, C., Monfort, A., Renault, E. and Trognon, A. (1987a): "Generalised Residuals", *Journal of Econometrics*, 34, 5-32.

Gourieroux, C., Monfort, A., Renault, E. and Trognon, A. (1987b): "Simulated Residuals", *Journal of Econometrics*, 34, 201-252.

Griliches, Z. (1986): "Economic Data Issues", in Griliches, Z. and Intriligator, M.D. (eds.), *Handbook of Econometrics*, Vol. III, Ch. 25, Amsterdam: North-Holland.

Guilkey, D.K. and Murphy, J.L. (1993): "Estimation and Testing in the Random Effects Probit Model", *Journal of Econometrics*, 59, 301-317.

Gumra, S. and Trivedi, P.K. (1992): "Recent Developments in Models of Event Counts: A Survey", *Working Paper*, Indiana-University.

Hajivassiliou, V.A. and McFadden, D. (1990): "The Method of Simulated Scores, with Application to Models of External Debt Crises", *Cowles Foundation Discussion Paper* 967, Yale University.

Hajivassiliou, V.A. (1993): "Simulation Estimation Methods for Limited Dependent Variable Models", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 19, Amsterdam: North-Holland.

Hamermesh, D.S. (1990): "Data Difficulties in Labor Economics", in Berndt, E.R. and Triplett, J.E. (eds.): *Fifty Years of Economic Measurement*, Chicago: The University of Chicago Press.

Hanau, A. (1928): "Die Prognose der Schweinepreise", *Vierteljahreshefte zur Konjunkturforschung*, Sonderheft 7, Berlin.

Hansen, L.P. (1982): "Large Sample Properties of Generalized Methods of Moments Estimators", *Econometrica*, 50, 1029-1055.

Härdle, W. (1990): *Applied Nonparametric Regression*, Cambridge: CUP.

Härdle, W., and Manski, C. (ed. 1993): "Nonparametric and Semiparametric Approaches in Discrete Response Analysis", *Journal of Econometrics*, 58.

Härdle, W., Hildenbrand, W. and Jerison, M. (1991): "Empirical Evidence on the Law of Demand", *Econometrica*, 59, 1525-1549.

Heckman, J.J. (1981): "Statistical Models for Discrete Panel Data", in Manski, C.F. and McFadden, D. (eds.), *Structural Analysis of Discrete Data,* Cambridge: MIT Press.

Heckman, J.J. (1984): "The Chi-square Goodness of Fit Statistic for Models with Parameters Estimated from Microdata", *Econometrica*, 52, 1543-1547.

Hendry, D.F. and Richard, J.-F. (1989): "Recent Developments in the Theory of Encompassing", in Cornet, B. and Tulkens, H. (eds.): *Contributions to Operations Research and Economics: The Twentieth Anniversary of CORE,* Cambridge, MA: MIT-Press.

Honoré, B. (1992): "Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects", *Econometrica*, 60, 533-565.

Honoré, B. (1993): "Orthogonality Conditions for Tobit Models with Fixed Effects and Lagged Dependent Variables", *Journal of Econometrics*, 59, 35-61.

Horowitz, J.L. (1992): "A Smoothed Maximum Score Estimator for the Binary Response Model", *Econometrica*, 60, 505-531.

Horowitz, J.L. (1993a): "Semiparametric and Nonparametric Estimation of Quantal Response Models", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 2, Amsterdam: North-Holland.

Horowitz J.L. (1993b): "Semiparametric Estimation of a Work-trip Mode Choice Model", *Journal of Econometrics*, 58, 49-70.

Horowitz, J.L. and Neumann, G.R. (1989): "Specification Testing in Censored Regression Models: Parametric and Semiparametric Methods", *Journal of Applied Econometrics*, 4, S61-S86.

Hotz, V.J. and Miller, R.A. (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models", *Review of Economic Studies*, 60, 497-529.

Hsiao, C. (1986): *The Analysis of Panel Data*, Cambridge: CUP.

Hsiao, C. (1992): "Logit and Probit Models", in Mátyás, L. and Sevestre, P. (eds.), *The Econometrics of Panel Data,* Ch. 11, Dordrecht: Kluwer.

