Review of the Integrated Conditional Moment Test and Its Implementation in EasyReg International

Herman J. Bierens
Pennsylvania State University

April 21, 2006

1 The ICM test

The ICM test is based on the following theorem:

THEOREM 1: *Let u be a random variable satisfying $E|u| < \infty$, and $P[E(u|x) = 0] < 1$, where $x \in \mathbb{R}^k$ is a bounded random vector.*
(a) *Let $w(u)$ be a complex or real valued function that is infinitely many times differentiable in $u = 0$ and satisfies the condition that $(d/du)^s w(u)\mid_{u=0} \neq 0$ for all but a finite number of natural numbers s. Then for every $\varepsilon > 0$ there exists a $\xi \in \mathbb{R}^k$ such that $E[u.w(\xi' x)] \neq 0$ and $\|\xi\| < \varepsilon$.*
(b) *If in addition $w(u)$ is a power series in an open neighborhood of $u = 0$, i.e., for some $\delta > 0$, $w(u) = \sum_{s=0}^{\infty} (\gamma_s/s!)\, u^s$ for $|u| < \delta$, where $\gamma_s = (d/du)^s w(u)\mid_{u=0}$, then the set $\left\{\xi \in \mathbb{R}^k : E[u.w(\xi' x)] = 0\right\}$ has Lebesgue measure zero and is nowhere dense.*

Proof: See Bierens (1982) for part (a) with $w(u) = \exp(i.u)$, Bierens (1990) for the case $w(u) = \exp(u)$, and Bierens and Ploberger (1997) for the general case. Examples of suitable functions $w(u)$ in the general case are $w(u) = \cos(u) + \sin(u)$, and $w(u) = 1/[1 + \exp(c - u)]$ for $c \neq 0$. See also Stinchcombe and White (1998) for further elaborations on this theorem, and Bierens (1994, Ch. 3) for the details of the proof of Theorem 1 for the cases $w(u) = \exp(i.u)$ and $w(u) = \exp(u)$.

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. See Bierens (1982, 1990, 1994, Ch. 3).

Theorem 1 suggests that, 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 $\Theta \subset \mathbb{R}^k$ is the parameter space, the null hypothesis

$$H_0\text{: There exists a } \theta_0 \in \Theta \text{ such that } P\left[E(y_j|x_j) = g(x_j, \theta_0)\right] = 1,$$

can be consistently tested against the general alternative hypothesis that the null hypothesis is false, i.e.,

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

on the basis of the Integrated Conditional Moment (ICM) statistic

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

In this integral,

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

where $\widehat{u}_j$ is the nonlinear least squares residual: $\widehat{u}_j = y_j - g(x_j, \widehat{\theta})$,with $\widehat{\theta}$ the nonlinear least squares estimator of θ_0, Φ is a bounded one-to-one mapping, $w(.)$ is a weight function satisfying the conditions of Theorem 1, and μ is a probability measure on a compact set $\Xi \subset \mathbb{R}^k$ with positive Lebesgue measure, which is absolute continuous with respect to Lebesgue measure.

This ICM statistic was proposed first by Bierens (1982), for the case $w(u) = \exp(i.u)$, Ξ a hypercube in $\mathbb{R}^k$, μ the Lebesgue measure on Ξ, and i.i.d. observations (y_j, x_j).

2 The asymptotic null distribution of the ICM statistic

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

$$u_j=y_j-g(x_j,\theta_0)=y_j-E[y_j|x_j]$$

and

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

Hence,

$$\widehat{z}(\xi)=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\widehat{u}_j w\left(\xi'\Phi(x_j)\right)=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}u_j\phi_j\left(\xi\right)+o_p(1),$$

where

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

with

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

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

It has been shown by Bierens (1990) and Bierens and Ploberger (1997) that under some mild regularity conditions (among which the assumption that the function $w(.)$ is real-valued), and the null hypothesis involved,

$$\widetilde{z}\left(\xi\right)=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}u_j\phi_j\left(\xi\right)$$

converges weakly to a zero-mean Gaussian process $z\left(\xi\right)$ on Ξ, with covariance function

$$\Gamma(\xi_1,\xi_2)=E\left[z(\xi_1)z(\xi_2)\right]=\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}E\left[u_j^2\phi_j\left(\xi_1\right)\phi_j\left(\xi_2\right)\right]\qquad(2)$$

