A TEST FOR MODEL SPECIFICATION OF NONLINEAR TIME SERIES REGRESSIONS

by

Herman J. Bierens*

Discussion Paper No. 81 - 156, November 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 a previous paper [3] we have proposed a test for model specification in the absence of alternative hypotheses. The asymptotic properties of this test have been derived under the i.i.d. assumption. In the present paper we shall extend and modify these results to simultaneously testing the truth of the functional form of the model and the absence of autocorrelation of the errors of nonlinear and nonstationary time series regressions.

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 our previous paper [3] we have introduced a new approach for testing the functional form of a regression model without using specified alternative hypotheses. The asymptotic properties of this test have been derived under the i.i.d. assumption on the distribution of the data. In the present paper we shall extend these results to nonlinear timeseries regressions, i.e. we shall consider nonlinear regression models where one or more explanatory variables are lagged dependent variables. However, dealing with a time series model not only the functional form of the model is important but also the structure of the error distribution, for it is well known that dynamic regression models with autocorrelated errors cannot be estimated consistently by least squares estimation. Therefore we shall modify our previous test to simultaneously testing the truth of the functional form and the absence of serial dependence of the errors. Moreover, we now shall also allow for nonstationarity of the exogenous variables (and hence of the dependent variables). This will complicate our analysis slightly, but the additional effort is counterbalanced by an obvious increase of generality.

2. PRELIMINARIES

For extending the results in our previous paper [3] to nonstationary nonlinear time series regressions we shall need the concepts of stochastic stability, nonstationary ϕ-mixing and proper convergence of distribution functions.

The stochastic stability concept is introduced in Bierens [1, Chapter 5]. In order to make this concept clear we consider a non-linear vector autoregression of the form

$$(2.1)\quad y_j = \Gamma(y_{j-1},y_{j-2},\ldots,y_{j-p},w_j) \quad , \quad -\infty < j < \infty,$$

where the y_j are random vectors in $\mathbb{R}^s$, the w_j are random vectors in $\mathbb{R}^q$ and Γ is a Borel measurable mapping from $\mathbb{R}^p \times \mathbb{R}^q$ into $\mathbb{R}^s$. By m-times backwards substitution of (2.1) we can write

$$(2.2)\quad y_j = \Gamma_m^*(w_j,w_{j-1},\ldots,w_{j-m},y_{j-m-1},\ldots,y_{j-m-p}), \quad -\infty < j < \infty.$$

Roughly speaking this stochastic process is a stable process (in a similar sense as for deterministic difference equations) if the impact of the initial values $y_{j-m-1},\ldots,y_{j-m-p}$ on the distribution of y_j vanishes if $m \to \infty$. In that case we may "approximate" y_j by a function of only $w_j,w_{j-1},\ldots,w_{j-m}$, for example by replacing the random vectors $y_{j-m-1},\ldots,y_{j-m-p}$ in (2.2) by appropriate constant vectors $c_{j-m-1},\ldots,c_{j-m-p}$, say, i.e. we then approximate y_j by

$$(2.3)\quad y_j^{(m)} = \Gamma_m^*(w_j,w_{j-1},\ldots,w_{j-m},c_{j-m-1},\ldots,c_{j-m-p}).$$

Another way of approximating y_j is to take the expectation of y_j conditional on $w_j,w_{j-1},\ldots,w_{j-m}$, thus

(2.4) $y_j^{(m)} = E(y_j \mid w_j, w_{j-1}, \ldots, w_{j-m})$.

Now if for such an approximation, only depending on $w_j, \ldots, w_{j-m}$, there exists a sequence (m_n) of positive integers such that together

(2.5) $m_n = o(n)$ for $n \to \infty$; $\lim_{n\to\infty} \frac{1}{n} \Sigma_{j=1}^{n} P[\| y_j - y_j^{(m_n)} \| > \delta] = 0$

for every $\delta > 0$,

then the stochastic process (y_j) is said to be stochastically stable with respect to the base (w_j) (See [1, Definitions 5.1.1 and 5.4.1]).

As is shown in [1], stochastic stability of linear autoregressions is closely related to the usual stability condition that the lag operator on y_j has roots all outside the unit circle. If for example the y_j's and w_j's are real valued and if Γ is a linear function, then (2.1) can be written as

(2.6) $y_j = \alpha_1 y_{j-1} + \alpha_2 y_{j-2} + \ldots + \alpha_p y_{j-p} + \beta w_j$

or equivalently

(2.7) $A(L) y_j = \beta w_j$,

where

(2.8) $A(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \ldots - \alpha_p L^p$

is a polynomial lag operator. If all the roots of $A(L)$ lie outside the unit circle then (y_j) is stochastically stable with respect to (w_j). (See [1, Theorem 5.1.2]).

A second concept we shall need is that of nonstationary ϕ-mixing. ϕ-Mixing processes are well known in the mathematical statistical literature (see for example Billingsley [4]), but it is usually employed together with stationarity. The nonstationary ϕ-mixing concept is used in [2] to prove weak consistency and asymptotic normality of nonlinear least squares estimators for the case that the data are generated by a stochastically stable process with respect to a nonstationary ϕ-mixing base. This concept can be defined as follows. Let $(v_j)_{-\infty}^{+\infty}$ be a nonstationary stochastic process in $\mathbb{R}^q$, let $\mathcal{F}_{-\infty}^{j}$ be the Borel field generated by $v_j, v_{j-1}, v_{j-2}, \ldots$ and let $\mathcal{F}_{j}^{\infty}$ be the Borel field generated by $v_j, v_{j+1}, v_{j+2}, \ldots$. The process (v_j) is said to be <u>nonstationary</u> ϕ-<u>mixing</u> if for each integer j $(-\infty < j < \infty)$ and each nonnegative integer i, $E_1 \in \mathcal{F}_{-\infty}^{j}$ and $E_2 \in \mathcal{F}_{j+i}^{\infty}$ together imply

$$(2.9)\quad |P(E_1 \cap E_2) - P(E_1)P(E_2)| \leq \phi(i,j)P(E_1),$$

where ϕ is a function such that

$$(2.10)\quad \lim_{i\to\infty} \phi(i,j) = 0 \text{ pointwise for all } j.$$

A third concept we shall need is that of proper convergence of distribution functions. This concept is well known (see Feller [5, p. 243]) but for convenience we will recall it here. Thus let $(F_n)_{n=1}^{\infty}$ be a sequence of distribution functions on $\mathbb{R}^k$. We say that F_n <u>converges properly</u> for $n \to \infty$ if there exists a distribution function F on $\mathbb{R}^k$ such that $\lim_{n\to\infty} F_n(x) = F(x)$ for every continuity point x of F, and we then write: $F_n \to F$ properly. Moreover we note that $F_n \to F$ properly implies the proper convergence of the marginal distributions of F_n to the corresponding marginal distributions of F (See [1, pp. 108-109]).

