A TEST FOR MODEL SPECIFICATION IN THE ABSENCE OF ALTERNATIVE HYPOTHESIS

by

Herman J. Bierens*

Discussion Paper No. 81 - 154, October 1981

* University of Amsterdam, The Netherlands. This paper was written while I was spending a three-month term at the University of Minnesota, whose kind hospitality is gratefully acknowledged. This visit is supported by a grant from the Netherlands Organisation for the Advancement of Pure Science (Z.W.O.).

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

Abstract

In recent papers ([6], [4]) several tests have been proposed for testing the truth of a nonlinear specification of a regression model against one or more well-specified alternative hypothesis. In this paper we first show by an example that these tests may be defective if all the hypotheses are false, and then we propose a new test that does not need alternative hypotheses in the sense that we only test whether or not a given specification is true.

Correspondence should be addressed to:

Herman J. Bierens
Stichting voor Economisch Onderzoek
der Universiteit van Amsterdam
Jodenbreestraat 23
1011 NH Amsterdam
The Netherlands

1. INTRODUCTION

In recent Econometrica articles Pesaran and Deaton [6] and Davidson and Mac Kinnon [4] propose several tests for testing a (non)linear specification of a regression model against one or more alternatives. These tests seem to perform well if either the null hypothesis or one of the alternative hypotheses is true. However, although these tests are able to reject all the hypotheses, it is possible that if all the hypotheses are false the least worst hypothesis will be accepted with relative high probability. We show this by applying the P-test of Davidson and Mac Kinnon [4] to the following example.

Consider an independent sample $\{(x_{1,1}, x_{2,1}, \varepsilon_1), \ldots, (x_{1,n}, x_{2,n}, \varepsilon_n)\}$ from the trivariate standard normal distribution and let for $j = 1,2,\ldots,n$,

$$y_j = x_{1,j} + x_{2,j} + x_{1,j} \cdot x_{2,j} + \varepsilon_j \quad .$$

Suppose we have specified this relationship as

$$H_o : y_j = \beta_1 \cdot x_{1,j} + \beta_2 \cdot x_{2,j} + \varepsilon_{o,j} = f(x_j, \beta) + \varepsilon_{o,j} \quad \text{, say,}$$

where $x_j^T = (x_{1,j}, x_{2,j})$ and $\beta^T = (\beta_1, \beta_2)$. In order to test the truth of this hypothesis by the P-test proposed in [4] we need a non-nested alternative hypothesis, "though not one in which we need have any faith" (quoting [4]), for example

$$H_1 : y_j = \gamma_1 x_{1,j}^3 + \gamma_2 x_{2,j}^3 + \varepsilon_{1,j} = g(x_j, \gamma) + \varepsilon_{1,j} \quad \text{, say,}$$

where x_j is as before and $\gamma^T = (\gamma_1, \gamma_2)$. Clearly both hypotheses are false, but H_o is closer to the truth than H_1 . Following [4], we denote:

$$\hat{f}^T = (f(x_1, \hat{\beta}), \ldots, f(x_n, \hat{\beta})) ,$$
$$\hat{g}^T = (g(x_1, \hat{\gamma}), \ldots, g(x_n, \hat{\gamma})) ,$$
$$\hat{M}_o = I - X(X^T X)^{-1} X^T ,$$
$$y^T = (y_1, \ldots, y_n) ,$$

where $\hat{\beta}$ and $\hat{\gamma}$ are the OLS estimates of β and γ, respectively, and X is the $2 \times n$ matrix with elements $x_{i,j}$ $(i = 1, 2; \; j = 1, 2, \ldots, n)$. The test statistic of the P-test is

$$\hat{t}_p = \frac{\sqrt{n}\,\hat{\alpha}_p}{\sqrt{\hat{\sigma}^2 \| \hat{M}_o(\hat{g} - \hat{f}) \|^{-2}}} , \tag{1.1}$$

where

$$\hat{\alpha}_p = \frac{(\hat{g} - \hat{f})^T \hat{M}_o (y - \hat{f})}{\| \hat{M}_o(\hat{g} - \hat{f}) \|^2} \tag{1.2}$$

and

$$\hat{\sigma}^2 = \frac{1}{n - k - 1} \| y - \hat{f} - \hat{\alpha}_p \hat{M}_o(\hat{g} - \hat{f}) \|^2 . \tag{1.3}$$

As is shown in [4] the test statistic $\hat{t}_p$ is asymptotically distributed as $N(0, 1)$ if H_o is true. In the case under review, however, neither H_o nor H_1 is true, but nevertheless we have

$$\hat{t}_p \to N(0, 4) \text{ in distribution as } n \to \infty , \tag{1.4}$$

as will be proved in Appendix 1. Thus if we proceed the test at the five percent significance level the probability that we accept the false null is

$$\int_{-1.96}^{1.96} \frac{e^{-\frac{1}{2}u^2/4}}{2\sqrt{2\pi}} \, du = \int_{-0.98}^{0.98} \frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}} \, du \approx 0.673 .$$

This example shows that the P - test proposed by Davidson and Mac Kinnon is not appropriate if there is a possibility that none of the hypotheses involved is true. The same applies to their J - test, because for the example under review the J - test and the P - test yield identical test statistics.

We have not checked through the performance of the test proposed by Pesaran and Deaton [6] for the above case, but it seems to be likely that we would find similar results.

Our argument makes clear that the problem of testing the functional form of a regression model is not yet conclusively solved, hence there is still a need for an alternative specification test that is more watertight than those we just have discussed. In this paper we shall propose such a test. This test has the advantage that it does not need any specified alternative hypothesis; we only test whether or not a given specification is true. In section 2 we shall give a review of this test procedure. In section 3 we shall deal with the asymptotic theory.