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

Imbens, G.W. (1992): "An Efficient Method of Moments Estimator for Discrete Choice Models with Choice Based Sampling", *Econometrica*, 60, 1187-1214.

Imbens, G.W. and Lancaster, T. (1991): "Combining Micro and Macro Data in Microeconometric Models", *Discussion Paper* 1578, Harvard University.

Joeng, J. and Maddala, G.S. (1993): "A Perspective on Application of Bootstrap Methods in Econometrics", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 21, Amsterdam: North-Holland.

Keane, M.P. (1993): "Simulation Estimation for Panel Data Models with Limited Dependent Variables", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 20, Amsterdam: North-Holland.

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

Kleibergen, F. and van Dijk, H.K. (1993): "Non-stationarity in GARCH Models: A Bayesian Analysis", *Journal of Applied Econometrics*, 8, S41-S61.

Klein, R.W. (1993): "Specification Tests for Binary Choice Models Based on Index Quantiles", *Journal of Econometrics*, 59, 343-375.

Klein, R.W. and Spady, R. (1993): "An Efficient Semiparametric Estimator for Binary Response Models" *Econometrica*, 61, 387-421.

Kodde, D.A., Palm, F.C. and Pfann, G.A. (1990): "Asymptotic Least-Squares Estimation Efficiency Considerations and Applications", *Journal of Applied Econometrics*, 5, 229-243.

Lafontaine, F. and White, K.J. (1986): "Obtaining any Wald Statistic You Want", *Economics Letters*, 21, 35-40.

Laisney, F. and Lechner, M. (1993): "Combing Panel Data and Macro Information: An Application to the Estimation of a Participation Model", *Discussion Paper*, Zentrum für Europäische Wirtschaftsforschung, 93-23.

Laisney, F., Lechner, M. and Pohlmeier, W. (1992): "Semi-Nonparametric Estimation of Binary Choice Models Using Panel Data", *Recherches Économiques de Louvain*, 58, 329-343.

Laisney, F., Lechner, M. and Strøm, S. (1991): "Lessons from Specification Tests for a Labour Supply Model", *Annales d'Économie et de Statistique*, 20/21, 193-217.

Lancaster, T. (1984): "The Covariance Matrix of the Information Matrix Test", *Econometrica*, 52, 1051-1053.

Lancaster, T. (1990): *The Econometric Analysis of Transition Data*, Cambridge: CUP.

Laroque, G. and Salanié, B. (1993): "Simulation-based Estimation of Models with Lagged Latent Variables", *Journal of Applied Econometrics*, 8, S119-S133.

Laroque, G. and Salinié, B. (1994): "Estimating the Canonical Disequilibrium Model: Asymptotic Theory and Finite Sample Properties", forthcoming in *Journal of Econometrics*.

Lechner, M. (1991): "Testing Logit Models in Practice", *Empirical Economics*, 16, 177-198.

Lechner, M. (1992): "Some Specification Tests for Static Limited Dependent Variable Models Estimated on Panel Data", *Beiträge zur Angewandten Wirtschaftsforschung*, 474-92, University of Mannheim.

Lechner, M. (1993a): "The Dynamics of Self-Employment in East Germany - An Empirical Analysis Using Panel Data and Allowing for State-Dependence and Endogenous Attrition", *Discussion Paper*, Zentrum für Europäische Wirtschaftsforschung, 93-24.

Lechner, M. (1993b): "Estimation of Limited Dependent Variable Habit Persistence Models on Panel Data with an Application to the Dynamics of Self-employment in the Former East Germany", in Bunzel, H., Jensen, P. and Westergård-Nielson, N. (eds.), *Panel Data and Labour Market Dynamics*, Amsterdam: North-Holland, 263-283.

Lee, L.F. and Maddala, G.S. (1985): "The Common Structure of Tests for Selectivity Bias, Serial Correlation and Non-Normality in the Tobit Model", *International Economic Review*, 26, 1-20.

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

Maddala, G.S. (1983): *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge: CUP.