These three concepts are the basic ingredients for extending our previous results in [3] to nonstationary time series regressions, as we now can use Theorem 2 in [2] instead of Theorem 2.3.4 in [1] for proving uniform convergence in probability, i.e.:

LEMMA 1. Let Y be a Euclidean space. Let (y_j) be a Y-valued stochastically stable process with respect to a nonstationary ϕ-mixing base, where ϕ satisfies

$$\lim_{n\to\infty} \frac{1}{n^2} \Sigma_{i=1}^{n}\Sigma_{j=1}^{n} \phi(i,j) = 0. \tag{2.11}$$

Let F_j be the distribution function of y_j and assume

$$\frac{1}{n} \Sigma_{j=1}^{n} F_j \to F \text{ properly.} \tag{2.12}$$

Let ψ be a continuous real function on $Y \times \Theta$, where Θ is a compact subset of a Euclidean space. Assume

$$\sup_{n\geq 1} \frac{1}{n} \Sigma_{j=1}^{n} E\{\sup_{\|y\| \leq \|y_j\|} \sup_{\theta\in\Theta} |\psi(y,\theta)|\}^{1+\delta} < \infty \tag{2.13}$$

for some $\delta > 0$. Then

$$\operatorname*{plim}_{n\to\infty} \sup_{\theta\in\Theta} \left|\frac{1}{n} \Sigma_{j=1}^{n}\psi(y_j,\theta) - \int_Y \psi(y,\theta)\,dF(y)\right| = 0.$$

Proof: Bierens [2, Theorem 2].

3. THE DATA

Our data set consists of n observations

$$\{(y_1,x_1^*),\ldots,(y_n,x_n^*)\}$$

on a vector time series process $\{(y_j,x_j^*)\}_{-\infty}^{+\infty}$ in $\mathbb{R}\times\mathbb{R}^q$. This process has the following form.

ASSUMPTION 1A. For $-\infty<j<\infty$ we have

$$(3.1)\quad y_j = g(y_{j-1},y_{j-2},\ldots,y_{j-p},x_j^*) + u_j = g(x_j) + u_j, \quad \text{say,}$$

where

$$(3.2)\quad x_j^T = (y_{j-1},y_{j-2},\ldots,y_{j-p},x_j^*) \in \mathbb{R}^k;\quad k = p+q,$$

g is a continuous real function on $\mathbb{R}^k$ not depending on j such that

$$(3.3)\quad \sup_{n\geq 1} \frac{1}{n}\Sigma_{j=1}^n E|g(x_j)|^{2+\delta} < \infty \quad \text{for some} \quad \delta > 0,$$

and (u_j) is a strictly stationary autoregressive stochastic process in $\mathbb{R}$ of the form

$$(3.4)\quad u_j = \psi(u_{j-1}) + \varepsilon_j,$$

with (ε_j) an i.i.d. process satisfying

(3.5) $E\varepsilon_j = 0, \quad E|\varepsilon_j|^{2+\delta} < \infty$ for some $\delta > 0$,

ε_j is independent of $x_j, x_{j-1}, x_{j-2}, \ldots,$

and ψ a <u>continuous</u> real function on $\mathbb{R}$ such that

(3.6) $E\psi(u_{j-1}) = 0, \quad E|\psi(u_{j-1})|^{2+\delta} < \infty$ for some $\delta > 0$.

Moreover we assume:

ASSUMPTION 1B. The sequence (x_j^*) of exogenous variables is a nonstationary ϕ-mixing process in $\mathbb{R}^q$, where ϕ satisfies:

$$\lim_{n\to\infty} \frac{1}{n^2} \Sigma_{i=1}^n \Sigma_{j=1}^n \phi(i,j) = 0.$$

ASSUMPTION 1C. The functions g and ψ are such that the process $\{(y_j, u_j)\}$ is stochastically stable with respect to the base $\{(x_j^*, \varepsilon_j)\}$.

ASSUMPTION 1D. The distribution functions H_j of $(x_j, u_j, x_{j-1}, u_{j-1})$

satisfy $\frac{1}{n}\Sigma_{j=1}^{n}H_j \to H$ properly.

These assumptions are assumptions on the distribution of the data, regardless how the functions g and ψ are parameterized. So they are assumed to hold for every hypothesis with respect to the functional form of g and ψ.

Although the error process (u_j) is stationary, the process (y_j) does not need to be stationary because the process (x_j^*) of exogenous variables is allowed to be nonstationary. The stationarity of the u_j implies that u_j can be written as

$$u_j = \varepsilon_j + \psi_\infty\,(\varepsilon_{j-1},\varepsilon_{j-2},\ldots)$$

where ψ_∞ is a Borel measurable function on $\mathbb{R}^\infty$. Since the ε_j are i.i.d. we then see that ε_j is independent of $u_{j-1},u_{j-2},\ldots$. The function ψ can now be interpreted as a conditional expectation function:

$$\psi(u_{j-1}) = E(u_j|u_{j-1}),$$

but $g(x_j)$ is <u>not</u> necessarily the expectation of y_j conditional on x_j, except in the case that $\psi(u_{j-1}) = 0$ a.s..

4. THE HYPOTHESES AND FURTHER ASSUMPTIONS

The Hypotheses

The specification of the model described by the equations (3.1) and (3.4) is twofold in the sense that both g and ψ have to be specified. However, the null hypothesis with respect to the latter function is quite simple, as our aim is to test the absence of autoregressive disturbances. Thus the null hypothesis involved is

$$(4.1) \quad \psi(u_{j-1}) = 0 \quad \text{a.s.}$$

or equivalently

$$(4.2) \quad u_j = \varepsilon_j \quad \text{a.s.}$$

and since (u_j) is a stationary process, (4.1) and (4.2) then hold for every j.