2. IDENTIFICATION OF THE NULL HYPOTHESIS AND SUMMARY OF THE TEST PROCEDURE

The Hypothesis

Consider an i.i.d. stochastic process $(y_1, x_1), \ldots, (y_n, x_n)$ in $\mathbb{R} \times \mathbb{R}^k$, where the y_j satisfy

$$(2.1) \qquad E|y_j| < \infty \ .$$

The condition (2.1) is sufficient for the existence of a Borel measurable function $g(\cdot)$ on $\mathbb{R}^k$ such that

$$(2.2) \qquad g(x_j) = E(y_j|x_j) \ \text{ a.s.} \ .$$

(See Chung [3, Theorem 9.1.2]. Defining $u_j = y_j - g(x_j)$ we then get the following tautological regression model

$$(2.3) \qquad y_j = g(x_j) + u_j, \quad j = 1, 2, \ldots, n, \ldots \ ,$$

where obviously the u_j satisfy

$$(2.4) \qquad E(u_j|x_j) = 0 \ \text{ a.s.}, \quad j = 1, 2, \ldots, n, \ldots \ .$$

Suppose we have specified the regression function $g(x)$ as $f(x, \theta_o)$, where $f(x, \theta)$ is a known real valued Borel measurable function on $\mathbb{R}^k \times \Theta$ and Θ is a parameter space containing the unknown parameter θ_o if this specification is true. Testing the truth of this specification we thus have to test the null hypothesis

$$(2.5) \qquad H_o : P[f(x_j, \theta_o) = g(x_j)] = 1 \quad \text{for some} \quad \theta_o \in \Theta$$

against the alternative hypothesis that H_o is false:

$$(2.6) \qquad H_1 : P[f(x_j, \theta) = g(x_j)] < 1 \quad \text{for all} \quad \theta \in \Theta \ .$$

Identification

But how can we identify H_o versus H_1 from the distribution of the data, or with other words, how do H_o and H_1, respectively, correspond with distinct characteristics of the distribution of (y_j, x_j) ? The answer to this question is suggested by the following fundamental theorem.

<u>THEOREM 1</u>. Let v be a real valued random variable satisfying $E|v| < \infty$ and let z be a random vector in $\mathbb{R}^k$. Then we have:

(I) $P[E(v|z) = 0] < 1$ if and only if $E\, v\, e^{it^T z} \neq 0$ for some nonrandom vector $t \in \mathbb{R}^k$.

(II) If in addition z is bounded then $P[E(v|z) = 0] < 1$ if and only if $E\, v\, e^{it_o^T z} \neq 0$ for some nonrandom vector $t_o \in \mathbb{R}^k$ in an arbitrarily small neighborhood of $t = 0$.

<u>Proof</u>: Appendix 2.

Thus from part I of theorem 1 it follows that $E[y_j - f(x_j, \theta)|x_j] = E(u_j|x_j) + E[g(x_j) - f(x_j, \theta)|x_j] = E[g(x_j) - f(x_j, \theta)|x_j] = 0$ with probability 1 for some $\theta_o \in \Theta$ if and only if $E(y_j - f(x_j, \theta_o))\, e^{it^T x_j} \equiv 0$ for all $t \in \mathbb{R}^k$, hence

<u>COROLLARY 1</u>: H_o is true if and only if for some $\theta_o \in \Theta$: $E(y_j - f(x_j, \theta_o))\, e^{it^T x_j} = 0$ for all $t \in \mathbb{R}^k$.

Also part II of theorem 1 can be used for identifying the hypotheses, as we now show. Let Φ be a bounded Borel measurable mapping from $\mathbb{R}^k$ into $\mathbb{R}^k$ such that x_j and $\Phi(x_j)$ generate the same Euclidean Borel field, for example

$$\Phi(x_j) = \begin{pmatrix} \mathrm{atg}\,(x_{1,j}) \\ \cdot \\ \cdot \\ \cdot \\ \mathrm{atg}\,(x_{k,j}) \end{pmatrix} \tag{2.7}$$

Then

$$(2.8) \qquad E[g(x_j) - f(x_j, \theta) | x_j] = E[g(x_j) - f(x_j, \theta) | \Phi(x_j)] \quad \text{a.s.},$$

hence, applying part II of theorem 1, we have:

<u>COROLLARY 2</u>: Let Φ be any bounded Borel measurable mapping from $\mathbb{R}^k$ into $\mathbb{R}^k$ such that $\Phi(x_j)$ and x_j generate the same Euclidean Borel field. Then H_1 is true if and only if for every $\theta \in \Theta$ there is a t_o in an arbitrarily small neighborhood of $t = 0$ such that $E(y_j - f(x_j, \theta))\, e^{it_o^T\Phi(x_j)} \neq 0$.

Of course, if H_o is true then for some $\theta_o \in \Theta, E(y_j - f(x_j, \theta_o))\, e^{it^T\Phi(x_j)} \equiv 0$ for all $t \in \mathbb{R}^k$, but it is sufficient to verify this only for an arbitrarily small neighborhood of $t = 0$.

A direct consequence of corollary 2 is:

<u>COROLLARY 3</u>: Let Φ be defined as in corollary 2. Let $N_o = \mathop{X}_{\ell=1}^{k} [-\varepsilon_\ell, \varepsilon_\ell]$, where $\varepsilon_\ell > 0$ ($\ell = 1, 2, \ldots, k$) are arbitrarily chosen. Put

$$(2.9) \qquad \eta(\theta) = \int_{N_o} |E(y_j - f(x_j, \theta))\, e^{it^T\Phi(x_j)}|^2 \, dt \quad .$$

Then H_o is true if $\eta(\theta_o) = 0$ for some $\theta_o \in \Theta$ and H_1 is true if $\inf_{\theta \in \Theta} \eta(\theta) > 0$.

<u>The Test</u>

This result suggests to test H_o against H_1 by using a test statistic of the form:

$$(2.10) \qquad \hat{\eta} = \int_{N_o} | \frac{1}{n} \sum_{j=1}^{n} (y_j - f(x_j, \hat{\theta}))\, e^{it^T\Phi(x_j)}|^2 \, dt \, ,$$

where N_o and Φ are as before and $\hat{\theta}$ is the nonlinear least squares estimator of θ_o , i.e.

$$\hat{\theta} \in \Theta \text{ a.s.}, \quad \sum_{j=1}^{n} (y_j - f(x_j, \hat{\theta}))^2 = \inf_{\theta \in \Theta} \sum_{j=1}^{n} (y_j - f(x_j, \theta))^2 \quad . \tag{2.11}$$

Note that if we put

$$z_j = \Phi(x_j), \quad z_j^T = (z_{1,j}, \ldots, z_{k,j}) \tag{2.12}$$

and

$$\hat{u}_j = y_j - f(x_j, \hat{\theta}), \tag{2.13}$$

$\hat{\eta}$ can be written as

$$\begin{aligned} \hat{\eta} &= \frac{1}{n^2} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \hat{u}_{j_1} \hat{u}_{j_2} \int_{N_o} e^{it^T(z_{j_1} - z_{j_2})} dt \\ &= 2^k \frac{1}{n^2} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \hat{u}_{j_1} \hat{u}_{j_2} \prod_{\ell=1}^{k} \frac{\sin[\varepsilon_\ell (z_{\ell,j_1} - z_{\ell,j_2})]}{z_{\ell,j_1} - z_{\ell,j_2}} \quad . \end{aligned} \tag{2.14}$$

