Diagnostic Tests for Models Based on Individual Data: A Survey

Pagan, Adrian and Frank Vella

Working Paper No. 162
October 1988

Diagnostic Tests for Models Based on Individual Data: A Survey*

Adrian Pagan and Frank Vella

University of Rochester

June 1988

Revised, September 1988

1. Introduction

The last twenty years have witnessed a movement in econometric work away from time series data to that based upon "individuals", where this term should be interpreted broadly as a set of data for which there is no natural ordering. Such a development is not surprising given the traditional focus of economic theory upon the behavior of individual agents, but it is also a concomitant of the emergence of large scale survey data that may be easily handled with modern computational and storage devices.

When investigators first came to analyse these data it became apparent that the regression model, which had been the cornerstone of econometrics almost since the beginning of the subject, was not always appropriate. For example it was frequently the case that some of the data were only qualitative, and therefore difficult to treat convincingly as observations associated with a continuous random variable, while in other situations it was known that the data had been censored, either at a known point or by the optimizing actions of agents. To deal with all of these issues of a plethora of models arose, and it is only now with books and articles such as Amemiya (1984,1985) and Maddala (1983) that some order has been brought to this literature.

---

*This research is partially supported by NSF Grant Number SES–8719520.

Most of the models advanced were easily handled by deriving the implied probability density function for the observations, whereupon estimates of any unknown parameters could be found by maximum likelihood. Computer programs such as LIMDEP and SST have now virtually automated estimation in this way, and applications of the techniques abound in most journals.

Even a small sampling of this literature reveals a big difference to the older tradition of data analysis with time series, in that almost no attention is paid to the quality of the models estimated. This lack of concern is quite remarkable, since it is the hallmark of these methods that they impose over–identifying restrictions upon the parameters, and there cannot be any presumption that such restrictions are valid. Thus, the assumption of normality is very common, and it is known that the estimators are sensitive to this assumption, yet very little attention is paid to an examination of whether this convention is compatible with the data or not.

This sorry state of affairs is not due to a lack of methods for doing such evaluation. Methods have been available for some time, as evidenced by the papers by Gourieroux et al. (1987a), McFadden (1987), Chesher and Irish (1987) and Blundell (1987). But to date these procedures have just not found widespread acceptance. Part of the reason seems to lie in the fact that the derivation of these tests in the original articles is quite complex, and it is easy to lose track of exactly what the test statistic is based upon. Moreover the impression is frequently given that special programming is required for any implementation. It is our contention that these tests are capable of being derived in a much simpler and transparent way than they have been; in particular, we feel that the concentration upon the Lagrange multiplier approach to their derivation, which is characteristic of the literature, makes the analysis more complex than it should be. Moreover, the bulk of the papers is occupied with the extraction of an asymptotic variance for the test statistic, and the simple ideas involved in the formulation of a test get lost in the details attending the derivation of its variance. Ideally one would want most attention paid to the *design* of the test, and the computation of a variance should be an automatic task.

In this paper we follow Newey (1985) and Tauchen (1985) and argue that the sole issue is how to formulate suitable orthogonality or conditional moment restrictions which should hold if a model is adequate. Using a general technology supplied by these authors, it is then possible to give a regression based procedure which enables the computation of the asymptotic variance of the test statistic testing if the orthogonality conditions are valid. Section 2 of the paper introduces this idea and gives a simplified treatment of Newey's and Tauchen's results.

Having separated out the problem of determining an asymptotic variance from the selection of a basis for a test statistic, in the remainder of the paper we are free to concentrate upon how the latter operation is done. Unification is therefore achieved by correlating different tests with different orthogonality conditions. These orthogonality conditions can be broadly categorized as deriving either from analogues to tests in the regression model or as checking the first order conditions from some estimator. In turn, the "regression analogue" formulations frequently simplify down to a situation involving functions of a set of "omitted regressors". Hence the same perspective as was used in Pagan (1984) for the regression model can also be employed here in order to emphasize the underlying unity of the test statistics available in the literature.

Sections 3, 4, 5 and 6 use this orientation to survey existing tests for the Tobit, discrete choice, selectivity and duration literatures. These model types account for a large fraction of the work done with individual data, but certainly not all of it. For example, we have ignored "event" or "count" data in which the dependent variable is an integer. Readers interested in diagnostic tests for these models can consult Cameron and Trivedi (1985). Our general strategy in the "regression analogue" approach is to write down a set of first order conditions that aim to test specific aspects of the conventions that underlie the models, and then to show that these correspond with the Lagrange multiplier tests that have been suggested by various authors. With the "first order conditions" method we have largely borrowed the ideas from the semi or non–parametric literature, which provides estimators that are robust to various aspects of distributional mis–specification in these models.

As well as surveying theoretical work in this area we also apply the tests to a number of published studies. Our aim in doing this is partly to emphasize the importance of engaging in such tests, but also to show that we can perform them by using a standard econometric package, LIMDEP. With the exception of the computation of some of the first order conditions for the non–parametric estimators, all computational work in the paper was done using the data transformation, matrix manipulation and econometric routines in LIMDEP (the PC version). Hence, it is hard to argue that the construction of these tests requires any special requirements, and in fact they can be computed reasonably simply by someone using such a package.

2. <u>The Diagnostic Tests in a Regression Framework</u>

It is well known that tests for specification error conducted within the linear regression model are normally based upon residuals, e.g. see Pagan and Hall (1983), but a brief review of this literature is useful in order to motivate the methods we use later.

Suppose the regression model is

$$y_i = x_i'\beta + u_i \qquad\qquad i = 1,\ldots,N \tag{1}$$

and define the residuals $\hat{u}_i = y_i - x_i'\hat{\beta}$, where $\hat{\beta}$ is the OLS estimator of β. The errors are traditionally assumed to have a number of properties.

(i) $E(z_i u_i) = 0$ i.e., the qx1 vector of variables z_i is incorrectly excluded from the regression.

(ii) $E(z_i(u_i^2 - \sigma^2)) = 0$ i.e., the errors are assumed to have constant variance σ^2 that is unrelated to the z_i's.

(iii) $E(u_i u_{i-j}) = 0$ $(j=1,...)$ i.e. the errors have no serial correlation.

(iv) $E(u_i^3) = 0$, $E(u_i^4) - 3\sigma^4 = 0$ i.e. the moments are those of a normally distributed random variable with mean zero and variance σ^2.

Since each one of these population moments should be zero it is natural to seek information on whether this is so by examining the magnitude of the sample analogues of (i) – (iv) viz. (i) $N^{-1}\Sigma z_i \hat{u}_i$, (ii) $N^{-1}\Sigma z_i(\hat{u}_i^2-\hat{\sigma}^2)$, (iii) $N^{-1}\Sigma \hat{u}_i\hat{u}_{i-j}$, (iv) $N^{-1}\Sigma \hat{u}_i^3$ and $N^{-1}\Sigma(\hat{u}_i^4-3\hat{\sigma}^4)$. These sample moments can then be used to test if the population moments are zero. In fact, equivalent information is available from quantities that are obtained as the product of the sample moments and any non–singular matrix. Thus, if $N^{-1}\Sigma \hat{u}_i\hat{u}_{i-j}$ is multiplied by $(N^{-1}\Sigma \hat{u}_{i-j}^2)^{-1}$, the outcome is the j'th order serial correlation coefficient of the residuals $\hat{u}_i$, which is the basis for most tests of j'th order serial correlation in the u_i. If $N^{-1}\Sigma z_i(\hat{u}_i^2-\hat{\sigma}^2)$ is multiplied by $(N^{-1}\Sigma z_i z_i')^{-1}$, the resulting quantity is the regression coefficient of z_i in the regression of $\hat{u}_i^2-\hat{\sigma}^2$ against z_i, which is the basis of the Lagrange Multiplier test for heteroskedasticity in Breusch and Pagan (1979) and Godfrey (1978). Obviously, from a theoretical viewpoint, it is a matter of indifference whether the sample means or their product with some non–singular quantity is used, and the alternatives need to be ranked by their degree of computational ease. We will work with the raw moments as this accords with our later use of diagnostic tests.

Each of the sample means above is a function of the variables $w_i' = (z_i', y_i, x_i')$ and a set of p parameters $\theta' = (\beta' \sigma^2)$ and this suggests that a suitable format for each population moment is $m(w_i,\theta)$. Consequently, we define $m(w_i,\theta)$ as a qx1 vector with elements $z_i u_i$, $z_i(u_i^2-\sigma^2)$ etc., making the restrictions (i) – (iv) have the generic form $N^{-1}\Sigma E(m(w_i,\theta_0)) = 0$, where θ_0 is the true value of θ. Then $\hat{\tau} = N^{-1}\Sigma m(w_i,\hat{\theta}) = N^{-1}\Sigma \hat{m}_i$ gives the vector of sample moments.

2.1 Computing the Asymptotic Variance of $N^{1/2}\hat{\tau}$

To build a test statistic for the population moment it is necessary to derive the distribution for $N^{1/2}\hat{\tau}$. Expanding $m(w_i,\hat{\theta})$ around θ_o by Taylor series yields

$$N^{1/2}[N^{-1}\Sigma m(w_i,\hat{\theta})] = N^{1/2}[N^{-1}\Sigma m(w_i,\theta_o) + (\operatorname*{plim}_{N\to\infty} N^{-1}\Sigma(\partial m_i/\partial\theta))(\hat{\theta}-\theta_o)] + o_p(1) \qquad (2)$$

where $m_i = m(w_i,\theta_o)$. This method of determining the asymptotic distribution of $N^{1/2}\hat{\tau}$, the δ-method, relies upon the mean of higher order derivatives converging to a constant and that $\hat{\theta}$ is $\sqrt{N}$ consistent, standard assumptions in most estimation contexts. If m_i was linear in θ the linear approximation in (2) is exact.

To proceed further it is necessary to invoke a central limit theorem for the m_i and to make precise what estimator $\hat{\theta}$ is being employed. We therefore adopt two further assumptions.

(A) The estimator $\hat{\theta}$ is a generalized method of moments estimator (GMM) as described in Hansen (1982) which derives from $E[g(\theta_o)] = E[\Sigma g_i(\theta_o)] = 0$ and which minimizes $g(\theta)'Wg(\theta)$ with respect to θ, where W is a suitable weighting matrix. Under the conditions set out in Hansen the optimal choice of W is $\left[E[g(\theta_o)g(\theta_o)']\right]^{-1}$ and $N^{1/2}(\hat{\theta} - \theta_o)$ is asymptotically distributed as $-(\operatorname*{plim}_{N\to\infty} N^{-1}G_\theta' W G_\theta)^{-1} N^{1/2} G_\theta' W g_\theta$ $= -Bg_\theta$ where $G_\theta = \partial g/\partial\theta$ and all quantities are evaluated at θ_o. If $\hat{\theta}$ is the MLE, $g(\theta) = d(\theta)$, the scores of the log likelihood L, $G_\theta = H_{\theta\theta}$, the Hessian of the log likelihood, while W would be the inverse of the asymptotic information matrix $\mathcal{J}_{\theta\theta} = \operatorname*{plim}_{N\to\infty} -N^{-1}H_{\theta\theta}$. Making these substitutions for the MLE, $B = -\mathcal{J}_{\theta\theta}$, $\operatorname*{plim}_{N\to\infty} N^{-1}G_\theta' W G_\theta = -\mathcal{J}$ and $N^{1/2}(\hat{\theta} - \theta_o) = \mathcal{J}_{\theta\theta}^{-1} N^{-1/2} d_\theta + o_p(1)$.

(*B*) *The vector of conditional moment restrictions* $m_i = m(w_i, \theta_o)$ *and the moments defining* $\hat{\theta}$, $g_i = g_i(\theta_o)$, *obey a central limit theorem such that*

$$N^{1/2}\begin{bmatrix} N^{-1}\Sigma m_i \\ N^{-1}\Sigma g_i \end{bmatrix} \xrightarrow{d} \mathcal{N}(0,V)$$

$$\textit{where } V = \begin{bmatrix} V_{mm} & V_{mg} \\ V_{gm} & V_{gg} \end{bmatrix}.$$

Substituting for $N^{1/2}(\hat{\theta}-\theta_o)$ in (*2*) gives

$$N^{1/2}\hat{\tau} = N^{-1/2}\Sigma m\,(w_i,\theta_o) - (\operatorname*{plim}_{N\to\infty} N^{-1}\Sigma(\partial m_i/\partial\theta))\,B(N^{-1/2}\Sigma g_i(\theta_o)) \tag{3}$$

$$= \begin{bmatrix} I & -(\operatorname*{plim}_{N\to\infty} N^{-1}\Sigma(\partial m_i/\partial\theta))B \end{bmatrix} \begin{bmatrix} N^{-1/2}\Sigma m(w_i,\theta_o) \\ N^{-1/2}\Sigma g_i(\theta_o) \end{bmatrix} \tag{4}$$

$$= A \begin{bmatrix} N^{-1/2}\Sigma m(w_i,\theta_o) \\ N^{-1/2}\Sigma g_i(\theta_o) \end{bmatrix}. \tag{5}$$

Applying the central limit theorem in assumption B to (4)

$$N^{1/2}\,\hat{\tau} \xrightarrow{d} \mathcal{N}(0,\ AVA'). \tag{6}$$

The covariance matrix AVA′ in expression (6) can be evaluated fairly easily with computer packages that do matrix manipulations, such as GAUSS, or using the matrix languages provided in LIMDEP and SHAZAM. One could then refer $(AVA')^{-1/2} N^{1/2}\tau$ to an $\mathcal{N}(0,I_q)$ density. But it is frequently desirable to be able to do calculations with a regression program. Provided $\hat{\theta}$ is the MLE, and under assumption

C, it is in fact possible to find an asymptotically equivalent form for which such computations are possible.[1]

Assumption C: The observations w_i are independently distributed random variables.

Assumption C means that the covariance matrix V can be consistently estimated by the sample means $N^{-1}\Sigma\hat{m}_i\hat{m}_i'$, $N^{-1}\Sigma\hat{m}_i\hat{d}_i'$, $N^{-1}\Sigma\hat{d}_i\hat{d}_i'$, where $\hat{m}_i = m(w_i,\hat{\theta})$, $\hat{d}_i = d_i(\hat{\theta})$. Similarly $-H_{\theta\theta}$ can be consistently estimated by $N^{-1}\Sigma\hat{d}_i\hat{d}_i'$ if the model is adequate because then $-E(H_{\theta\theta}) = E(N^{-1}\Sigma d_i d_i')$. This means that only the element in A, $\underset{N\to\infty}{\text{plim}}\ N^{-1}\Sigma\partial m_i/\partial\theta$, remains to be simplified. Tauchen (1985) and Newey (1985) have done this by what Tauchen refers to as the generalized information equality which we state and prove here as a lemma (see also Beran (1977)).

Lemma: $E(\partial m_i/\partial\theta) = -E(m_i(w_i,\theta_o)d_i'(\theta_o))$ (7)

Proof: Under the hypothesis of correct specification,

$E(m(w_i,\theta_o)) = 0$ i.e.

$$\int_{-\infty}^{\infty} m(w_i,\theta_o)f_i(w_i,\theta_o)dw_i = 0 \tag{8}$$

using independence and defining the density of w_i as $f_i(w_i,\theta_o)$. Differentiating (8) gives

$$\int_{-\infty}^{\infty} \frac{\partial m(w_i,\theta_o)}{\partial\ \theta} f_i dw_i + \int_{-\infty}^{\infty} m(w_i,\theta_o) \frac{\partial \log f_i(w_i,\theta_o)'}{\partial\theta} f_i\ dw_i = 0$$

from which (7) follows since $d_i = \partial \log f_i(w_i,\theta_o)/\partial\theta$. □

---

[1]In fact it is only necessary that the m_i and d_i be martingale differences, but in the context of individual data that extension does not seem important.

