The Integrated Conditional Moment Test

Herman J. Bierens

Abstract

In this lecture I will review the integrated conditional moment (ICM) test for functional form of a conditional expectation model. This is a consistent test: the ICM test has asymptotic power 1 against all deviations from the null hypothesis. Moreover, this test has non-trivial $\sqrt{n}$ local power.

I will focus on the mathematical foundations of the ICM approach, in particular the consistency proof and the derivation of the asymptotic null distribution.

1 The Fourier transform of a Borel measurable function

Let $g(x)$ be a Borel measurable real function on $\mathbb{R}^k$. The Fourier transform of $g(x)$ relative to a probability measure $\mu_X(.)$ on the Borel sets in $\mathbb{R}^k$ is defined by

$$\varphi(\xi) = \int \exp(i.\xi' x)\, g(x) d\mu_X(x),\ i = \sqrt{-1},$$

provided that $\int |g(x)| d\mu_X(x) < \infty$.

LEMMA 1: *Let* $g_1(x)$ *and* $g_2(x)$ *be Borel measurable functions on* $\mathbb{R}^k$ *satisfying* $\int |g_1(x)| d\mu_X(x) < \infty$, $\int |g_2(x)| d\mu_X(x) < \infty$, *with Fourier transforms* $\varphi_1(\xi)$, $\varphi_2(\xi)$, *respectively, relative to a probability measure* $\mu(.)$ *on the Borel sets in* $\mathbb{R}^k$. *Then* $g_1(x) = g_2(x)$ *a.s.* $\mu_X(.)$, *i.e.,*
$B_0 = \left\{x \in \mathbb{R}^k : g_1(x) - g_2(x) = 0\right\} \Rightarrow$
$\mu_X(B_0) = 1$, *if and only if* $\varphi_1(\xi) \equiv \varphi_2(\xi)$.

Proof: Suppose $\varphi_1(\xi) \equiv \varphi_2(\xi)$ and $\mu_X(B_0) < 1$. Let

$$\begin{aligned} r_1(x) &= \max(0, g_1(x) - g_2(x)), \\ r_2(x) &= \max(0, -g_1(x) + g_2(x)) \end{aligned}$$

Then $g_1(x) - g_2(x) = r_1(x) - r_2(x)$ and

$$\int \exp(i.\xi' x)\, r_1(x) d\mu_X(x)$$
$$- \int \exp(i.\xi' x)\, r_2(x) d\mu_X(x)$$
$$= \varphi_1(\xi) - \varphi_2(\xi) \equiv 0$$

Substituting $\xi = 0$ yields

$$\int r_1(x) d\mu_X(x) = \int r_2(x) d\mu_X(x) = c \geq 0.$$

If $c = 0$ then $r_1(x) = r_2(x) = 0$ a.s. $\mu(.)$, hence $g_1(x) = g_2(x)$ a.s. $\mu(.)$.

Therefore, assume that $c > 0$. Then we can define the probability measures

$$\upsilon_m(B) = \frac{1}{c} \int_B r_m(x) d\mu_X(x),\ m = 1, 2,$$

with corresponding characteristic functions

$$\begin{aligned}\psi_m(\xi) &= \int \exp(i.\xi' y)\, d\upsilon_m(y) \\ &= \frac{1}{c}\int \exp(i.\xi' x)\, r_m(x) d\mu_X(x)\end{aligned}$$

for $m = 1, 2$. But I have just established that $\int \exp(i.\xi' x)\, r_1(x) d\mu_X(x) \equiv \int \exp(i.\xi' x)\, r_2(x) d\mu_X(x)$, hence $\psi_1(\xi) \equiv \psi_2(\xi)$, which by the uniqueness of characteristic functions implies that $\upsilon_1(B) = \upsilon_2(B)$ for all Borel sets $B \subset \mathbb{R}^k$. It is now an easy (ECON 501) exercise to verify that the latter implies $r_1(x) = r_2(x)$ a.s. $\mu_X(.)$, hence $g_1(x) = g_2(x)$ a.s. $\mu_X(.)$.