In section 3 we show that under H_o $n\hat{\eta}$ converges in distribution to a nonnegative random variable with first moment μ, say, and that we can estimate this moment consistently by

$$\hat{\mu} = 2^k \hat{\sigma}^2 \{ \prod_{\ell=1}^{k} \varepsilon_\ell - \mathrm{tr}\, (\hat{A}^{-1} \hat{B}) \} , \tag{2.15}$$

where $\hat{\sigma}$ is the usual estimate of the variance of the u_j,

$$\hat{A} = \frac{1}{n} \sum_{j=1}^{n} \{(\partial/\partial \theta^T) f(x_j, \hat{\theta})\} \{(\partial/\partial \theta) f(x_j, \hat{\theta})\} \tag{2.16}$$

and

$$\hat{B} = \frac{1}{n^2} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \{(\partial/\partial \theta^T) f(x_{j_1}, \hat{\theta})\} \{(\partial/\partial \theta) f(x_{j_2}, \hat{\theta})\} \cdot \prod_{\ell=1}^{k} \frac{\sin[\varepsilon_\ell (z_{\ell,j_1} - z_{\ell,j_2})]}{z_{\ell,j_1} - z_{\ell,j_2}} \quad . \tag{2.17}$$

So by applying Chebishev's inequality (for first moments) we conclude that for any $\alpha \in (0, 1)$, (2.18) $$\limsup_{n\to\infty} P[n\hat{\eta} > \frac{1}{\alpha}\hat{\mu}] \le \alpha \quad \text{if } H_0 \text{ is true.}$$

Moreover, for every $\alpha > 0$ we have

(2.19) $$\lim_{n\to\infty} P[n\hat{\eta} > \frac{1}{\alpha}\hat{\mu}] = 1 \quad \text{if } H_1 \text{ is true,}$$

as also will be shown in the next section. Thus, proceeding the test at the $\alpha \cdot 100$ percent confidence level we accept H_0 if $n\hat{\eta} \le \frac{1}{\alpha}\hat{\mu}$ and we reject H_0 if not.

The result (2.18) is somewhat crude, as it is based on Chebishev's inequality, which is not a very sharp inequality. It would, of course, be better to base the test directly on the limiting distribution of $n\hat{\eta}$, but this limiting distribution appears to have a too complicated structure for that. However, if we assume that the distribution of the u_j is i.i.d. $N(0,\sigma^2)$ and that the sequence (u_j) is independent of the sequence (x_j) we can by Monte Carlo simulation estimate a sharper lower bound than the bound $\frac{1}{\alpha}\hat{\mu}$ used in (2.18). We shall discuss this point too in the next section.

The above test is only one example out of a large class of similar tests that can be derived on the basis of theorem 1. For example, we also may use a test statistic of the form

$$\hat{\hat{\eta}} = \int_{\mathbb{R}^k} \left| \frac{1}{n} \sum_{j=1}^{n} (y_j - f(x_j, \hat{\theta}))\, e^{it^T x_j} \right|^2 w(t)\, dt \quad ,$$

where $w(t)$ is a positive weight function, for example a k-variate normal density, and for this test statistic we can, on basis of part I of theorem 1, derive similar results as (2.18) and (2.19).

3. ASYMPTOTIC THEORY

The Assumptions

In this section we shall set forth conditions such that (2.18) and (2.19) are true. The conditions involved are just those for weak consistency and asymptotic normality of the nonlinear least squares estimator as can be found in Jennrich [5] and Bierens [1, 2].

Our first assumption concerns the distribution of the data.

ASSUMPTION 1: The observations $(y_1, x_1), \ldots, (y_n, x_n)$ are i.i.d. random vectors in $\mathbb{R} \times \mathbb{R}^k$. The y_j satisfy $E|y_j|^{2+\delta} < \infty$ for some $\delta > 0$.

The i.i.d. assumption is, of course, very restrictive and rules out most of the econometric applications. It is, however, merely made for convenience. Using the results in [1] and [2] it is possible to extend the results in this paper to nonstationary and nonlinear time series regressions, but this will be out of the scope of the present paper. The condition on the absolute moments of y_j is stronger than (2.1), but needed in order to assure that the errors u_j defined by

$$u_j = y_j - E(y_j | x_j) = y_j - g(x_j) \tag{3.1}$$

have finite absolute moments of order slightly larger than 2. Thus we have

$$E(u_j \mid x_j) = 0 \text{ a.s.}, \quad Eu_j^2 = \sigma^2 < \infty, \quad E|u_j|^{2+\delta} < \infty \text{ for some } \delta > 0. \tag{3.2}$$

Moreover, we shall limit our attention to specifications $f(x, \theta)$ satisfying the following conditions.

<u>ASSUMPTION 2</u>: The function $f(x, \theta)$ and its first and second partial derivatives $(\partial/\partial\theta_i) f(x, \theta)$ and $(\partial/\partial\theta_{i_1})(\partial/\partial\theta_{i_2}) f(x, \theta)$, $(i, i_1, i_2 = 1, 2, \ldots, m)$, are continuous real functions on $R^k \times \Theta$, where Θ is a convex compact subset of R^m. If H_o is true then θ_o is an interior point of Θ.

This assumption implies that if H_o is true then the true regression function g is continuous on R^k, but under H_1 g may be any Borel measurable function on R^k.