Now suppose we have specified g as

$$(4.3) \quad g(x_j) = f(x_j,\theta_0) \quad \text{a.s.}, \quad -\infty < j < \infty,$$

where

ASSUMPTION 2. $f(x,\theta)$ and its 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 $\mathbb{R}^k \times \Theta$ with Θ a compact convex subset of $\mathbb{R}^m$.

In the nonstationary case the distributions of the x_j's may differ for each j, hence the hypothesis (4.3) consists in fact of an infinite number of distinct hypotheses and for testing each of these hypotheses we have only one observation x_j. However, if (4.3) holds for every $j \geq 1$ then for each $n \geq 1$

$$(4.4)\quad g(\cdot) = f(\cdot,\theta_0) \text{ a.s. with respect to } \frac{1}{n}\Sigma_{j=1}^{n} H_j^{(1)},$$

where $H_j^{(1)}$ is the marginal distribution of x_j, i.e.

$$(4.5)\quad H_j^{(1)}(x) = \lim_{\substack{u\to\infty \\ x_*^T \to (\infty,\ldots,\infty) \\ u_* \to \infty}} H_j(x,u,x_*,u_*)\ ,$$

and because of assumption 1D the same holds for the corresponding marginal distribution of H, i.e.

$$(4.6)\quad g(\cdot) = f(\cdot,\theta_0) \text{ a.s. with respect to } H^{(1)},$$

where

$$(4.7)\quad H^{(1)}(x) = \lim_{\substack{u\to\infty \\ x_*^T \to (\infty,\ldots,\infty) \\ u_* \to \infty}} H(x,u,x_*,u_*).$$

Obviously in the stationary case (4.3) and (4.6) are equivalent because then $H^{(1)}$ is just the common distribution function of the x_j's, but in the nonstationary case (4.6) does not imply (4.3), for (4.6) still holds if for example (4.3) fails to hold for a finite number of j's. Nevertheless, from a practical point of view (4.6) is the only feasible form of a null hypothesis with respect to the functional form of g, and therefore we shall adopt it. Realizing that (4.6) is equivalent with

$$\int_{\mathbb{R}^k} I[g(x) = f(x,\theta_0)]dH^{(1)}(x) = 1,$$

where $I(\cdot)$ is the well known indicator function, the null hypothesis can now be stated as follows:

(4.8) H_0: 1) $P[\psi(u_{j-1}) = 0] = 1$, <u>and</u>,

2) $\int_{\mathbb{R}^k} I[g(x) = f(x,\theta_0)]dH^{(1)}(x) = 1$ for a point $\theta_0 \in \Theta$.

The alternative hypothesis is that H_0 is false, i.e.

(4.9) H_1: 1) $P[\psi(u_{j-1}) = 0] < 1$, <u>and/or</u>

2) $\int_{\mathbb{R}^k} I[g(x) = f(x,\theta)]dH^{(1)}(x) < 1$ for all $\theta \in \Theta$.

<u>Further Assumptions</u>

The further assumptions we shall need are somewhat similar to the assumptions 3 through 5 in [3]. The differences are mainly due to the fact that we now have to account for nonstationarity and that under H_1, u_j and x_j need not to be independent. Thus:

ASSUMPTION 3. There exists a $\delta > 0$ such that

$$\sup_{n\geq 1} \frac{1}{n} \Sigma_{j=1}^{n} E\{\sup_{\|x\|\leq\|x_j\|} \sup_{\theta\in\Theta} |f(x,\theta)|\}^{2+\delta} < \infty,$$

$$\sup_{n\geq 1} \frac{1}{n} \Sigma_{j=1}^{n} E\{\sup_{\|x\|\leq\|x_j\|} \sup_{\theta\in\Theta} |(\partial/\partial\theta_i) f(x,\theta)|\}^{2+\delta} < \infty, \ (i=1,2,\ldots,m),$$

$$\sup_{n\geq 1} \frac{1}{n} \Sigma_{j=1}^{n} E\{\sup_{\|x\|\leq\|x_j\|} \sup_{\theta\in\Theta} |(\partial/\partial\theta_{i_1})(\partial/\partial\theta_{i_2}) f(x_j,\theta)\}^{2+\delta} < \infty,$$

$$(i_1,i_2=1,2,\ldots,m).$$

ASSUMPTION 4A. If H_0 is true then θ_0 is a unique interior point of Θ such that

$$0 = \int_{\mathbb{R}^k} \{g(x) - f(x,\theta_0)\}^2 dH^{(1)}(x) = \inf_{\theta\in\Theta} \int_{\mathbb{R}^k} \{g(x) - f(x,\theta)\}^2 dH^{(1)}(x).$$

(We recall that $H^{(1)}$ is defined by (4.7)).

ASSUMPTION 4B. If H_1 is true then there exists a unique point θ_* in Θ such that

$$0 < \int_{\mathbb{R}^k\times\mathbb{R}} \{\psi(u_*) + g(x) - f(x,\theta_*)\}^2 dH^{(2)}(x,u_*) =$$

$$= \inf_{\theta\in\Theta} \int_{\mathbb{R}^k\times\mathbb{R}} \{\psi(u_*) + g(x) - f(x,\theta)\}^2 dH^{(2)}(x,u_*),$$

where

$$(4.10) \quad H^{(2)}(x,u_*) = \lim_{\substack{u \to \infty \\ x_*^T \to (\infty,\ldots,\infty)}} H(x,u,x_*,u_*).$$

ASSUMPTION 5. The matrix

$$(4.11) \quad A(\theta) = \int_{\mathbb{R}^k} \{(\partial/\partial\theta^T)f(x,\theta)\}\{(\partial/\partial\theta)f(x,\theta)\}dH^{(1)}(x)$$

is nonsingular for $\theta = \theta_0$ as well as for $\theta = \theta_*$.

The modifications of the assumptions 3 and 5, compared with the corresponding assumptions in [3], are due to nonstationarity. Assumption 4, however, has now been split apart. Assumption 4A is quite obvious, but assumption 4B might need some comment. In order to see what assumption 4B means it is helpful to assume for the moment that (x_j) is stationary. In that case, $H^{(2)}(x,u_*)$ is just the distribution function of (x_j,u_{j-1}), so that assumption 4B then becomes:

$$0 < E\{\psi(u_{j-1}) + g(x_j) - f(x_j,\theta_*)\}^2 = \inf_{\theta \in \Theta} E\{\psi(u_{j-1}) + g(x_j) - f(x_j,\theta)\}^2$$

where $\theta_* \in \Theta$ is unique. We now see why we have assumed that the infimum is positive, for otherwise $\psi(u_{j-1}) + g(x_j) - f(x_j,\theta_*) = 0$ with probability 1 and hence $y_j = f(x_j,\theta_*) + \varepsilon_j$ with probability 1, which is just the case described by the null hypotheses, whereas assumption 4B aims to deal with the case that H_0 is false.