Corollary:

LEMMA 2: *Let U be a random variable satisfying $E[|U|] < \infty$, and let $X \in \mathbb{R}^k$ be a random vector. If $P[E(U|X) = 0] < 1$ then there exists a $\xi \in \mathbb{R}^k$ such that $E[U \exp(i.\xi' X)] \neq 0$.*

Question: Where to look for such a ξ ?

LEMMA 3: *If* X *is bounded then under the conditions of Lemma 2, for each* $\varepsilon > 0$ *there exists a* ξ *satisfying* $\|\xi\| < \varepsilon$ *such that* $E\left[U\exp\left(i.\xi'X\right)\right] \neq 0$.

Proof: Let $X \in \mathbb{R}$. Then

$$\begin{aligned} E\left[U\exp\left(i.\xi X\right)\right] &= E\left[U\sum_{m=0}^{\infty}\frac{i^m\xi^m X^m}{m!}\right] \\ &= \sum_{m=0}^{\infty}\frac{i^m\xi^m}{m!}E\left[U.X^m\right] \end{aligned}$$

Since $E\left[U\exp\left(i.\xi X\right)\right] \neq 0$ for some ξ we must have that $E\left[U.X^m\right] \neq 0$ for some integer $m \geq 0$. Let m_0 be the smallest m for which $E\left[U.X^m\right] \neq 0$. Then

$$\left.\frac{d^{m_0}E\left[U\exp\left(i.\xi X\right)\right]}{\left(d\xi\right)^{m_0}}\right|_{\xi=0} = i^{m_0}E\left[U.X^{m_0}\right] \neq 0$$

which implies that $E\left[U\exp\left(i.\xi X\right)\right] \neq 0$ for $\xi \neq 0$ arbitrarily close to zero.

LEMMA 4: *Under the conditions of Lemma 3, the set* $S_0 = \left\{\xi \in \mathbb{R}^k : E\left[U.\exp\left(i.\xi' X\right)\right] = 0\right\}$ *has Lebesgue measure zero and is nowhere dense.*

Proof: Let $k = 1$ and $\xi_0 \in S_0$. Define $U_0 = U\exp\left(i.\xi_0 X\right)$. Then $P\left(E[U_0|X] = 0\right) < 1$, hence for an arbitrarily small $\varepsilon > 0$ there exists a $\xi \in (-\varepsilon, 0) \cup (0, \varepsilon)$ such that

$$E\left[U\exp\left(i.\xi_0 X\right)\exp\left(i.\xi X\right)\right] \neq 0.$$

By continuity it follows now that for each $\xi_0 \in S_0$ there exists an $\varepsilon > 0$ such that

$\xi \notin S_0$ for all $\xi \in (\xi_0 - \varepsilon, \xi_0) \cup (\xi_0, \xi_0 + \varepsilon)$.

Consequently, in the case $k = 1$, the set S_0 is countable and is nowhere dense. In the general case $k \geq 1$, S_0 has Lebesgue measure zero and is nowhere dense.

More generally, we have:

LEMMA 5: *Let $w(u)$ be a real or complex valued function of the type*

$$w(u) = \sum_{s=0}^{\infty} (\gamma_s/s!)\, u^s$$

where $|\gamma_s| < \infty$ and at most a finite number of γ_s's are zero. Then under the conditions of Lemma 3, the set

$$S_0 = \left\{\xi \in \mathbb{R}^k : E\left[U.w\left(\xi' X\right)\right] = 0\right\}$$

has Lebesgue measure zero and is nowhere dense.

For example, let $w(u) = \cos(u) + \sin(u)$, or $w(u) = \exp(u)$.

The condition that the random vector X is bounded can be get rid of by replacing X with $\Phi(X)$, where Φ is a Borel measurable bounded one-to-one mapping, because the σ-algebra generated by X is then the same as the σ-algebra generated by $\Phi(X)$, hence conditioning on $\Phi(X)$ is equivalent to conditioning on X.

THEOREM 1: *Let U be a random variable satisfying $E[|U|] < \infty$ and let $X \in \mathbb{R}^k$ be a random vector. Denote*

$$S = \left\{\xi \in \mathbb{R}^k : E\left[U.w\left(\xi'\Phi(X)\right)\right] = 0\right\},$$

where $w(.)$ is defined in Lemma 5, and $\Phi(.)$ is a Borel measurable bounded one-to-one mapping. If $P[E(U|X) = 0] < 1$ then S has Lebesgue measure zero and is nowhere dense, whereas if $P[E(U|X) = 0] = 1$ then $S = \mathbb{R}^k$.