Furthermore, following [1] we shall impose some weak conditions on the moments of $f(x_j, \theta)$ and its first and second partial derivatives to θ, i.e.,

<u>ASSUMPTION 3</u>: There exists a $\delta > 0$ such that:

$$E\left\{\sup_{\|x\|\le\|x_j\|}\ \sup_{\theta\in\Theta} |f(x_j, \theta)|\right\}^{2+\delta} < \infty \quad ,$$

$$E\left\{\sup_{\|x\|\le\|x_j\|}\ \sup_{\theta\in\Theta} |(\partial/\partial\theta_i) f(x_j,\theta)|\right\}^{2+\delta} < \infty \ (i = 1, 2, \ldots, m) \quad ,$$

$$E\left\{\sup_{\|x\|\le\|x_j\|}\ \sup_{\theta\in\Theta} |(\partial/\partial\theta_{i_1})(\partial/\partial\theta_{i_2}) f(x_j, \theta)|\right\}^{2+\delta} < \infty \ , \quad (i_1, i_2 = 1, 2, \ldots, m) \ .$$

In order to estimate the true parameter θ_o consistently by least squares estimation it should be unique in the sense that it is the only point in Θ such that

$$(3.3) \qquad 0 = E\{g(x_j) - f(x_j, \theta_o)\}^2 = \inf_{\theta\in\Theta} E\{g(x_j - f(x_j, \theta)\}^2 ,$$

provided H_o is true. But even if H_o is false we can define a point θ_o in Θ by the right equality in (3.3), and if such a point is unique it can be estimated consistently by least squares. Therefore we assume:

ASSUMPTION 4: Under H_o as well as under H_1 there exists a unique point θ_* in Θ such that

$$E\{g(x_j) - f(x_j, \theta_*)\}^2 = \inf_{\theta \in \Theta} E\{g(x_j) - f(x_j, \theta)\}^2 \quad .$$

Of course, under H_o we have $\theta_o = \theta_*$. Moreover we note that by assumption 4 the alternative hypothesis can be restated as

$$(3.4) \qquad H_1 : P[g(x_j) = f(x_j, \theta_*)] < 1 .$$

From the assumptions 1 through 4 and Theorem 3.1.3 in [1] it now follows that the least squares estimator $\hat{\theta}$ is weakly consistent:

$$(3.5) \qquad \operatorname*{plim}_{n\to\infty} \hat{\theta} = \theta_o \quad \text{under } H_o ,$$

but using the argument in Section 3.1 in [1] it is not hard to show that also

$$(3.6) \qquad \operatorname*{plim}_{n\to\infty} \hat{\theta} = \theta_* \quad \text{under } H_1 .$$

Next suppose:

ASSUMPTION 5: The matrix

$$(3.7) \qquad A = E\{(\partial/\partial\theta^T) f(x_j, \theta_o)\} \{(\partial/\partial\theta) f(x_j, \theta_o)\}$$

is positive definite and $E(u_j^2|x_j) = \sigma^2$ a.s..

From the argument in Section 3.1.3 in [1] it then follows that under H_o

$$(3.8) \qquad \operatorname*{plim}_{n\to\infty} \left\| \sqrt{n}(\hat{\theta} - \theta_o) - \frac{1}{\sqrt{n}} \sum_{j=1}^{n} u_j A^{-1}(\partial/\partial\theta^T) f(x_j, \theta_o) \right\| = 0 \quad ,$$

so that by the central limit theorem

$$(3.9) \qquad \sqrt{n}(\hat{\theta} - \theta_o) \to N_m[0, \sigma^2 A^{-1}] \quad \text{in distribution} \quad .$$

Asymptotic Theory under H_o

Before we continue our argument we shall first introduce some additional notation. Thus we denote

$$(3.10) \qquad \hat{\xi}^{(1)}(t) = \frac{1}{n}\sum_{j=1}^{n} (y_j - f(x_j, \hat{\theta}))\, e^{it^T z_j} \quad ,$$

$$(3.11) \qquad \hat{\xi}^{(2)}(t) = \frac{1}{n}\sum_{j=1}^{n} \{u_j + (\theta_o - \hat{\theta})^T (\partial/\partial\theta^T) f(x_j, \theta_o)\}\, e^{it^T z_j} \quad ,$$

$$(3.12) \qquad \hat{\xi}^{(3)}(t) = \frac{1}{n}\sum_{j=1}^{n} u_j\{e^{it^T z_j} - (\partial/\partial\theta) f(x_j, \theta_o) A^{-1} \hat{b}_o(T)\} \quad ,$$

$$(3.13) \qquad \hat{\xi}^{*}(t) = \frac{1}{n}\sum_{j=1}^{n} u_j\{e^{it^T z_j} - (\partial/\partial\theta) f(x_j, \theta_o) A^{-1} b(t)\} = \frac{1}{n}\sum_{j=1}^{n} u_j \rho_j(t) \quad ,$$

say, where z_j is defined by (2.12),

$$(3.14) \qquad \rho_j(t) = e^{it^T z_j} - (\partial/\partial\theta) f(x_j, \theta_o) A^{-1} b(t) \quad ,$$

$$(3.15) \qquad \hat{b}_o(t) = \frac{1}{n}\sum_{j=1}^{n} (\partial/\partial\theta^T) f(x_j, \theta_o)\, e^{it^T z_j}$$

and

$$(3.16) \qquad b(t) = E\,\hat{b}_o(t) = E(\partial/\partial\theta^T) f(x_j, \theta_o)\, e^{it^T z_j} \quad .$$

Note that from (2.10) and (3.10),

$$(3.17) \qquad \hat{\eta} = \int_{N_o} |\hat{\xi}^{(1)}(t)|^2\, dt \quad .$$

Now put

$$(3.18) \qquad \hat{\eta}^{*} = \int_{N_o} |\hat{\xi}^{*}(t)|^2\, dt \quad .$$

We then have

THEOREM 2. If H_o is true and if the assumptions 1 through 5 hold then $\operatorname{plim}_{n\to\infty} |n\hat{\eta} - n\hat{\eta}^{*}| = 0$.

Proof: Using the mean value theorem we can write

$$(3.19)\qquad \hat{\xi}^{(1)}(t) = \frac{1}{n}\sum_{j=1}^{n}[u_j + f(x_j,\theta_o) - f(x_j,\hat{\theta})]e^{it^Tz_j}$$

$$= \frac{1}{n}\sum_{j=1}^{n}[u_j - (\hat{\theta}-\theta_o)^T(\partial/\partial\theta^T)f(x_j,\tilde{\theta}(t)]e^{it^Tz_j} ,$$

where $\tilde{\theta}(t)$ is a mean value satisfying

$$(3.20)\qquad \|\tilde{\theta}(t)-\theta_o\| \le \|\hat{\theta}-\theta_o\| \quad \text{a.s. for all} \quad t \in \mathbb{R}^k .$$