Maddala, G.S. (1986): "Disequilibrium, Self-selection and Switching Models", in Griliches, Z. and Intriligator, M.D. (eds.), *Handbook of Econometrics*, Vol. III, Ch. 28, Amsterdam: North-Holland.

Maddala, G.S. (1987): "Limited Dependent Variable Models Using Panel Data", *Journal of Human Resources*, 22, 307-338.

Maddala, G.S. (1993): "Estimation of Limited Dependent Variable Models Under Rational Expectations", in Maddala, G.S. and Vinod H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 7, Amsterdam: North-Holland.

Maddala, G.S., Rao, C.R. and Vinod, H.D. (1993): *Handbook of Statistics, Vol. 11: Econometrics*, Amsterdam: North-Holland.

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

Manski, C.F. (1985): "Semiparametric Analysis of Discrete Response - Asymptotic Properties of the Maximum Score Estimator", *Journal of Econometrics*, 27, 313-333.

Manski, C.F. (1987): "Semiparametric Analysis of Random Effects Linear Models from Binary Panel Data", *Econometrica*, 55, 357-362.

Manski, C.F. (1988a): *Analog Estimation Methods in Econometrics*, New York: Chapman and Hall.

Manski, C.F. (1988b): "Identification of Binary Response Models", *Journal of the American Statistical Association*, 83, 729-738.

Manski, C.F. (1989): "The Anatomy of the Selection Problem", *Journal of Human Ressources*, 24, 343-360.

Manski, C.F. (1991): "Nonparametric Estimation of Expectations in the Analysis of Discrete Choice Under Uncertainty", in Barnett, W.A., Powell, J. and Tauchen, G. (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge: CUP.

Manski, C.F. (1993a): "Dynamic Choice in Social Settings - Learning from the Experiences of Others", *Journal of Econometrics*, 58, 121-136.

Manski, C.F. and Thompson, T.S. (1986): "Operational Characteristics of Maximum Score Estimation", *Journal of Econometrics*, 32, 85-108.

Manski, C.F. and Thompson, T.S. (1989): "Estimation of Best Predictors of Binary Response", *Journal of Econometrics*, 40, 97-123.

Mariano, R.S. and Brown, B.W. (1993): "Stochastic Simulations for Inference in Nonlinear Errors-in-Variables Models", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 22, Amsterdam: North-Holland.

McFadden, D. (1981): "Econometric Models of Probabilistic Choice", in Manski, C.F. and McFadden, D., *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge: MIT-Press.

McFadden, D. (1984): "Econometric Analysis of Qualitative Response Models", in Griliches, Z. and Intriligator, M.D. (eds.), *Handbook of Econometrics*, Vol. II, Ch. 24, Amsterdam: North-Holland.

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

Meyer, J.R. and Küh, E. (1957): "*The Investment Decision - An Empirical Study*",Cambridge: Harvard University Press.

Mincer, J. (1958): "Investment in Human and Personal Income Distribution", *Journal of Political Economy*, 66, 281-302.

Mincer, J. (1974): *Schooling, Experience and Earnings*, New York: NBER.

Mizon, G. and Richard, J.-F. (1986): "The Encompassing Principle and its Application to Non-Nested Hypotheses", *Econometrica*, 54, 657-678.

Mofitt, R. (1993): "Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross Sections", *Journal of Econometrics*, 59, 99-123.

Moon, C.G. (1989): "A Monte Carlo Comparison of Semiparametric Tobit Estimators", *Journal of Applied Econometrics*, 4, 361-382.

Muthén, B. (1987): "LISCOMP-Analysis of Linear Equations Using a Comprehensive Measurement Model", *Scientific Software*, Mooresville.

Nelson, C.R. and Savin, N.E. (1990): "The Danger of Extrapolating Asymptotic Local Power", *Econometrica*, 58, 977-981.

Nerlove, M. (1971): "Further Evidence on the Estimation of Dynamic Economic Relations From a Time Series of Cross Sections", *Econometrica*, 39, 359-383.