2 The ICM test

Given a random sample (Y_j, X_j), $j = 1, .., n$, $X_j \in \mathbb{R}^k$, and a conditional expectation model

$$E(Y_j|X_j) = g(X_j, \theta_0),\ \theta_0 \in \Theta,$$

where $\Theta \subset \mathbb{R}^m$ is the parameter space, Theorem 1 suggests to test the correctness of the functional specification of this model on the basis of following ICM statistic:

$$\int \left|\widehat{z}(\xi)\right|^2 d\mu\left(\xi\right)$$

In this expression, $\mu\left(\xi\right)$ is an absolutely continuous (w.r.t. Lebesgue measure) probability measure with compact support $\Xi \subset \mathbb{R}^k$, and

$$\widehat{z}(\xi) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \widehat{U}_j w\left(\xi' \Phi(X_j)\right).$$

where $\widehat{U}_j = Y_j - g(X_j, \widehat{\theta})$, with $\widehat{\theta}$ the NLLS estimator of θ_0, Φ is a bounded one-to-one mapping, and $w(.)$ is a weight function satisfying the conditions of Theorem 1.

More formally, the null hypothesis to be tested is that

H_0: There exists a $\theta_0 \in \Theta$ such that
$$P\left[E(y_j|x_j) = g(x_j, \theta_0)\right] = 1,$$
and the alternative hypothesis is that H_0 is false:

H_1: For all $\theta \in \Theta$,
$$P\left[E(y_j|x_j) = g(x_j, \theta)\right] < 1,$$

Under the null hypothesis and standard regularity conditions,

$$\sqrt{n}\left(\widehat{\theta}-\theta_0\right) = A^{-1}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left.\frac{\partial g(X_j,\theta)}{\partial\theta'}\right|_{\theta=\theta_0} U_j + o_p(1)$$

where

$$A = p\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\left(\frac{\partial g(X_j,\theta)}{\partial\theta'}\right)\left.\left(\frac{\partial g(X_j,\theta)}{\partial\theta'}\right)'\right|_{\theta=\theta_0}$$

Hence, by the uniform law of large numbers,

$$\begin{aligned}\widehat{z}(\xi) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\widehat{U}_j w\left(\xi'\Phi(X_j)\right)\\ &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}U_j w\left(\xi'\Phi(X_j)\right)\\ &-\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(g(X_j,\widehat{\theta})-g(X_j,\theta_0)\right)w\left(\xi'\Phi(X_j)\right)\\ &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}U_j\phi_j\left(\xi\right)+o_p(1),\end{aligned}$$

say, where

$$\begin{aligned}\phi_j(\xi) &= w(\xi'\Phi(X_j)) \\ &\quad -b(\xi)' A^{-1} \left.\frac{\partial g(X_j,\theta)}{\partial \theta'}\right|_{\theta=\theta_0}\end{aligned}$$

with

$$b(\xi) = p \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \left.\frac{\partial g(X_j,\theta)}{\partial \theta'}\right|_{\theta=\theta_0} w(\xi'\Phi(X_j)),$$

and $o_p(1)$ is uniform in $\xi \in \Xi$.

THEOREM 2: *Under the null hypothesis and some regularity conditions* (*one of these conditions is that* Ξ *is compact*),

$$\widehat{z}(\xi) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \widehat{U}_j w\left(\xi' \Phi(X_j)\right) \Rightarrow z(\xi) \text{ on } \Xi,$$

where $z(\xi)$ *is a zero-mean Gaussian process on* Ξ, *with covariance function*

$$\Gamma(\xi_1, \xi_2) = E\left[z(\xi_1) z(\xi_2)\right].$$

Hence by the continuous mapping theorem,

$$\int \left|\widehat{z}(\xi)\right|^2 d\mu\left(\xi\right) \to_d \int \left|z(\xi)\right|^2 d\mu\left(\xi\right).$$