Using Theorem 2.3.4 in [1] it is not hard to show

$$(3.21)\qquad \operatorname*{plim}_{n\to\infty}\sup_{t\in N_o}|\sqrt{n}\,\hat{\xi}^{(1)}(t) - \sqrt{n}\,\hat{\xi}^{(2)}(t)| = 0$$

and combining (3.8) and (3.11) we see that

$$(3.22)\qquad \operatorname*{plim}_{n\to\infty}\sup_{t\in N_o}|\sqrt{n}\,\hat{\xi}^{(2)}(t) - \sqrt{n}\,\hat{\xi}^{(3)}(t)| = 0 .$$

Again using Theorem 2.3.4 in [1] we may conclude that

$$(3.23)\qquad \operatorname*{plim}_{n\to\infty}\sup_{t\in N_o}|\hat{b}_o(t) - b(t)| = 0$$

and consequently that

$$(3.24)\qquad \operatorname*{plim}_{n\to\infty}\sup_{t\in N_o}|\sqrt{n}\,\hat{\xi}^{(3)}(t) - \sqrt{n}\,\hat{\xi}^{*}(t)| = 0 .$$

Combining (3.21), (3.22) and (3.24) we obtain

$$(3.25)\qquad \operatorname*{plim}_{n\to\infty}\sup_{t\in N_o}|\sqrt{n}\,\hat{\xi}^{(1)}(t) - \sqrt{n}\,\hat{\xi}^{*}(t)| = 0 ,$$

and applying Lemma 3.3.1 in [1] we conclude from (3.25) that also

$$(3.26)\qquad \operatorname*{plim}_{n\to\infty}\sup_{t\in N_o}\left|\,|\sqrt{n}\,\hat{\xi}^{(1)}(t)|^2 - |\sqrt{n}\,\hat{\xi}^{*}(t)|^2\right| = 0 .$$

This proves the theorem.

Q.E.D.

Theorem 2 implies that under H_o $n\hat{\eta}$ and $n\hat{\eta}^*$ have the same limiting distribution. But what is the limiting distribution of $n\hat{\eta}^*$? If we substitute (3.13) in (3.18) we get

$$(3.27) \qquad n\hat{\eta}^* = \frac{1}{n} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} u_{j_1} u_{j_2} \int_{N_o} \rho_{j_1}(t) \overline{\rho_{j_2}(t)}\, dt$$

where $\bar{\rho}_j$ is the complex conjugate of ρ_j. If we would be able to write $\int_{N_o} \rho_{j_1}(t)\overline{\rho_{j_2}(t)}\, dt$ as a product of i.i.d. random variables r_{j_1} and r_{j_2}, say, then $n\hat{\eta}^*$ would converge in distribution to χ_1^2 times $\sigma^2 E\, r_j^2$, but it appears impossible to split up the integral involved in this way. So the limiting distribution of $n\hat{\eta}^*$ is probably of an unknown type. There are, however, two ways out of this problem. The first way out is to compute the expectation of $n\hat{\eta}^*$ and to apply Chebishev's inequality. This expectation is:

$$(3.28) \qquad \mu = E\, n\hat{\eta}^* = \frac{1}{n} \sum_{j=1}^{n} E\, u_j^2 E \int_{N_o} \rho_j(t)\, \bar{\rho}_j(t)\, dt = \sigma^2 E \int_{N_o} \rho_j(t)\, \bar{\rho}_j(t)\, dt \quad ,$$

which is obviously independent of the sample size n. We then have by Chebishev's inequality

$$(3.29) \qquad P[n\hat{\eta}^* > \frac{1}{\alpha}\mu] \le \frac{E\, n\hat{\eta}^*}{\frac{1}{\alpha}\mu} = \alpha$$

and consequently by theorem 2

$$(3.30) \qquad \limsup_{n\to\infty} P[n\hat{\eta} > \frac{1}{\alpha}\mu] \le \alpha \quad .$$

So if we can find a consistent estimate $\hat{\mu}$, say, of μ then

$$(3.31) \qquad \limsup_{n\to\infty} P[n\hat{\eta} > \frac{1}{\alpha}\hat{\mu}] \le \alpha \quad .$$

We can construct such an estimate $\hat{\mu}$ as follows. Put

$$(3.32) \qquad \hat{\delta}(t) = \frac{1}{n} \sum_{j=1}^{n} (\partial/\partial\theta^T) f(x_j, \hat{\theta})\, e^{it^T z_j}$$

and

$$(3.33)\qquad \hat{\rho}_j(t) = e^{it^T z_j} - (\partial/\partial\theta) f(x_j, \hat{\theta}) \hat{A}^{-1} \hat{b}(t),$$

where $\hat{A}$ is defined by (2.16). Then

$$(3.34)\qquad \sup_{t\in N_o} \left| \frac{1}{n} \sum_{j=1}^{n} \hat{\rho}_j(t) \overline{\hat{\rho}_j(t)} - \frac{1}{n} \sum_{j=1}^{n} \rho_j(t) \overline{\rho_j(t)} \right|$$
$$\leq \sup_{t\in N_o} \left| \hat{b}(t)^T \hat{A}^{-1} \overline{\hat{b}(t)} - b(t)^T \hat{A}^{-1} \overline{b(t)} \right| \to 0 \text{ in prob.}$$

because

$$(3.35)\qquad \operatorname*{plim}_{n\to\infty} \hat{A} = A,$$

where A is defined in assumption 5, and

$$(3.36)\qquad \operatorname*{plim}_{n\to\infty} \sup_{t\in N_o} |\hat{b}(t) - b(t)| = 0,$$

as is not hard to verify by using Theorem 2.3.4 in [1]. Moreover,

$$(3.37)\qquad \operatorname*{plim}_{n\to\infty} \sup_{t\in N_o} \left| \frac{1}{n} \sum_{j=1}^{n} \rho_j(t) \overline{\rho_j(t)} - E \rho_j(t) \overline{\rho_j(t)} \right| = 0$$

as is also easily verified by using Theorem 2.3.4 in [1]. So if we combine (3.34) and (3.37) we then may conclude

$$(3.38)\qquad \operatorname*{plim}_{n\to\infty} \int_{N_o} \frac{1}{n} \sum_{j=1}^{n} \hat{\rho}_j(t) \overline{\hat{\rho}_j(t)}\, dt = \int_{N_o} E\{\rho_j(t) \overline{\rho_j(t)}\}\, dt = E \int_{N_o} \rho_j(t) \overline{\rho_j(t)}\, dt$$

and hence

$$(3.39)\qquad \hat{\mu} = \hat{\sigma}^2 \frac{1}{n} \sum_{j=1}^{n} \int_{N_o} \hat{\rho}_j(t) \overline{\hat{\rho}_j(t)}\, dt$$

is a consistent estimator of $\hat{\mu}$. We leave it to the reader to verify that the estimator $\hat{\mu}$ defined by (3.39) can be written as (2.15).