Newey, W.K. (1985a): "Maximum Likelihood Specification Testing and Conditional Moment Tests", *Econometrica*, 53, 1047-1070.

Newey, W.K. (1985b): "Generalised Method of Moments Specification Testing", *Journal of Econometrics*, 29, 229-256.

Newey, W.K. (1990a): "Efficient Instrumental Variables Estimation of Nonlinear Models", *Econometrica*, 59, 809-837.

Newey, W.K. (1990b): "Semiparametric Efficiency Bounds", *Journal of Applied Econometrics*, 5, 99-135.

Newey, W.K. (1991): "Uniform Convergence in Probability and Stochastic Equicontinuity", *Econometrica*, 59, 1161-1167.

Newey, W.K. (1993): "Efficient Estimation of Models with Conditional Moment Restrictions", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 16, Amsterdam: North-Holland.

Newey, W.K. and Powell, J.L. (1993): "Efficiency Bounds for Some Semiparametric Selection Models", *Journal of Econometrics*, 58, 169-184.

Newey, W.K., Powell, J.L., and Walker, J.R. (1990): "Semiparametric Estimation of Selection Models: Some Empirical Results", *American Economic Review*, Papers and Proceedings, 324-328.

Nijman, T. and Verbeek, M. (1992): "Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function", *Journal of Applied Econometrics*, 7, 243-257.

Orme, C. (1988): "The Calculation of the Information Matrix Test for Binary Data Models", *The Manchaster School*, 56, 370-376.

Orme, C. (1990): "The Small Sample Performance of the Information Matrix Test for Binary Data Models", *Journal of Econometrics*, 46, 309-331.

Pagan, A.R. and Vella, F. (1989): "Diagnostic Tests for Models Based on Individual Data", *Journal of Applied Econometrics*, 4, 529-559.

Pagan, A.R. and Pak, Y. (1993): "Testing for Heteroskedasticity", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 18, Amsterdam: North-Holland.

Pakes, A. and Pollard, D. (1989): "Simulation and the Asymptotics of Optimisation Estimators", *Econometrica*, 57, 1027-1057.

Pesaran, M.H. (1987): "Global and Partial Non-Nested Hypotheses and Asymptotic Local Power", *Econometric Theory*, 3, 69-97.

Pesaran, M.H. and Pesaran, B. (1993): "A Simulation Approach to the Problem of Computing Cox's Statistic for Testing Nonnested Models", *Journal of Econometrics*, 57, 377-392.

Pinske, C.A.P. (1993): "On the Computation of Semiparametric Estimates in Limited Dependent Variable Models", *Journal of Econometrics*, 58, 185-205.

Pohlmeier, W. (1994): "Count Data Models Estimated on Panel Data", *this volume.*

Poirier, R.J. (1980): "A Lagrange Multiplier Test for Skewness in Binary Logit Models", *Economics Letters*, 5, 141-143.

Pollak, R.A. and Wales, T.J. (1991): "The Likelihood Dominance Criterion", *Journal of Econometrics*, 47, 227-242.

Pötscher, B.M. and Prucha, I.R. (1994): "Generic Uniform Convergence and Equicontinuity Concepts for Random Functions: An Exploration of the Basic Structure", *Journal of Econometrics*, 60, 23-63.

Powell, J.L. (1984): "Least Absolute Deviations Estimation for the Censored Regression Model", *Journal of Econometrics*, 25, 303-325.

Powell, J.L. (1986): "Symmetrically Trimmed Least Squares Estimation for Tobit Models", *Econometrica*, 54, 1435-1460.

Powell, J.L. (1989): "Semiparametric Estimation of Censored Selection Models", University of Wisconsin, *mimeo*.

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

Pudney, S. (1989): *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes*, Oxford: Basil Blackwell.

Raj, B. and Baltagi, B.H. (eds., 1992): *Panel Data Analysis*, Heidelberg: Physica.

Robinson, P.M. (1988): "Semiparametric Econometrics: A Survey", *Journal of Applied Econometrics*, 3, 35-51.

Ronning, G. (1991): *Mikroökonometrie*, Berlin: Springer.