In view of the above argument we see that the hypotheses H_0 and H_1 described by (4.8) and (4.9), respectively, may now be restated as follows:

(4.12) H_0: $\int_{\mathbb{R}^k \times \mathbb{R}} I[\psi(u_*) + g(x) - f(x,\theta_0) = 0]dH^{(2)}(x,u_*) = 1$

for some $\theta_0 \in \Theta$,

(4.13) H_1: $\int_{\mathbb{R}^k \times \mathbb{R}} I[\psi(u_*) + g(x) - f(x,\theta) = 0]dH^{(2)}(x,u_*) < 1$ for

all $\theta \in \Theta$.

5. IDENTIFICATION

In this section we show that Theorem 1 in [3] can also be used for identifying the present hypotheses H_0 and H_1. In our argument the following easy lemma will play a role.

LEMMA 2. Put

$$(5.1)\quad w_j^T = (y_{j-p-1}, x_{j-1}^{*T}) \qquad (w_j \in \mathbb{R}^{q+1})$$

Then for every $\theta \in \Theta$, $(x_j, w_j, u_{j-1} + g(x_{j-1}) - f(x_{j-1}, \theta))$ and (x_j, w_j, u_{j-1}) generate the same Borel field.

Proof: This lemma follows easily by observing from (3.2) that x_{j-1} can be constructed from x_j and w_j, and that (x_j, w_j) can be constructed from (x_j, x_{j-1}), so that (x_j, w_j, u_{j-1}) can be constructed from $(x_j, w_j, u_{j-1} + g(x_{j-1}) - f(x_{j-1}, \theta))$ and visa versa, given the functions g and f. Q.E.D.

Now let Φ_1 be a bounded continuous mapping from $\mathbb{R}^k$ into $\mathbb{R}^k$ such that $\Phi_1(x)$ and x $(x \in \mathbb{R}^k)$ generate the same Euclidean Borel field, and define the mappings Φ_2 $(\mathbb{R}^{q+1} \to \mathbb{R}^{q+1})$ and $\Phi_3(\mathbb{R} \to \mathbb{R})$ similarly. For example, Φ_1, Φ_2 and Φ_2 may be vector functions with components $\operatorname{atg}(\cdot)$. Moreover, put

$$(5.2)\quad \hat{\xi}_j(t_1,t_2,t_3,\theta) =$$

$$= \{y_j - f(x_j,\theta)\} \cdot e^{it_1^T\Phi_1(x_j) + it_2^T\Phi_2(w_j) + it_3\Phi_3(y_{j-1} - f(x_{j-1},\theta))}$$

$$= \{\varepsilon_j + \psi(u_{j-1}) + g(x_j) - f(x_j,\theta)\} \cdot$$

$$\cdot\, e^{it_1^T\Phi_1(x_j) + it_2^T\Phi_2(w_j) + it_3\Phi_3(u_{j-1} + g(x_{j-1}) - f(x_{j-1},\theta))},$$

where $(t_1,t_2,t_3) \in \mathbb{R}^k \times \mathbb{R}^{q+1} \times \mathbb{R}$ is nonrandom, and assume for the moment that the data are generated by a strictly stationary process. From Theorem 1 in [3] it follows that for given $\theta \in \Theta$,

$$E\hat{\xi}_j(t_1,t_2,t_3,\theta) = E[\{\psi(u_{j-1}) + g(x_j) - f(x_j,\theta)\} \cdot$$

$$\cdot\, e^{it_1^T\Phi_1(x_j) + it_2^T\Phi_2(w_j) + it_3\Phi_3(u_{j-1} + g(x_{j-1}) - f(x_{j-1},\theta))}] = 0$$

for every (t_1,t_2,t_3) in an arbitrarily small neighborhood of $(0,0,0)$ if and only if

$$(5.3)\quad E[\psi(u_{j-1}) + g(x_j) - f(x_j,\theta)\,|\,\Phi_1(x_j),\Phi_2(w_j),\Phi_3(u_{j-1} + g(x_j) -$$

$$- f(x_{j-1},\theta))] = 0 \quad \text{a.s.}$$

But $(\Phi_1(x_j),\Phi_2(w_j),\Phi_3(u_{j-1} + g(x_{j-1}) - f(x_{j-1},\theta))$ generates the same Euclidean Borel field as $(x_j,w_j,u_{j-1} + g(x_{j-1}) - f(x_{j-1},\theta))$ does, and by lemma 2 the latter generates the same Euclidean Borel field as (x_j,w_j,u_{j-1}) does. Thus (5.3) is equivalent with

$$(5.4)\quad E\{\psi(u_{j-1}) + g(x_j) - f(x_j,\theta)\,|\,x_j,w_j,u_{j-1}\} = 0 \quad \text{a.s.}$$

which in its turn is equivalent with

(5.5) $\psi(u_{j-1}) + g(x_j) - f(x_j,\theta) = 0$ a.s..

Comparing (5.5) with (4.12) and (4.13) and realizing that we have assumed stationarity, we now see that (5.5) can only hold for some $\theta \in \Theta$ if H_0 is true, and consequently $E\hat{\xi}_j(t_1,t_2,t_3,\theta) = 0$ for every (t_1,t_2,t_3) in an arbitrarily small neighborhood of $(0,0,0)$ if and only if H_0 is true and $\theta = \theta_0$.

Now drop the stationarity assumption and denote:

(5.6) $H_j^{(3)}(x,w,x_*,u_*) = P(x_j \leq x, w_j \leq w, x_{j-1} \leq x_*, u_{j-1} \leq u_*)$.