So we have proved by now

<u>THEOREM 3</u>. If H_o is true and if the assumptions 1 through 5 hold then for every $\alpha \in (0, 1)$, $\limsup_{n\to\infty} P[n\hat{\eta} > \frac{1}{\alpha}\hat{\mu}] \leq \alpha$.

This is our first way out of the problem that the limiting distribution of $n\hat{\eta}^*$ is unknown. The second way is the following.

If we make the additional assumption that the sequence (u_j) of disturbances $(u_j = y_j - E(y_j|x_j))$ is independent of the sequence (x_j) of regressors and moreover that these u_j's are distributed as $N(0, \sigma^2)$, then by replacing the u_j's in (3.13) by other independent random drawings from $N(0, \sigma^2)$ the resulting integral of the type (3.18) has the same distribution as the original one. Thus if we draw an artificial random sample $\{W_1, \ldots, W_n\}$ from the standard normal distribution and if we put

$$(3.40) \qquad \tilde{\xi}^*(t) = \frac{1}{n}\sum_{j=1}^{n} \sigma W_j \rho_j(t),$$

$$(3.41) \qquad \tilde{\eta}^*(t) = \int_{N_o} |\tilde{\xi}^*(t)|^2 \, dt,$$

then $n\hat{\eta}^*$ and $n\tilde{\eta}^*$ have the same distribution. However, σ and $\rho_j(t)$ are not observable. Therefore we replace σ in (3.40) by its estimate $\hat{\sigma}$ and $\rho_j(t)$ by $\hat{\rho}_j(t)$ (defined by (3.33)). Thus, we put:

$$(3.42) \qquad \tilde{\xi}^{**}(t) = \frac{1}{n}\sum_{j=1}^{n} \hat{\sigma} W_j \hat{\rho}_j(t),$$

$$(3.43) \qquad \tilde{\eta}^{**} = \int_{N_o} |\tilde{\xi}^{**}(t)|^2 \, dt \quad .$$

Now similar to theorem 2 it can be shown that

$$(3.44) \qquad \operatorname{plim}_{n\to\infty} |n\tilde{\eta}^{**} - n\tilde{\eta}^*| = 0$$

and hence that $n\tilde{\eta}^{**}$ and $n\hat{\eta}^*$ have the same limiting distribution. Therefore by computing (3.43) for a large number of artificially drawn random samples $\{W_1, \ldots, W_n\}$ from the standard normal distribution we can establish a (random) number $\tilde{p}_\alpha$ such that $\alpha \times 100$ percent of the $\tilde{\eta}^{**}$'s are larger than $\tilde{p}_\alpha$. We then have approximately $P[n\hat{\eta} > \tilde{p}_\alpha] \approx \alpha$ for large n, provided H_o is true.

Asymptotic Theory under H_1

We have seen that if H_1 is true then under the assumptions 1 through 4, $\text{plim}\, \hat{\theta} = \theta_*$, where θ_* is defined by assumption 4. Therefore, if we put

$$(3.45) \qquad \hat{\xi}_*^{(1)}(t) = \frac{1}{n} \sum_{j=1}^{n} (y_j - f(x_j, \theta_*))\, e^{it^T z_j}$$

then from (3.10), (3.6) and (3.45)

$$(3.46) \qquad \underset{n\to\infty}{\text{plim}} \sup_{t \in N_o} |\hat{\xi}^{(1)}(t) - \hat{\xi}_*^{(1)}(t)| = 0 \quad ,$$

and moreover,

$$(3.47) \qquad \underset{n\to\infty}{\text{plim}} \sup_{t \in N_o} |\hat{\xi}_*^{(1)}(t) - \xi_*(t)| = 0 \quad ,$$

where

$$(3.48) \qquad \xi_*(t) = E\,\hat{\xi}_*^{(1)}(t) = E[g(x_j) - f(x_j, \theta_*)]\, e^{it^T z_j} \quad ,$$

as is not hard to verify by applying Theorem 2.3.4 in [1]. Thus we have under H_1

$$(3.49) \qquad \underset{n\to\infty}{\text{plim}} \sup_{t \in N_o} |\hat{\xi}^{(1)}(t) - \xi_*(t)| = 0$$

and consequently

$$(3.50) \qquad \underset{n\to\infty}{\text{plim}}\, \hat{\eta} = \eta_* = \int_{N_o} |\xi_*(t)|^2\, dt \quad .$$

But from corollary 3 it follows that $\eta_* > 0$, hence:

<u>THEOREM 4</u>. If H_1 is true and if the assumptions 1 through 4 hold then $\operatorname{plim}_{n\to\infty} n\hat{\eta} = \infty$.

Since it is not hard to show that also under H_1 the estimator $\hat{\mu}$ converges in probability to a finite number, Theorem 4 implies that (2.19) holds.

APPENDIX 1

Proof of (1.4)

In addition to the notation introduced in section 1 we put

(A1.1) $$u^T = (\varepsilon_1 + x_{1,1} \cdot x_{2,1}, \ \ldots\ldots, \ \varepsilon_n + x_{1,n} \cdot x_{2,n}) \quad .$$

(A1.2) $$Z = \begin{pmatrix} x_{1,1}^3, & x_{2,1}^3 \\ - \ - & - \ - \\ x_{1,n}^3, & x_{2,n}^3 \end{pmatrix} , \quad .$$

(A1.3) $$m_q = \int_{-\infty}^{+\infty} \varepsilon^q \frac{e^{-\frac{1}{2}\varepsilon^2}}{\sqrt{2\pi}} \, d\varepsilon \quad , \quad q = 1,2,\ldots \quad .$$

Using the notations (A1.1) and (A1.2) and realizing that $\hat{M}_o \hat{f} = \hat{M}_o X \hat{\beta} = 0$ identically we may rewrite (1.1) through (1.3) as:

(A1.4) $$\hat{t}_p = \frac{\hat{\alpha}_p}{\sqrt{\hat{\sigma}^2 \|\hat{M}_o Z \hat{\gamma}\|^{-2}}} \quad ,$$

