

VU Research Portal

A consistent Hausman-type model specification test

Bierens, H.J.

1987

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Bierens, H. J. (1987). *A consistent Hausman-type model specification test*. (Serie Research Memoranda; No. 1987-2). Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

05348

1987

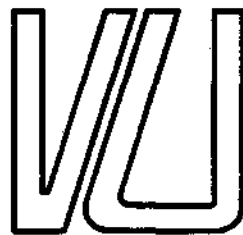
SERIE RESEARCH MEMORANDA

A CONSISTENT HAUSMAN-TYPE MODEL
SPECIFICATION TEST

Herman J. Bierens

Research Memorandum 1987-2

Jan. '87



**VRIJE UNIVERSITEIT
FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
A M S T E R D A M**



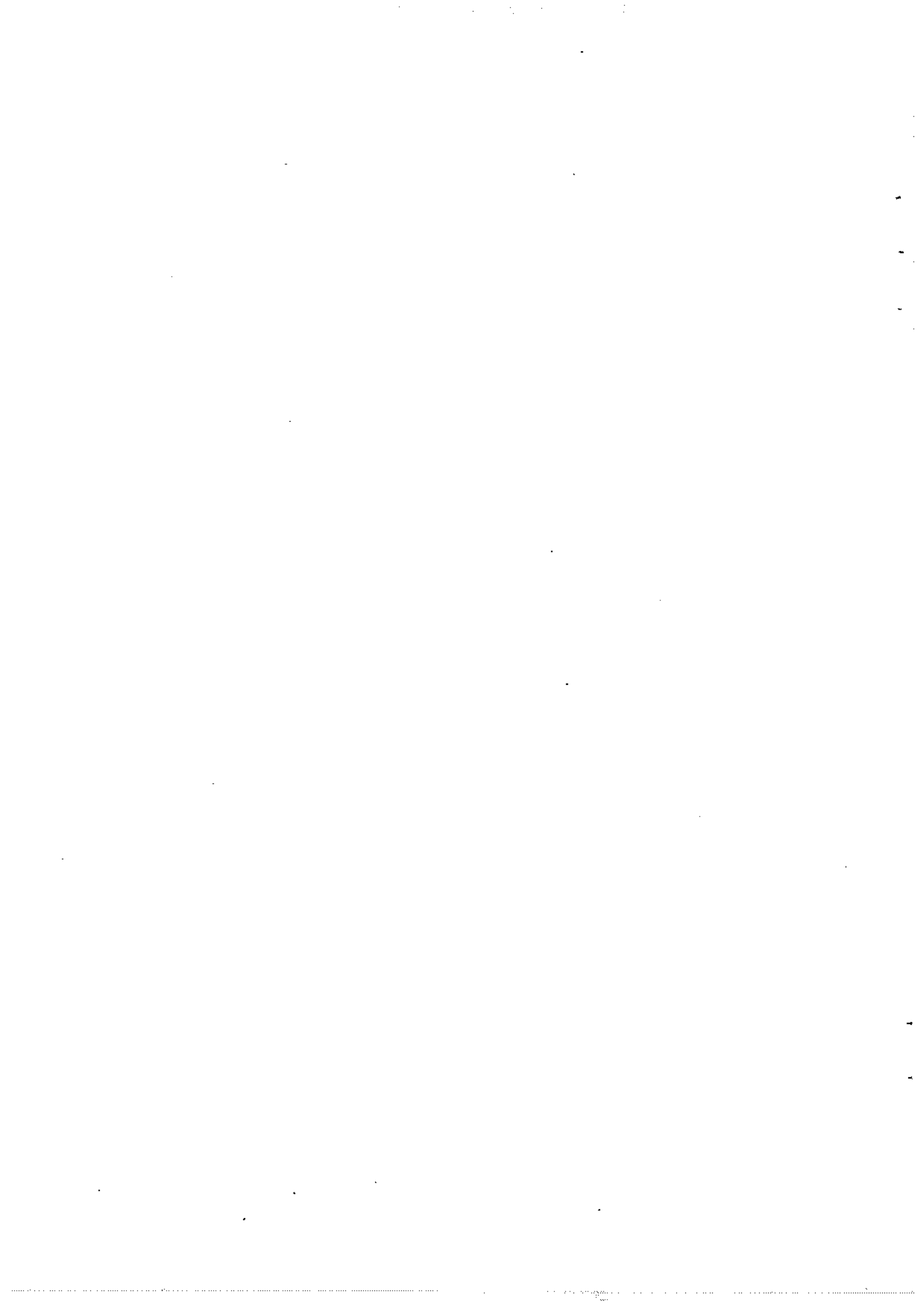
A CONSISTENT HAUSMAN-TYPE MODEL SPECIFICATION TEST

By Herman J. BIERENS

Free University
Faculty of Economics & Econometrics
Department of Econometrics
P.O.Box 7161
1007 MC Amsterdam

Abstract: In this paper we propose a modification of White's (1981) version of Hausman's (1978) model specification test such that the test has asymptotic power 1 against any deviation from the null hypothesis that the model is correctly specified.