Consequently, under the null hypothesis

$$\int |\widehat{z}(\xi)|^2 \, d\mu(\xi) \to \int |z(\xi)|^2 \, d\mu(\xi)$$

in distribution, whereas under the general alternative that the null is false,

$$\widehat{z}(\xi)/\sqrt{n} \to \eta(\xi) = p \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \widehat{u}_j w\left(\xi' \Phi(x_j)\right)$$

in probability, uniformly on Ξ, where $\eta(\xi) \neq 0$ except on a set with zero Lebesgue measure. Thus, under the alternative hypothesis,

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

a.s.

Moreover, Bierens and Ploberger (1997) have shown that the asymptotic null distribution of the ICM statistic is of the type

$$\int |z(\xi)|^2 \, d\mu(\xi) = \sum_{i=1}^{\infty} \lambda_i \varepsilon_i^2,$$

where the ε_i's are i.i.d. $N(0,1)$ and the λ_i's are the eigenvalues of the covariance function Γ, and that

$$\int \sigma^2(\xi) d\mu(\xi) = \sum_{i=1}^{\infty} \lambda_i,$$

where

$$\sigma^2(\xi) = \Gamma(\xi, \xi),$$

is the variance function of $z(\xi)$. Furthermore, they have shown that

$$\frac{\int |z(\xi)|^2 \, d\mu(\xi)}{\int \sigma^2(\xi) d\mu(\xi)} = \frac{\sum_{i=1}^{\infty} \lambda_i \varepsilon_i^2}{\sum_{i=1}^{\infty} \lambda_i} \leq \sup_{m\geq 1} \frac{1}{m} \sum_{i=1}^{m} \varepsilon_i^2 = \overline{T},$$

say, so that asymptotic critical values can be derived from the latter distribution.

As to the choice of the probability measure μ, Boning and Sowell (1999) have shown that the uniform probability measure on Ξ is optimal. The actual test statistic of the ICM test is therefore

$$\widehat{T}_{ICM} = \frac{\int_\Xi |\widehat{z}(\xi)|^2 d\xi}{\int_\Xi \widehat{\sigma}^2(\xi) d\xi}, \tag{3}$$

where $\widehat{\sigma}^2(\xi)$ is a consistent estimator of the variance function $\sigma^2(\xi)$, uniformly on Ξ.

The asymptotic null distribution of $\widehat{T}_{ICM}$ is case-dependent, because the eigenvalues λ_i depend on the distribution of (y_t, x_t) and the conditional expectation model $g(x_t, \theta_0)$, but is dominated by the distribution of $\overline{T}$. Thus, denoting the $1-\alpha$ quantile of $\overline{T}$ by T_α, i.e., $P(\overline{T} \geq T_\alpha) = \alpha$, the null hypothesis is rejected at the $\alpha \times 100\%$ significance level if $\widehat{T}_{ICM} \geq T_\alpha$. The values of T_α for $\alpha = 0.10$ and $\alpha = 0.05$ are:

$$T_{0.10} = 3.23$$
$$T_{0.05} = 4.26$$

3 The ICM test of the martingale difference hypothesis

The Bierens-Ploberger version of the ICM test allows for consistently testing of linear and nonlinear ARX models, but not for ARMAX models, because, given a $k+1$-variate vector time series process

$$z_t = (y_t, x'_{t+1})' \in \mathbb{R} \times \mathbb{R}^{k-1}, \tag{4}$$

an ARMAX model represents the conditional expectation of the dependent variable y_t relative to *all* lagged z_t's. Bierens (1984) and De Jong (1996) have, in different ways, extended the ICM test to the case where z is infinite dimensional, i.e., $z = (z'_{t-1}, z'_{t-2}, ...)'$, in order to accommodate conditioning on the infinite past of a k-variate time series process z_t. In this paper I shall review the approach of De Jong (1996).