The lemma means that $\operatorname*{plim}_{N\to\infty} N^{-1}\Sigma\partial m_i/\partial\theta$ can be consistently estimated by $-N^{-1}\Sigma\hat{m}_i\hat{d}_i'$. All of these expressions can be given a convenient matrix form. Let M be the Nxq matrix with $\hat{m}_i'$ as i'th element and D be the Nxp matrix with $\hat{d}_i'$ as i'th element. Then $N^{-1}\Sigma\hat{m}_i\hat{m}_i' = N^{-1}M'M$, $N^{-1}\Sigma\hat{m}_i\hat{d}_i' = N^{-1}M'D$ etc. and it follows that $A = [I : -M'D(D'D)^{-1}]$, and $V_{mm} = N^{-1}M'M, V_{md}' = V_{dm} = N^{-1}D'M, V_{dd} = N^{-1}D'D$ so that $AVA' = N^{-1}(M'M-M'D(D'D)^{-1}D'M)$. Now the standard test statistic for the population moments being zero based on (6) will be $N[\hat{\tau}'(AVA')^{-1}\hat{\tau}] = N^2\hat{\tau}'(M'M-M'D(D'D)^{-1}D'M)^{-1}\hat{\tau} = \iota'M(M'M-M'D(D'D)^{-1}D'M)^{-1}M'\iota$ where ι is the (Nx1) vector of units. This version is easier to compute than AVA' because $N^{-1}(M'M-M'D(D'D)^{-1}D'M)$ is the estimated covariance matrix of the residuals from the regression of M against D.

But further simplification is possible owing to the fact that $\Sigma\,\hat{d}_i = 0$ making $\iota'D = 0$. Consider the regression of M against ι and D. Clearly the fact that $\iota'D = 0$ means that the estimated coefficient of ι is $(\iota'\iota)^{-1}\iota'M = \hat{\tau}$. From standard SUR theory the variance of $\hat{\tau}$ will be $(\iota'\iota \otimes \hat{\Sigma}^{-1})^{-1}$, where $\hat{\Sigma}$ is the estimated covariance matrix of the errors in each of the q equations that has $\hat{m}_i$ as the dependent variable and the same set of regressors, unity and $\hat{d}_i$. By definition $\hat{\Sigma} = N^{-1}[M'(I - (\iota\ D)\begin{pmatrix}\iota'\iota & 0\\ 0 & D'D\end{pmatrix}^{-1}\begin{pmatrix}\iota'\\ D'\end{pmatrix})M] = N^{-1}[M'M-N\hat{\tau}\hat{\tau}' - M'D(D'D)^{-1}D'M] = N^{-1}(M'M-M'D(D'D)^{-1}D'M) - \hat{\tau}\hat{\tau}'$. Under the null hypothesis $\hat{\tau} \overset{p}{\to} 0$ so that $\hat{\Sigma}$ can be estimated by $N^{-1}(M'M-M'D(D'D)^{-1}D'M)$ and therefore the test statistic that the intercepts are zero in the SUR system is $N^2\hat{\tau}'(M'M-M'D(D'D)^{-1}D'M)^{-1}\hat{\tau}$, which is just $N[\hat{\tau}'(AVA')^{-1}\hat{\tau}]$ as required.

Thus there is a very simple way to compute a test statistic that the conditional moments are actually zero–regress $m(w_i,\hat{\theta})$ against unity and $d_i(\hat{\theta})$ and test if the coefficients on the intercepts are zero. Because the regressors are the same in every equation, if a joint test is not desired the test based on $\hat{\tau}$ can be done by use of an

OLS package. Note that this treatment splits up the problem of choosing the moment condition, $m_i(\theta)$, that is to form the basis of the specification test, from the computation of its variance, AVA'. Thus effort can be put where it should be, namely in the design of suitable ways of assessing model adequacy.

There are both advantages and disadvantages to simplifying AVA'. The advantage is that it provides an estimator of the covariance matrix of $N^{1/2}\hat{\tau}$ that is easy to compute. But that facility incurs a potential cost, in that it has a number of disadvantages. First, as observed by Wooldridge (1987), it is not robust to certain types of mis–specification, e.g. heteroskedasticity in the scores and non–normality in (say) the u_i of equation (1). This is because it assumes that $-E(H_{\theta\theta}) = E(\Sigma d_i d_i')$ and that $E(\partial m_i/\partial\theta) = -E(m_i d_i')$, which is only true when the complete density is correctly specified. To compute a robust estimator one could evaluate (6) directly, which is reasonably easy with programs such as GAUSS or any program with a good matrix language.

A more serious problem that has been observed by a number of authors has been a poor correspondence between nominal and actual sizes for test statistics formed in this simplified fashion (see Taylor (1987), Chesher and Spady (1988) and Kennan and Neumann (1988)). The last two sets of authors have traced the problem to the randomness induced into the test statistic by utilizing the outer product forms as estimators of the components of A and V, i.e., to the problem of using sample moments as estimators of population expectations. It appears from this work that the discrepencies between these two values can be very large when higher order moments need to be estimated, and this discrepency is magnified in the test statistic since AVA' appears in the denominator. However, we suspect that the fact that the procedure is so easy to compute will make it attractive to many investigators and for this reason we have generally adopted it in the following sections of this paper.

2.2 Formulating Suitable Conditional Moment Restrictions

Our motivating examples for choice of $m_i(\theta)$ above were derived from potential errors in the moments of y_i associated with the regression model and, indeed, later analogues of these will be employed as one procedure for generating diagnostic tests. Another will be what we refer to as the "first order condition" method. A popular strategy for constructing specification tests, particularly in the literature to be surveyed later, is to follow Hausman (1978) and compare two different estimators of θ, $\hat{\theta}$ and $\tilde{\theta}$, both of which are consistent estimators of θ_o if the chosen specification is adequate, but which will have different probability limits if it is not. To incorporate them into the framework above let $\Sigma\psi_i(\tilde{\theta}) = 0$ be the "first order conditions" defining the estimator of $\tilde{\theta}$. Then define $m(w_i,\hat{\theta}) = \psi_i(\hat{\theta}) - \psi_i(\tilde{\theta})$. Clearly, $N^{-1/2}\Sigma m(w_i,\hat{\theta}) = (N^{-1}\Sigma\ \partial\psi_i/\partial\theta)N^{1/2}(\hat{\theta}-\tilde{\theta}) + o_p(1)$ after an expansion of $\psi_i(\hat{\theta})$ around $\tilde{\theta}$ (under the null of no specification error). Hence, since $N^{-1}\Sigma\partial\psi_i/\partial\theta$ will generally be non–singular, a test based on $N^{1/2}(\hat{\theta} - \tilde{\theta})$ will be asymptotically identical, under the null, to one based on $N^{-1/2}\Sigma(\psi_i(\hat{\theta})-\psi_i(\tilde{\theta})) = N^{-1/2}\Sigma\psi_i(\hat{\theta})$ because $\Sigma\psi_i(\tilde{\theta}) = 0$ (see Ruud(1984) for this idea). To give an example from the regression model (1), let $\tilde{\theta}$ be the instrumental variable estimator $(\Sigma z_i x_i')^{-1}\Sigma z_i y_i$ so that $\psi_i(\hat{\theta}) = z_i(y_i - x_i'\hat{\beta})$. This constitutes a test for the exogeneity of x_i, as $\hat{\beta}$ is inconsistent if x_i is not exogenous whereas $\tilde{\beta}$ will be consistent if the instrument is valid.

Later we will encounter first order conditions that only hold "asymptotically" in the sense that $N^{-1/2}\Sigma\psi_i$ does not have zero expectation in any finite sample but the expectation converges to zero in infinite samples. Frequently these first order conditions can be written as $\psi_i = \psi_i^1 + \psi_i^2$ where $E(\psi_i^1) = 0$ and $N^{-1/2}\Sigma\psi_i^2$ is $o_p(1)$ when the null hypothesis of an adequate model is correct. From the derivation of the limiting distribution above it is apparent that a test based upon $\Sigma\psi_i(\hat{\theta})$ will still have the same limit distribution as described earlier.

There is one set of diagnostic tests that cannot be immediately put into either framework above. These focus upon the use of ***prediction errors*** for detecting a poor model. In this case the conditional moment restrictions for (1) would be of the form $E\left[n^{-1}\sum_{i=N+1}^{N+n} z_i(y_i - x_i'\beta)\right] = 0$, where N+1,...,N+n are the individuals whose behavior is being predicted — the "prediction period". In principle z_i could be many possible variables, but setting z_i to x_i focuses directly upon whether the orthogonality conditions associated with least squares hold out-of-sample. Because $\Sigma x_i(y_i - x_i'\beta)$ are proportional to the scores when the u_i are normally distributed, an obvious extension of this prediction idea to models estimated by MLE is to examine the average score in the prediction period, that is to form the test $\hat{\tau} = n^{-1}\sum_{i=N+1}^{N+n} \hat{d}_i$, since the mean post-sample scores should be close to zero for a good model. Defining $k = n/N$, and exploiting the fact that the observations are independent, the asymptotic variance is easily seen to be $(N+n)\mathcal{J}_{\theta\theta} = N(1+k)\mathcal{J}_{\theta\theta}$, Pagan and Hoffman (1988), and therefore the test statistic will be $(1+k)^{-1}N^{-1}\hat{\tau}'\hat{\mathcal{J}}_{\theta\theta}^{-1}\hat{\tau}$. Notice that it is envisaged in this approach that n tends to infinity as N does, and so n needs to be quite large. This requirement should not pose any major problems as N is typically extremely large.

3. Diagnostic Tests for the Tobit Model

3.1 Regression Diagnostic Analogues

The Tobit model has been a popular model for the study of individual data, and even for some time series problems, as seen from Amemiya's (1984) survey. As normally motivated it is presumed that there exists a latent model corresponding to (1)

$$y_i^* = x_i'\beta + u_i, \tag{9}$$

but observations $\{y_i\}$ are made on $\{y_i^*\}$ only if $y_i^* > 0$. For $y_i^* \le 0$, $y_i = 0$ is observed. Extensions to allow for a different censoring threshold than zero are also possible.

If y_i^* was observed an interesting set of conditional moment restrictions would be (i) – (iv) discussed in section 2 earlier. Consider the first of these – $N^{-1}\Sigma E(z_i(y_i^*-x_i'\beta)) = 0$. Now $E[z_i(y_i^*-x_i'\beta)] = E[E(z_i(y_i^*-x_i'\beta)|y_i)] = E[z_i(E(y_i^*|y_i) - x_i'\beta)] = 0$, and repeating this conditioning argument for each of (i) – (iv) in section 2 yields

(i) $$N^{-1}\sum_{i=1}^{N} E[z_i(E(y_i^*|y_i) - x_i'\beta)] = 0 \tag{10a}$$

(ii) $$N^{-1}\sum_{i=1}^{N} E[z_i(E(u_i^2|y_i) - \sigma^2)] = 0 \tag{10b}$$

(iii) $$N^{-1}\sum_{i=1}^{N} E[E(u_i u_{i-j}|y_i,y_{i-j})] = 0 \tag{10c}$$

(iv) $$N^{-1}\sum_{i=1}^{N} E[E(u_i^3|y_i)] = 0,\ E[E(u_i^4|y_i) - 3\sigma^4] = 0. \tag{10d}$$

It is well known (see Amemiya (1985, p. 367)) that $E(y_i^*|y_i = 0) = x_i'\beta-\sigma\lambda_i$, where λ_i is the ratio of $\phi(x_i'\beta/\sigma)$ to $1-\Phi(x_i'\beta/\sigma)$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are respectively the density and cumulative density of the standard normal random variable. For positive observations $E(y_i^*|y_i>0) = y_i$. Accordingly, when testing for an omitted regressor z_i, the sample moment corresponding to (i) above would be $N^{-1}\sum_{i=1}^{N} z_i\{-\hat{\sigma}(1-I_i)\hat{\lambda}_i+I_i\hat{u}_i\}$, with I_i being the binary variable that is zero if $y_i = 0$ and unity if $y_i > 0$. By comparing this to $N^{-1}\Sigma z_i\hat{u}_i$ in the linear regression case we are led to describe $\hat{\eta}_i = -\hat{\sigma}(1-I_i)\hat{\lambda}_i + I_i\hat{u}_i$ as a ***generalized residual*** (in the sense of Cox and Snell (1968)), making specification tests for a correct conditional mean in the Tobit model a function of the sample moments of z_i with the generalized residuals $\hat{\eta}_i$. This is the essence of the diagnostic tests proposed by Chesher and Irish (1987) and

Gourieroux et al. (1987a), although the latter derive it as an LM test by showing that $N^{-1}\Sigma z_i \eta_i$ is proportional to the score of the log likelihood with respect to the γ in the expanded model $y_i^* = x_i'\beta + z_i'\gamma + u_i$.

The sample moments corresponding to (i) – (iv) above when β is replaced by $\hat{\beta}$ have all been proposed as diagnostic tests for particular types of model inadequacy. Chesher and Irish and Gourieroux et al. propose the second as a test for heteroskedasticity. Under the null hypothesis that u_i is independent of u_j, (iii) would be $N^{-1}\sum_{i=1}^{N} E(u_i|y_i)E(u_{i-j}|y_{i-j})$ with sample analogue $N^{-1}\Sigma\hat{\eta}_i\hat{\eta}_{i-j}$. Robinson et al. (1985) used the first order serial correlation coefficient of the $\hat{\eta}_i$ as a test for serial correlation in the Tobit model, and this is proportional to $N^{-1}\Sigma\hat{\eta}_i\hat{\eta}_{i-j}$. Bera, Jarque and Lee (1984) use (iv) as tests for normality of the errors, deriving it from an expanded model in which the alternative density is a member of the Pearson family. Chesher and Irish (1987) give a different derivation but also settle upon (iv) as a test for the Tobit assumption of normality in the errors of the latent variable model. Perhaps the main difficulty with implementing (i) – (iv) is the need to evaluate $E(u_i^{\ell}|y_i = 0)$, but Lee and Maddala (1985, p.4) give a recursion for the moments of $(y_i - x_i'\beta)/\sigma$ which yields

$$E(u_i^{\ell+1}|y_i=0) = \sigma^2\ell\ E\ (u_i^{\ell-1}|y_i=0) - \sigma(-x_i'\beta)^{\ell}\lambda_i, \ell \geq 1$$

and $$E(u_i|y_i=0) = -\sigma\lambda_i.$$

All of the existing tests for omission of variables etc. described above were generally derived from the LM or score test approach, sometimes after quite complex argument. Our analysis shows that the same outcome could be obtained by stating the appropriate conditional moment restriction associated with the latent model, and then replacing it with a moment restriction based upon the observable random variables. Application of this principle highlights the fact that all of the standard tests used in the basic regression model can be extended directly to the Tobit model, e.g. putting $z_i = (x_i'\hat{\beta})^2, (x_i'\hat{\beta})^3$ for non-zero y_i and the squares and cubes of the estimated $pr(y_i=0)$ for zero y_i would yield the analogue of Ramsey's (1969) RESET test.

Because of the connection with the LM test, diagnostic tests based upon the moments described earlier provide an asymptotically locally most powerful test of the alternative hypothesis that they were constructed with reference to. Thus the generalized residuals may well be the best instrument for specification analysis. Used as parts of a sample moment this is true, but sometimes a good deal is learned from a graphical analysis of residuals, and a number of authors have pointed out that the generalized residuals can be hard to interpret. For this reason it may be useful to develop classes of residuals that resemble ordinary residuals more closely but which can be used in the same way as generalized residuals.