Under the alternative that the null is false, $\widehat{z}(\xi)/\sqrt{n} \to_p \eta(\xi)$ *uniformly on* Ξ, *where* $\eta(\xi) \neq 0$ *except on a set with zero Lebesgue measure, so that*

$$(1/n) \int \left|\widehat{z}(\xi)\right|^2 d\mu\left(\xi\right) \to_p \int \left|\eta(\xi)\right|^2 d\mu\left(\xi\right) > 0,$$

provided that $\mu\left(\xi\right)$ *is absolutely continuous w.r.t. Lebesgue measure and its support* Ξ *has positive Lebesgue measure.*

3 The null distribution of the ICM test

If we choose the weight function w real-valued, for example $w(u) = \cos(u) + \sin(u)$, then $z(\xi)$ is a real-valued zero-mean Gaussian process on Ξ, with real-valued covariance function

$$\begin{aligned} \Gamma(\xi_1, \xi_2) &= E\left[z(\xi_1)z(\xi_2)\right] \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} E\left[U_j^2 \phi_j\left(\xi_1\right) \phi_j\left(\xi_2\right)\right]. \end{aligned}$$

This covariance function is symmetric and positive semidefinite, in the following sense:

$$\int\int \psi(\xi_1)\Gamma(\xi_1, \xi_2)\psi(\xi_2) d\mu\left(\xi_1\right) d\mu\left(\xi_2\right) \geq 0$$

for all Lebesgue integrable functions $\psi(\xi)$ on Ξ. Such functions have non-negative eigenvalues and corresponding orthonormal eigenfunctions:

THEOREM 3: *The functional eigenvalue problem* $\int \Gamma(\xi_1,\xi_2)\psi(\xi_2)d\mu\left(\xi_2\right) = \lambda.\psi(\xi_1)$ *a.e. on* Ξ *has a countable number of solutions*

$$\int \Gamma(\xi_1,\xi_2)\psi_j(\xi_2)d\mu\left(\xi_2\right) = \lambda_j.\psi_j(\xi_1) \text{a.e. on } \Xi,$$

$$j = 1, 2, ...$$

where

$$\lambda_j \geq 0, \ \sum_{j=1}^{\infty} \lambda_j < \infty,$$

$$\int_{\Xi} \psi_{j_1}(\xi)\psi_{j_2}(\xi)d\mu\left(\xi\right) \ \begin{cases} = 0 \ \ if \ \ j_1 \neq j_2 \\ = 1 \ \ if \ \ j_1 = j_2 \end{cases}$$

Moreover,

THEOREM 4 (Mercer's Theorem): *The covariance function* $\Gamma(\xi_1,\xi_2)$ *can be written as* $\Gamma(\xi_1,\xi_2) = \sum_{j=1}^{\infty} \lambda_j\psi_j(\xi_1)\psi_j(\xi_2)$.

The sequence $\{\psi_t(\xi)\}_{t=1}^{\infty}$ is an orthonormal basis for a Hilbert space $\mathcal{H}(\mu)$ of Lebesgue integrable functions on Ξ, with inner product

$$\langle f, g\rangle = \int f(\xi)\, g(\xi)\, d\mu(\xi).$$

so that every function f in $\mathcal{H}(\mu)$ can be written as

$$f(\xi) = \sum_{t=1}^{\infty} \gamma_t \psi_t(\xi), \quad \sum_{t=1}^{\infty} \gamma_t^2 < \infty, \text{ where}$$
$$\gamma_t = \langle f, \psi_t\rangle, \ t = 1, 2, 3, ...$$

It can be shown that $z(\xi)$ is a random element of $\mathcal{H}(\mu)$, hence

$$z(\xi) = \sum_{t=1}^{\infty} z_t \psi_t(\xi), \text{ where}$$
$$z_t = \int_{\Xi} z(\xi)\psi_t(\xi) d\mu(\xi), \ t = 1, 2, 3, ...$$

Consequently,

$$\int |z(\xi)|^2 \, d\mu(\xi) = \int \left(\sum_{t=1}^{\infty} z_t \psi_t(\xi) \right)^2 d\mu(\xi)$$

$$= \sum_{t_1=1}^{\infty} \sum_{t_2=1}^{\infty} z_{t_1} z_{t_2} \int \psi_{t_1}(\xi) \psi_{t_2}(\xi) d\mu(\xi)$$

$$= \sum_{t_1=1}^{\infty} \sum_{t_2=1}^{\infty} z_{t_1} z_{t_2} I(t_1 = t_2) = \sum_{t=1}^{\infty} z_t^2$$