January 1987



1. INTRODUCTION

In an influential paper in *Econometrica*, Hausman (1978) proposed a test for model misspecification. This test is based on the idea that for a true model the difference between an efficient estimator and an inefficient estimator, times the square root of the sample size, will converge in distribution to the normal with zero mean, whereas under the alternative that the model is misspecified it is likely that the two estimators will deviate in the limit. White's (1981) version of this test compares the ordinary nonlinear least squares estimator with the weighted nonlinear least squares estimator. See also Ruud (1984) for a review of Hausman-type model specification tests.

The power of the Hausman test depends heavily on the choice of the inefficient estimator or, in White's case, on the weight function. Bierens (1982, p.128) illustrates the latter by the following example. Let the true model be

$$y_j = x_{1j} + x_{2j} + x_{1j}x_{2j} + u_j,$$

where the x_{1j} 's and the u_j 's are independently $N(0,1)$ distributed. If this model is specified as

$$y_j = \theta_1 x_{1j} + \theta_2 x_{2j} + u_j$$

and if White's version of Hausman's test is conducted with weight function

$$w(x_{1j}, x_{2j}) = x_{1j}^2 + x_{2j}^2,$$

then the test statistic involved converges in distribution to $(13/2)\chi^2(2)$. The asymptotic power of the test is therefore less than 1. Although this example is rather artificial and does not bear much relevance to econometrics, it shows at least that the test is not watertight.

The problem of lack of consistency is not typical for Hausman's test, but is common to virtually all model specification tests. Tests which are designed to test a null against a specific alternative may have low power

against other alternatives. Although Hausman's test is not designed to test against a specific alternative there exists in fact an implicit alternative, depending on the choice of the inefficient estimator, against which it has maximal power. Again the test may have low power against deviations from this implicit alternative. See Holly (1982).

To the best of our knowledge the only consistent model specification tests are those of Bierens (1982, 1984, 1986a,b). These tests test the null hypothesis that the response function of the model equals the conditional expectation of the dependent variable relative to the regressors (and all lagged dependent variables and lagged regressors in the time series case). The alternative against which these tests have asymptotic power 1 is the general alternative that the null is false. Thus, any deviation from the null will be (asymptotically) detected. The tests of Bierens (1982, 1984), however, have some serious drawbacks. First, the distribution of the test statistic under the null is of an unknown form, so that the critical level had to be determined by using Chebishev's inequality. Second, these tests are quite laborious. The tests of Bierens (1986a,b) are consistent standard normal tests, but they rely on the assumption that the data generating process is rational-valued or is truncated rational-valued. Although in practice economic data are always rational-valued, we shall not employ this property here.

In the present paper it will be shown that White's version of Hausman's test can be converted into a consistent test, i.e., we shall propose a class of weight functions such that the test has asymptotic power 1 against all deviations from the null.

In Section 2 we outline the test and the ideas behind it. In Section 3 we set forth (mild) maintained hypotheses for the consistency of the test. Basic to our approach is a generalisation of Theorems 1 and 2 of Bierens (1982), which is stated as a Lemma in Section 3 and proved in the Appendix. In Section 4 we briefly discuss the asymptotic distribution of the test statistic.

Throughout the paper we assume an i.i.d. data generating process. Using the approach in Bierens (1981, Ch.3, 1984, 1986a,b), however, it seems feasible to extend the results to the heterogeneous and/or time series case. These extensions are left for future research.

2. THE TEST

Let $\{(y_1, x_1), \dots, (y_n, x_n)\}$ be a sample from a $k+1$ -variate distribution, where $y_j \in \mathbb{R}$ (and thus $x_j \in \mathbb{R}^k$). We consider a nonlinear regression model:

$$y_j = f(x_j, \theta_0) + u_j, \quad \theta_0 \in \Theta,$$

where Θ is a compact and convex subset of \mathbb{R}^m and $f(x, \theta)$ is for each $\theta \in \Theta$ a Borel measurable real function on \mathbb{R}^k and for each k -vector x a continuously differentiable real function on Θ .

In order that this model is correctly specified the errors u_j should satisfy

$$E(u_j | x_j) = 0 \text{ a.s.}$$

This is the null hypothesis to be tested, or equivalently

$$(1) H_0: P\{E(y_j | x_j) = f(x_j, \theta_0)\} = 1 \text{ for some } \theta_0 \in \Theta.$$

The alternative hypothesis is that the null is false, i.e.,

$$(2) H_1: P\{E(y_j | x_j) = f(x_j, \theta)\} < 1 \text{ for all } \theta \in \Theta.$$

In this paper we propose a consistent Hausman-type test of H_0 against H_1 , by comparing the nonlinear least squares estimator $\hat{\theta}$ for θ_0 with a weighted nonlinear least squares estimator $\hat{\theta}(t)$, where $t \in \mathbb{R}^k$ is a vector of nuisance parameters of the weight function. Thus we consider White's version of Hausman's test, but with a weight function guaranteeing consistency. Moreover our null is slightly more general in that H_0 does not imply that the u_j 's are independent of x_j or that $E(u_j^2 | x_j) = E(u_j^2)$ a.s.