There is a temptation to adopt $\hat{\nu}_i = y_i - x_i'\hat{\beta}$ for this purpose, but the random variable $\nu_i = y_i - x_i'\beta$ does not have a zero mean, and correcting for that fact reproduces η_i. Nevertheless there are other ways to obtain a series of errors that do have the correct mean of zero (correct that is if the model is adequate). Powell (1986) has developed a "symmetrically trimmed least squares" estimator of β for the Tobit model that trims the original data set so that the resulting error term has an expected value of zero provided the underlying latent density is symmetric around $x_i'\beta$. He does this by first deleting all observations for which $x_i'\hat{\beta}<0$, producing the set of observations $i=1,..,L$, and secondly by setting $y_i = 2x_i'\hat{\beta}$ for all y_i that exceed $2x_i'\hat{\beta}(i=1,..,L)$. In terms of the latent model, when $x_i'\beta \geq 0$, $y_i^* \geq 2x_i'\beta$ if $u_i^* \geq x_i'\beta$, while $y_i^* \leq 0$ if $u_i^* \leq -x_i'\beta$. Hence, by deleting the negative $x_i'\beta$ and matching the $y_i = 0$ with $y_i = 2x_i'\beta$, the resulting $y_i(i=1,..,L)$ are symmetrically distributed over $(0, 2x_i'\beta)$; that is, the errors $y_i - x_i'\beta(i=1,..,L)$ are symmetrically distributed over $-x_i'\beta$ to $x_i'\beta$.

Formally, the symmetrically trimmed residuals are defined as $\xi_i = I(x_i'\hat{\beta}>0)[\min(y_i, 2x_i'\hat{\beta}) - x_i'\hat{\beta}]$, where I is the indicator function. A plot of these residuals may be helpful in detecting model inadequacy, although it is apparent that their utility is limited by the number of observations with $x_i'\hat{\beta} \geq 0$ in the sample and the number that exceed $2x_i'\hat{\beta}$. If either event occurs often, a good deal of information can be lost if the specification errors happen to be mainly manifest in those data points. Nevertheless, the symmetrically trimmed residuals are easily computed and a check on their patterns would seem to be worthwhile. In fact, because the ξ_i are heteroskedastic it is probably more informative to examine the standardized residuals.

If interest centers mainly upon distributional issues, it is possible to estimate the distribution function of the underlying errors with the Kaplan–Meier (1958) "product limit" estimator applied to the residuals $y_i - x_i'\hat{\beta}$. This may then be compared with the cumulative normal distribution and an assessment of the validity of the normality assumption made. Chesher et al. (1985) illustrate the power of this graphical procedure in determining departures from normality. Formal tests for the difference between the estimated and postulated distributio functions can be done as conditional moment tests following Heckman (1984), Andrews (1987) and Tauchen (1985).

Some other diagnostic tests for the Tobit model also appear in the literature. Smith and Blundell (1986) were concerned with whether the regressors of the latent model were exogenous. They add to (9) a variable w_i which is related to other variables z_i in a linear fashion, making a two equation system

$$y_i^* = w_i'\gamma + x_i'\beta + u_i \tag{11a}$$

$$w_i = z_i'\pi + v_i, \tag{11b}$$

where the errors u_i and v_i are bivariate normal.

When w_i is weakly exogenous $E(u_i v_i) = 0$, so they propose to test for this by adding the OLS residuals $\hat{v}_i = w_i - z_i'\hat{\pi}$ to 11(a), applying the Tobit MLE and checking if the coefficient of $\hat{v}_i$ is zero. Under the null hypothesis the conditional moment restriction $E(u_i v_i) = 0$ can be written as $E[E(u_i|y_i)v_i] = 0$, exploiting the independence of w_i and v_i, i.e. a suitable test for exogeneity of w_i could be constructed from the sample covariance of the generalized residuals $\hat{\eta}_i$ from (11a) and the OLS residuals $\hat{v}_i$. But this is easily seen to be asymptotically equivalent, under the null hypothesis, to the Smith and Blundell test. Since Smith and Blundell apply the Tobit MLE to (11a) with $\hat{v}_i$ as an additional regressor, let us call their estimator of γ, β and α (the coefficient of $\hat{v}_i$), $\tilde{\theta}$. Their "t-statistic" is proportional to the score for α, and hence to $N^{-1}\Sigma\hat{v}_i\eta_i$, from which it is evident that the difference in the two approaches can only be in the evaluation of η_i. The method of this paper evaluates η_i with $\hat{\theta}$, the MLE of θ under the null hypothesis, whereas Smith and Blundell use $\tilde{\theta}$, an estimator that is consistent under the alternative hypothesis that $E(u_iv_i) \neq 0$. It is worth noting that they evaluate the variance of $\hat{\alpha}$ by setting $\alpha = 0$, so that their actual test statistic is a mixture of $\hat{\theta}$ and $\tilde{\theta}$ evaluations. To get the variance of $N^{1/2}\hat{\tau} = N^{-1/2}\Sigma\hat{\eta}_i\hat{v}_i$ one need only apply the regression approach detailed in section 2.1, so that both motivation and derivation of the exogeneity test is aided consiuerably by recognizing its origins as a conditional moment restriction.

An "omnibus" test for model inadequacy is the information matrix test formulated in White (1982) and extensively applied to the Tobit model by Chesher et al. (1985). This is a conditional moment restriction with $m_i(\theta) = \mathrm{vech}(E(\partial^2 L_i/\partial\theta\partial\theta') + d_i d_i')$ where L_i is the value of the log likelihood and d_i is the score vector at the i-th point. Chesher et al. decompose $m_i(\theta)$ into three components that test for heteroskedasticity and normality.

3.2 Diagnostic Tests from First Order Conditions

Section 2 also noted the possibility of constructing diagnostic tests by using the first order conditions for an alternative root N consistent estimator of θ, i.e. $\psi(\tilde{\theta}) = 0$ and $N^{1/2}(\tilde{\theta} - \theta_o)$ is $O_p(1)$. Forming $\psi(\hat{\theta})$ and testing if this is zero is equivalent to a test for whether $\hat{\theta}$ and $\tilde{\theta}$ are different. Quite a few proposals of this type have been made. Newey (1986) took $\tilde{\theta}$ as Powell's (1986) symmetrically trimmed least squares estimator, for which $\psi_i(\theta) = N^{-1}\Sigma x_i\xi_i(\theta)$, where ξ_i were the symmetrically trimmed errors defined earlier. Hence, we can interpret Newey's work as examining $N^{-1}\Sigma x_i\hat{\xi}_i$[2].

There are other sets of first order conditions from the semi-parametric estimation literature which can be utilized as the basis of a diagnostic test. Ichimura (1987) showed that the estimator of β which solved the first order conditions $\Sigma(y_i-\hat{g}_i)(\partial\hat{g}_i/\partial\beta) = 0$, where $\hat{g}_i$ is an estimate of $g_i = E(y_i|x_i'\beta)$, was asymptotically normal with normalizing factor $N^{1/2}$. To get this result he estimated the conditional expectation g_i by kernel regression as $\sum_{j\neq i} y_jK((x_i'\beta-x_j'\beta)/h)/\sum_{j\neq i} K((x_i'\beta-x_j'\beta)/h)$, where K(.) is a function with the properties that it is non-negative, symmetric and integrates to unity, while h is a window-width. There are many choices that could be made for K, mainly from the class of probability density functions for continuous random variables, but h needs to be set as proportional to a number between $N^{-1/3}$ and $N^{-1/4}$, since it was only for this range that the estimator $\tilde{\beta}$ solving these first order conditions was shown to be root-N consistent; to interpret this test as one based upon the difference between $\hat{\beta}$ and $\tilde{\beta}$ the window width choice needs to be circumscribed. It is also of interest to note that the expectation of $\psi_i = (y_i-g_i)(\partial\hat{g}_i/\partial\beta)$ above is not zero in finite samples, but it can be written as $\Sigma[(y_i-g_i)(\partial\hat{g}_i/\partial\beta) + \psi_i^2]$, where the expectation of the first term is clearly zero due to the i≠j restriction in the construction of the kernel, while the second term when normalized by $N^{-1/2}$ is $o_p(1)$. Hence, as mentioned earlier in section 2, the distribution of the test statistic is the same as if the expectation of ψ_i had been zero in finite samples. In later use of this test we set $h = N^{-7/24}$ and use the Gaussian kernel $K(z) = (2\pi)^{-1/2}\exp(-(1/2)z'z)$. One advantage of Ichimura's estimator is that it utilizes all the data and therefore it might have greater power as a diagnostic test than that based on Powell's estimator even though it is much more

---

[2]In fact Newey also suggested obtaining $\tilde{\theta}$ by solving $N^{-1}\Sigma x_i\rho(\xi_i(\tilde{\theta})) = 0$ where $\rho(\cdot)$ is any odd function, as the conditional symmetry of ξ_i ensures that $E(x_i\xi_i(\theta_o)) = 0$.

cumbersome to compute the first order conditions. A disadvantage is that it is consistent only under distributional mis-specification whereas Powell's estimator is also robust to heteroskedasticity in the u_i.

Nelson (1980) found a $\bar{\theta}$ from a rather odd binary formulation of the Tobit model (see Ruud (1984) for details), and Ruud (1984) suggested that a more appropriate procedure would be to apply the Probit estimator to the observations, i.e. replace $y_i > 0$ by $y_i = 1$, to find $\tilde{\theta}$. Hence his test compares the Probit ($\tilde{\theta}$) and Tobit ($\hat{\theta}$) MLE's and can be regarded as focusing upon $N^{-1}\Sigma x_i((\hat{\Phi}_i(1-\hat{\Phi}_i))^{-1}(y_i-\hat{\Phi}_i)\hat{\Phi}_i$. Of course what appears in the first order conditions for the Probit model is (β/σ), and therefore only $(\hat{\beta}/\sigma)$ is substituted into the $\psi(\cdot)$ function.

The variance of each of these statistics may be found via the Newey–Tauchen method outlined in section 2. It should be mentioned that in the case of the conditional moment restrictions from the symmetrically trimmed least squares estimator, neither Newey nor Tauchen's paper strictly provides the requisite asymptotic theory. Because ξ_i is not differentiable in θ, the δ–method described in section 2 does not apply, although Tauchen showed that one could allow non–differentiability if observations were identically distributed. However, it is not too hard to verify that the results of section 2 apply in this instance by adapting the limit theorems in Powell (1986).

3.3 Prediction Error Tests

Exploiting prediction errors for diagnostic purposes has not been as popular with the Tobit model as for the basic regression model. Anderson (1987) is a notable exception. Essentially he advocated a comparison of the log likelihood over the sample period with that when the model was fitted using both the sample and post–sample data. Although easy to compute, this test is not very informative in the event of a rejection, since it is not clear which of the first order conditions defining the estimators has been violated by the post–sample data. Anderson (1987) tried to implement Salkever's (1976) dummy variable technique, but ran into difficulties with constructing a test statistic. Hoffman and Pagan (1988) applied the prediction test based on the post–sample scores discussed earlier to Fair's (1978) Tobit model of the number of extra–marital affairs engaged in by individuals. This is an example that we now turn to. However before doing so it probably should be noted that the dependent variable y_i is integer valued and therefore a Tobit model is not really appropriate.

3.4 An Example

Fair's (1978) model explained the number of times an individual engaged in extra–marital intercourse as determined by seven variables — sex, age, number of years married, the presence of children, the degree of religious attachment, level of education and occupation. He fitted this model to two data sets; we chose the 601 observation set from Psychology Today.[3] Table 1 presents the diagnostic tests in (10) for this model, obtained by regressing the conditional moment restrictions against an intercept and the scores. It is immediately clear from these results that the model is seriously mis–specified.

As well as the diagnostic tests stemming from the testing of the moment conditions reported in Table 1, Fair's model was also evaluated with the non–parametric first order conditions described earlier. The first method employed was to insert Fair's maximum likelihood estimates into the first order conditions for Powell's symmetrically trimmed least squares estimator. These values were then regressed against the scores from Fair's model and an intercept; the t–statistics on the intercept in these regressions are tests of whether the first order conditions sum to zero

---

[3]We are grateful to Ray Fair as well as all the other authors who supplied their data.

TABLE 1

Tests of Misspecification in Fair's (1978) Tobit Model of Extra-Marital Affairs

Test	Moment Restriction	t-Statistic
RESET[1]	$E(PRED^{2}*\eta)=0$	6.953
	$E(PRED^{3}*\eta)=0$	5.948
Heteroskedasticity[2]	$E(z_1(E(u^2\|y)-\sigma^2))=0$	12.238
	$E(z_2(E(u^2\|y)-\sigma^2))=0$	17.391
	$E(z_3(E(u^2\|y)-\sigma^2))=0$	13.923
	$E(z_4(E(u^2\|y)-\sigma^2))=0$	13.685
	$E(z_5(E(u^2\|y)-\sigma^2))=0$	16.490
	$E(z_6(E(u^2\|y)-\sigma^2))=0$	20.032
	$E(z_7(E(u^2\|y)-\sigma^2))=0$	18.052
	$E(z_8(E(u^2\|y)-\sigma^2))=0$	16.771
Normality	$E(E(u^3\|y))=0$	14.726
	$E(E(u^4\|y)-3\sigma^4)=0$	14.231
Prediction[3]	E(post-sample scores)=0	61.48 (χ^2_8)

Notes:

1. In the RESET test PRED are the predictions $x_i'\hat{\beta}$ for the non-zero observations and $\Phi(x_i'\hat{\beta})$ for the zero observations. η are the generalized errors.

2. z_1 = sex, z_2 = age, z_3 = number of years married, z_4 = presence of children, z_5 = degree of religious attachment, z_6 = education level, z_7 = occupation. See Fair (1978) for a complete description of these variables.

3. Test that the average scores for i = 541,..,601 are zero after random selection of observations. The test is χ^2 with eight degrees of freedom as there is a score for σ^2. See Pagan and Hoffman (1988) for further details.

and are reported in Table 2. An examination of this table reveals that, in many instances, the requirement that the first order conditions have probability limit of zero is most likely violated. The evidence for inadequacy is however much weaker than for the parametric tests, although this might be explained by the fact that only 47 observations remain after the trimming. That there was a potential problem with the method stemming from such a loss of information was alluded to earlier, but this provides a dramatic illustration of the point.

Table 2

First Order Conditions for the Symmetrically Trimmed Least Squares Estimator Evaluated at Fair's Tobit MLE's

Variable	t–stat	Variable	t–stat
Intercept	2.538	z1	.624
z2	2.065	z3	1.401
z4	1.781	z5	.982
z6	2.183	z7	1.320
z8	2.216		

Fair's model was also examined by evaluating the values of the first order conditions of Ichimura's estimator implied by Fair's maximum likelihood estimates. As with the symmetric least squares estimator this was done by regressing the implied first order conditions against an intercept and the scores from the Tobit model. The t–tests on the intercept being equal to zero are reported in Table 3. These results support the earlier results that the model is misspecified. In three instances the requirement that the first order conditions is equal to zero is clearly violated. However, in

Table 3

First Order Conditions for Ichimura's Single Index Estimator

Evaluated at Fair's Tobit M.L.E.'s

Variable	t-stat (abs. value)	Variable	t-stat (abs. value)
z_1	1.870	z_2	1.458
z_3	.417	z_4	.948
z_5	2.778	z_6	2.299
z_7	.802	z_8	2.089

light of the earlier results it appears that this non-parametric procedure may be weak in detecting misspecification, even with the quite large sample available here.