It then follows from assumption 1D that

(5.7) $\frac{1}{n}\Sigma_{j=1}^n H_j^{(3)} \to H^{(3)}$ properly,

where $H^{(3)}$ is a marginal distribution of H corresponding to H in a similar way as $H_j^{(3)}$ corresponds to H_j. Moreover, let

$$(5.8)\quad \xi(t_1,t_2,t_3,\theta) = \int_{\mathbb{R}^k \times \mathbb{R}^{q+1} \times \mathbb{R}} \{\psi(u_*) + g(x) - f(x,\theta)\} \cdot$$

$$\cdot\ e^{it_1\Phi_1(x) + it_2\Phi_2(w) + it_3\Phi_3(u_* + g(x_*) - f(x_*,\theta))} \cdot$$

$$\cdot\ dH^{(3)}(x,w,x_*,u_*) .$$

From the above argument for the stationary case it is now clear that the following theorem holds.

THEOREM 1. If the assumptions 1 through 4 are satisfied then $\xi(t_1,t_2,t_3,\theta) = 0$ for all (t_1,t_2,t_3) in an arbitrarily small neighborhood of $(0,0,0)$ if and only if H_0 is true and $\theta = \theta_0$.

Notice that if we choose arbitrary positive numbers $\varepsilon_1,\ldots,\varepsilon_k,\varepsilon_{k+1},\ldots,\varepsilon_{k+q+1},\varepsilon_{k+q+2}$ and if we put

$$(5.9)\quad N_0^{(1)} = \mathop{\mathrm{X}}_{\ell=1}^{k} [-\varepsilon_\ell,\varepsilon_\ell];\ N_0^{(2)} = \mathop{\mathrm{X}}_{\ell=k+1}^{k+q+1} [-\varepsilon_\ell,\varepsilon_\ell];\ N_0^{(3)} = [-\varepsilon_{k+q+2},\varepsilon_{k+q+2}],$$

$$(5.10)\quad \eta(\theta) = \int_{N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} |\xi(t_1,t_2,t_3,\theta)|^2 dt_1 dt_2 dt_3 ,$$

then theorem 1 implies:

COROLLARY 1: If the assumptions 1 through 4 are satisfied then H_0 is true if $0 = \eta(\theta_0) = \inf_{\theta\in\Theta} \eta(\theta)$ and H_1 is true if $\inf_{\theta\in\Theta} \eta(\theta) > 0$.

6. THE TEST

The procedure for testing the truth of H_0 is similar as in our previous paper [3]. Thus let $\hat{\theta}$ be the nonlinear least squares estimator of θ_0 and put

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

(6.2) $$\hat{z}_j^T = (\Phi_1(x_j)^T, \Phi_2(w_j)^T, \Phi_3(\hat{u}_{j-1})) = (\hat{z}_{1,j},\ldots,\hat{z}_{k*,j}),\ k* = k+q+2,$$

(6.3) $$t^T = (t_1^T, t_2^T, t_3),\quad t \in \mathbb{R}^{k*},$$

(6.4) $$N_0 = N_0^{(1)} \times N_0^{(2)} \times N_0^{(3)},$$

(6.5) $$\hat{\hat{\eta}} = \int_{N_0} \left|\frac{1}{n}\Sigma_{j=1}^n \hat{u}_j e^{it^T\hat{z}_j}\right|^2 dt =$$

$$= 2^{k*}\frac{1}{n^2}\Sigma_{j_1=1}^n \Sigma_{j_2=1}^n \hat{u}_{j_1}\hat{u}_{j_2} \prod_{\ell=1}^{k*} \frac{\sin[\varepsilon_\ell(\hat{z}_{\ell,j_1} - \hat{z}_{\ell,j_2})]}{\hat{z}_{\ell,j_1} - \hat{z}_{\ell,j_2}} ,$$

and

(6.6) $$\hat{\hat{\mu}} = 2^{k*}\hat{\sigma}^2\{\prod_{\ell=1}^{k*} \varepsilon_\ell - tr[\hat{A}^{-1}\hat{B}]\} ,$$

where $\hat{\sigma}^2$ is the usual estimate of the variance of the disturbance u_j,

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

and

$$(6.8)\quad \hat{\hat{B}} = \frac{1}{n^2}\Sigma_{j_1=1}^{n}\Sigma_{j_2=1}^{n}\{(\partial/\partial\theta^T)f(x_{j_1},\hat{\theta})\}\{(\partial/\partial\theta)f(x_{j_2},\hat{\theta})\cdot$$

$$\cdot \prod_{\ell=1}^{k*} \frac{\sin[\varepsilon_\ell(\hat{z}_{\ell,j_1} - \hat{z}_{\ell,j_2})]}{\hat{z}_{\ell,j_1} - \hat{z}_{\ell,j_2}}$$

We then have:

THEOREM 2. If the assumptions 1 through 5 are satisfied and if $(d/du)\Phi_3(u)$ is continuous and uniformly bounded then for every $\alpha \in (0,1)$,

$$(6.9)\quad \limsup_{n\to\infty} P[n\hat{\hat{\eta}} > \frac{1}{\alpha}\hat{\hat{\mu}}] \leq \alpha \quad \text{if } H_0 \text{ is true}$$

and

$$(6.10)\quad \lim_{n\to\infty} P[n\hat{\hat{\eta}} > \frac{1}{\alpha}\hat{\hat{\mu}}] = 0 \quad \text{if } H_1 \text{ is true.}$$

Proof: Section 7.

Thus proceeding the test at the $\alpha.100$ % significance level we accept H_0 if $n\hat{\hat{\eta}} \leq \frac{1}{\alpha}\hat{\hat{\mu}}$ and we reject H_0 if not.

Remarks:

1) Also now the statistic $\hat{\hat{\mu}}$ is a consistent estimator of the first moment of the limiting distribution of $n\hat{\hat{\eta}}$ in the case that H_0 is true, so that (6.9) is simply a consequence of Chebishev's inequality for first absolute moments. In our previous paper [3] we have proposed a Monte Carlo approach for deriving a sharper bound than $\frac{1}{\alpha}\hat{\hat{\mu}}$.

In the present case, however, this is no longer feasible, because the Monte Carlo approach requires that the sequence (u_j) of disturbances is independent of the sequence (x_j) of regressors, whereas in the present case we only have that under H_0 $u_j(=\varepsilon_j)$ is independent of $x_j, x_{j-1}, x_{j-2}, \ldots$ for each j but not that x_j is independent of $u_j, u_{j-1}, u_{j-2}, \ldots$ for each j, for x_j contains lagged y_j's as components.