The space $(\Xi, \|\cdot\|)$ defined in De Jong (1996) is given as follows. For two infinite sequences of points in $\mathbb{R}^k \times \mathbb{R}^\infty$, ξ and ζ, given by $\xi = (\xi'_1, \xi'_2, ...)'$ and $\zeta = (\zeta'_1, \zeta'_2, ...)'$, where $\xi_j, \zeta_j \in \mathbb{R}^k$, define the norm

$$\|\xi - \zeta\| = \sqrt{\sum_{j=1}^{\infty} j^2 |\xi_j - \zeta_j|^2}, \tag{5}$$

where $|\xi_j - \zeta_j|$ is the Euclidean norm on $\mathbb{R}^k$. Next, define the space Ξ as

$$\Xi = \{\xi \in \mathbb{R}^{k+1} \times \mathbb{R}^{\infty} : a_j \leq \xi_j \leq b_j, \forall j \geq 1\}, \tag{6}$$

where $a_j < b_j$ and $|a_j|, |b_j| \leq cj^{-2}$ for some constant $c > 0$. Note that Ξ has finite Lebesgue measure. With this definition $(\Xi, \|\cdot\|)$ is a compact metric space, and therefore it is totally bounded.

Following Bierens (1990), De Jong now proposes to use the weight function

$$w_t(\xi) = \exp\left(\sum_{j=1}^{t-1} \xi_j' \Phi(z_{t-j})\right),$$

and the Lebesgue measure on Ξ as the measure μ. However, in view of Theorem 1, De Jong's results carry over to the more general case

$$w_t(\xi) = w\left(\sum_{j=1}^{t-1} \xi_j' \Phi(z_{t-j})\right), \tag{7}$$

where $w(\cdot)$ is a real-valued function satisfying the conditions of Theorem 1(b), and Φ is a bounded one-to-one mapping. If $w(\cdot)$ is real valued but only satisfies the conditions of Theorem 1(a), we have to choose $a_j < 0 < b_j, \forall j$. In this case the results of Theorem 1(b) read:

THEOREM 2: *Let u_t be a random variable satisfying $E|u_t| < \infty$, and let z_t be a k-variate time series process, such that (u_t, z_t) is stationary. Let $(\Xi, \|\cdot\|)$ be defined by (5) and (6), and let*

$$\overline{w}_t(\xi) = w\left(\sum_{j=1}^{\infty} \xi_j' \Phi(z_{t-j})\right), \tag{8}$$

where $w(\cdot)$ satisfies the conditions of Theorem 1. Then

$$P\left[E(u_t | z_{t-1}, z_{t-2}, ...) = 0\right] < 1$$

if and only if the set $\{\xi \in \Xi : \ E\left(u_t \overline{w}_t(\xi)\right) = 0\}$ has Lebesgue measure zero and is nowhere dense in Ξ.

Proof: De Jong (1997).

The actual test statistic is now similar to (3), and the upper bounds of the critical values still apply.

4 The ICM test in EasyReg International

4.1 Cross-section data

If your model is a (nonlinear) cross-section regression model

$$y_j = g(x_j, \theta_0) + u_j,\ j = 1, ..., n, \tag{9}$$

the default instrumental variables are the components of x_j. You may remove some of these components from the list of instrumental variables, or add other variables to the list. Once you have confirmed the list of instrumental variables, they will be standardized by taking them in deviation of their sample means, and then dividing them by their sample standard errors. Thus, each component $x_{i,j}$ of x_j is standardized as

$$\widetilde{x}_{i,j} = \left(x_{i,j} - \overline{x}_i\right)/s_i,$$

where

$$\overline{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{i,j},\ s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(x_{i,j} - \overline{x}_i\right)^2.$$

Next, replace x_t in (1) with $\widetilde{x}_t = (\widetilde{x}_{1,t}, ..., \widetilde{x}_{k,t})'$:

$$\widehat{z}(\xi) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\widehat{u}_t w\left(\xi'\Phi(\widetilde{x}_t)\right).$$

The reason will be given below, after discussing the choice of Φ.

The components of the bounded one-to-one mapping Φ in EasyReg are arctan(.). Thus

$$\Phi(\widetilde{x}_t) = \begin{pmatrix} \arctan(\widetilde{x}_{1,t}) \\ \vdots \\ \arctan(\widetilde{x}_{k,t}) \end{pmatrix}.$$

Now without standardization, $\arctan(x_{i,t}) \approx \pi/2$ if all the $x_{i,t}$ take large positive values, or $\arctan(x_{i,t}) \approx -\pi/2$ if all the $x_{i,t}$ take large negative values, which would destroy the consistency of the ICM test.