Rust, J.P. (1991): "Estimation of Dynamic Structural Models: Problems and Prospects. Part I: Discrete Decision Process", *SSRI Workshop Series No. 9106*, University of Wisconsin-Madison, presented at the 6th World Congress of the Econometric Society, Barcelona, August, 1990.

Rust, J.P. (1994): "Structural Estimation of Markov Decision Processes", forthcoming in Engle, R. and McFadden, D. (eds.), *Handbook of Econometrics*, Vol. IV, Amsterdam: North-Holland.

Ruud, P.A. (1983): "Sufficient Conditions for the Consistency of Maximum Likelihood Estimation Despite Misspecification of Distribution", *Econometrica*, 51, 225-228.

Ruud, P.A. (1986): "Consistent Estimation of Limited Dependent Variable Models Despite Misspecification of Distribution", *Journal of Econometrics*, 32, 157-187.

Shephard, N. (1993): "Fitting Non-Linear Time-Series Models with Applications to Stochastic Variance Models", *Journal of Applied Econometrics*, 8, S135-S152.

Smith, A.A. (1993): "Estimating Non-Linear Time-Series Models Using Simulated Vector Autoregressions", *Journal of Applied Econometrics*, S63-S84.

Smith, R.J. (1989): "On the Use of Distributional Mis-Specification Checks in Limited Dependent Variable Models", *The Economic Journal*, 99, 178-192.

Smith, R.J. (1992): "Non-Nested Tests for Competing Models Estimated by Generalized Method of Moments", *Econometrica*, 60, 973-980.

Stern, S. (1992): "A Method of Smoothing Simulated Moments of Discrete Probabilities in Multinomial Probit Models", Econometrica, 60, 943-952.

Stone R. (1954): "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand", *Economic Journal*, 64, 511-527.

Tauchen, T. (1985): "Diagnostic Testing and Evaluation of Maximum Likelihood Methods", *Journal of Econometrics*, 30, 415-443.

Thomas, J.M. (1993): "On Testing the Logistic Assumption in Binary Dependent Variable Models", *Empirical Economics*, 18, 381-392.

Thompson, T.S. (1993): "Some Efficiency Bounds for Semiparametric Discrete Choice Models", *Journal of Econometrics*, 58, 257-274.

Tobin, J. (1958): "Estimation of Relationships for Limited Dependent Variables", *Econometrica*, 26, 24-36.

Ullah, A. and Vinod, H.D. (1993): "General Nonparametric Regression Estimation and Testing in Econometrics", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 4, Amsterdam: North-Holland.

Veall, M.R. (1990): "Testing for a Global Maximum in an Econometric Context", *Econometrica*, 58, 1459-1465.

Verbeek, M. (1992): "Pseudo Panel Data", in Mátyás, L. and Sevestre, P. (eds.), *The Econometrics of Panel Data*, Ch. 14, Dordrecht: Kluwer.

Verbeek, M. and Nijman, T. (1993): "Minimum MSE estimation of a Regression Model with Fixed Effects From a Series of Cross Sections", *Journal of Econometrics*, 59, 125-136.

Vinod, H.D. (1993): "Bootstrap Methods: Applications in Econometrics", in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Statistics, Vol. 11: Econometrics*, Ch. 23, Amsterdam: North-Holland.

Vuong, Q.H. (1989): "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses", *Econometrica*, 57, 307-333.

Whang, Y. and Andrews, D.W.K. (1993): "Tests of Specification for Parametric and Semiparametric Models", *Journal of Econometrics*, 57, 277-318.

White, H. (1982): "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, 50, 1-25.

White, H. (1983): "Corrigendum", *Econometrica*, 51, 513.

White, H. (ed., 1992): *Artificial Neural Networks*, Cambridge: Basil Blackwell.

Wooldridge, J.M. (1991): "Specification Testing and Quasi-Maximum Likelihood Estimation", *Journal of Econometrics*, 48, 29-55.

Zabel, J.E. (1993): "A Comparison of Nonnested Tests for Misspecified Models Using the Method of Approximate Slopes", *Journal of Econometrics*, 57, 205-232.