The estimator $\hat{\theta}(t)$ is obtained by minimizing

$$(3) \hat{Q}(\theta, t) = (1/n) \sum_{j=1}^n [y_j - f(x_j, \theta)]^2 \exp(t' \Phi(x_j))$$

over the parameter space Θ , where Φ is a bounded Borel measurable one-to-one mapping from \mathbb{R}^k into \mathbb{R}^k . For example, we may choose

$$\Phi(z) = \Phi(z_1, \dots, z_k) = (\text{atan}(z_1), \dots, \text{atan}(z_k))'; \quad z = (z_1, \dots, z_k)' \in \mathbb{R}^k.$$

Thus $\hat{\theta}(t)$ is for given non-random $t \in \mathbb{R}^k$ a measurable solution of

$$\hat{\theta}(t) \in \Theta: \hat{Q}(\hat{\theta}(t), t) = \inf_{\theta \in \Theta} \hat{Q}(\theta, t).$$

Obviously, the ordinary nonlinear least squares estimator $\hat{\theta}$ equals $\hat{\theta}(0)$.
Now let $\theta(t)$ be the probability limit of $\hat{\theta}(t)$:

$$\theta(t) = \text{plim}_{n \rightarrow \infty} \hat{\theta}(t),$$

and let S be the following subset of \mathbb{R}^k :

$$(4) \quad S = \{ t \in \mathbb{R}^k: \theta(t) = \theta(0) \}.$$

We shall set forth mild conditions such that under H_1 this set S has *Lebesgue measure zero*, or more generally that S is a *null set* with respect to any probability measure μ on \mathbb{R}^k induced by an absolutely continuous k -variate distribution:

$$\mu(S) = 0.$$

On the other hand, if H_0 is true then clearly

$$\mu(S) = 1,$$

and moreover we then have

$$\sqrt{n} (\hat{\theta}(t) - \hat{\theta}(0)) \rightarrow N_m[0, \Omega(t)] \text{ in distr.}$$

for every $t \in \mathbb{R}^k$. Furthermore, we shall propose an estimator $\hat{\Omega}(t)$ such that

$$\text{plim}_{n \rightarrow \infty} \hat{\Omega}(t) = \Omega(t)$$

under H_0 as well as under H_1 , although in the latter case $\Omega(t)$ is no longer the variance matrix of the asymptotic distribution of $\hat{\theta}(t)$. Now define

$$\hat{W}(t) = n(\hat{\theta}(t) - \hat{\theta}(0))' \hat{\Omega}(t)^{-1} (\hat{\theta}(t) - \hat{\theta}(0))$$

and assume that $\Omega(t)$ is nonsingular for $t \neq 0$. Then

$$\hat{W}(t) \rightarrow \chi_m^2 \quad \text{in distr. if } H_0 \text{ is true and } t \neq 0,$$

whereas

$$\text{plim}_{n \rightarrow \infty} \hat{W}(t)/n = (\theta(t) - \theta(0))' \Omega(t)^{-1} (\theta(t) - \theta(0)) > 0$$

if H_1 is true and $t \notin S$.

The latter result implies, of course, that under H_1 ,

$$\text{plim}_{n \rightarrow \infty} \hat{W}(t) = \infty \text{ if } t \in S.$$

This result establishes the 'almost' consistency of the test under review, i.e., consistency fails only for t in a set S with Lebesgue measure zero.

In practice the set S is unknown, and although S is 'almost' empty we will never be sure whether a given fixed t is in S or not (except $t = 0$). But even if we would be sure that $t \notin S$ so that the asymptotic power of the test equals 1, the small sample power of the test may vary substantially with t . Thus, we might think of maximizing $\hat{W}(t)$ over some subset T of R^k , and using the resulting \hat{t} instead of t . However, $\hat{W}(\hat{t})$ will then no longer be asymptotically χ_m^2 distributed under H_0 . So optimizing the small sample power of the test seems not feasible. The only way to guarantee that t lies outside S is to draw t randomly from an absolutely continuous k -variate distribution, for example the k -variate standard normal distribution, as then $P\{t \in S\} = 0$. Denoting such a random drawing by \hat{t} we then have, similarly to the results in Bierens (1986a,b),

$$\hat{W}(\hat{t}) \rightarrow \chi_m^2 \text{ in distr. under } H_0$$

and

$$\text{plim}_{n \rightarrow \infty} \hat{W}(\hat{t}) = \infty \text{ under } H_1.$$

Thus by replacing the fixed t with this random \hat{t} we get a χ^2 test which has asymptotic power 1 against any deviation from the null hypothesis.