2) If we are convinced that the disturbances u_j of the model (3.1) are i.i.d., but not that the specification $f(x,\theta)$ of $g(x)$ is true, we may use the above test for only testing this specification by replacing $\hat{z}_j$ by $z_j = \Phi_1(x_j)$ and k^* by k.

3) The above test can also be modified for simultaneously testing the truth of the specification of g and the absence of autocorrelation of the form

$$(6.11)\quad u_j = \psi(u_{j-1}, u_{j-2}, \ldots, u_{j-r}) + \varepsilon_j.$$

We then have to replace w_j^T (defined by (5.1)) by

$$(6.12)\quad w_j^{*T} = (y_{j-p-1}, \ldots, y_{j-p-r}, x_{j-1}^*, \ldots, x_{j-r}^*),$$

Φ_2 by a similar mapping Φ_2^* from $\mathbb{R}^{r\cdot q+r}$ into $\mathbb{R}^{r\cdot q+r}$ and $\hat{z}_j$ by

$$(6.13)\quad \hat{z}_j^{*T} = (\Phi_1(x_j)^T, \Phi_2^*(w_j^*)^T, \Phi_3(\hat{u}_{j-1}), \ldots, \Phi_3(\hat{u}_{j-r})) =$$

$$= (\hat{z}_{1,j}^*, \ldots, \hat{z}_{k^*,j}^*),$$

where now k^* becomes $k^* = k + 2r + r.q.$

7. PROOF OF THEOREM 2

For proving theorem 2 we need the following lemma.

LEMMA 3. If the assumptions 1 through 4 are satisfied then

(i) $$\operatorname{plim} \hat{\theta} = \begin{cases} \theta_0 & \text{under } H_0, \\ \theta_* & \text{under } H_1. \end{cases}$$

If in addition assumption 5 is satisfied and if H_0 is true then

(ii) $$\operatorname{plim} \left\| \sqrt{n}\,(\hat{\theta} - \theta_0) - \frac{1}{\sqrt{n}} \Sigma_{j=1}^{n} u_j A_0^{-1} (\partial/\partial\theta^T) f(x_j, \theta_0) \right\| = 0,$$

(iii) $\sqrt{n}\,(\hat{\theta} - \theta_0) \to N_m(0, \sigma^2 A_0^{-1})$ in distr.,

where $\sigma^2 = E\varepsilon_j^2$ and $A_0 = A(\theta_0)$.

Proof:

Part (i). Put

(7.1) $$\hat{Q}(\theta) = \frac{1}{n} \Sigma_{j=1}^{n} \{y_j - f(x_j, \theta)\}^2 =$$
$$= \frac{1}{n} \Sigma_{j=1}^{n} \{\varepsilon_j + \psi(u_{j-1}) + g(x_j) - f(x_j, \theta)\}^2$$

(7.2) $$Q(\theta) = \sigma^2 + \int_{\mathbb{R}^k \times \mathbb{R}} \{\psi(u_*) + g(x) - f(x,\theta)\}^2 dH^{(2)}(x, u_*)$$

where $H^{(2)}$ is defined by (4.10). From the assumptions 1A (the parts (3.3), (3.5) and (3.6)), 1D, 2 and 3 it follows that the random function $\{\varepsilon_j + \psi(u_{j-1}) + g(x_j) - f(x_j,\theta)\}^2$ satisfies conditions of the type (2.12) and (2.13), and is continuous as a function of $(\varepsilon_j, u_{j-1}, x_j)$ and θ. Moreover, $\{(\varepsilon_j, u_{j-1}, x_j)\}$ is a stochastically stable process with respect to a nonstationary ϕ-mixing base. Thus lemma 1 is applicable:

$$(7.3)\quad \operatorname*{plim}_{n\to\infty} \sup_{\theta \in \Theta} |\hat{Q}(\theta) - Q(\theta)| = 0$$

Furthermore, from the assumption 4A and 4B it follows that $Q(\theta)$ takes a unique infinum at $\theta = \theta_0$ if H_0 is true and at $\theta = \theta_*$ if H_1 is true. Therefore part (i) of lemma 3 follows from [1, Lemma 3.1.8].

Part (ii). This part of Lemma 3 follows from the argument in Section 3.1.3 in [1] by replacing all the references to Theorem 2.3.3 by references to Lemma 1.

Part (iii) is proved in [2, Theorem 3]. Q.E.D.

Lemma 3 shows that the results (3.5), (3.6), (3.8) and (3.9) in [3] carry over to the present case.

Now let

$$(7.4)\quad z_j^T = (\Phi_1(x_j)^T, \Phi_2(w_j)^T, \Phi(u_{j-1})) = (z_{1,j}, \ldots, z_{k^*,j}).$$

$$(7.5)\quad \hat{\eta} = \int_{N_0} \left|\frac{1}{n}\Sigma_{j=1}^n \hat{u}_j e^{it^T z_j}\right| dt$$

and

$$(7.6)\quad \hat{\mu} = 2^{k^*}\hat{\sigma}^2\{\pi_{\ell=1}^{k^*} \varepsilon_\ell - tr(\hat{A}^{-1}\hat{B}),$$

where

$$(7.7)\quad \hat{B} = \frac{1}{n^2} \Sigma_{j_1=1}^{n} \Sigma_{j_2=1}^{n} \{(\partial/\partial\theta^T) f(x_{j_1},\hat{\theta})\}\{(\partial/\partial\theta) f(x_{j_2},\hat{\theta})\} \cdot$$

$$\cdot \prod_{\ell=1}^{k*} \frac{\sin[\varepsilon_\ell (z_{\ell,j_1} - z_{\ell,j_2})]}{z_{\ell,j_1} - z_{\ell,j_2}}$$

Thus $\hat{\eta}$ and $\hat{\mu}$ are formed from $\hat{\hat{\eta}}$ and $\hat{\hat{\mu}}$ respectively, by replacing $\hat{z}_j$ by z_j. But $\hat{\eta}$ and $\hat{\mu}$ are now the same as in [3], provided we replace k in the formulas (2.14) and (2.17) in [3] by k*. It is not hard to verify that the argument in the proofs of the Theorems 2 and 3 in [3] carry over if we replace all the references to Theorem 2.3.4 of Bierens [1] by references to lemma 1. Thus we conclude:

LEMMA 4. If the assumptions 1 through 5 are satisfied and if H_0 is true then for every $\alpha \in (0,1)$, $\limsup_{n\to\infty} P[n\hat{\eta} > \frac{1}{\alpha}\hat{\mu}] \leq \alpha$.