(A1.5) $$\hat{\alpha}_p = \frac{\hat{\gamma}^T Z^T \hat{M}_o u}{\|\hat{M}_o Z \hat{\gamma}\|^2} \quad ,$$

(A1.6) $$\hat{\sigma} = (1/(n-k-1)) \| y - X\hat{\beta} - \hat{\alpha}_p \hat{M}_o Z \hat{\gamma} \|^2 \quad .$$

We shall need the following results, which are not hard to prove:

(A1.7) $$\operatorname*{plim}_{n\to\infty} (1/n) \, X^T X = I \ ,$$

(A1.8) $$\operatorname*{plim}_{n\to\infty} (1/n) \, Z^T Z = m_6 I = 15I \ ,$$

(A1.9) $$\operatorname*{plim}_{n\to\infty} (1/n) \, Z^T X = m_4 I = 3I \ ,$$

(A1.10) $$\operatorname*{plim}_{n\to\infty} (1/n) \, X^T y = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad ,$$

(A1.11) $$\operatorname*{plim}_{n\to\infty} (1/n) \, Z^T y = m_4 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad ,$$

and consequently

$$\text{(A1.12)} \qquad \underset{n\to\infty}{\text{plim}}\ \hat{\beta} = \begin{pmatrix}1\\1\end{pmatrix} ,$$

$$\text{(A1.13)} \qquad \underset{n\to\infty}{\text{plim}}\ \hat{\gamma} = \frac{m_4}{m_6}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{5}\begin{pmatrix}1\\1\end{pmatrix} .$$

These results will now be used for showing that $(1/\sqrt{n})\,\hat{\gamma}^T Z^T \hat{M}_o u$ is asymptotically normally distributed with zero mean. First, observe that

$$\text{(A1.14)} \qquad (1/\sqrt{n})\,\hat{\gamma}^T Z^T \hat{M}_o\, u = (\hat{\gamma}^T, -\hat{\gamma}^T((1/n)\,Z^T X)((1/n)\,X^T X)^{-1})\begin{pmatrix}(1/\sqrt{n})\,Z^T u\\ (1/\sqrt{n})\,X^T u\end{pmatrix} .$$

Since $E\,Z^T u = E\,X^T u = 0$ and since $\begin{pmatrix}Z^T u\\ X^T u\end{pmatrix}$ is a sum of i.i.d. random vectors in $\mathbb{R}^4$ we have by the central limit theorem

$$\text{(A1.15)} \qquad (1/\sqrt{n})\begin{pmatrix}Z^T u\\ X^T u\end{pmatrix} \to N_4\ (0,\ \Omega) \text{ in distribution,}$$

where

$$\text{(A1.16)} \qquad \Omega = \text{var}\left[(1/\sqrt{n})\begin{pmatrix}Z^T u\\ X^T u\end{pmatrix}\right] = \frac{1}{n}E\begin{pmatrix}(Z^T u)^T\,(Z^T u),\ (Z^T u)^T\ (X^T u)\\ (X^T u)^T\ (Z^T u),\ (X^T u)^T\ (X^T u)\end{pmatrix}$$

$$= m_2\begin{pmatrix} m_6+m_8, & 0, & m_4+m_6, & 0\\ 0, & m_6+m_8, & 0, & m_4+m_6\\ m_4+m_6, & 0, & m_2+m_4, & 0\\ 0, & m_4+m_6, & 0, & m_2+m_4\end{pmatrix}$$

$$= \begin{pmatrix} 120, & 0, & 18, & 0\\ 0, & 120, & 0, & 18\\ 18, & 0, & 4, & 0\\ 0, & 18, & 0, & 4\end{pmatrix} .$$

Moreover it follows from (A1.8), (A1.9) and (A1.13) that

$$\text{(A.17)} \qquad \underset{n\to\infty}{\text{plim}}\ (\hat{\gamma}^T, -\hat{\gamma}^T((1/n)\,Z^T X)\ ((1/n)\,X^T X))^{-1}) = \left(\frac{1}{5}, \frac{1}{5}, -\frac{3}{5}, -\frac{3}{5}\right) = \delta ,$$

say. Thus from (A1.14) through (A1.17) it follows

$$(A1.18)\qquad (1/\sqrt{n})\,\hat{\gamma}^T Z^T \hat{M}_o u \to N(0, \delta^T \Omega \delta) = N\left(0, \frac{96}{25}\right) \text{ in distribution.}$$

Next we observe that we may write

$$(A1.19)\qquad \|\hat{M}_o Z \hat{\gamma}\|^2 = \hat{\gamma}^T((1/n) Z^T Z)\hat{\gamma} - \hat{\gamma}^T((1/n) Z^T X)((1/n) X^T X)^{-1}((1/n) X^T Z)\hat{\gamma} \ ,$$

so that from (A1.7), (A1.8), (A1.9) and (A1.13) we have

$$(A1.20)\qquad \operatorname*{plim}_{n\to\infty} (1/n)\|\hat{M}_o Z \hat{\gamma}\|^2 = (1, 1)\begin{pmatrix}1\\1\end{pmatrix}\cdot\begin{Bmatrix}m_4^2 & - & m_4^4\\ \overline{m_6^2} & & \overline{m_6^2}\end{Bmatrix} = \frac{12}{25} \ .$$

Combining (A1.5), (A1.18) and (A1.20) we now conclude:

$$(A1.21)\qquad \sqrt{n}\,\hat{\alpha}_p \to N\left[0, \frac{96}{25}\Big/\left(\frac{12}{25}\right)^2\right] \text{ in distribution.}$$

Moreover, since now $\operatorname*{plim}_{n\to\infty} \hat{\alpha}_p = 0$, we have

$$(A1.22)\qquad \operatorname*{plim}_{n\to\infty} \hat{\sigma}^2 = \operatorname*{plim}_{n\to\infty} (1/n)\|y - X\hat{\beta}\|^2 = \operatorname*{plim}_{n\to\infty} (1/n) u^T u = m_2 + m_2^2 = 2 \ ,$$

and combining this result with (A1.20) we get

$$(A1.23)\qquad \operatorname*{plim}_{n\to\infty}\sqrt{\hat{\sigma}^2\|\hat{M}_o Z \hat{\gamma}\|^{-2}}\Big/\sqrt{n} = \sqrt{2\cdot\frac{25}{12}} \ .$$

The desired result (1.4) follows now easily from (A1.4), (A1.21) and (A1.23).

Q.E.D.