4. Discrete Choice Models

4.1 Univariate Models

4.1.1 Regression Diagnostic Analogues

In discrete choice models only the fact that a choice has been made in the negative ($y_i = 0$) or the positive ($y_i = 1$) is available. One way to place this in the context of the linear regression model is to proceed with the latent variable model (9) used in the Tobit model discussion, identifying the outcomes $y_i = 0$ with $y_i^* \leq 0$ and $y_i = 1$ with $y_i^* > 0$. Diagnostic tests for omitted variables, heteroskedasticity then have exactly the same format as for the Tobit model, except that $E(y_i^* | y_i > 0) \neq y_i^*$ any longer. The generalized residuals therefore differ from those for the Tobit model, but only to the extent of the substitution of $E(y_i^* | y_i = 1)$ for y_i^* in the formula. When the latent errors are normal we get the Probit model, and $E(y_i^* | y_i = 1)$ is given by $\phi_i \Phi^{-1}$ (Amemiya (1985, p. 269)).

Although the latent variable approach can be a useful way of thinking about the origin of the data $\{y_i\}$, discrete choice modeling more often proceeds by focusing directly upon $\{y_i\}$ and regarding it as realizations of a binary random variable with $Pr[y_i=1] = F_i$, $Pr[y_i=0] = 1-F_i$. With the latent variable interpretation and normality of the u_i, F_i would equal $\int_{-\infty}^{x_i'\beta} \phi(z)dz$, and this constitutes the Probit model. Other formulations of F_i are possible, with the most popular alternative being the Logit model which has $F_i = \exp(x_i'\beta)/(1 + \exp(x_i'\beta))$.

Suppose y_i is a binary random variable with true probability function $\bar{F}_i$. It is easily seen that $E(y_i) = \bar{F}_i$, $var(y_i) = \bar{F}_i(1-\bar{F}_i)$. By definition therefore,

$$y_i \;=\; E(y_i) + \bar{v}_i \;=\; \bar{F}_i + \bar{v}_i \tag{12}$$

where $E(\bar{v}_i) = 0$, $E(\bar{v}_i^2) = \bar{F}_i(1-\bar{F}_i)$. Now let $\bar{F}_i$ be a function of the data, the parameters of the chosen model θ, and a set of parameters γ that characterizes an alternative. Setting $\gamma = 0$ yields the chosen model, which is defined as

$$y_i \;=\; \bar{F}_i(\theta,\gamma=0) + v_i \;=\; F_i+v_i, \tag{13}$$

and v_i has zero mean and variance $F_i(1-F_i)$ under the null that $\gamma=0$.
Expanding $\bar{F}_i$ in (12) around $\gamma=0$ by Taylor series gives[4]

$$y_i \;=\; \bar{F}_i(\theta,\gamma=0) + F_i'\gamma + F_i''\gamma^2 +.....+\bar{v}_i, \tag{14}$$

where the primes indicate differentiation with respect to γ, and $F_i' = \bar{F}_i'(\theta,\gamma=0)$, $F_i'' = \bar{F}_i''(\theta,\gamma=0)$ etc. (14) embodies the conditional moment restrictions $E[N^{-1}\Sigma F_i'\bar{v}_i] = 0$, $E[N^{-1}\Sigma F_i''\bar{v}_i] = 0,......$. Under the null hypothesis $\gamma=0$, $\bar{v}_i = v_i$, and the restrictions are $E[N^{-1}\Sigma F_i'v_i] = 0$, $E[N^{-1}\Sigma F_i''v_i] = 0,.....$.

[4]We will assume γ is a scalar in the exposition below for simplicity.

If θ was known, the sample moments corresponding to the population moments would be $N^{-1}\Sigma F_i' v_i, N^{-1}\Sigma F_i'' v_i, \ldots$. Now, rather than use $N^{-1}\Sigma F_i' v_i$ as the basis of a test, it was observed in section 2 that an equivalent test for whether $E(N^{-1}\Sigma F_i' v_i) = 0$ could be formed from $[(N^{-1}\Sigma(F_i')^2]^{-1}[N^{-1}\Sigma F_i' v_i]$ whenever $N^{-1}\Sigma(F_i')^2 \neq 0$. This is the OLS estimate of the coefficient of F_i' in the regression of v_i against F_i'. But the errors in such a regression, under the null, are heteroskedastic with variance $F_i(1-F_i)$, and therefore a GLS estimator of this coefficient will be more efficient. Thus it is more appropriate to regress $v_i^* = F_i^{-1/2}(1-F_i)^{-1/2}v_i$ against $(F_i')^* = F_i^{-1/2}(1-F_i)^{-1/2}F_i'$, pointing to $E(N^{-1}\Sigma(F_i')^* v_i^*) = E(N^{-1}\Sigma F_i'(F_i(1-F_i))^{-1}v_i)$ as the relevant population moment condition. Taking cognizance of the fact that θ needs to be replaced by its MLE $\hat{\theta}$, the sample moment condition to be used for the construction of diagnostic tests in discrete choice models will therefore be

$$\hat{\tau} \;=\; N^{-1}\Sigma \hat{F}_i'(\hat{F}_i(1-\hat{F}_i))^{-1}(y_i-\hat{F}_i). \tag{15}$$

Many diagnostic tests for the Probit and Logit model are based upon (15), differing only in the choice of F_i'. To appreciate this it is useful to write F_i as a function of a variable μ_i so that $F_i(\mu_i=-\infty) = 0$, $F_i(\mu_i=\infty) = 1$, allowing F_i' to be expressed as $f_i\partial\mu_i/\partial\gamma = f_i z_i$, where f_i is $(\partial F_i/\partial\mu_i)$ and would be the density ϕ_i if $\mu_i = x_i'\beta+z_i'\gamma$ in the Probit case. Hence, $\hat{\tau} = N^{-1}\Sigma z_i f_i(\hat{F}_i(1-\hat{F}_i))^{-1}(y_i-\hat{F}_i)$ and the elements in the sum are $-z_i\hat{f}_i(1-\hat{F}_i)^{-1}$(if $y_i=0$) and $z_i\hat{f}_i\hat{F}_i^{-1}$(if $y_i=1$). For the probit model $\hat{f}_i = \hat{\phi}_i$, $\hat{F}_i = \hat{\Phi}_i$ and $\hat{\tau}$ is seen to be the covariance between z_i and the series $-\hat{\phi}_i(1-\hat{\Phi}_i)^{-1}$(if $y_i=0$) and $\hat{\phi}_i\hat{\Phi}_i^{-1}$(if $y_i=1$), the last two terms being identical to the generalized residuals $E(y_i^*|y_i) - x_i'\hat{\beta}$. Accordingly, $\hat{\tau} = N^{-1}\Sigma z_i(E(y_i^*|y_i) - x_i'\hat{\beta})$. Because of this relationship in the Probit model, $\hat{f}_i(\hat{F}_i(1-\hat{F}_i))^{-1}(y_i-\hat{F}_i)$ will be termed the generalized residuals for *any* discrete choice model. It is also worth noting that, from Davidson and MacKinnon (1984) and Newey (1985), τ would also be proportional to the score vector for γ, making $\hat{\tau}$ equivalent to the score test of $\gamma=0$.

Many suggestions are available for z_i and Newey (1985, p.1062) lists these for the Probit model. When $\mu_i = x_i'\beta+z_i'\gamma$, z_i being an omitted regressor, $\partial\mu_i/\partial\gamma = z_i$ and this tests for mis-specification in the conditional mean of the latent variable model. When $\mu_i = x_i'\beta(1+z_i'\gamma)$, $\partial\mu_i/\partial\gamma = z_i(x_i'\beta)$ (under the null that $\gamma = 0$), and one has a test for heteroskedasticity. To test for normality as the underlying distribution for the latent variable model, one can follow Ruud (1984) and set $F_i = \Phi(\mu_i+\gamma_1\mu_i^2+\gamma_2\mu_i^3)$, where $\mu_i = x_i'\beta$, whence it follows that $z_i = [(x_i'\hat{\beta})^2\ (x_i'\hat{\beta})^3]$, and this would provide a RESET-like test for normality in the Probit model (Taylor (1985) investigated the power of this test to detect some types of mis-specification in Probit models and found it had quite low power for conditional mean mis-specification, but he did not concentrate upon distributional mis-specification).

One interesting feature to emerge from the above analysis is that tests for higher order moment mis-specification in the latent variable model are capable of being detected from the correlation of z_i with the generalized residuals, and powers of them are not needed. This is because, unlike the regression model, no strict demarcation can really be drawn between the effects of mis-specification upon different moments of the y_i, since a mis-specification in any part of F_i affects $E(y_i)$ as well as higher order moments.

It is natural to ask why the moment conditions $E(N^{-1}\Sigma F_i'{}'\bar{v}_i) = 0$ etc. were ignored above. In fact, this broader set of orthogonality relations could be used as the basis of diagnostic tests, but they are generally discarded because of the following "local alternatives" argument. It is a standard procedure to compare test statistics by their ability to detect specification errors which disappear as $N\to\infty$, as this constitutes a fairly stringent test of their quality. Under this argument $\gamma = \delta/\sqrt{N}$ and from (14)

$$[N^{1/2}(N^{-1}\Sigma F_i' v_i)] \;=\; [N^{-1/2}\Sigma F_i'(F_i'\gamma + F_i''\gamma^2 + \ldots + \bar{v}_i)] \tag{16}$$

$$= \; N^{-1/2}(\Sigma(F_i')^2)\gamma + N^{-1/2}(\Sigma F_i' F_i'')\gamma^2 + \ldots + N^{-1/2}\Sigma F_i' \bar{v}_i \tag{17}$$

$$= \; (N^{-1}\Sigma(F_i')^2\delta) + N^{-1/2}(N^{-1}\Sigma F_i' F_i'')\delta^2 + \ldots + N^{-1/2}\Sigma F_i' \bar{v}_i \tag{18}$$

indicating that terms in the expansion above first order have no asymptotic effect when alternatives are local, thereby justifying a concentration upon the moment restriction involving F_i' and v_i only.

Essentially the diagnostic tests described above examine the residuals $y_i - F_i$ for specification errors, taking into account the fact that the errors are heteroskedastic. Unfortunately, even when studentized, these residuals tend to be hard to interpret when plotted, as seen from the graphs in Gourieroux et al. (1987b). When a discrete choice model can be given a latent variable interpretation, these authors have proposed a class of "simulated residuals" which are better suited to graphical analysis. Taking the Probit model as an illustration, they would simulate y_i^* by adding on to $x_i'\hat{\beta}$ the first $\mathcal{N}(0,1)$ random number generated on a computer that made $y_i^* \leq 0$ (if y_i=0) or $y_i^* >$ 0 (if y_i=1). Designating the simulated y_i^* as $\tilde{y}_i^*$, they regress $\tilde{y}_i^*$ against x_i to obtain the simulated residuals $\tilde{u}_i$. Of course the $\tilde{u}_i$ rarely look much like the u_i, in contrast to the OLS residuals which at least converge to u_i. Moreover the variance of $\tilde{u}_i$ involves a rather complicated formula, although this feature does not seem very important unless one wished to construct test statistics from them, and it is hard to see why such residuals would be preferred over the generalized residuals in this context. How effective simulated residuals are in detecting specification errors remains to be seen. One of their best uses may be in the detection of outliers in the data, and this was an example employed by Gourieroux et al. Isolation of outliers in the Logit model was also Pregibon's (1981) concern, and he has suggested some measures of the importance of the i–th point to the parameter estimates, based on the concepts of leverage that have been popular in linear regression.

A number of other conditional moment restrictions have been proposed as a way of generating diagnostic tests which are not functions of the generalized residuals. A procedure that has found favor in the mode or brand choice literature is to compare the average predicted probabilities of choosing a particular mode or brand for a class of individuals with the observed relative frequency of that choice. Thus the test statistic is

$$\hat{\tau} = N_j^{-1} \sum_{i \epsilon I_j} \hat{F}_i - N_j^{-1} \sum_{i \epsilon I_j} y_i \qquad (19)$$

where I_j indicates the j–th class of individuals within a sample of i=1,..,N and N_j is the number of individuals in that class. Defining the dummy variable δ_i to be the value unity if $i \epsilon I_j$ and zero otherwise,

$$\hat{\tau} = N_j^{-1} \sum_{i=1}^{N} \delta_i(\hat{F}_i - y_i) = N^{-1} \sum_{i=1}^{N} (N_j^{-1}N)\delta_i(\hat{F}_i - y_i) \qquad (20)$$

and this would be the sample equivalent of the conditional moment restriction of $E[N_j^{-1}N\delta_i(F_i - y_i)] = 0$. Because $E(y_i) = F_i$ under the maintained hypothesis of an adequate model this is indeed so.

Horowitz (1985) formalized the above test, finding its limiting distribution after a lengthy derivation. But, because it fits into the Newey–Tauchen framework, the variance can be found much more easily from the regression of the $N_j^{-1}N\delta_i(F_i - y_i)$ against a constant term and the scores for β, as described earlier. (Note that the asymptotic theory requires that $N_j \to \infty$ as $N \to \infty$ and for N_j/N to tend to some constant.) Horowitz found that the power of the test was fairly weak when compared to that of a likelihood ratio test (LRT), except when the postulated alternative needed to define the LRT departed substantially from the true model.

Rivers and Vuong (1988) test for exogeneity of regressors in a Probit model in the same way as was discussed earlier in connection with Smith and Blundell's (1986) work on the Tobit model. Their test would involve the correlation of the generalized residuals from the Probit model with the reduced form residuals, exactly as described earlier. Note that, if the simple method of computing the asymptotic variance of $N^{1/2}\hat{\tau}$ set out in section 2 is to be used, the scores for both fitted equations need to appear as regressors.

4.1.2 Diagnostic Tests from First Order Conditions

Just as for the Tobit model one might work with diagnostic tests in which the MLE of β is substituted into the first order conditions of another estimator which is consistent under a wider range of alternatives than the MLE is. Hence Ichimura's estimator discussed in the Tobit case would also generate a consistent estimator of β for any underlying density for the u_i (although there are some restrictions on the admissible class coming from the theory of non–parametric estimation). By examining the first order conditions for this estimator, evaluated at the MLE, one could develop a test that might be expected to yield some information about specification errors in the density for u_i. Note that not all semi–parametric estimators could be used in this way. Manski's (1975) maximum score estimator does not yield a root–N consistent estimator of θ, and therefore cannot be meaningfully compared to the MLE of θ from the maintained model.

Two recent papers have focused upon testing for exogeneity of regressors in the probit model utilizing first order conditions. The first of these, Even (1988), adapts the Hausman test along the lines of that suggested by Nelson for the Tobit model. In Even's test the consistent estimator under the alternative is Amemiya's simultaneous equation probit estimator which is inefficient compared to the standard probit estimator under the null hypothesis of exogeneity. Application of the "first order" idea here would see the substitution of the probit MLE's into the equations defining Amemiya's estimator, and this may be more appealing then the direct Hausman test since it does not require estimation under the alternative. Rivers and Vuong (1988) find a simpler two stage estimator under the alternative whose defining first order conditions might also be exploited.

4.2 Multivariate Models

4.2.1 Regression Diagnostic Analogues

Consider now the situation when there is a set C of K alternative actions, and the probability that the i–th individual will chose alternative k from this set is $\bar{P}_{ik}$ which becomes the equivalent of $\bar{F}_i$ in the unvariate case. The data consists of observations on which alternative is chosen by the i–th individual. Define K variables $y_{ik}(k = 1,..,K)$ which take the value unity if alternative k is selected by the i–th individual and zero otherwise. Then, exactly as for the binary alternative situation, $E(y_{ik}) = \bar{P}_{ik}$ and the binary random variable has variance $\bar{P}_{ik}(1-\bar{P}_{ik})$, while if the maintained model is adequate $E(y_{ik}) = P_{ik}$ with variance $P_{ik}(1-P_{ik})$, where $P_{ik} = \bar{P}_{ik}(\theta,\gamma{=}0)$. Consequently, the equivalent of (12) when there are multiple choices is

$$y_{ik} = E(y_{ik}) + \bar{v}_{ik} = \bar{P}_{ik} + \bar{v}_{ik} \tag{20}$$

while that for (14) is

$$y_{ik} = P_{ik} + P'_{ik}\gamma + ... + \bar{v}_{ik}. \tag{21}$$

As in the analysis of (12) and (14) the heteroskedasticity present in the $\bar{v}_{ik}$(v_{ik} when $\gamma = 0$) means that it is more efficient to look at the conditional moment restriction

$$E(N^{-1}\sum_{i=1}^{N} P'_{ik}(P_{ik}^{-1}(1 - P_{ik})^{-1})(y_{ik} - P_{ik})) = 0 \tag{22}$$

than to consider the covariance between P'_{ik} and $(y_{ik} - P_{ik})$.