The sequence z_t is a zero-mean Gaussian process, with variance function

$$E\left[z_t^2\right] = E\left[\left(\int z(\xi) \psi_t(\xi) d\mu(\xi) \right)^2 \right]$$

$$= \int \int E\left[z(\xi_1) z(\xi_2)\right] \psi_t(\xi_1) \psi_t(\xi_2) d\mu(\xi_1) \, d\mu(\xi_2)$$

$$= \int \int \left(\sum_{j=1}^{\infty} \lambda_j \psi_j(\xi_1) \psi_j(\xi_2) \right) \psi_t(\xi_1) \psi_t(\xi_2)$$

$$\times d\mu(\xi_1) \, d\mu(\xi_2)$$

where the latter equality follows from Mercer's theorem.

Thus by the orthonormality of the eigenfunctions,

$$\begin{aligned} E\left[z_t^2\right] &= \sum_{j=1}^{\infty} \lambda_j \left(\int \psi_j(\xi)\psi_t(\xi) d\mu\left(\xi\right) \right)^2 \\ &= \sum_{j=1}^{\infty} \lambda_j I\left(j = t\right) = \lambda_t. \end{aligned}$$

Moreover, by a similar argument it follows that

$$\begin{aligned} E\left[z_{t_1} z_{t_2}\right] &= \sum_{j=1}^{\infty} \lambda_j I\left(j = t_1\right) I\left(j = t_2\right) \\ &= 0 \text{ if } t_1 \neq t_2. \end{aligned}$$

Hence, denoting $\varepsilon_t = z_t/\sqrt{\lambda_t}$ if $\lambda_t > 0$, we have:

THEOREM 5: $\int \left|z(\xi)\right|^2 d\mu\left(\xi\right) = \sum_{t=1}^{\infty} \lambda_t \varepsilon_t^2$, *where the* ε_t*'s are i.i.d.* $N(0,1)$ *and the* λ_t*'s are the eigenvalues of the covariance function* Γ.

4 Critical values

The problem is that the eigenvalues λ_t are case-dependent: They depend on the distribution of the regressors, the functional form of the NLLS model, and the conditional variance of the errors. Therefore, the distribution of $\int |z(\xi)|^2 \, d\mu(\xi)$ cannot be tabulated. A possible way to get around this problem is to bootstrap this distribution. However, a convenient way to get around this problem is to derive upper bounds of the critical values, as follows.

Without loss of generality we may assume that the λ_t's are positive and arranged in decreasing order. Moreover, it follows from Mercer's theorem that

$$\begin{aligned} \int \Gamma(\xi, \xi) d\mu(\xi) &= \sum_{j=1}^{\infty} \lambda_j \int \psi_j(\xi)^2 d\mu(\xi) \\ &= \sum_{j=1}^{\infty} \lambda_j \end{aligned}$$

THEOREM 6: *Denoting* $p_t = \lambda_t / \sum_{j=1}^{\infty} \lambda_j$, *we have*

$$\frac{\int |z(\xi)|^2 \, d\mu(\xi)}{\int \Gamma(\xi,\xi) d\mu(\xi)} = \sum_{t=1}^{\infty} p_t \varepsilon_t^2$$

$$\leq \sup_{p_1 \geq p_2 \geq .....,\sum_{t=1}^{\infty} p_t = 1} \sum_{t=1}^{\infty} p_t \varepsilon_t^2$$

$$= \sup_{m \geq 1} \frac{1}{m} \sum_{i=1}^{m} \varepsilon_i^2 = \overline{T},$$

say,

so that asymptotic critical values can be derived from the latter distribution. The actual test statistic of the ICM test is therefore

$$\widehat{T}_{ICM} = \frac{\int |\widehat{z}(\xi)|^2 \, d\mu(\xi)}{\int \widehat{\Gamma}(\xi,\xi) d\mu(\xi)},$$

where $\widehat{\Gamma}(\xi_1, \xi_2)$ is a consistent estimator of $\Gamma(\xi_2, \xi_2)$, uniformly on $\Xi \times \Xi$.