APPENDIX 2

Proof of Theorem 1

Proof of Part I

From Chung [3, Theorem 9.1.2] it follows that there exists a Borel measurable real function r, say, on R^k such that

(A2.1) $$E(v|z) = r(z) \quad \text{a.s.} .$$

Put

(A2.2) $$r_1(\cdot) = \max\{r(\cdot), 0\}, \quad r_2(\cdot) = \max\{-r(\cdot), 0\} .$$

Then obviously r_1 and r_2 are nonnegative Borel measurable real functions on R^k satisfying

(A2.3) $$r = r_1 - r_2 \ .$$

Now assume for the moment

(A2.4) $$c_1 = E\, r_1(z) > 0 , \quad c_2 = E\, r_2(z) > 0 .$$

Then we can define probability measures μ_1 and μ_2 on the Euclidean Borel field $\mathcal{B}$ as follows:

(A2.5) $$\mu_j(B) = \int_B r_j(x)\, d\mu(x)/c_j , \quad j = 1, 2 ,$$

where μ is the probability measure generated by the random vector z and B is an arbitrary Borel set in $\mathcal{B}$. Then we may write:

(A2.6) $$E\, v e^{it^T z} = E\, r(z)\, e^{it^T z} = \int r(x)\, e^{it^T x}\, d\mu(x)$$
$$= \int r_1(x)\, e^{it^T x} d\mu(x) - \int r_2(x)\, e^{it^T x}\, d\mu(x)$$
$$= c_1 \int e^{it^T x}\, d\mu_1(x) - c_2 \int e^{it^T x}\, d\mu_2(x) = c_1 \eta_1(t) - c_2 \eta_2(t) ,$$

say, where

(A2.7) $$\eta_j(t) = \int e^{it^T x}\, d\mu_j(x)/c_j\ , \quad (j = 1, 2)$$

are the characteristic functions of the probability measures μ_j $(j = 1, 2)$, respectively.

If $E\, v\, e^{it^T z} \equiv 0$ for every $t \in \mathbb{R}^k$ then it follows from (A2.6) that

(A2.8) $$c_1\eta_1(t) - c_2\eta_2(t) = 0 \text{ for every } t \in \mathbb{R}^k .$$

Hence, substituting $t = 0$, we get

(A2.9) $$c_1\eta_1(0) - c_2\eta_2(0) = c_1 - c_2 = 0 ,$$

so that from (A2.4), (A2.8) and (A2.9)

(A2.10) $$\eta_1(t) = \eta_2(t) \text{ for every } t \in \mathbb{R}^k .$$

But (A2.10) implies that the probability measures μ_1 and μ_2 are equal, i.e.,

(A2.11) $$\mu_1(B) = \mu_2(B) \text{ for every Borel set } B .$$

From (A2.5), (A2.9) and (A2.11) we now obtain

(A2.12) $$\int_B r_1(x)\, d\mu(x) = \int_B r_2(x)\, d\mu(x) \text{ for every Borel set } B$$

and consequently

(A2.13) $$\int_B r(x)\, d\mu(x) = 0 \text{ for every Borel set } B .$$

But

(A2.14) $$B_1 = \{x \in \mathbb{R}^k : r(x) > 0\}$$

is a Borel set, and thus:

(A2.15) $$0 = \int_{B_1} r(x)\, d\mu(x) ,$$

which is only possible if B_1 is a null set with respect to μ. Similarly we conclude that the Borel set

(A2.16) $$B_2 = \{x \in \mathbb{R}^k : r(x) < 0\}$$

is a null set with respect to μ, and hence

(A2.17) $$B_1 \cup B_2 = \{x \in \mathbb{R}^k : r(x) \neq 0\}$$

is a null set with respect to μ. This means that $r(z) = 0$ a.s. . Thus we have proved by now that if (A2.4) holds and if $E\,v\,e^{it^Tz} = 0$ for all $t \in \mathbb{R}^k$ then $E(v|z) = 0$ a.s. . However, if (A2.4) does not hold then our conclusion still holds, as is not hard to prove. This completes the "only if" part of part I of theorem 1.

Since the "if" part is trivial, part I of theorem 1 is proved by now.

Q.E.D.

<u>Proof of Part II</u>

Since now z is bounded we may write

(A2.18) $$E\,v\,e^{it^Tz} = E\,v \sum_{j=0}^{\infty} \frac{i^j}{j!}(t^Tz)^j = \sum_{j=0}^{\infty} \frac{i^j}{j!} E\,v\,(t^Tz)^j \quad .$$

So if $E\,v\,e^{it_*^Tz} \neq 0$ for some $t_* \in \mathbb{R}^k$ then there exists a nonnegative integer j_* such that

(A2.19) $$E\,v\,(t_*^Tz)^{j_*} \neq 0 \quad .$$

Assuming that j_* is minimal, we therefore have

(A2.20) $$\lim_{\lambda \downarrow 0} (\partial/\partial\lambda)^{j_*} E\,v\,e^{i\lambda t_*^Tz} = i^{j_*} E\,v\,(t_*^Tz)^{j_*} \neq 0,$$

which implies that $E\,v\,e^{i\lambda t_*^Tz} \neq 0$ for an arbitrarily small $\lambda > 0$, say λ_*. Putting $t_o = \lambda_* t_*$, this proves part II of the theorem.

Q.E.D.

REFERENCES

[1] BIERENS, H. J.: Robust Methods and Asymptotic Theory in Nonlinear Econometrics, Lecture Notes in Economics and Mathematical Systems, vol. 192, Heidelberg: Springer-Verlag, 1981.

[2] BIERENS, H. J.: "A Uniform Weak Law of Large Numbers under φ - Mixing with Application to Nonlinear Least Squares Estimation," Statistica Neerlandica, 36 (1982) (to appear).

[3] CHUNG, K. L.: A Course in Probability Theory, New York: Academic Press, 1974.

[4] DAVIDSON, R. AND L. G. MACKINNON: "Several Tests for Model Specification in the Presence of Alternative Hypotheses," Econometrica, 49 (1981), 781 - 793.

[5] JENNRICH, R. I.: "Asymptotic Properties of Nonlinear Least Squares Estimators," The Annals of Mathematical Statistics, 40 (1969), 633 - 643.

[6] PESARAN, M. H. AND A. S. DEATON: "Testing Non-nested Nonlinear Regression Models," Econometrica, 46 (1978), 677 - 694.