3. CONSISTENCY OF THE TEST

Throughout this section we assume that the alternative hypothesis (3) is true and thus that the model is misspecified. Moreover, we shall state a number of other assumptions which are to hold under the null as well. These assumptions are parts of the maintained hypothesis and are quite mild. The first one guarantees that $\hat{Q}(\theta, t)$ converges uniformly on θ to its mathematical expectation.

ASSUMPTION 1. Let $E y_1^2 < \infty$ and $E \sup_{\theta \in \Theta} f(x_1, \theta)^2 < \infty$.

Then it follows from Theorem 2 of Jennrich (1969) that for each $t \in R^k$,

$$\sup_{\theta \in \Theta} |\hat{Q}(\theta, t) - Q(\theta, t)| \rightarrow 0 \text{ a.s.,}$$

where

$$Q(\theta, t) = E \hat{Q}(\theta, t) = E \{y_1 - f(x_1, \theta)\}^2 \exp(t' \Phi(x_1)).$$

Clearly this limit function $Q(\theta, t)$ is continuous in both arguments. Since Θ is compact there exists for each t a $\theta(t)$ such that

$$\theta(t) \in \Theta : Q(\theta(t), t) = \inf_{\theta \in \Theta} Q(\theta, t).$$

Now we assume:

ASSUMPTION 2. For each $t \in \mathbb{R}^k$, $\theta(t)$ is unique.

Then it follows from Lemma 3.1.8. of Bierens (1981) that for each $t \in \mathbb{R}^k$

$$\hat{\theta}(t) \rightarrow \theta(t) \text{ a.s.}$$

Moreover, it can even be shown that the a.s. convergence involved is uniform on any compact subset of \mathbb{R}^k .

Next, assume:

ASSUMPTION 3. For each $t \in \mathbb{R}^k$ let $\theta(t)$ be an interior point of Θ .

Then the first order condition for a minimum of $Q(\theta, t)$ applies, i.e.

$$(\partial/\partial\theta')Q(\theta, t) \Big|_{\theta=\theta(t)} = 0,$$

hence

$$(5) \quad E[y_1 - f(x_1, \theta(t))] (\partial/\partial\theta') f(x_1, \theta(t)) \exp(t' \Phi(x_1)) = 0.$$

Now let S^* be the set

$$(6) \quad S^* = \{t \in \mathbb{R}^k : E[y_1 - f(x_1, \theta(0))] (\partial/\partial\theta') f(x_1, \theta(0)) \exp(t' \Phi(x_1)) = 0 \}.$$

Suppose that for some $t \notin S^*$, $\theta(t) = \theta(0)$. Then (5) contradicts with $t \notin S^*$, hence $S \subset S^*$, where S is defined in (4). Thus for proving $\mu(S) = 0$ it suffices to show $\mu(S^*) = 0$, and for that we need the following fundamental lemma:

LEMMA. Let v be a random variable satisfying $E|v| < \infty$ and let x be a bounded random vector in \mathbb{R}^k such that

$$P\{E(v|x) = 0\} < 1.$$

Consider the set

$$S = \{ t \in \mathbb{R}^k : E v \cdot \exp(t'x) = 0 \}.$$

For every probability measure μ induced by an absolutely continuous distribution on \mathbb{R}^k we have:

$$\mu(S) = 0.$$

Proof: Appendix.

Since the model is misspecified, we have

$$P\{E[y_1 - f(x_1, \theta(0)) | x_1] = 0\} < 1,$$

hence,

$$P\{E[(y_1 - f(x_1, \theta(0))) (\partial/\partial \theta') f(x_1, \theta(0)) | x_1] = 0\} < 1.$$

Moreover, since Φ is a one-to-one mapping conditioning on x_1 is equivalent to conditioning on the bounded random vector $\Phi(x_1)$, hence

$$(7) \quad P\{E[(y_1 - f(x_1, \theta(0))) (\partial/\partial \theta') f(x_1, \theta(0)) | \Phi(x_1)] = 0\} < 1.$$

It follows now from (7) and the Lemma that $\mu(S^*)=0$ and thus:

THEOREM 1. Under Assumptions 1-3 and H_1 the set $S = \{ t \in \mathbb{R}^k : \theta(t) \neq \theta(0) \}$ is a null set with respect to any probability measure μ induced by an absolutely continuous k -variate distribution.