We are interested in the question of whether some z_i has been invalidly omitted from P_{ik}, i.e., we wish to test if $\gamma = 0$. For the multinomial logit model the expanded probabilities are $\bar{P}_{ik} = e^{x'_{ik}\beta + z_{ik}\gamma}/(\sum_{j \epsilon C} e^{x'_{ij}\beta + z_{ij}\gamma})$ so that $P'_{ik} = z_{ik}P_{ik} - z_{iC}P_{ik}$, where $z_{iC} = \sum_{j \epsilon C} z_{ij}P_{ij}$ and $P_{ik} = e^{x'_{ik}\beta}/\sum_{j \epsilon C} e^{x'_{ij}\beta}$ For those individuals selecting the first alternative, $i \epsilon I_1$, after substituting for P'_{ik} in (22) and noting that $y_{i1} = 1$, their contribution to (22) would be

$$E[N^{-1}\sum_{i \epsilon I_1} (z_{i1} - z_{iC})] = 0. \tag{23}$$

The same relation holds for those individuals opting for alternative 2 but with z_{i1} replaced by z_{i2}. Because each individual must select one alternative the data set is composed of the union of the mutually exclusive sets $I_1, I_2, \ldots, I_K$. Hence, aggregating over all choices, (23) will be

$$E[N^{-1} \sum_{i=1}^{N} \sum_{k=1}^{K} y_{ik}(z_{ik} - z_{iC})] = 0, \tag{24}$$

which is just the score of the log likelihood $N^{-1} \sum_{i=1}^{N} \sum_{k=1}^{K} y_{ik}\log\bar{P}_{ik}$ with respect to γ evaluated at $\gamma = 0$, allowing the interpretation that the score test for $\gamma = 0$ is a test of whether P'_{ik} is correlated with the errors $y_{ik} - P_{ik}$ (for $y_{ik} = 1$). Because of the origin of this test in the linear relation (20), errors for the multinomial probability models are probably best defined as $(y_{ik} - P_{ik})$, or perhaps $P_{ik}^{-1/2}(1 - P_{ik})^{-1/2}(y_{ik} - P_{ik})$ if it is desirable that they have unit variance. McFadden (1987) opts for $P_{ik}^{-1/2}(y_{ik} - P_{ik})$ as his definition of a generalized error, which seems peculiar. Essentially, the reason for a lack of agreement about what should be termed an "error" stems from the fact that we cannot express P'_{ik} as a product of x_{ik} with another quantity, the error; in the linear regression, Tobit and binary choice models the generalized errors were always defined in this way. One advantage of the interpretation and definition offered above is that it applies to any model that specifies a form for P_{ik}, so the use of $\hat{\tau} = N^{-1} \sum_{i=1}^{N} \sum_{k=1}^{K} y_{ik}\hat{P}'_{ik}(\hat{P}_{ik}^{-1}(1-\hat{P}_{ik}))^{-1}(y_{ik} - \hat{P}_{ik})$ as a diagnostic statistic applies to the multinomial probit as well.

4.2.2 Diagnostic Tests from First Order Conditions

Perhaps the most controversial assumption in these models is that known as "Independence from Irrelevant Alternatives" (IIA) which characterizes the multinomial logit model. For this reason investigators have set out to test the restriction, and a number of proposals have been made that can be interpreted as substituting estimates from one set of first order conditions into another.

Hausman and McFadden (1984) compared the estimates of β obtained when the "irrelevant choice" appears in the model and when it is dropped. If the IIA assumption is correct both estimators are consistent, and under an alternative they will generally converge to different probability limits. The "first order" approach would therefore substitute the M.L. estimates obtained from one set of choices into the first order conditions (24) defining the other. A related idea is Horowitz's (1981) and Small and Hsiao's (1985) proposal that the two sets of estimates be found from different halves of the sample and then the log likelihood ratio test be applied to test for "structural stability". Denoting $\hat{\theta}_u$ and $\hat{\theta}_R$ as the unrestricted and restricted estimates of the common set of parameters and $L(\hat{\theta}_u)$ and $L(\hat{\theta}_R)$ as the estimated log likelihoods, Horowitz's test is based on $L(\hat{\theta}_u) - L(\hat{\theta}_R)$. Since the estimators are formed from different samples we can condition upon either one and then $L(\hat{\theta}_u) - L(\hat{\theta}_R)$ will have the standard distributional properties. This situation contrasts with the case when $\hat{\theta}_u$ and $\hat{\theta}_R$ come from the same sample; Hsiao and Small (1985) point out that the likelihood ratio test would then be biased.

4.3 Some Examples

To illustrate the applicability of these test procedures to discrete choice models, two empirical studies involving probit and logit models were chosen. The probit model is that estimated by Cymrot and Dunlevy (1987) in their study of the migration behavior of baseball players, while the logit model examined is that estimated by Volker (1983) dealing with the use of credit cards.

The equation replicated from Cymrot and Dunlevy for this study is that reported in their footnote 14. From that Probit equation, intended to explain moving or staying behavior of baseball players, they calculate the Mills ratio which subsequently enters their wage equation in order to correct for any selectivity bias. It is well known that the auxiliary equation predicting mobility must be correctly specified, otherwise the "selection factor" employed in the wage equation will be invalid and this will lead to incorrect estimates of the determinants of wages (Olsen, 1982). The auxiliary equation was estimated for the two types of players comprising their data set, denoted "eligible" and "ineligible" as reflecting free agency status. Diagnostic tests were performed on the equations for each group and these are presented in Table 4.

Beginning with the "eligibles", the first two rows represent a test for omitted regressors SAL80 and DSA79, respectively salary in 1980 and the difference between the player's and the team's slugging average in 1979, and is the covariance between these variables and the generalized residuals. It might be expected that these variables could have some influence upon the decision to move teams, but neither was included in Cymrot and Dunlevy's equation. Both variables appear to be legitimately excluded. Rows 3 and 4 give the RESET–type test discussed earlier, which was motivated by the possibility of a mis–specified density for the underlying latent variable governing mobility. In this case the test is not passed at conventional levels of significance. Tests for heteroskedasticity provide some evidence that the second moment of the underlying density is not constant. This is important since, unlike the regression

Table 4

Conditional Moment Tests for the Probit Model of Cymrot and Dunlevy (1987)

Test	Moment Restriction	t–Statistic[4] Eligible	t–Statistic[4] Ineligible
Omitted[1,2]	$E(SAL80*\eta)=0$	1.427	1.145
Variables	$E(DSA79*\eta)=0$	1.101	1.365
RESET[3]	$E(PRED^2*\eta)=0$	2.404	2.600
	$E(PRED^3*\eta)=0$	2.287	.525
Heteroskedasticity	$E(SAL80*PRED*\eta)=0$	1.364	.891
	$E(MOVES*PRED*\eta)=0$	2.879	7.670
	$E(RACE*PRED*\eta)=0$	1.090	1.855
	$E(LSA79*PRED*\eta)=0$	1.647	1.349
	$E(MLE*PRED*\eta)=0$	1.449	5.869
	$E(ATBATS*PRED*\eta)=0$	2.258	1.043
	$E(ADJS*PRED*\eta)=0$	2.194	.083
	$E(DFN*PRED*\eta)=0$	1.981	4.626
	$E(BYR*PRED*\eta)=0$	2.197	2.595
	$E(YSRM*PRED*\eta)=0$	2.541	3.557

Notes:

1. For a full discussion of all variables see Cymrot and Dunlevy. Variable names come from that article.
2. η denotes the generalized error.
3. PRED denotes $x_i'\hat{\beta}$.
4. Absolute values.

model, the presence of heteroskedastic errors means inconsistent estimators of the parameters of the Probit model.

The major flaws that arise with the "eligibles" equation appear to also exist in the "ineligibles" equation. The t–statistics on the RESET type tests imply that the assumption of normality required by probit is unlikely to be correct. Furthermore, the tests for heteroskedasticity show major defects in this area. Overall, the diagnostics indicate that both equations are clearly mis–specified.

As with Fair's Tobit model of extra–marital affairs, the estimates of Cymrot and Dunlevy's two probit equations were inserted into Ichimura's first order conditions and evaluated. These are reported in Table 5. Let us again begin with the "eligibles". Somewhat surprisingly, in the light of Table 4, on the basis of the "parametric" based tests there are only two cases where the non–parametric first order conditions do not seem to have their expected value of zero. An examination of the ineligible equation also provides only a small number of rejections. Again this is unexpected given that the earlier parametric tests identified serious difficulties with the model. The evidence in Table 5 therefore, when combined with that in Table 4, may be revealing that these non–parametric tests are not very powerful.

Volker (1983) fitted an equation to determine what influences the decision of an individual to hold the credit card issued by Australian banks, Bankcard, and the equation we subject to test is that reported in column (i) of his Table 1, although we used thirty–seven more observations to give a complete sample of N=1547 . Results of the diagnostic tests are reported in Table 6. The first two rows give the RESET type tests, and the logit assumption seems satisfactory. In fact it is only in a few of the "heteroskedasticity" tests that there is any evidence of problems with the model. In particular, it seems as if there may be different densities for females and for card holders who live in Brisbane and Sydney. Perhaps it should be emphasised that the tests here, although designed to look for a particular type of distributional

Table 5

First Order Conditions For Ichimura's

Single Index Estimator Evaluated at

Cymrot and Dunlevy's MLE's

Variable	Eligible	Ineligible
MOVES	.130	1.732
RACE	.011	2.347
LSAFG	.835	.781
MLE	2.409	.142
ATBATS	.637	2.986
ADJSA79	.760	.770
DEN	.328	.227
BYR	2.602	.028
YJRM	.019	2.300

mis-specification, could also be picking up problems in the specification of the determinants. It may be that multiplicative terms involving sex and geographical location should be included amongst the x_i and the ommision of these is what the diagnostic tests are detecting.

5. Selectivity

Perhaps the most pervasive of all the problems in the analysis of individual data sets is that coming from the self-selection of individuals into or out of a sample. The simplest representation of self-selection is the Type I Tobit model (Amemiya, 1985),

$$z_i^* = w_i'\alpha + v_i \qquad i=1,\ldots,N_1 \tag{25}$$

$$y_i = x_i'\beta + u_i \qquad i=1,\ldots,N_1 \tag{26}$$

Table 6

Conditional Moment Tests for the Logit Model of Volker (1983)

Test	Moment Restriction	t–Statistic[2]
RESET[1,3]	$E(PRED^2 * \eta)=0$	.330
	$E(PRED^3 * \eta)=0$	.068
Heteroskedasticity[4]	$E(AGE16\text{–}19 * PRED * \eta)=0$	.377
	$E(AGE20\text{–}24 * PRED * \eta)=0$	.354
	$E(AGE35\text{–}44 * PRED * \eta)=0$	1.042
	$E(AGE45\text{–}54 * PRED * \eta)=0$	.149
	$E(AGE55 * PRED * \eta)=0$	.946
	$E(PROF * PRED * \eta)=0$	1.360
	$E(SKIL1 * PRED * \eta)=0$	.707
	$E(UNSKILL * PRED * \eta)=0$	.149
	$E(FEMALE * PRED * \eta)=0$	2.982
	$E(SINGLE * PRED * \eta)=0$	.490
	$E(WID/DIV * PRED * \eta)=0$	.765
	$E(SYD * PRED * \eta)=0$	2.572
	$E(BRIS * PRED * \eta)=0$	3.707
	$E(ADEL * PRED * \eta)=0$	.697
	$E(PERTH * PRED * \eta)=0$	.731

Notes:

1. PRED denotes $x_i'\hat{\beta}$.
2. Absolute values.
3. η denotes the generalized errors.
4. For a full discussion of the variables employed see Volker (1983). The variable names are taken from that article.

where z_i^* is a latent variable with associated binary indicator z_i. Observations on y_i and x_i are only available for those individuals i=1,...,N who participate in the sample (z_i=1), and hence the sample is selected according to values of w_i and v_i. Whenever the u_i and v_i are correlated the error term in (26), for i=1,...,N, will have a non-zero mean dependent upon the level of w_i. Consequently, the OLS estimator of β formed by regressing y_i against x_i is inconsistent.

Estimation of the parameters has traditionally proceeded under the maintained assumption of bivariate normality of u_i and v_i. From the properties of the conditional density of u_i given v_i , $u_i = \rho v_i + e_i$, where $E(e_i v_i) = 0$. Sustituting for u_i in (26)

$$y_i \;=\; x_i'\beta + \rho v_i + e_i \,, \qquad i=1,....,N_1 \qquad (27)$$

from which $E(y_i|z_i=1) = x_i'\beta + \rho E(v_i|z_i=1)$ and (27) has the regression form

$$y_i \;=\; x_i'\beta + \rho E(v_i|z_i=1) + \epsilon_i \,. \qquad i=1,....,N \qquad (28)$$

5.1 Regression Diagnostics Analogues

Selectivity bias in the OLS estimator of β arises because of the presence of $\rho E(v_i|z_i=1)$ in the error term of (28), being absent if $\rho = 0$. Since ρ is proportional to $E(u_i v_i)$, repeating the same analysis as for the exogeneity of regressors in a Tobit model given earlier indicates that a suitable moment condition to examine is $E(u_i E(v_i|z_i=1)) = E(u_i \eta_i) = 0$, with η_i being the generalized errors for the Probit model (25). The sample moment will be $\hat{\tau} = N^{-1}\Sigma \hat{u}_i \hat{\eta}_i$, where unknown parameters β and α are replaced with their MLE's. For the Type I Tobit models, under the null hypothesis of $\rho = 0$, the MLE of β is the OLS estimator, while the MLE of α is the single equation Probit estimator. For generalizations of (25) such as the Type 2 Tobit model, in which z_i^* is observed when positive, the MLE of β and α cannot be obtained so easily, although one might ignore the fact that more than the sign of z_i^* is observed. In fact, in order to implement the test statistic of section 2 such information could be discarded without affecting the computation of the asymptotic variance of $N^{1/2}\hat{\tau}$, since the simplifications just require that a likelihood approach is being used.

The standard test in the literature for selection bias, Heckman (1979), is the t–test that the coefficient of $\hat{\eta}_i$ is zero in the regression of y_i against x_i and $\hat{\eta}_i$ (the $\hat{\eta}_i=\hat{\phi}_i/\hat{\Phi}_i$, where $\hat{\phi}_i$ and $\hat{\Phi}_i$ are evaluated with $w_i'\hat{\alpha}$). This t–statistic is proportional to $N^{-1}\Sigma(y_i-x_i'\hat{\beta}^*-\hat{\eta}_i\hat{\rho}^*)\hat{\eta}_i$, and so differs from $\hat{\tau}$ only in the use of the two step estimators of β and ρ, β^* and ρ^*, rather than the MLE of β and $\rho = 0$. But under the null hypothesis this substitution has no effect asymptotically.

As noted in the introduction the majority of the diagnostic tests developed in this literature have employed the Lagrange multiplier framework and the method consequently requires the evaluation of the information matrix. This is precisely the strategy adopted by Lee and Maddala (1985) in their examination of several issues concerning selectivity. Their approach is to estimate a relatively simple model and then test up in the direction of selectivity. In fact, their approach is based upon conditional moment restrictions defined by the fact that the conditional expectation of the score with respect to ρ of more complex models should have zero expectation if ρ=0.

