



Working Paper 06-41
Economics Series 11
June 2006

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624 98 75

A CONSISTENT SPECIFICATION TEST FOR MODELS DEFINED BY CONDITIONAL MOMENT RESTRICTIONS

Manuel A. Domínguez and Ignacio N. Lobato¹

Abstract

This article addresses statistical inference in models defined by conditional moment restrictions. Our motivation comes from two observations. First, generalized method of moments, which is the most popular methodology for statistical inference for these models, provides a unified methodology for statistical inference, but it yields inconsistent statistical procedures. Second, consistent specification testing for these models has abandoned a unified approach by regarding as unrelated parameter estimation and model checking. In this article, we provide a consistent specification test, which allows us to propose a simple unified methodology that yields consistent statistical procedures. Although the test enjoys optimality properties, the asymptotic distribution of the considered test statistic depends on the specific data generating process. Therefore, standard asymptotic inference procedures are not feasible. Nevertheless, we show that a simple original wild bootstrap procedure properly estimates the asymptotic null distribution of the test statistic.

Keywords and Phrases: Generalized Method of Moments, Identification, Unconditional Moments, Marked Empirical Process, Integrated Regression Function.

JEL Classification numbers: C12 and C52.

Acknowledgements: We thank seminar participants at Universidad de Alicante, Universidad de Bilbao, Universidad de Navarra, Universidad Carlos III de Madrid, Universidad Autónoma de Barcelona, Universidad Pompeu Fabra and CEMFI for useful comments. Part of this research was carried out while Lobato was visiting Universidad Carlos III de Madrid thanks to the Spanish Secretaría de Estado de Universidades e Investigación, Ref. no. SAB2004-0034. Domínguez acknowledges financial support from Asociación Mexicana de Cultura. Lobato acknowledges financial support from Asociación Mexicana de Cultura.

¹ Instituto Tecnológico Autónomo de México (ITAM) Av. Camino Sta. Teresa 930. Col. Héroes de Padierna, México D.F. 10700, Mexico , e-mail: ilobato@itam.mx

A CONSISTENT SPECIFICATION TEST FOR MODELS DEFINED BY CONDITIONAL MOMENT RESTRICTIONS ¹

Manuel A. Domínguez and Ignacio N. Lobato

Instituto Tecnológico Autónomo de México (ITAM)

Av. Camino a Santa Teresa #930

Col. Héroes de Padierna

10700 México, D.F., MEXICO

E-mail: ilobato@itam.mx

Abstract

This article addresses statistical inference in models defined by conditional moment restrictions. Our motivation comes from two observations. First, generalized method of moments, which is the most popular methodology for statistical inference for these models, provides a unified methodology for statistical inference, but it yields inconsistent statistical procedures. Second, consistent specification testing for these models has abandoned a unified approach by regarding as unrelated parameter estimation and model checking. In this article, we provide a consistent specification test, which allows us to propose a simple unified methodology that yields consistent statistical procedures. Although the test enjoys optimality properties, the asymptotic distribution of the considered test statistic depends on the specific data generating process. Therefore, standard asymptotic inference procedures are not feasible. Nevertheless, we show that a simple original wild bootstrap procedure properly estimates the asymptotic null distribution of the test statistic.

Keywords and Phrases: Generalized Method of Moments, Identification, Unconditional Moments, Marked Empirical Process, Integrated Regression Function.

JEL classification numbers: C12 and C52

¹We thank seminar participants at Universidad de Alicante, Universidad de Bilbao, Universidad de Navarra, Universidad Carlos III de Madrid, Universidad Autónoma de Barcelona, Universidad Pompeu Fabra and CEMFI for useful comments. Part of this research was carried out while Lobato was visiting Universidad Carlos III de Madrid thanks to the Spanish Secretaría de Estado de Universidades e Investigación, Ref. no. SAB2004-0034. Domínguez acknowledges financial support from Asociación Mexicana de Cultura. Lobato acknowledges financial support from Asociación Mexicana de Cultura.

1 Introduction

A successful approach to statistical inference in econometrics models is based on the use of a compatibility index between the model and the data. Formally, this compatibility index is expressed in terms of an Objective Function (OF, hereinafter) that takes an ideal value when there is full agreement between the model assumptions and the data. Once the OF is defined, all inferential procedures are related to it. Parameter estimators are the parameter values that make the OF closest to the ideal value. Tests for correct specification are based on the difference between the ideal value of the OF and the value it takes on the model. Tests for parameter restrictions are based on the change in the OF derived from the imposition of these restrictions.