4. FURTHER ASYMPTOTIC THEORY

The asymptotic theory of the test statistic $\hat{W}(t)$ under H_0 is quite standard (see White 1981) and will therefore be discussed only briefly. Throughout we assume that Assumptions 1-3 hold. Then under H_0

$$\theta(t) = \theta(0) = \theta_0 \text{ for all } t \in \mathbb{R}^k$$

and consequently,

$$\hat{\theta}(t) \rightarrow \theta_0 \text{ a.s., for all } t \in \mathbb{R}^k.$$

Now denote

$$\hat{\Omega}_1(\theta, t) = (1/n) \sum_{j=1}^n \{(\partial/\partial\theta')f(x_j, \theta)\} \{(\partial/\partial\theta)f(x_j, \theta)\} \exp(t'\Phi(x_j)),$$

$$\hat{\Omega}_2(\theta, t) = (1/n) \sum_{j=1}^n (y_j - f(x_j, \theta))^2 \{(\partial/\partial\theta')f(x_j, \theta)\} \{(\partial/\partial\theta)f(x_j, \theta)\} \\ \times \exp(t'\Phi(x_j)),$$

$$\hat{\Gamma}(\theta, t) = (1/n) \sum_{j=1}^n (y_j - f(x_j, \theta)) (\partial/\partial\theta) (\partial/\partial\theta') f(x_j, \theta) \exp(t'\Phi(x_j))$$

and let the following assumption hold.

ASSUMPTION 4. For $i_1, i_2 = 1, \dots, m$

$$E \sup_{\theta \in \Theta} | \{(\partial/\partial\theta_{i_1})f(x_1, \theta)\} \{(\partial/\partial\theta_{i_2})f(x_1, \theta)\} | < \infty,$$

$$E \sup_{\theta \in \Theta} | (y_1 - f(x_1, \theta))^2 \{(\partial/\partial\theta_{i_1})f(x_1, \theta)\} \{(\partial/\partial\theta_{i_2})f(x_1, \theta)\} | < \infty,$$

$$E \sup_{\theta \in \Theta} | (y_1 - f(x_1, \theta)) (\partial/\partial\theta_{i_1}) (\partial/\partial\theta_{i_2}) f(x_1, \theta) | < \infty.$$

Then it follows from Jennrich (1969, Theorem 2) that for each $t \in \mathbb{R}^k$,

$$\hat{\Omega}_1(\theta, t) \rightarrow E \hat{\Omega}_1(\theta, t) \text{ a.s., uniformly on } \Theta,$$

$$\hat{\Omega}_2(\theta, t) \rightarrow E \hat{\Omega}_2(\theta, t) \text{ a.s., uniformly on } \Theta,$$

$$\hat{\Gamma}(\theta, t) \rightarrow E \hat{\Gamma}(\theta, t) \text{ a.s., uniformly on } \Theta.$$

Thus:

$$\begin{aligned} (\partial/\partial\theta)(\partial/\partial\theta')\hat{Q}(\theta, t) &= 2\hat{\Omega}_1(\theta, t) - 2\hat{\Gamma}(\theta, t) \\ &\rightarrow E (\partial/\partial\theta)(\partial/\partial\theta')\hat{Q}(\theta, t) = 2 E \hat{\Omega}_1(\theta, t) - 2 E \hat{\Gamma}(\theta, t) \text{ a.s.} \end{aligned}$$

uniformly on Θ . Moreover, by the central limit theorem we have

$$(8) \quad \sqrt{n} \begin{bmatrix} (\partial/\partial\theta')\hat{Q}(\theta_0, t) \\ (\partial/\partial\theta')\hat{Q}(\theta_0, 0) \end{bmatrix} \rightarrow N_{2m} \left[0, \begin{bmatrix} E\hat{\Omega}_2(\theta_0, 2t) & E\hat{\Omega}_2(\theta_0, t) \\ E\hat{\Omega}_2(\theta_0, t) & E\hat{\Omega}_2(\theta_0, 0) \end{bmatrix} \right]$$

Consequently, if

ASSUMPTION 5. $E\hat{\Omega}_1(\theta_0, t)$ is nonsingular, for all $t \in R^k$,

then by the standard Taylor expansion approach and the fact that

$$E \hat{\Gamma}(\theta_0, t) = 0$$

it follows

$$\sqrt{n} \begin{bmatrix} \hat{\theta}(t) - \theta_0 \\ \hat{\theta}(0) - \theta_0 \end{bmatrix} \rightarrow N_{2m} \left[0, \begin{bmatrix} \Omega_1(t)^{-1} \Omega_2(2t) \Omega_1(t)^{-1}, \Omega_1(t)^{-1} \Omega_2(t) \Omega_1(0)^{-1} \\ \Omega_1(0)^{-1} \Omega_2(t) \Omega_1(t)^{-1}, \Omega_1(0)^{-1} \Omega_2(0) \Omega_1(0)^{-1} \end{bmatrix} \right],$$

where

$$(9) \quad \Omega_1(t) = E\hat{\Omega}_1(\theta(0), t) \quad , \quad \Omega_2(t) = E\hat{\Omega}_2(\theta(0), t)$$

It is now easy to show:

THEOREM 2. Under Assumptions 1-5 and H_0 we have

$$\sqrt{n}(\hat{\theta}(t) - \hat{\theta}(0)) \rightarrow N_m[0, \Omega(t)] \text{ in distr.}$$

where

$$\begin{aligned} \Omega(t) = & \Omega_1(t)^{-1} \Omega_2(2t) \Omega_1(t)^{-1} - \Omega_1(t)^{-1} \Omega_2(t) \Omega_1(0)^{-1} \\ & - \Omega_1(0)^{-1} \Omega_2(t) \Omega_1(t)^{-1} + \Omega_1(0)^{-1} \Omega_2(0) \Omega_1(0)^{-1} \end{aligned}$$

with $\Omega_1(t)$ and $\Omega_2(t)$ defined in (9).