As to the function $w(.)$, EasyReg provides two options,

$$w(.) = \cos(.) + \sin(.), \tag{10}$$

which is the default option, and

$$w(.) = \exp(.), \tag{11}$$

which was used in Bierens (1990).

Finally, the only option for the set Ξ in (3) is

$$\Xi(c) = \times_{\ell=1}^{k} [-c, c], \tag{12}$$

where $c > 0$ has to be chosen.

4.2 Computation of the ICM test statistic

Note that

$$\begin{aligned}
& \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} E\left[u_j^2 \phi_j(\xi)^2\right] \\
= & \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} E\left[u_j^2 \left(w\left(\xi' \Phi(\widetilde{x}_j)\right) - b(\xi)' A^{-1} \left. \frac{\partial g(x_j,\theta)}{\partial \theta'} \right|_{\theta=\theta_0} \right)^2 \right] \\
= & \, p \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} u_j^2 w\left(\xi' \Phi(\widetilde{x}_j)\right)^2 \\
& -2b(\xi)' A^{-1} \left(p \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} u_j^2 \left. \frac{\partial g(x_j,\theta)}{\partial \theta'} \right|_{\theta=\theta_0} w\left(\xi' \Phi(\widetilde{x}_j)\right) \right) \\
& +b(\xi)' A^{-1} \left(p \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} u_j^2 \left. \frac{\partial g(x_j,\theta)}{\partial \theta'} \right|_{\theta=\theta_0} \left. \frac{\partial g(x_j,\theta)}{\partial \theta} \right|_{\theta=\theta_0} \right) A^{-1} b(\xi) \\
= & \, p \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} u_j^2 w\left(\xi' \Phi(\widetilde{x}_j)\right)^2 + b(\xi)' A^{-1} B A^{-1} b(\xi) - 2b(\xi)' A^{-1} c(\xi)
\end{aligned}$$

where

$$\begin{aligned}
c(\xi) &= p \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} u_j^2 \left. \frac{\partial g(x_j,\theta)}{\partial \theta'} \right|_{\theta=\theta_0} w\left(\xi' \Phi(x_j)\right) \\
B &= p \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} u_j^2 \left. \frac{\partial g(x_j,\theta)}{\partial \theta'} \right|_{\theta=\theta_0} \left. \frac{\partial g(x_j,\theta)}{\partial \theta} \right|_{\theta=\theta_0}
\end{aligned}$$

Thus,

$$\begin{aligned}\int_{\Xi}\sigma^2(\xi)d\xi &= \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}E\left[u_j^2\int_{\Xi}\phi_j\left(\xi\right)^2 d\xi\right] \\ &= p\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}u_j^2\int_{\Xi}w\left(\xi'\Phi(\widetilde{x}_j)\right)^2 d\xi \\ &\quad +\int_{\Xi}b\left(\xi\right)'A^{-1}BA^{-1}b\left(\xi\right)d\xi-2\int_{\Xi}b\left(\xi\right)'A^{-1}c\left(\xi\right)d\xi\end{aligned}$$

Denoting,

$$\widehat{X}_j=\left.\frac{\partial g(x_j,\theta)}{\partial\theta'}\right|_{\theta=\widehat{\theta}},$$

the matrices A and B can be consistently estimated by

$$\widehat{A}=\frac{1}{n}\sum_{j=1}^{n}\widehat{X}_j\widehat{X}_j',\ \ \widehat{B}=\frac{1}{n}\sum_{j=1}^{n}\widehat{u}_j^2\widehat{X}_j\widehat{X}_j',$$

respectively, $b\left(\xi\right)$ can be consistently estimated by

$$\widehat{b}\left(\xi\right)=\frac{1}{n}\sum_{j=1}^{n}\widehat{X}_j w\left(\xi'\Phi(x_j)\right),$$

and $c\left(\xi\right)$ can be consistently estimated by

$$\widehat{c}\left(\xi\right)=\frac{1}{n}\sum_{j=1}^{n}\widehat{u}_j^2\widehat{X}_j w\left(\xi'\Phi(x_j)\right).$$