Consequently part (6.9) of theorem 2 follows from:

LEMMA 5. If the conditions of theorem 2 are satisfied and if H_0 is true then

(i) $\operatorname*{plim}_{n\to\infty} |n\hat{\hat{\eta}} - n\hat{\eta}| = 0,$

(ii) $\operatorname*{plim}_{n\to\infty} |\hat{\hat{\mu}} - \hat{\mu}| = 0.$

<u>Proof of part (i)</u>

Put

(7.8) $$\tilde{\xi}(t_1,t_2,t_3,\theta_1,\theta_2) =$$

$$= \frac{1}{n}\Sigma_{j=1}^{n}\{y_j - f(x_j,\theta_1)\}e^{it_1^T\Phi_1(x_j)+it_2^T\Phi_2(w_j)+it_3\Phi_2(y_{j-1}-f(x_{j-1},\theta_2))}$$

Then

(7.9) $$\hat{\hat{\eta}} = \int_{N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} |\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\hat{\theta})|^2 dt_1 dt_2 dt_3 \ ,$$

(7.10) $$\hat{\eta} = \int_{N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} |\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\theta_0)|^2 dt_1 dt_2 dt_3.$$

So part (i) of lemma 5 is proved by showing that

(7.11) $$\operatorname*{plim}_{n\to\infty} \sup_{(t_1,t_2,t_3)\in N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} | \sqrt{n}\ \tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\hat{\theta}) -$$

$$- \sqrt{n}\ \tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\theta_0)| = 0.$$

Take the real part of $\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\hat{\theta})$ and apply the mean value theorem in the following way:

(7.12) $$\mathrm{Re}\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\hat{\theta}) = \mathrm{Re}\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\theta_0) + $$
$$+ (\hat{\theta} - \theta_0)^T(\partial/\partial\theta_2^T)\mathrm{Re}\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\tilde{\theta}(t_1,t_2,t_3))$$

where $\tilde{\theta}(t_1,t_2,t_3)$ is a mean value satisfying

(7.13) $$\|\tilde{\theta}(t_1,t_2,t_3) - \theta_0\| \leq \|\hat{\theta} - \theta_0\| \text{ uniformly in } (t_1,t_2,t_3).$$

We show that

(7.14) $$\operatorname*{plim}_{n\to\infty} \sup_{(t_1,t_2,t_3)\in N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} \|(\partial/\partial\theta^T)\mathrm{Re}\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\tilde{\theta}(t_1,t_2,t_3))\| = 0,$$

so that then by (7.12) and lemma 3 (iii),

(7.15) $$\operatorname*{plim}_{n\to\infty} \sup_{(t_1,t_2,t_3)\in N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} |\sqrt{n}\ \mathrm{Re}\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\hat{\theta}) - $$
$$- \sqrt{n}\ \mathrm{Re}\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\theta_0)| = 0.$$

Observe that under H_0

(7.16) $$(\partial/\partial\theta_2^T)\mathrm{Re}\tilde{\xi}(t_1,t_2,t_3,\theta_1,\theta_2) =$$
$$= (\partial/\partial\theta_2^T)\ \frac{1}{n}\ \Sigma_{j=1}^n(y_j-f(x_j,\theta_1))\cos[t_1^T\Phi_1(x_j)+t_2^T\Phi_2(w_j)+t_3\Phi_3(y_{j-1}-f(x_{j-1},\theta_2))] =$$
$$= \frac{1}{n}\ \Sigma_{j=1}^n[u_j + f(x_j,\theta_0) - f(x_j,\theta_1)]\ \cdot$$
$$\cdot\ \sin[t_1^T\Phi_1(x_j)+t_2^T\Phi_2(w_j)+t_3\Phi_3(u_{j-1}+f(x_{j-1},\theta_0)-f(x_{j-1},\theta_2))]\ \cdot$$
$$\cdot\ t_3\Phi_3'(u_{j-1}+f(x_{j-1},\theta_0)-f(x_{j-1},\theta_2))(\partial/\partial\theta_2^T)f(x_{j-1},\theta_2) =$$

$$= \frac{1}{n} \Sigma_{j=1}^{n} \gamma(x_j, w_j, u_j, x_{j-1}, u_{j-1}, t_1, t_2, t_3, \theta_1, \theta_2),$$

say, where the vector function γ is continuous in all its arguments. Moreover, $\{(x_j, w_j, u_j, x_{j-1}, u_{j-1})\}$ is a stochastically stable process with respect to the nonstationary ϕ-mixing base $\{(\varepsilon_j, \varepsilon_{j-1}, x_j^*, x_{j-1}^*)\}$ and putting

$$(7.17) \quad H_j^*(x,w,u,x_*,u_*) = P(x_j \leq x, w_j \leq w, u_j \leq u, x_{j-1} \leq x_*, u_{j-1} \leq u_*)$$

it follows from assumption 1D that

$$(7.18) \quad \frac{1}{n} \Sigma_{j=1}^{n} H_j^* \to H^* \text{ properly.}$$

Also a condition of the form (2.13) is satisfied, so that Lemma 1 is applicable to (7.16). Thus we have

$$(7.19) \quad \operatorname*{plim}_{n\to\infty} \sup_{(t_1,t_2,t_3,\theta_1,\theta_2) \in N_0^{(1)} \times N_0^{(2)} \times N_0^{(3)} \times \Theta \times \Theta} \|(\partial/\partial\theta^T)\mathrm{Re}\tilde{\xi}(t_1,t_2,t_3,\theta_1,\theta_2) - \Gamma_*(t_1,t_2,t_3,\theta_1,\theta_2)\| = 0,$$

where

$$(7.20) \quad \Gamma_*(t_1,t_2,t_3,\theta_1,\theta_2) = \int_{\mathbb{R}^k \times \mathbb{R}^{q+1} \times \mathbb{R} \times \mathbb{R}^k \times \mathbb{R}} \gamma(x,w,u,x_*,u_*,t_1,t_2,t_3,\theta_1,\theta_2) \cdot dH^*(x,w,u,x_*,u_*) =$$

$$= \lim_{n\to\infty} E(\partial/\partial\theta^T)\mathrm{Re}\tilde{\xi}(t_1,t_2,t_3,\theta_1,\theta_2) \cdot$$

It is not hard to verify that

$$(7.21) \quad \Gamma_*(t_1,t_2,t_3,\theta_0,\theta_0) = 0 \text{ for every } (t_1,t_2,t_3)$$

and moreover that (7.13), (7.20), (7.21) and the consistency of $\hat{\theta}$ imply that (7.14) holds and hence that (7.15) holds.