5 Local power of the ICM test

Consider the local alternative hypothesis

$$H_1^L : E[Y_j|X_j] = g(X_j, \theta_0) + \frac{h(X_j)}{\sqrt{n}} \text{ a.s.,}$$

where $h(X_j)$ is not constant:

$$P[h(X_j) = E(h(X_j))] < 1.$$

Then under H_1^L,

$$\widehat{z}(\xi) \Rightarrow z(\xi) + \omega(\xi) \text{ on } \Xi,$$

where $z(\xi)$ is the same zero-mean Gaussian process on Ξ as before, and $\omega(\xi)$ is a deterministic mean function satisfying $0 < \int \omega(\xi)^2 d\mu(\xi) < \infty$.

Similar to the case under the null hypothesis, we can write

$$z(\xi) + \omega(\xi) = \sum_{t=1}^{\infty} z_t \psi_t(\xi)$$

where now

$$z_t = \varepsilon_t \sqrt{\lambda_t} + \omega_t$$

with ε_t i.i.d. $N(0,1)$ and $\omega_t = \int \omega(\xi)\psi_t(\xi)d\mu(\xi)$.

Hence,

$$\int \left|\widehat{z}(\xi)\right|^2 d\mu\left(\xi\right) \;\rightarrow_d \int \left|z(\xi)+\omega(\xi)\right|^2 d\mu\left(\xi\right)$$
$$= \sum_{t=1}^{\infty}\left(\varepsilon_t\sqrt{\lambda_t}+\omega_t\right)^2$$

THEOREM 7: *The ICM test has nontrivial $\sqrt{n}$-local power, in the sense that for every $K >$ 0,*

$$P\left[\sum_{t=1}^{\infty}\left(\varepsilon_t\sqrt{\lambda_t}+\omega_t\right)^2 \leq K\right]$$
$$< P\left[\sum_{t=1}^{\infty}\lambda_t\varepsilon_t^2 \leq K\right]$$

Proof:

Let

$$C_n = \sum_{t=1}^{\infty} \left(\varepsilon_t \sqrt{\lambda_t} + \omega_t\right)^2 - \left(\varepsilon_n \sqrt{\lambda_n} + \omega_n\right)^2$$

and suppose that $\omega_n \neq 0$. Then

$$\begin{aligned}
& P\left[\sum_{t=1}^{\infty} \left(\varepsilon_t \sqrt{\lambda_t} + \omega_t\right)^2 \leq K\right] \\
& = P\left[\left(\varepsilon_n \sqrt{\lambda_n} + \omega_n\right)^2 \leq K - C_n \; and \; C_n \leq K\right] \\
& = P\left[-\sqrt{K - C_n} \leq \varepsilon_n \sqrt{\lambda_n} + \omega_n \leq \sqrt{K - C_n} \right. \\
& \qquad \left. and \; C_n \leq K\right] \\
& < P\left[-\sqrt{K - C_n} \leq \varepsilon_n \sqrt{\lambda_n} \leq \sqrt{K - C_n} \right. \\
& \qquad \left. and \; C_n \leq K\right] \\
& = P\left[\varepsilon_n^2 \lambda_n + C_n \leq K \; and \; C_n \leq K\right] \\
& = P\left[\varepsilon_n^2 \lambda_n + C_n \leq K\right]
\end{aligned}$$

where the inequality is due to the symmetry and unimodality of the $N(0, \lambda_n)$ distribution. The result of Theorem 7 follows now by induction.

6 Bibliography

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. (1990): ”A Consistent Conditional Moment Test of Functional Form”, *Econometrica* 58, 1443-1458.

Bierens, H. J. and W. Ploberger (1997): ”Asymptotic Theory of Integrated Conditional Moment Tests”, *Econometrica* 65, 1129-1151.

De Jong, R. M. (1996): ”On the Bierens Test Under Data Dependence”, *Journal of Econometrics* 72, 1-32.

Stinchcombe, M. B., and H. White (1998): ”Consistent Specification Testing with Nuisance Parameters Present Only Under the Alternative”, *Econometric Theory* 14, 295-325.