Hence,

$$\begin{aligned}\int_{\Xi(c)}\widehat{\sigma}^2(\xi)d\xi &= \frac{1}{n}\sum_{j=1}^{n}\widehat{u}_j^2\int_{\Xi(c)}w\left(\xi'\Phi(x_j)\right)^2 d\xi \\ &\quad +\int_{\Xi(c)}\widehat{b}\left(\xi\right)'\widehat{A}^{-1}\widehat{B}\widehat{A}^{-1}b\left(\xi\right)d\xi-2\int_{\Xi(c)}\widehat{b}\left(\xi\right)'\widehat{A}^{-1}\widehat{c}\left(\xi\right)d\xi.\end{aligned}$$

Next, denote

$$\omega_{j_1,j_2}\left(c\right)=\left(2c\right)^{-k}\int_{\Xi(c)}w\left(\xi'\Phi(\widetilde{x}_{j_1})\right)w\left(\xi'\Phi(\widetilde{x}_{j_2})\right)d\xi$$

Then the ICM test statistic takes the form

$$\widehat{T}_{ICM}(c) = \frac{\widehat{T}_1(c)}{\widehat{T}_2(c)} \tag{13}$$

where

$$\widehat{T}_1(c) = \frac{\int_{\Xi(c)} |\widehat{z}(\xi)|^2 \, d\xi}{\int_{\Xi(c)} d\xi} = \frac{1}{n} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \widehat{u}_{j_1} \widehat{u}_{j_2} \omega_{j_1,j_2}(c) \tag{14}$$

and

$$\begin{aligned} \widehat{T}_2(c) &= \frac{\int_{\Xi(c)} \widehat{\sigma}^2(\xi) d\xi}{\int_{\Xi(c)} d\xi} = \frac{1}{n} \sum_{j=1}^{n} \widehat{u}_j^2 \omega_{j,j}(c) \\ &+ \frac{1}{n^2} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \widehat{X}_{j_1}' \widehat{A}^{-1} \widehat{B} \widehat{A}^{-1} \widehat{X}_{j_2} \omega_{j_1,j_2}(c) \\ &- 2 \frac{1}{n^2} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \widehat{X}_{j_1}' \widehat{A}^{-1} \widehat{X}_{j_2} \widehat{u}_{j_2}^2 \omega_{j_1,j_2}(c) . \end{aligned} \tag{15}$$

To derive $\omega_{j_1,j_2}(c)$ let

$$\xi' \Phi(\widetilde{x}_j) = \sum_{\ell=1}^{k} \xi_\ell \phi_{\ell,j}.$$

Then it is easy to verify that

$$\begin{aligned} \omega_{j_1,j_2}(c) &= \prod_{\ell=1}^{k} \frac{\sin(c(\phi_{\ell,j_1} - \phi_{\ell,j_2}))}{c(\phi_{\ell,j_1} - \phi_{\ell,j_2})} \text{ if } w(.) = \cos(.) + \sin(.), \\ \omega_{j_1,j_2}(c) &= \prod_{\ell=1}^{k} \frac{\exp(c(\phi_{\ell,j_1} + \phi_{\ell,j_2})) - \exp(-c(\phi_{\ell,j_1} + \phi_{\ell,j_2}))}{2.c(\phi_{\ell,j_1} + \phi_{\ell,j_2})} \\ &\quad \text{if } w(.) = \exp(.) . \end{aligned}$$

Note that in these cases $\omega_{j_1,j_2}(0) = 1$, so that

$$\widehat{T}_1(0) = \frac{1}{n} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \widehat{u}_{j_1} \widehat{u}_{j_2} = \left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \widehat{u}_j \right)^2 ,$$

$$\widehat{T}_2(0) = \frac{1}{n}\sum_{j=1}^{n}\widehat{u}_j^2 + \left(\frac{1}{n}\sum_{j=1}^{n}\widehat{X}_j\right)'\widehat{A}^{-1}\widehat{B}\widehat{A}^{-1}\left(\frac{1}{n}\sum_{j=1}^{n}\widehat{X}_j\right)$$
$$-2\left(\frac{1}{n}\sum_{j=1}^{n}\widehat{X}_j\right)'\widehat{A}^{-1}\left(\frac{1}{n}\sum_{j=1}^{n}\widehat{X}_j\widehat{u}_j^2\right),$$

If the model contains a constant term then by the first-order conditions for NLLS, $\sum_{j=1}^{n}\widehat{u}_j = 0$, hence $\widehat{T}_1(0) = 0$ and thus $T_{ICM}(0) = 0$. Therefore, c should not be chosen too close to 0.