Along the same lines we can prove

$$(7.22)\quad \text{plim} \sup_{(t_1,t_2,t_3)\in N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} |\sqrt{n}\ \text{Im}\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\hat{\theta}) - \sqrt{n}\ \text{Im}\tilde{\xi}(t_1,t_2,t_3,\hat{\theta},\theta_0)| = 0.$$

Combining (7.15) and (7.22) yields (7.11), and so part (i) of lemma 5 is proved by now.

Proof of part (ii) of Lemma 5

It suffices to prove

$$(7.23)\quad \text{plim}_{n\to\infty} (\hat{\hat{B}} - \hat{B}) = 0,$$

for by applying Lemma 1 it can be shown that

$$(7.24)\quad \text{plim}_{n\to\infty} \hat{A} = A(\theta_0)$$

and

$$(7.25)\quad \text{plim}_{n\to\infty} \hat{\sigma}^2 = \sigma^2.$$

Now put

$$(7.26)\quad \hat{\hat{b}}(t) = \frac{1}{n}\Sigma_{j=1}^{n}(\partial/\partial\theta^T)f(x_j,\hat{\theta})e^{it^T\hat{z}_j}$$

$$(7.27)\quad \hat{b}(t) = \frac{1}{n}\Sigma_{j=1}^{n}(\partial/\partial\theta^T)f(x_j,\hat{\theta})e^{it^T z_j}$$

We then see that

$$(7.28)\quad \hat{\hat{B}} = \int_{N_0} \hat{\hat{b}}(t)\overline{\hat{\hat{b}}(t)}^T dt;\ \hat{B} = \int_{N_0} \hat{b}(t)\overline{\hat{b}(t)}^T dt,$$

where the bar means the complex conjugate, hence for proving (7.23) it suffices to prove

$$(7.29)\quad \operatorname*{plim}_{n\to\infty} \sup_{t\in N_0} \|\hat{\hat{b}}(t) - \hat{b}(t)\| = 0.$$

Observe from (6.2), (7.4), (7.26) and (7.27) that

$$(7.30)\quad \sup_{t\in N_0} \|\hat{\hat{b}}(t) - \hat{b}(t)\| \le \frac{1}{n}\Sigma_{j=1}^{n}\|(\partial/\partial\theta^T)f(x_j,\hat{\theta})\|\cdot\|\hat{z}_j - z_j\|\cdot\sup_{t\in N_0}\|t\| \le$$

$$\le \sup_{t\in N_0}\|t\|\cdot\frac{1}{n}\Sigma_{j=1}^{n}\|(\partial/\partial\theta^T)f(x_j,\hat{\theta})\|\cdot|\Phi_3(u_{j-1}+g(x_j)-f(x_j,\hat{\theta}))-\Phi_3(u_{j-1})| =$$

$$= \sup_{t\in N_0}\|t\|\cdot\frac{1}{n}\Sigma_{j=1}^{n}\psi(x_j,\hat{\theta}),$$

say. Now by applying lemma 1 to $\frac{1}{n}\Sigma_{j=1}^{n}\psi(x_j,\theta)$ and realizing that under H_0 $\psi(x_j,\theta_0) = 0$, we see that $\operatorname*{plim}_{n\to\infty}\frac{1}{n}\Sigma_{j=1}^{n}\psi(x_j,\hat{\theta}) = 0$, hence (7.29) holds. This proves part (ii) of lemma 5. Q.E.D.

We have by now completed the proof of part (6.9) of Theorem 2.

Proof of part (6.10) of Theorem 2

We can write

$$(7.31)\quad \hat{\hat{\eta}} = \int_{N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} \left|\frac{1}{n}\Sigma_{j=1}^{n}\tilde{\xi}_j(t_1,t_2,t_3,\hat{\theta})\right|^2 dt_1dt_2dt_3$$

where $\hat{\xi}_j$ is defined by (5.2). Applying lemma 1 we see that

$$(7.32)\quad \operatorname*{plim}_{n\to\infty} \sup_{(t_1,t_2,t_3,\theta)\in N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}\times\Theta} \left|\frac{1}{n}\Sigma_{j=1}^{n}\tilde{\xi}_j(t_1,t_2,t_3,\theta) - \xi(t_1,t_2,t_3,\theta)\right| = 0$$

where $\xi(t_1,t_2,t_3,\theta)$ is defined by (5.8). Since by lemma 3, plim $\hat{\theta} = \theta_*$, (7.32) implies:

$$(7.33)\quad \underset{n\to\infty}{\text{plim}} \sup_{(t_1,t_2,t_3)\in N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} \left|\frac{1}{n}\Sigma_{j=1}^{n}\hat{\xi}_j(t_1,t_2,t_3,\hat{\theta}) - \xi(t_1,t_2,t_3,\theta_*)\right| = 0$$

and in its turn (7.33) implies

$$(7.34)\quad \underset{n\to\infty}{\text{plim}}\ \hat{\hat{\eta}} = \int_{N_0^{(1)}\times N_0^{(2)}\times N_0^{(3)}} |\xi(t_1,t_2,t_3,\theta_*)|^2 dt_1 dt_2 dt_3 ,$$

which by Corollary 1 is positive. Consequently we have under H_1

$$(7.35)\quad \underset{n\to\infty}{\text{plim}}\ n\hat{\hat{\eta}} = \infty.$$

We leave it to the reader to verify that also under H_1 the probability limit of $\hat{\hat{\mu}}$ is finite. This completes the proof of part (6.10) of Theorem 2. 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] BIERENS, H.J.: "A Test for Model Specification in the Absence of Alternative Hypotheses", Discussion Paper No. 81-154, Centre for Economic Research, University of Minnesota (1981).

[4] BILLINGSLEY, P.: Convergence of Probability Measures, New York-London: John Wiley (1968).

[5] FELLER, W.: An Introduction to Probability Theory and Its Applications, Vol. II, New York-London: John Wiley (1966).