Moreover, denoting

$$\begin{aligned} \hat{\Omega}(t) = & \hat{\Omega}_1(\hat{\theta}(0), t)^{-1} \hat{\Omega}_2(\hat{\theta}(0), 2t) \hat{\Omega}_1(\hat{\theta}(0), t)^{-1} \\ & - \hat{\Omega}_1(\hat{\theta}(0), t)^{-1} \hat{\Omega}_2(\hat{\theta}(0), t) \hat{\Omega}_1(\hat{\theta}(0), 0)^{-1} \\ & - \hat{\Omega}_1(\hat{\theta}(0), 0)^{-1} \hat{\Omega}_2(\hat{\theta}(0), t) \hat{\Omega}_1(\hat{\theta}(0), t)^{-1} \\ & + \hat{\Omega}_1(\hat{\theta}(0), 0)^{-1} \hat{\Omega}_2(\hat{\theta}(0), 0) \hat{\Omega}_1(\hat{\theta}(0), 0)^{-1} \end{aligned}$$

we have:

THEOREM 3. Under Assumptions 1-5, $\hat{\Omega}(t) \rightarrow \Omega(t)$ a.s., for each $t \in R^k$.

Note that this result holds under H_0 as well as under H_1 .

Next, assume:

ASSUMPTION 6. $\Omega(t)$ is nonsingular for $t \neq 0$.

A sufficient condition for this is that the asymptotic variance matrix in (8) is nonsingular for each $t \neq 0$ and that Assumption 5 holds. This assumption is of a high level and therefore not quite satisfactory. On the other hand it is not implausible an assumption and moreover it is basically the same assumption as the one in White (1981, Theorem 4.1).

It follows now from Theorems 1, 2 and 3:

THEOREM 4. Under Assumptions 1-6 we have

$$(10) \hat{W}(t) \rightarrow \chi_m^2 \text{ in distr. if } H_0 \text{ is true and } t \neq 0,$$

whereas

$$(11) \text{plim}_{n \rightarrow \infty} \hat{W}(t)/n = (\theta(t) - \theta(0))' \Omega(t)^{-1} (\theta(t) - \theta(0)) > 0$$

if H_1 is true and $t \notin S$.

Finally, let \hat{t} be a random drawing from an absolutely continuous k -variate distribution with density $h(t)$. Then under H_0 the characteristic function of $\hat{W}(\hat{t})$ will converge to the characteristic function of χ_m^2 :

$$\begin{aligned} E \exp(i\zeta \hat{W}(\hat{t})) &= E(E\{\exp(i\zeta \hat{W}(\hat{t})) | \hat{t}\}) = \int E \exp(i\zeta \hat{W}(t)) h(t) dt \\ &\rightarrow E \exp(i\zeta \chi_m^2) \end{aligned}$$

by bounded convergence and (10). Moreover, denoting the right hand side of (11) by $\eta(t)$ we have under H_1 ,

$$\hat{W}(\hat{t})/n \rightarrow \eta(\hat{t}) \text{ in distr. and } P[\eta(\hat{t}) > 0] = P[\hat{t} \notin S] = 1,$$

hence

$$\text{plim}_{n \rightarrow \infty} \hat{W}(\hat{t}) = \infty$$

Thus:

THEOREM 5. Let \hat{t} be a random drawing from an absolutely continuous k -variate distribution. Moreover, let Assumptions 1-6 hold. Then

$$\hat{W}(\hat{t}) \rightarrow \chi_m^2 \text{ in distr. under } H_0,$$

and

$$\hat{W}(\hat{t}) \rightarrow \infty \text{ in prob. under } H_1.$$

APPENDIX: Proof of the Lemma

First, let $k=1$. According to Theorem 2 of Bierens (1982) there exists a nonnegative integer m such that $\text{Ev}.x^m \neq 0$, hence

$$\begin{aligned} (d/dt)^m \text{Ev}.e^{tx} &= \sum_{j=0}^{\infty} (d/dt)^m \text{Ev}(tx)^j / j! = \sum_{j=m}^{\infty} t^{j-m} \text{Ev}.x^j / (j-m)! \\ &\rightarrow \text{Ev}.x^m \neq 0 \text{ as } t \rightarrow 0. \end{aligned}$$

This implies that $\text{Ev}.\exp(tx) \neq 0$ in a neighborhood of zero. Now let t_0 be such that $\text{Ev}.\exp(t_0x) = 0$. Since

$$P(E[v.\exp(t_0x) | x] = 0) = P(E(v|x) = 0) < 1$$

it follows from the above argument, with v replaced by $v.\exp(t_0x)$, that $\text{Ev}.\exp(t_0x+tx) \neq 0$ in a neighborhood of $t = 0$, hence $\text{Ev}.\exp(tx) \neq 0$ in a neighborhood of $t = t_0$. This implies

$$(A1) \inf_{t \in S, t \neq t_0} |t - t_0| > 0 \text{ if } t_0 \in S$$

and in its turn (A1) implies that S is *countable*. Since a countable set has zero probability with respect to any continuous distribution, the lemma follows for the case $k = 1$.