Probably the most popular application of this approach are the Generalized Method of Moments (GMM, henceforth) procedures employed in models defined by Conditional Moment Restrictions (CMR, hereinafter). In these procedures, the OF is a function of a finite number of unconditional moment restrictions implied by the infinite restrictions that define the model. The ideal value of the OF is zero and larger values for the OF indicate larger discrepancies between the model assumptions and the data. Parameters are estimated by the value that minimizes the OF. Specification testing is carried out using the overidentifying restriction test that rejects when the minimized value of the OF is statistically different from zero.

The GMM approach has been criticized because CMR cannot be fully imposed by a finite number of unconditional restrictions. This problem affects any aspect of statistical inference, and in particular, both specification testing and estimation. Regarding specification testing, the problem was early noticed, see Newey (1985) or Tauchen (1985), and it implies that the overidentifying restriction test is inconsistent. To address this problem, Bierens (1982) proposed an alternative specification test based on a compatibility index that targets to impose in the data an infinite number of unconditional moments that are equivalent to the CMR that define the model. Bierens' test is the first example of a consistent specification test for CMR models. Since then, a great variety of specification tests have been proposed based on the same idea, see for instance, Bierens and Ploberger (1997), Stute (1997), Carrasco and Florens (2000) and references therein. An important and common feature of all these tests is that the parameters of the model are considered nuisance parameters, which are typically substituted by some GMM estimator.

Regarding estimation, this problem has been overlooked until recently. Domínguez and

Lobato (2004, DL hereinafter) have shown that it may result in inconsistency of the GMM estimators. Specifically, DL showed that consistency of GMM estimators depends on the particular model and on additional assumptions to the model, such as assumptions on the distribution of the conditioning variables. As an alternative to GMM, DL have considered the compatibility index employed by the consistent tests referenced in the previous paragraph as an OF and defined the parameter estimator as its minimizing value. The resulting estimator, which we call Consistent Method of Conditional Moment estimator, is always consistent, irrespective of the model and of any additional assumption to the model.

The purpose of this article is to present a global methodology for performing consistent statistical inference on CMR models by extending the results in DL. For model checking we propose to use the value of their OF at its minimum. In this way, we recover the unified approach to inference, and relate in a natural way both parts of inference, estimation and diagnostic testing. Note that the consistency of both the estimators and the specification test derives from the fact that the OF considers an infinite number of unconditional restrictions that fully impose the CMR.

In addition, the resulting specification test presents two advantages over the existing ones. The first one is its simplicity: the test is a by-product of the estimation procedure. The second advantage concerns the behavior of the test under the null hypothesis: the proposed test properly controls the type I error without further assumptions. Note that all specification tests regard the model parameters as nuisance, and need to replace them by consistent estimators. However, the existing tests are very careful imposing the full model definition only at the model checking stage, and not at the estimation stage. As a result, the estimators are consistent only under additional assumptions. If these assumptions do not hold, the tests will not control the type I error.

Concerning the behavior under the alternative, it is a common belief that more powerful tests are obtained by replacing the nuisance parameters by efficient estimators. This would suggest that more powerful tests could be constructed by evaluating our OF at the efficient GMM estimators rather than at the estimator proposed in DL. However, notice that efficiency of the estimators holds just under the null hypothesis, whereas power is a property under the alternative. Therefore, in general, such a test would not be more powerful (besides this test would be computationally more costly and would not properly control the type I error).

In summary, this article complements the results in DL and establishes an approach that produces consistent estimators and tests that control the type I error and are simultaneously

consistent for CMR models. With our approach we recover the traditional econometric spirit of basing all inference on just one compatibility index, linking naturally estimation and hypothesis testing, see for instance, Davidson (2000)².

The plan of this note is the following. In Section 2 we present the testing framework, introduce our test statistic and comment on related tests. Section 3 states the asymptotic properties of the test. Since the asymptotic distribution of the proposed test is case dependent, it cannot be automatically implemented. In Section 4 we propose a feasible implementation of the test that employs critical values obtained by a simple bootstrap procedure. Section 5 concludes and establish some directions for further research.