As evident from the above discussion normality plays a key role in the estimation of selectivity models, making it critical to determine how accurate the assumption of normality is. Oddly enough there appears to be very little written upon checking this assumption. There are a number of ways that one could proceed to a test for normality of the u_i and v_i. The easiest is to apply the standard diagnostic tests for normality in the probit (or Tobit if it is Type 2) model (25) previously set out in sections 3 and 4, and this gives a test for the marginal density of v_i being normal. It is much harder to do a test for marginal normality of the u_i since the error term in the two–step regression is ϵ_i, not u_i, and this is certainly not normal. For this reason it seems useful to borrow some insights from the semi–nonparametric literature and to design a test that might be usefully applied.

The essence of the two–step procedure is to find $E(u_i|z_i=1)$, for which a preliminary step is to obtain $E(u_i|v_i)$. Hence, an expression for this conditional expectation when the density of u_i and v_i is not normal is needed. We borrow an idea from Gallant and Nychka (1987) of replacing the unknown density by the approximation $f_{uv} = (\sum_{k=0}^{K} \sum_{j=0}^{J} \gamma_{kj} u^k v^j)\phi_{uv}$, where ϕ_{uv} is the bivariate normal density of u_i and v_i , and $\gamma_{00} = 1$. With this density

$$E(u|v) = \int u f_{u|v} du = \int u(f_{uv}|f_v)du = \sum_k \sum_j b^{-1}\gamma_{kj}(\int u^{k+1}\phi_{u|v}du)v^j \tag{29}$$

where $f_{u|v}$ is the conditional density of u_i given v_i, $\phi_{u|v}$ is the conditional normal and $b = f_v/\phi_v$. Setting $K = 0$,

$$E(u|v) = \sum_j b^{-1}\gamma_{0j}(\int u\phi_{u|v}du)v^j \tag{30}$$

$$= \sum_j b^{-1}\gamma_{0j}\rho v^{j+1} \tag{31}$$

and therefore, under the null hypothesis,

$$E(u_i|z_i=1)=E(\ v_i|z_i=1) + \gamma_{01}E(\ v_i^2|z_i=1) + \ldots + \gamma_{0J}E(v_i^J|z_i=1), \tag{32}$$

as $b = 1$ when the errors are distributed bivariate normal because $f_v = \phi_v$.

Testing if γ_{kj} equals zero therefore is a test for normality. Amemiya (1973) gives expressions for $E(v^j|z_i=1)$ (j=1,2,3,4), and these are proportional to ϕ_i/Φ_i, $(1-w_i'\alpha(\phi_i/\Phi_i))$, $(\phi_i/\Phi_i)[(w_i'\alpha)^2+2]$, $[3-3w_i'\alpha(\phi_i/\Phi_i)-(w_i'\alpha)^3(\phi_i/\Phi_i)]$ respectively. Hence a test for the normality assumption is to add on the variables $(w_i'\hat{\alpha})^j\hat{\phi}_i/\hat{\Phi}_i$, j=1,2,3 to the two step estimator and test if these are jointly zero. It is interesting to observe that this is like a RESET test in which the predictions from the selection equation are powered up, although they are weighted by a function of the Mill's ratio. It would also be possible to allow K to be non–zero, and it may be that setting K and L to be equal to one another (as was done by Gallant and Nychka) would provide a better test. However from the viewpoint of a diagnostic test the null distribution remains the same regardless of the value of K that is chosen. Further investigation of this question is needed, but at least the addition of the above regressors provides a simple way to do some testing for the normality assumption.

A similar idea to that advanced above has been put forth by Lee (1984). He approximates the underlying bivariate density as the product of a normal density and a series of Hermite polynomials, i.e. $f_{uv} = \phi_{uv}[1 + \sum\sum_{r+s\geq 3} \alpha_{rs}H_{rs}(u,v)]$, where $H_{rs}(u,v)$ are bivariate Hermite polynomials. When $\alpha_{rs} = 0$, one obtains normality, and Lee tests if $\alpha_{rs} = 0$ using the Lagrange Multiplier test associated with the approximating density f_{uv}. Because it is bivariate normality which is being tested, the scores are quite complex and this test seems to have had little application. Moreover, the fact that most of this literature utilizes the two–step estimator suggests that augmenting that equation with variables as in (32) is likely to be more appealing. Lee, in fact, restricts r+s to be a maximum of four and observes (his eq. (2.15)) that $E(u|v) = \rho v + K_{12}[(v^2-1)/2] + K_{13}[(v^3-3v)/6]$ when his approximating density is employed. With normality K_{12} and K_{13} are zero, and it is clear from (31) that one would obtain the same test for normality following his approach as what we derived above from Gallant and Nychka's work.

5.2 Diagnostic Tests from First Order Conditions

A final approach to generating diagnostic tests that also emphasises the question of correct density is to use the first order conditions from Powell's recent (1987) semi–parametric estimator for bivariate latent variable models. Powell solves for β

from the first order conditions $\psi_i = \sum_{j=i+1}^{N} (\bar{z}_i - \bar{z}_j) K_{ij}[(y_i - y_j) - (x_i' - x_j')\beta]$, where $K_{ij} = h^{-1}K((w_i'\hat{\alpha} - w_j'\hat{\alpha})/h)$, K(.) is a kernel, h a bandwidth, and $\bar{z}_i$ are p instruments constructed from the w_i. Powell shows that the estimator of β from these first order conditions is root–N consistent provided $\hat{\alpha}$ is, h goes to zero as $N^{-1/7}$, and K(.) is a kernel with moments up to third order equal to zero. To achieve the latter requirement it is necessary to work with a "bias reducing higher order" kernel. Robinson's (1987) method of generating one is employed later with the base kernel being Gaussian, and h set to the product of the standard deviation of $w_i'\hat{\alpha}$ and $N^{-1/7}$. The test statistic $\hat{\tau} = N^{-1}\Sigma\, \hat{\psi}_i$ is evaluated with maximum likelihood estimates of β and α obtained from LIMDEP. It might have been preferable to use the two–step estimator of β but then it is not possible to adopt the general method of computing the asymptotic variance given in section 2.

5.3 An Example

As Cymrot and Dunlevy (1987) also estimate wage equations by employing the two step sample selection procedure we again analyze their data to estimate the applicability of the tests discussed above. The first of these is the RESET type test for normality discussed in 5.1 based on the null hypothesis that the powered up values of the predictions weighted by the Mills ratio are correctly excluded from the equation. The tests were performed for all four groups comprising Cymrot and Dunlevy's sample and are reported in Table 7. The included variables are Pred*Mills ratio, Pred^2*Mills ratio and Pred^3*Mills ratio, and the F value reported is for the null that their coefficients are jointly zero. The results in Table 7 are surprising as they do not reflect any signs of serious misspecification. In fact, in only one instance, eligible movers, was the F value even close to resulting in a rejection of the null at conventional levels. This result is somewhat disturbing given that the parametric based tests revealed severe misspecification in this model earlier. The results in Table 7 may reflect that the non-parametric tests are less powerful or may in part indicate that larger samples than those being employed are necessary.

TABLE 7

"RESET" Test of Cymrot and Dunlevy's Selectivity Model

Group	Pred*W (1)	Pred^2*W (2)	Ped^3*W (3)	(4)
Ineligible Movers	1.009	1.165	.211	.618
Ineligible Non–Movers	.654	.278	.413	.301
Eligible Movers	.771	.803	.529	.434
Eligible Non–Movers	.375	.849	.603	.475

Notes:

1. For movers W is equal to ϕ(MOVE)/Φ(MOVE) while for non-movers W equals $-\phi$(MOVE)/1$-\Phi$(MOVE) where MOVE denotes the values from the auxiliary equation.
2. Columns (1), (2), and (3) contain the absolute value of the t–statistics for the null that the coefficient is equal to zero.
3. Column (4) contains the significance level of the χ^2 value for the null that the parameters corresponding to cols 1,2,3 are zero.
4. All tests are adjusted for heteroskedasticity.

The model reported by Cymrot and Dunlevy was also examined by evaluating the first–order conditions of Powell's semi–parametric estimator implied by Cymrot and Dunlevy's estimates. To perform this evaluation it would be preferable to utilize maximum likelihood estimates of the parameters, but the iterative algorithms in LIMDEP were unable to reach a maximum of the likelihood. The two step estimates were therefore substituted instead and the estimated first order conditions were regressed against an intercept. This procedure produces a bias in the estimated test statistics since it ignores the sampling variablility in the estimated coefficients. The absolute value of the t–test for the null that the intercept is equal to zero is reported in Table 8. These tests were only performed for the non–movers as the large number of regressors in this model severely reduced the degrees of freedom. The results reported in Table 8 are consistent with the findings of section 4.3, with the model being mis–specified for both groups.

TABLE 8

First Order Conditions for Powell's Semi–Parametric Estimator Evaluated at Cymrot and Dunlevy's Estimates

Variable	Group	
	Eligible Non–Movers	Ineligible Non–Movers
MOVES	5.988	6.559
RACE	5.415	6.104
LSA79	8.617	5.419
MLE	5.647	5.086
ATBATS79	3.040	5.802
POP3	0.483	6.424
GR3	.216	6.314
W79	1.345	1.202
ADJSA79	7.002	6.585
DSA79	4.259	7.505
DFN	3.636	4.661
BYR	1.809	5.490
Observations	122	88

6. Duration Models

A further family of models which has experienced popularity with the advent of increased availability of unit record data is that based upon duration data (see Kiefer (1988) for example). The core of these models is the hazard function h(t) of a random variable T, the length of time spent by an individual in a given state, that has cumulative distribution function $F(t) = 1-\Pr(T>t)$ and density f(t). As discussed in Amemiya (1985, p.435), this function h(t) is the conditional probability of exiting from a specified state during an interval $t+\Delta t$ given that the individual was in another state at t. For an observed set of data upon completed spell lengths t_i, $i=1,\ldots,N$ it is easy to write down the likelihood as $\prod_i f(t_i)$ provided $f(t_i)$ can be specified. Equivalently, formulation of an $h(t_i)$ will suffice to determine an $f(t_i)$ and this literature usually proceeds to specify the hazard function.

A popular assumption for $h(t_i)$ is that it varies with characteristics of individuals as $\exp(x_i'\beta)$. Straightforward integration then gives $f(t_i) = \exp(x_i'\beta)\exp[-\exp(x_i'\beta)t_i]$ and the log likelihood will be

$$L = \Sigma x_i'\beta - \Sigma t_i \exp(x_i'\beta), \tag{32}$$

where t_i is the spell length experienced by individual i. Differentiating (32) with respect to β gives the score vector as $\Sigma x_i(1-t_i\exp(x_i'\beta))$, and arguing by an analogy with regression, it is reasonable to claim that $\nu_i = (1-t_i\exp(x_i'\beta))$ is a generalized error. In fact Lancaster (1985) has referred to $t_i\exp(x_i'\beta)$ as the generalized error, although it seems sensible to subtract off its expectation of unity so as to agree with the notion that an error should have a zero mean. When spells are incomplete the log likelihood in (32) has to be modified in that only the the second part of it is relevant for those with incomplete spells. Hence, the generalized residuals for individuals with incomplete spells are $-t_i\exp(x_i'\beta)$ and the moment conditions need to be varied to reflect this.

6.1 Regression Diagnostic Analogues

Although the assumption that h depends upon characteristics in the exponential form is a convenient one, mainly because it forces non-negativity upon estimates of h, there is no certainty that this functional relation is correct, or that the x_i capture all of the determinants. Failure of these assumptions might be expected to show up as patterns in the generalized errors, and this suggests that the moment conditions $E(z_i\nu_i) = 0$ be checked, where z_i are variables chosen in light of the postulated deficiency. Amemiya (1985, p.440) shows that, for data on completed spells, when $h_i = \exp(x_i'\beta)$, $E(t_i) = \exp(-x_i'\beta)$. It follows that a RESET-like test can be constructed by choosing z_i as the powers of $\exp(-x_i'\hat{\beta})$.

The hazard function adopted above had the characteristic that it was not a function of t, that is it exhibited no duration dependence. In practice an important characteristic of actual data is the fact that the probability of exiting from a state is a function of the amount of time spent in it, making the testing for a relation between h_i and t_i a concern of a number of authors. Kiefer (1985), and later Sharma (1987), modified the conditional density $f(t_i)$ described above to $f^*(t_i) = f(t_i)[1 + \sum_{j=2}^{n} \gamma_j L_{ji}]$, where $L_{ji} = L_j(t_i\exp(x_i'\beta))$ and $L_j(\omega)$ is the j'th Laguerre polynomial. Kiefer (1985, p. 153) gives the first nine of these polynomials, but it is worth noting that $L_j(\omega)$ is a j'th order polynomial in $(1-\omega)$. With this expanded density the likelihood is now $\Pi f^*(t_i)$ and a test for the absence of duration dependence is available from the Lagrange multiplier test that $\gamma_2,\ldots,\gamma_n$ are zero. Since the scores for γ_j, under the null hypothesis that $\gamma_j=0$, $j=2,\ldots,n$, are L_{ji}, this amounts to testing the conditional moment restriction $E(L_{ji}) = 0$. But L_j is a j-th order polynomial in $\omega = 1-\exp(x_i'\beta)t_i$, showing that this conditional moment restriction effectively tests if the moments of the generalized errors have their predicted values when there is no duration dependence. To illustrate this, take j=2, giving $L_2 = 1/2[\omega^2-4\omega+2] = 1/2[\{(\omega-1)^2-1\}-2(\omega-1)]$.

From Lancaster (1985, p.157, eq.(10)), when there is no duration dependence $E(\omega^j) = j!$, and therefore $E[(\omega-1)^2] = 1$, $E(\omega-1) = 0$.

Because the Laguerre polynomials encompass many popular specifications of duration dependence, including the lognormal, gamma, Weibull and Pareto densities, the specification test should be a useful one. Kiefer (1985) found that this was so when analysing some data from the Denver Income Maintenance Experiment. Perhaps the main difficulty he experienced was in deriving the variance of L_{ij} when it is evaluated at $\hat{\beta}$. Kiefer (1985) ignores the fact that β has been estimated, arguing in footnote 12 that this yields a conservative test. Actually, this is incorrect. The situation is exactly that analysed by Durbin (1970), and the true variance of the test statistic is smaller than the computed variance. The method of calculating standard errors in section 2 obviates this problem. However, closer inspection of L_{ij} reveals that $E(\partial L_{ij}/\partial\beta) = 0$ when the $\gamma_j(j=2,..,n)$ are all zero, since the derivative can be written as a linear combination of the j Laguerre polynomials, and each of these has zero expectation. This can be illustrated for j=3 when the $L_j = (1/6)(-\omega^3+9\omega^2-18\omega+6)$ and $\partial L_j/\partial\beta = (1/6)(-3\omega^2+18\omega-18)(\partial\omega/\partial\beta) = (1/6)(-3\omega^2+18\omega-18)x_i(1-\omega)$, where we have dropped the i subscripts from ω. Now this has zero expectation if $E(-\omega^3+6\omega^2-6\omega) = 0$. But the term in brackets can be written as $(\{-\omega^3+9\omega^2-18\omega+6\}-\{3\omega^2-12\omega+6\}) = (1/6)L_3-(3/2)L_2$, and thus has zero expectation under the null hypothesis. Consequently, the information matrix is block diagonal between β and γ and it is possible to compute the variance of the test as if the β were known. Kiefer's formula is therefore correct and the scores could be excluded from the regression of section 2. This result is unlikely to extend out of the exponential hazard framework, but in other instances the effects of estimated parameters can be allowed for as in section 2.