Next, consider the general case $k \geq 1$. Let

$$x = (x_1, x_2, \dots, x_k)'; \quad t = (t_1, t_2, \dots, t_k)'; \quad \hat{t} = (\hat{t}_1, \hat{t}_2, \dots, \hat{t}_k)',$$

where \hat{t} is a random drawing from μ . Again it follows from Theorem 2 of Bierens (1982) that there exist nonnegative integers m_1, \dots, m_k such that

$$E v.x_1^{m_1} x_2^{m_2} \dots x_k^{m_k} \neq 0$$

and similarly to the case $k = 1$ this implies that there exists a t_* close to the origin of R^k such that

$$E v_* \exp(t_*' x) \neq 0.$$

Now let $v_* = v_* \exp(t_*' x)$. Since $E v_* \neq 0$ we have

$$P(E[v_* | x_1, \dots, x_\ell] = 0) < 1 \text{ for } \ell = 1, \dots, k.$$

Let for $\ell = 1, \dots, k$

$$S_\ell^* = \{(t_1, \dots, t_\ell)' \in R^\ell : E v_* \exp(t_1 x_1 + \dots + t_\ell x_\ell) = 0\}.$$

We show now that if the lemma holds for $k = \ell$ it holds too for $k = \ell + 1$, i.e.

$$P((\hat{t}_1, \dots, \hat{t}_{\ell+1})' \in S_{\ell+1}^*) = 0.$$

By induction, we then have

$$(A2) \quad P((\hat{t}_1, \dots, \hat{t}_k)' \in S_k^*) = 0,$$

and replacing \hat{t} by $\hat{t} - t_*$ we see that then also

$$P((\hat{t}_1, \dots, \hat{t}_k)' \in S) = 0,$$

which is the desired result.

Thus, assume

$$P((\hat{t}_1, \dots, \hat{t}_\ell)' \in S_\ell^*) = 0.$$

Then

$$\begin{aligned}
 & P\{(\hat{t}_1, \dots, \hat{t}_{\ell+1})' \in S_{\ell+1}^*\} \\
 &= P\{(\hat{t}_1, \dots, \hat{t}_{\ell+1})' \in S_{\ell+1}^* \text{ and } (\hat{t}_1, \dots, \hat{t}_\ell)' \in S_\ell^*\} \\
 &+ P\{(\hat{t}_1, \dots, \hat{t}_{\ell+1})' \in S_{\ell+1}^* \text{ and } (\hat{t}_1, \dots, \hat{t}_\ell)' \notin S_\ell^*\} \\
 &= P\{(\hat{t}_1, \dots, \hat{t}_{\ell+1})' \in S_{\ell+1}^* \text{ and } (\hat{t}_1, \dots, \hat{t}_\ell)' \notin S_\ell^*\} \\
 &= P\{\hat{t}_{\ell+1} \in S_{\ell+1}^{**}(\hat{t}_1, \dots, \hat{t}_\ell) \text{ and } (\hat{t}_1, \dots, \hat{t}_\ell)' \notin S_\ell^*\},
 \end{aligned}$$

where

$$S_{\ell+1}^{**}(\hat{t}_1, \dots, \hat{t}_\ell) = \{t \in \mathbb{R} : E[v \cdot \exp(t_1 x_1 + \dots + t_\ell x_\ell) | \exp(t \cdot x_{\ell+1}) = 0]\}.$$

Since by the argument for the case $k = 1$ it follows that

$$S_{\ell+1}^{**}(\hat{t}_1, \dots, \hat{t}_\ell) \text{ is countable if } (\hat{t}_1, \dots, \hat{t}_\ell)' \notin S_\ell^*,$$

and since by the absolute continuity of the distribution of $(\hat{t}_1, \dots, \hat{t}_k)'$ the conditional distribution of $\hat{t}_{\ell+1}$ relative to $\hat{t}_1, \dots, \hat{t}_\ell$ is continuous, the desired result (A2) easily follows. Q.E.D.

Remark: Observe that the result of the lemma under review implies that the set S has Lebesgue measure zero.