4.3 Time series data

If your data consist of time series, EasyReg automatically assumes that you want to use De Jong's (1996) version of the ICM test. The default instrumental variables are then all the lagged dependent variables and the current and lagged X variables, if any. If there are X variables in your model, for example if your model is a (nonlinear) ARX(p) model,

$$y_t = g(y_{t-1,\ldots,}y_{t-p}, x_t, \theta_0) + u_t, \tag{16}$$

or a (nonlinear) ARMAX(p, q) model,

$$y_t = g(y_{t-1}, \cdots, y_{t-p}, x_t, \theta_0) + u_t - \gamma_1 u_{t-1} - \cdots - \gamma_q u_{t-q}, \tag{17}$$

where u_t is a martingale difference process, and $x_t \in \mathbb{R}^k$ is a vector of exogenous (X) variables, then the default instrumental variables are $y_{t-j}, j \geq 1$, and all (lagged) x_t. However, only the latter are shown in the list of default instrumental variables. You may remove some or all of the components of x_t from the list of instrumental variables, or add other variables to the list of instrumental variables, but you cannot remove the lagged dependent variables. Once you have confirmed the choice of instruments other than the lagged dependent variables, you have to specify their minimum lags. In the cases (16) and (17) I recommend that you adopt the default instruments x_t, and specify zero as the minimum lag. Then the ICM test tests the martingale difference hypothesis

$$P\left(E[u_t|z_{t-1}, z_{t-2}, z_{t-3}, \ldots\ldots\ldots] = 0\right) = 1,$$

where

$$z_{t-1} = (y_{t-1}, x_t')'.$$

Similarly to the cross-section case, the z_{t-j}'s in (7) are standardized as $\widetilde{z}_{t-j}$ by taking them in deviation of the sample mean, and then dividing them by the sample standard error. Moreover, the $\widetilde{z}_t$'s that are not observable are set equal to zero vectors, say for $t < 1$ Then the actual weight functions are

$$w_t(\xi) = w\left(\sum_{j=1}^{t-1} \xi_j' \Phi(\widetilde{z}_{t-j})\right),$$

with Φ the same as in the cross-section case. Moreover, the set Σ is now infinite dimensional,

$$\Xi\left(c\right) = \times_{j=1}^{\infty}\left\{\times_{\ell=1}^{k}\left[-cj^{-2}, cj^{-2}\right]\right\}$$

where $c > 0$ has to be chosen.

Finally, note that the same limiting null distribution and critical values as in the cross-section case apply.

4.4 Computational options

It is clear from (14) and (15) that if the sample size n is large, the computation of the integrals involved will take a long time on a PC. Therefore EasyReg offers the option (Option 1) to compute (14) and (15) by Monte Carlo integration:

$$\widehat{T}_1(c) \approx \frac{1}{M}\sum_{j=1}^{M}\left|\widehat{z}(\xi_j)\right|^2, \ \widehat{T}_2(c) \approx \frac{1}{M}\sum_{j=1}^{M}\widehat{\sigma}^2(\xi_j),$$

where the ξ_j are random drawings from the uniform distribution $\Xi(c)$. However, if you choose the option of computing the ICM test exactly (Option 2), you can specify a time period after which you will get a message asking whether you if want to continue, or to switch to Monte Carlo integration.

REFERENCES

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. (1994): *Topics in Advanced Econometrics: Estimation, Testing, and Specification of Cross-Section and Time Series Models*, Cambridge: Cambridge University Press

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

Boning, W. B. and F. Sowell (1999): "Optimality for the Integrated Conditional Moment Test", *Econometric Theory* 15, 710-718.

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.