In fact, it is common for some assumption to be made about duration dependence, and the relevant question is really whether the allowance made has been sufficient. A popular specification intended to allow for dependence is to form the proportional hazards model with function $h_i = \alpha t_i^{\alpha-1}\exp(x_i\beta)$, where the term $\alpha t_i^{\alpha-1}$ comes from a Weibull density. The log likelihood corresponding to (32) is then

$$\Sigma[\ln \alpha + (\alpha-1)\ln t_i + x_i'\beta - t_i^{\alpha}\exp(x_i'\beta)]. \tag{33}$$

There are now two types of scores. Those for β remain as they were earlier but there is also one for α. Hence, there can be no unambiguous definition of a generalized error. Nevertheless, it should be true that $E(z_i d_i) = 0$ for suitably chosen z_i, where d_i are the vector of scores with respect to β and α. When an intercept term appears among the β this implies that the covariance of the z_i and the earlier definition of generalized errors should be zero. But clearly there is a broader set of orthogonality conditions that must be satisfied in this expanded model. Sharma (1987) constructs tests for duration dependence not captured in a Weibull formulation by multipling the density for the proportional hazards model with Laguerre polynomials, just as was done for the exponential hazard model, except that the argument in his Laguerre polynomials is $\omega = t_i\exp(x_i'\beta)^{1/\alpha}$.

A last item of concern in the duration literature has been that of neglected heterogeneity. In this the hazard is replaced by $\exp(x_i'\beta+e_i)$ where e_i is an i.i.d .error term that has mean zero and variance σ_e^2 and is distributed independently of the x_i. As is well known neglected heterogeneity induces duration dependence unless σ_e^2 is zero. Lancaster (1985) and Kiefer (1985) both develop a test statistic by noting that, conditional upon e_i, the density of the t_i is $\bar{f}(t_i|u_i) = \exp(x_i'\beta+e_i)\exp(-t_i\exp(x_i'\beta+e_i))$. They expand this around $e_i = 0$ (implying $\sigma_e^2 = 0$) retaining only the first two terms. The unconditional density of t_i is then $f_i^* = f_i\{1+f_i^{-1}\{(1/2)\sigma_e^2(\partial^2 f(t_i|u_i=0)/\partial u_i^2)\}$ and they propose to test if σ_e^2 is zero by evaluating the score for σ_e^2 with f_i^* as the density making up the likelihood. Obviously this density looks very close to the Laguerre approximation and, indeed, Sharma (1987) shows that the second term in the curly brackets is in fact proportional to L_{i2}. So, as might be expected given the connection between heterogeneity and duration dependence, the tests found by these different methods are actually equivalent.

Although the emphasis above was on likelihoods, in a number of the models it is possible to regard the determinants of duration as being capable of formulation as a regression, albeit generally censored. Horowitz and Neumann (1988) point out that one can formulate many of these models as $t_i = \min[\exp(x_i'\beta)v_i, t^*]$, where t^* is the right censoring point and v_i is an error term with a density to be described. After taking logs an equivalent form would be $\log t_i = \min[x_i'\beta + \log v_i, \log t^*] = \min[x_i'\beta + \nu_i, \log t^*]$ and this is a censored regression model of the Tobit type. To estimate, the density of γ_i must be prescribed, for example as Weibull, and it is possible to then compare the empirical density function of the residuals $\hat{\nu}_i = \log t_i - x_i'\hat{\beta}$ with the postulated theoretical one using the Kaplan–Meier estimate, just as described earlier for the Tobit model. They illustrate the utility of this method by applying it to the strike duration data in Kennan (1985). Of course, many of the tests of model adequacy set out for the Tobit model can also be applied in this context as well, e.g., the information matrix test, RESET–type tests, etc.

6.2 Diagnostic Tests from First Order Conditions

A range of alternative estimators of β is available for the censored regression model, and the first order conditions from each of these might be applied to give diagnostic test statistics. Examples would be quantile estimators or Powell's (1984) least absolute deviations (LAD) estimator. Horowitz and Neumann canvass these and illustrate their use with Kennan's strike data.

7. Conclusion

This paper has attempted to survey the wide variety of tests for specification error that is emerging for models based on individual data. Because this literature is very diverse, both in the type of model and the nature of the tests that are appropriate for each, a simplifying framework was needed in order to place some order upon it. One approach would be to work through the Lagrange Multiplier perspective, but we believe that this sometimes obscures the simple rationale for many of the tests. In particular, we argued that the questions of how to formulate a basis for a test and how to compute its variance should be divorced, as it is the tight linking of these issues in most existing surveys of the area which is primarily responsible for their complexity. Consequently, we borrowed an idea from work by Newey (1985) and Tauchen (1985) which argues that the basis of all tests is the formulation of suitable conditional moment restrictions. Their treatment of how to get the variance was also exposited, as it can be a useful procedure in some instances.

Having disposed of the problem of determining a variance the remainder of the paper systematically sets out the relevent conditional moment restrictions for censored regression models (Tobit, selectivity and duration) and discrete choice models (Probit, Logit and their multinomial versions). Three categories are used as the organising framework. First, we argue that many of the restrictions can be thought of as analogues of tests for specification error, heteroskedasticity, and normality in the linear regression model. A second grouping can be thought of as arising from the substitution of the parameter estimates obtained by (say) maximum likelihood into the first order conditions of another estimator which is consistent under certain mis-specifications. Under this heading we propose a number of new tests derived from the burgeoning literature on non-parametric estimation of individual data models. Finally, there is the possibility of exploiting some of the sample for validation purposes, and we refer to this group of tests as "prediction" tests. As well as the theoretical work we carry a number of examples, drawn from recent literature, through the paper, in order to show that there is a serious need to subject these estimated models to scrutiny.

REFERENCES

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

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

Amemiya, T. (1985), Advanced Econometrics (Harvard University Press, Cambridge, Mass).

Anderson, G. (1987), "Prediction Tests in Limited Dependent Variable Models," Journal of Econometrics, 34, 253–262.

Andrews, D.W.K. (1987), "Chi–Square Diagnostic Tests for Econometric Models: Theory," Econometrica (forthcoming).

Beran, R. (1977), "Adaptive Estimates for Autoregressive Processes," Annals of the Institute of Mathematical Statistics, 28, 77–89.

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

Breusch, T. S. and A. R. Pagan (1979), "A Simple Test for Heteroscedasticity and Random Coefficient Variation," Econometrica, 47, 1287–1294.

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

Beran, R. (1977), "Adaptive Estimates for Autoregressive Processes," Annals of the Institute of Statistical Mathematics, 28, 77–89.

Cameron, C. and P. Trivedi (1985), "Regression Based Tests for Overdispersion", Stanford Department of Economics Technical Paper No. 9.

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

Chesher, A. and M. Irish (1987), "Residual Analysis in the Grouped Data and Censored Normal Linear Model," Journal of Econometrics, 34, 33–62.

Chesher,A. and R. Spady (1988), " Asymptotic Expansions of the Information Matrix Test Statistic", (mimeo, Bristol University)

Chesher, A., T. Lancaster and M. Irish (1985), "On Detecting the Failure of Distributional Assumptions," Annales de l'Insee, 59/60, 7–44.

Cox, D.R and E.J. Snell (1968), "A General Definition of Residuals", Journal of the Royal Statistical Society, Series B, 39, 248–275.

Cymrot, D. and J. Dunlevy (1987), "Are Free Agents Perspicacious Peregrinators?" Review of Economics and Statistics, 69, 50–58.

Davidson, J. and MacKinnon (1984), "Convenient Specification Tests for Logit and Probit Models," Journal of Econometrics, 25, 241–262.

Durbin, J. (1970), "Testing for Serial Correlation in Least Squares Regression when Some of the Regressors are Lagged Dependent Variables," Econometrica, 38, 410–421.

Even, W. (1988), "Testing Exogeneity in a Probit Model," Economic Letters, 26, 125–128.

Fair, R. F. (1978), "A Theory of Extramarital Affairs," Journal of Political Economy, 86, 45–61.

Gallant, A. R. and D. W. Nychka (1987), "Semi–Nonparametric Maximum Likelihood Estimation," Econometrica, 55, 363–390.

Godfrey, L. G. (1978), "Testing for Multiplicative Heteroscedasticity," Journal of Econometrics, 8, 227–236.

Gourieroux, C., A. Monfort, E. Renault and A. Trognon (1987a), "Generalized Residuals," Journal of Econometrics, 34, 5–32.

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

Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," Econometrica, 50, 1029–1054.

Hausman, J. J. (1978), "Specification Tests in Econometrics," Econometrica, 46, 1251–1272.

Hausman, J. J. and D. McFadden (1984), "Specification Tests for the Multinomial Logit Model," Econometrica, 52, 1219–1240.

Heckman, J. J. (1974), "Shadow Prices, Market Wages and Labor Supply," Econometrica, 42, 679–694.

Heckman, J. J. (1979), "Sample Selection Bias as a Specification Error," Econometrica, 47, 153–161.

Heckman, J. J. (1984), "The χ^2 Goodness of Fit Statistic for Models with Parameters Estimated from Microdata," Econometrica, 53, 1543–1547.

Hoffman, D. and A. R. Pagan (1988), "Post–Sample Prediction Tests for Generalized Method of Moment Estimators," Rochester Center for Economic Research Working Paper No. 129, University of Rochester.

Horowitz, J. L. (1981), "Identification and Diagnosis of Specification Errors in the Multinomial Logit Model," Transportation Science, Vol. 15b, 345–360.

Horowitz, J. L. (1985), "Testing Probabilistic Discrete Choice Models of Travel Demand by Comparing Predicted and Observed Aggregate Choice Shares," Transportation Science, Vol. 19b, 17–38.

Horowitz, J. L. (1987), "Specification Tests for Nested Logit Models," Environment and Planning, Vol. 19, 395–402.

Horowitz, J. L. (198?), "Specification Tests for Probabilistic Discrete Choice Models of Consumer Behavior," forthcoming in Behavioral Modelling in Geography and Planning," Eds., H. Timmermans and R. Colledge.

Horowitz, J. L. and G. R. Neumann (1988), "Specification Testing in Censored Regression Models: Parametric and Semiparametric Models" (mimeo, University of Iowa).

Ichimura, H. (1987), "Consistent Estimation of Index Model Coefficients," (mimeo, M.I.T.).

Kaplan, E. L. and P. Meier (1958), "Nonparametric Estimation from Incomplete Observations," Journal of the American Statistical Association, 53, 457–481.

Kennan, J. (1985), "The Duration of Contract Strikes in U.S. Manufacturing," Journal of Econometrics, 28, 5–28.

Kennan, J. and G. R. Neumann (1988), "Why Does the Information Matrix Test Reject Too Often," Working Paper No. 88–4, University of Iowa.

Kiefer, N. M. (1984), "A Simple Test for Heterogeneity in Exponential Models of Duration," Journal of Labor Economics, 2, 539–546.

Kiefer, N. M. (1985), "Specification Diagnostics Based on Laguerre Alternatives for Econometric Models of Duration," Journal of Econometrics, 28, 135–154.

Kiefer, N. M. (1988), "Economic Duration Data and Hazard Functions," Journal of Economic Literature, 26, 646–679.

Lancaster, T. (1985), "Generalized Residuals and Heterogenous Duration Models," Journal of Econometrics, 28, 155–169.

Lee, L. F. (1984), "Tests for the Bivariate Normal Distribution in Econometric Models with Selectivity," Econometrica, 52, 843–863.

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

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

McFadden, D. (1987), "Regression Based Specification Tests for the Multinomial Logit Model," Journal of Econometrics, 34, 63–82.

Maddala, G. S. (1983), Limited Dependent and Qualitative Variables in Econometrics (Cambridge University Press, New York).

Nelson, F. D. (1980), "A Test for Misspecification in the Censored Normal Model," Econometrica, 48, 1317–1329.

Newey, W. (1985), "Maximum Likelihood Specification Testing and Conditional Moment Tests," Econometrica, Vol. 53, 1047–1070.

Newey, W. (1986), "Specification Tests for Distributional Assumptions in the Tobit Model," Journal of Econometrics, 34, 125–146.

Olsen, R. (1982), "Distributional Tests for Selectivity Bias and a More Robust Likelihood Estimator," International Economic Review, Vol. 23, 223–240.

Pagan, A. (1984), "Model Evaluation by Variable Addition," in D. F. Hendry and K. F. Wallis (eds) Econometrics and Quantitative Economics, (Blackwell, Oxford) 103–133.

Pagan, A. R. and A. D. Hall, (1983), "Diagnostic Tests as Residual Analysis," Econometric Reviews, 2, 159–218.

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. (1987), "Semiparametric Estimation of Bivariate Latent Variable Models," (Mimeo, University of Wisconsin).

Pregibon, D. (1981), "Logistic Regression Diagnostics," The Annals of Statistics, Vol. 9, pp. 705–724.

Ramsey, J. B. (1974), "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," Journal of the Royal Statistical Society, B, 31, 350–371.

Rivers, D. and Q. Vuong (1988), "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models," Journal of Econometrics, forthcoming.

Robinson, P. M. (1987), "Root–N Consistent Semiparametric Regression," Econometrica, forthcoming.

Robinson, P. M. (1988), "Semi–Parametric Econometrics: A Survey," Journal of Applied Econometrics, 3, 35–52.

Robinson, P., A. Bera and C. Jarque (1985), "Tests for Serial Dependence in Limited Dependent Variable Models," International Economic Review, Vol. 26, 629–638.

Ruud, P. (1984), "Tests of Specification in Econometrics," Econometric Reviews, 3, 211–242.

Salkever, D. S. (1976), "The Use of Dummy Variables to Compute Predictions, Prediction Errors and Confidence Intervals," Journal of Econometrics, 4, 353–397.

Sharma, S. (1987), "Specification Diagnostics for Econometric Models of Durations," U.C.L.A. Working Paper.

Small, K. and C. Hsiao (1985), "Multinomial Logit Specification Tests," International Economic Review, 26, 619–627.

Smith, R. and R. Blundell (1986), "An Exogeneity Test for a Simultaneous Equation Tobit Model with an application to Labor Supply," Econometrica, Vol. 54, 679–685.

Tauchen, G. (1985), "Diagnostic Testing and Evaluation of Maximum Likelihood Models," Journal of Econometrics, 30, 415–443.

Taylor, L. W. (1985), "Comparing RESET and Hausman's Procedure for Detecting Misspecification in the Probit Model," (paper presented at the Fifth World Congress of the Econometric Society).

Taylor, L. W. (1987), "The Size Bias of White's Information Matrix Test," Economics Letters, 24, 63–87.

Volker, P. (1983), "A Note on Factors Influencing the Utilization of Bankcard," Economic Record, 59, 281–289.

White, H. (1982), "Maximum Likelihood Estimation of Misspecified Models," Econometric, 50, 1–26.

Wooldridge, J. (1987), "A Unified Approach to Robust, Regression–Based Specification Tests," (mimeo, M.I.T.).