REFERENCES

BIERENS, H.J. (1981): *Robust Methods and Asymptotic Theory in Nonlinear Econometrics*. New York: Springer

BIERENS, H.J. (1982): "Consistent Model Specification Tests", *Journal of Econometrics*, 20, 105-134

BIERENS, H.J. (1984): "Model Specification Testing of Time Series Regressions", *Journal of Econometrics*, 26, 323-353

BIERENS, H.J. (1986a): "ARMA Memory Index Modeling of Economic Time Series", *Econometric Theory* (forthcoming)

BIERENS, H.J. (1986b): "ARMAX Model Specification Testing, with an Application to Unemployment in the Netherlands", *Annals of Econometrics* (forthcoming)

HAUSMAN, J.A. (1978): "Specification Testing in Econometrics", *Econometrica*, 46, 1251-1271

HOLLY, A. (1982): "A Remark on Hausman's Specification Test", *Econometrica*, 50, 749-760

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

RUUD, P.A. (1984): "Tests of Specification in Econometrics", *Econometric Reviews*, 3, 211-242

WHITE, H. (1981): "Consequences and Detection of Misspecified Nonlinear Regression Models", *Journal of the American Statistical Association*, 76, 419-433

1981-1	E. Vogelving	A quarterly econometric model for the Price Formation of Coffee on the World Market	1981-18	F.C. Palm and J.M. Sneek	Some econometric Applications of the exact Distribution of the ratio of Two Quadratic Forms in Normal Variates
1981-2	H.P. Smit	Demand and Supply of Natural Rubber, Part I	1981-19	P. Nijkamp and P. Rietveld	Soft Econometrics as a Tool for Regional Discrepancy Analysis
1981-3	R. Vos	The political Economy of the Republic of Korea; A proposal for a model framework of an open economy in the ESCAP-region, with emphasis on the Role of the State	1981-20	H. Blommestein and P. Nijkamp	Soft Spatial Econometric Causality Models
1981-4	F.C. Palm	Structural Econometric Modeling and Time Series Analysis - Towards an Integrated Approach	1981-21	P. Nijkamp and P. Rietveld	Ordinal Econometrics in Regional and Urban Modeling
1981-5	P. Nijkamp in co-op. with H. v. Handenboven and R. Janssen	Urban Impact Analysis in a Spatial Context: Methodologie and Case Study	1981-22	F. Brouwer and P. Nijkamp	Categorical Spatial Data Analysis
1981-6	R. Ruben	Primaire exporten en economische ontwikkeling	1981-23	A. Kleinknecht	Prosperity, Crises and Innovation Patterns: Some more Observations on neo-Schumpeterian Hypotheses
1981-7	D.A. Kodde	Het genereren en evalueren van voorspellingen van omzet en netto winst: een toegepast kwantitatieve benadering	1981-24	Hidde P. Smit	World Tire and Rubber Demand
1981-8	B. Out	Financiële vraagstukken onder onzekerheid	1982-1	Peter Nijkamp	Long waves or catastrophes in Regional Development
1981-9	P. van Dijk and H. Verbruggen	A Constant-Market-Shares Analysis of ASEAN Manufactured Exports to the European Community	1982-2	J.M. Sneek	Some Approximations to the Exact Distribution of Sample Autocorrelations for Autoregressive-moving Average Models
1981-10	P. Nijkamp, H. de Graaff and E. Siga	A Multidimensional Analysis of Regional Infrastructure and Economic Development	1982-3	F.E. Schippers	Empirisme en empirische toetsing in de wetenschapsfilosofie en in de economische wetenschap
1981-11	P. Nijkamp	International Conflict Analysis	1982-4	Piet van Helsdingen maart 1982	Mantelprojekt 'Management & Politiek': Productiebeleid en overheid; een onderzoek naar de invloed van de overheid op het produktbeleid in de verpakkingsmiddelenindustrie
1981-12	A.J. Mathot	L'Utilisation 'du Cr�dit lors de l'Achat d'une Voiture	1982-5	Peter Nijkamp Jaap Spronk	Integrated Policy Analysis by means of Interactive Learning Models
1981-13	S.P. van Duin en P.A. Cornelis	Onderzoek naar levensomstandigheden en opvattingen over arbeid bij mensen zonder werk, deel I	1982-6	Ruerd Ruben (ed.)	The Transition Strategy of Nicaragua
1981-14	W. van Lierop and P. Nijkamp	Disaggregate Models of Choice in a Spatial Context	1982-7	H.W.M. Jansen mei 1982	Een alternatieve modellering van het gedrag van een besluitvormer: 'satisficing' i.p.v. 'maximizing'
1981-15	Hidde P. Smit	The World Vehicle Market	1982-8	J. Klaassen and H. Schreuder mei	Confidential revenue and Profit Forecasts by management and financial analysts: some first results
1981-16	F.C. Palm	Structural Econometric Modeling and Time Series Analysis: An Integrated Approach	1982-9	F. Brouwer and P. Nijkamp	Multiple Rank Correlation Analysis
1981-17	F.C. Palm and Th.E. Nijman	Linear Regression Using Both Temporally Aggregated and Temporally Disaggregated Data			