Rochester Center for Economic Research
University of Rochester
Department of Economics
Rochester, NY 14627

1987–88 DISCUSSION PAPERS

WP#68 RECURSIVE UTILITY AND OPTIMAL CAPITAL ACCUMULATION, I: EXISTENCE, by Robert A. Becker, John H. Boyd III, and Bom Yong Sung, January 1987

WP#69 MONEY AND MARKET INCOMPLETENESS IN OVERLAPPING-GENERATIONS MODELS, by Marianne Baxter, January 1987

WP#70 GROWTH BASED ON INCREASING RETURNS DUE TO SPECIALIZATION by Paul M. Romer, January 1987

WP#71 WHY A STUBBORN CONSERVATIVE WOULD RUN A DEFICIT: POLICY WITH TIME-INCONSISTENT PREFERENCES by Torsten Persson and Lars E.O. Svensson, January 1987

WP#72 ON THE CONTINUUM APPROACH OF SPATIAL AND SOME LOCAL PUBLIC GOODS OR PRODUCT DIFFERENTIATION MODELS by Marcus Berliant and Thijs ten Raa, January 1987

WP#73 THE QUIT-LAYOFF DISTINCTION: GROWTH EFFECTS by Kenneth J. McLaughlin, February 1987

WP#74 SOCIAL SECURITY, LIQUIDITY, AND EARLY RETIREMENT by James A. Kahn, March 1987

WP#75 THE PRODUCT CYCLE HYPOTHESIS AND THE HECKSCHER-OHLIN-SAMUELSON THEORY OF INTERNATIONAL TRADE by Sugata Marjit, April 1987

WP#76 NOTIONS OF EQUAL OPPORTUNITIES by William Thomson, April 1987

WP#77 BARGAINING PROBLEMS WITH UNCERTAIN DISAGREEMENT POINTS by Youngsub Chun and William Thomson, April 1987

WP#78 THE ECONOMICS OF RISING STARS by Glenn M. MacDonald, April 1987

WP#79 STOCHASTIC TRENDS AND ECONOMIC FLUCTUATIONS by Robert King, Charles Plosser, James Stock, and Mark Watson, April 1987

WP#80 INTEREST RATE SMOOTHING AND PRICE LEVEL TREND-STATIONARITY by Marvin Goodfriend, April 1987

WP#81 THE EQUILIBRIUM APPROACH TO EXCHANGE RATES by Alan C. Stockman, revised, April 1987

WP#82 INTEREST-RATE SMOOTHING
by Robert J. Barro, May 1987

WP#83 CYCLICAL PRICING OF DURABLE LUXURIES
by Mark Bils, May 1987

WP#84 EQUILIBRIUM IN COOPERATIVE GAMES OF POLICY FORMULATION
by Thomas F. Cooley and Bruce D. Smith, May 1987

WP#85 RENT SHARING AND TURNOVER IN A MODEL WITH EFFICIENCY UNITS OF HUMAN CAPITAL
by Kenneth J. McLaughlin, revised, May 1987

WP#86 THE CYCLICALITY OF LABOR TURNOVER: A JOINT WEALTH MAXIMIZING HYPOTHESIS
by Kenneth J. McLaughlin, revised, May 1987

WP#87 CAN EVERYONE BENEFIT FROM GROWTH? THREE DIFFICULTIES
by Herve' Moulin and William Thomson, May 1987

WP#88 TRADE IN RISKY ASSETS
by Lars E.O. Svensson, May 1987

WP#89 RATIONAL EXPECTATIONS MODELS WITH CENSORED VARIABLES
by Marianne Baxter, June 1987

WP#90 EMPIRICAL EXAMINATIONS OF THE INFORMATION SETS OF ECONOMIC AGENTS
by Nils Gottfries and Torsten Persson, June 1987

WP#91 DO WAGES VARY IN CITIES? AN EMPIRICAL STUDY OF URBAN LABOR MARKETS
by Eric A. Hanushek, June 1987

WP#92 ASPECTS OF TOURNAMENT MODELS: A SURVEY
by Kenneth J. McLaughlin, July 1987

WP#93 ON MODELLING THE NATURAL RATE OF UNEMPLOYMENT WITH INDIVISIBLE LABOR
by Jeremy Greenwood and Gregory W. Huffman

WP#94 TWENTY YEARS AFTER: ECONOMETRICS, 1966-1986
by Adrian Pagan, August 1987

WP#95 ON WELFARE THEORY AND URBAN ECONOMICS
by Marcus Berliant, Yorgos Y. Papageorgiou and Ping Wang, August 1987

WP#96 ENDOGENOUS FINANCIAL STRUCTURE IN AN ECONOMY WITH PRIVATE INFORMATION
by James Kahn, August 1987

WP#97 THE TRADE-OFF BETWEEN CHILD QUANTITY AND QUALITY: SOME EMPIRICAL EVIDENCE
by Eric Hanushek, September 1987

WP#98 SUPPLY AND EQUILIBRIUM IN AN ECONOMY WITH LAND AND PRODUCTION
by Marcus Berliant and Hou-Wen Jeng, September 1987

WP#99 AXIOMS CONCERNING UNCERTAIN DISAGREEMENT POINTS FOR 2-PERSON BARGAINING PROBLEMS
by Youngsub Chun, September 1987

WP#100 MONEY AND INFLATION IN THE AMERICAN COLONIES: FURTHER EVIDENCE ON THE FAILURE OF THE QUANTITY THEORY
by Bruce Smith, October 1987

WP#101 BANK PANICS, SUSPENSIONS, AND GEOGRAPHY: SOME NOTES ON THE "CONTAGION OF FEAR" IN BANKING
by Bruce Smith, October 1987

WP#102 LEGAL RESTRICTIONS, "SUNSPOTS", AND CYCLES
by Bruce Smith, October 1987

WP#103 THE QUIT-LAYOFF DISTINCTION IN A JOINT WEALTH MAXIMIZING APPROACH TO LABOR TURNOVER
by Kenneth McLaughlin, October 1987

WP#104 ON THE INCONSISTENCY OF THE MLE IN CERTAIN HETEROSKEDASTIC REGRESSION MODELS
by Adrian Pagan and H. Sabau, October 1987

WP#105 RECURRENT ADVERTISING
by Ignatius J. Horstmann and Glenn M. MacDonald, October 1987

WP#106 PREDICTIVE EFFICIENCY FOR SIMPLE NONLINEAR MODELS
by Thomas F. Cooley, William R. Parke and Siddhartha Chib, October 1987

WP#107 CREDIBILITY OF MACROECONOMIC POLICY: AN INTRODUCTION AND A BROAD SURVEY
by Torsten Persson, November 1987

WP#108 SOCIAL CONTRACTS AS ASSETS: A POSSIBLE SOLUTION TO THE TIME-CONSISTENCY PROBLEM
by Laurence Kotlikoff, Torsten Persson and Lars E. O. Svensson, November 1987

WP#109 EXCHANGE RATE VARIABILITY AND ASSET TRADE
by Torsten Persson and Lars E. O. Svensson, Novmeber 1987

WP#110 MICROFOUNDATIONS OF INDIVISIBLE LABOR
by Vittorio Grilli and Richard Rogerson, November 1987

WP#111 FISCAL POLICIES AND THE DOLLAR/POUND EXCHANGE RATE: 1870-1984
by Vittorio Grilli, November 1987

WP#112 INFLATION AND STOCK RETURNS WITH COMPLETE MARKETS
by Thomas Cooley and Jon Sonstelie, November 1987

WP#113 THE ECONOMETRIC ANALYSIS OF MODELS WITH RISK TERMS
by Adrian Pagan and Aman Ullah, December 1987

WP#114 PROGRAM TARGETING OPTIONS AND THE ELDERLY
by Eric Hanushek and Roberton Williams, December 1987

WP#115 BARGAINING SOLUTIONS AND STABILITY OF GROUPS
by Youngsub Chun and William Thomson, December 1987

WP#116 MONOTONIC ALLOCATION MECHANISMS
by William Thomson, December 1987

WP#117 MONOTONIC ALLOCATION MECHANISMS IN ECONOMIES WITH PUBLIC GOODS
by William Thomson, December 1987

WP#118 ADVERSE SELECTION, AGGREGATE UNCERTAINTY, AND THE ROLE FOR MUTUAL INSURANCE COMPANIES
by Bruce D. Smith and Michael J. Stutzer, February 1988

WP#119 INTEREST ON RESERVES AND SUNSPOT EQUILIBRIA: FRIEDMAN→S PROPOSAL RECONSIDERED
by Bruce D. Smith, February 1988

WP#120 INTERNATIONAL FINANCIAL INTERMEDIATION AND AGGREGATE FLUCTUATIONS UNDER ALTERNATIVE EXCHANGE RATE REGIMES
by Jeremy Greenwood and Stephen D. Williamson, February 1988

WP#121 FINANCIAL DEREGULATION, MONETARY POLICY, AND CENTRAL BANKING
by Marvin Goodfriend and Robert G. King, February 1988

WP#122 BANK RUNS IN OPEN ECONOMIES AND THE INTERNATIONAL TRANSMISSION OF PANICS
by Peter M. Garber and Vittorio U. Grilli, March 1988

WP#123 CAPITAL ACCUMULATION IN THE THEORY OF LONG RUN GROWTH
by Paul M. Romer, March 1988

WP#124 FINANCIAL INTERMEDIATION AND ENDOGENOUS GROWTH
by Valerie R. Bencivenga and Bruce D. Smith, March 1988

WP#125 UNEMPLOYMENT, THE VARIABILITY OF HOURS, AND THE PERSISTENCE OF "DISTURBANCES": A PRIVATE INFORMATION APPROACH
by Bruce D. Smith, March 1988

WP#126 WHAT CAN BE DONE WITH BAD SCHOOL PERFORMANCE DATA?
by Eric Hanushek and Lori Taylor, March 1988

WP#127 EQUILIBRIUM MARKETING STRATEGIES: IS THERE ADVERTISING, IN TRUTH?
by Ignatius Horstmann and Glenn MacDonald, revised, March 1988

WP#128 REAL EXCHANGE RATE VARIABILITY UNDER PEGGED AND FLOATING NOMINAL EXCHANGE RATE SYSTEMS: AN EQUILIBRIUM THEORY
by Alan C. Stockman, April 1988

WP#129 POST-SAMPLE PREDICTION TESTS FOR GENERALIZED METHOD OF MOMENT ESTIMATORS
by Dennis Hoffman and Adrian Pagan, April 1988

WP#130 GOVERNMENT SPENDING IN A SIMPLE MODEL OF ENDOGENOUS GROWTH
by Robert J. Barro, May 1988

WP#131 FINANCIAL DEVELOPMENT, GROWTH, AND THE DISTRIBUTION OF INCOME
by Jeremy Greenwood and Boyan Jovanovic, May 1988

WP#132 EMPLOYMENT AND HOURS OVER THE BUSINESS CYCLE
by Jang-Ok Cho and Thomas F. Cooley, May 1988

WP#133 A REFINEMENT AND EXTENSION OF THE NO-ENVY CONCEPT
by Dimitrios Diamantaras and William Thomson, May 1988

WP#134 NASH SOLUTION AND UNCERTAIN DISAGREEMENT POINTS
by Youngsub Chun and William Thomson, May 1988

WP#135 NON-PARAMETRIC ESTIMATION AND THE RISK PREMIUM
by Adrian Pagan and Y. Hong, May 1988

WP#136 CHARACTERIZING THE NASH BARGAINING SOLUTION WITHOUT PARETO-OPTIMALITY
by Terje Lensberg and William Thomson, May 1988

WP#137 SOME SIMULATION STUDIES OF NON-PARAMETRIC ESTIMATORS
by Y. Hong and A. Pagan, June 1988

WP#138 SELF-FULFILLING EXPECTATIONS, SPECULATIVE ATTACKS AND CAPITAL CONTROLS
by Harris Dellas and Alan C. Stockman, June 1988

WP#139 APPROXIMATING SUBOPTIMAL DYNAMIC EQUILIBRIA: AN EULER EQUATION APPROACH
by Marianne Baxter, June 1988

WP#140 BUSINESS CYCLES AND THE EXCHANGE RATE SYSTEM: SOME INTERNATIONAL EVIDENCE
by Marianne Baxter and Alan C. Stockman, June 1988

WP#141 RENT SHARING IN AN EQUILIBRIUM MODEL OF MATCHING AND TURNOVER
by Kenneth J. McLaughlin, June 1988

WP#142 CO-MOVEMENTS IN RELATIVE COMMODITY PRICES AND INTERNATIONAL CAPITAL FLOWS: A SIMPLE MODEL
by Ronald W. Jones, July 1988

WP#143 WAGE SENSITIVITY RANKINGS AND TEMPORAL CONVERGENCE
by Ronald W. Jones and Peter Neary, July 1988

WP#144 FOREIGN MONOPOLY AND OPTIMAL TARIFFS FOR THE SMALL OPEN ECONOMY
by Ronald W. Jones and Shumpei Takemori, July 1988

WP#145 THE ROLE OF SERVICES IN PRODUCTION AND INTERNATIONAL TRADE: A THEORETICAL FRAMEWORK
by Ronald W. Jones and Henryk Kierzkowski, July 1988

WP#146 APPRAISING THE OPTIONS FOR INTERNATIONAL TRADE IN SERVICES
by Ronald W. Jones and Frances Ruane, July 1988

WP#147 SIMPLE METHODS OF ESTIMATION AND INFERENCE FOR SYSTEMS CHARACTERIZED BY DETERMINISTIC CHAOS
by Mahmoud El-Gamal, August 1988

WP#148 THE RICARDIAN APPROACH TO BUDGET DEFICITS
by Robert J. Barro, August 1988

WP#149 A MODEL OF NOMINAL CONTRACTS
by Bruce D. Smith, August 1988

WP#150 A BUSINESS CYCLE MODEL WITH PRIVATE INFORMATION
by Bruce D. Smith, August 1988

WP#151 ASYMPTOTIC LIKELIHOOD BASED PREDICTION FUNCTIONS
by Thomas F. Cooley, August 1988

WP#152 MORAL HAZARD, IMPERFECT RISK-SHARING, AND THE BEHAVIOR OF ASSET RETURNS
by James A. Kahn, August 1988

WP#153 SPECIALIZATION, TRANSACTIONS TECHNOLOGIES, AND MONEY GROWTH
by Harold Cole and Alan C. Stockman, August 1988

WP#154 SAVINGS, INVESTMENT AND INTERNATIONAL CAPITAL FLOWS
by Linda L. Tesar, August 1988

WP#155 THE INFLATION TAX IN A REAL BUSINESS CYCLE MODEL
by Thomas F. Cooley and Gary D. Hansen, August 1988

WP#156 RAW MATERIALS, PROCESSING ACTIVITIES AND PROTECTIONISM
by Ronald W. Jones and Barbara J. Spencer, September 1988

WP#157 A TEST OF THE HARRIS ERGODICITY OF STATIONARY DYNAMICAL ECONOMIC MODELS
by Ian Domowitz and Mahmoud El-Gamal, September 1988

WP#158 SYMMETRIC STOCHASTIC GAMES OF RESOURCE EXTRACTION: THE EXISTENCE OF NON-RANDOMIZED STATIONARY EQUILIBRIUM
by Mukul Majumdar and Rangarajan Sundaram, September 1988

WP#159 SOME CONSEQUENCES OF CREDIT RATIONING IN AN ENDOGENOUS GROWTH MODEL
by Valerie Bencivenga and Bruce Smith, September 1988

WP#160 COMPETITIVE DIFFUSION
by Boyan Jovanovic and Glenn MacDonald, September 1988

WP#161 EXCHANGE RATES, THE CURRENT ACCOUNT, AND MONETARY POLICY
by Alan C. Stockman, October 1988

WP#162 DIAGNOSTIC TESTS FOR MODELS BASED ON INDIVIDUAL DATA: A SURVEY
by Adrian Pagan and Frank Vella, October 1988

To order a copy of the above papers complete the attached form and return to Mrs. Ellen Bennett, or call (716) 275-8396. The first three papers requested will be provided free of charge. Each additional paper will require a $3.00 service fee which <u>must be enclosed with your order</u>.

Requestor's Name ______________________________

Requestor's Address ______________________________

____

____

____

Please send me the following papers free of charge.

WP# ____ **WP#** ____ **WP#** ____

I understand there is a $3.00 fee for each additional paper. Enclosed is my check or money order in the amount of $ ________. Please send me the following papers.

WP# ____ **WP#** ____ **WP#** ____
WP# ____ **WP#** ____ **WP#** ____
WP# ____ **WP#** ____ **WP#** ____
WP# ____ **WP#** ____ **WP#** ____