2 Framework

In this section we will formally introduce our consistent specification test and compare it with related tests procedures. We will follow the notation in DL as close as possible. That is, for all t , Z_t is a time series vector and $\{Y_t, X_t\}$ are two subvectors of Z_t (that could have common coordinates), where Y_t is a k -dimensional time series vector that may contain endogenous and exogenous variables and a finite number of these variables lagged and X_t is a d -dimensional time series vector that contains the exogenous variables (again, a finite number of these variables lagged can be included). The coordinates of Z_t are related by an econometric model which establishes that the true distribution of the data satisfies the following conditional moment restrictions

$$E(h(Y_t, \theta_0) \mid X_t) = 0, \quad a.s. \tag{1}$$

for a unique value $\theta_0 \in \Theta$, where $\Theta \subset \mathbb{R}^m$. Equation (1) defines the parameter θ_0 which is unknown to the econometrician. The function h that maps $\mathbb{R}^k \times \Theta$ into \mathbb{R}^l is supposed to be known. In general, $h(Y_t, \theta_0)$ can be understood as the errors in a multivariate nonlinear dynamic regression model; for instance, $h(Y_t, \theta_0)$ are called generalized residuals in Wooldridge (1990). In this paper, for simplicity, we will consider the case where $l = 1$. This model has been repeatedly considered in the econometrics literature and several estimators have been proposed, see references in DL.

²"All the usual optimization estimators share the feature that the value of the expected criterion function at the minimum is an indicator of goodness of fit" (Davidson, 2000, p.221).

In this article, we consider testing whether model (1) is correctly specified. Specifically, we consider as null hypothesis

$$H_0 : E(h(Y_t, \theta_0) | X_t) = 0, \quad a.s.$$

for a unique value $\theta_0 \in \Theta$, where $\Theta \subset \mathbb{R}^m$; and the alternative hypothesis is that for any θ

$$H_A : P(E(h(Y_t, \theta) | X_t) = 0) < 1 \quad a.s..$$

As mentioned in the introduction, it is well known that the GMM overidentifying restriction test is not consistent for our null hypothesis because it just tests the validity of an arbitrary finite number of unconditional restrictions (from the infinite implied by the conditional expectation (1)). Note that the problem is not the selection of some arbitrary (vs. optimal) instruments, as the examples in DL show, but the use of a finite number of instruments. In order to avoid this problem, Bierens (1982, 1990), Bierens and Ploberger (1997) or Stute (1997), among others, proposed tests which employ an infinite number of unconditional moments. However, note that those references do not consider inference as a whole, but they just focus on the model check stage. Then, since the parameters of the model are nuisance for model checking, they propose to replace these parameters by consistent estimators, without discussing carefully the estimation stage. As a result, the proposed tests may not control properly the type I error, as we show next. Assume that H_0 holds but the true value θ_0 is estimated with an inconsistent estimator $\hat{\theta}$ which converges in probability to the random variable S . Then, these tests check whether $E(h(Y_t, S) | X_t) = 0, \quad a.s.$, which may be false, although H_0 holds. Therefore, under H_0 , these tests will reject asymptotically more often than the specified theoretical level. In particular, the examples presented by DL can be worked out further to show that the GMM estimator converges to θ_0 with probability p and to some other parameter values $\theta \neq \theta_0$ with probability $1 - p$. In this particular case, the asymptotic type I error of the tests that employ the GMM estimator is $\alpha p + (1 - p)$ where α is the desired nominal size. This example illustrates that, unless additional assumptions are imposed, these tests do not control properly the level.

Next, we describe our test procedure. We propose a testing procedure that uses the whole information about θ_0 contained in expression (1). Let P_X be the probability law of X_t and let $I(X_t \leq x)$ denote the indicator function that equals 1 when each component in X_t is less or equal than the corresponding component in x ; and equals 0 otherwise. DL used the compatibility index

$$Q(\theta) = \int_{\mathbb{R}^d} E(h(Y_t, \theta) I(X_t \leq x) | X_t)^2 dP_X(x) \quad (2)$$

which is 0 at θ_0 only if the conditional moment restrictions hold. DL proposed estimating θ_0 by

$$\hat{\theta} = \arg \min_{\theta \in \Theta} Q_n(\theta),$$

where $Q_n(\theta)$ is the sample analog of $Q(\theta)$, namely,

$$Q_n(\theta) = \frac{1}{n^3} \sum_{\ell=1}^n \left(\sum_{t=1}^n h(Y_t, \theta) I(X_t \leq X_\ell) \right)^2.$$

DL showed that this estimator is consistent. Following the discussion in the introduction, a natural goodness of fit test procedure is based on evaluating the distance of the previously minimized objective function to zero (note that zero is the value of the population analogue of $Q_n(\theta)$ if the model is correctly specified). Hence, in order to avoid a degenerate asymptotic null distribution, the proposed test statistic is $T_n = nQ_n(\hat{\theta})$.

As an additional point, notice that instead of plugging in $\hat{\theta}$, that is, a consistent, but inefficient, estimator, one could propose to employ as test statistic $T_n^{(2)} = nQ_n(\tilde{\theta})$, where $\tilde{\theta}$ is an efficient estimator, such as the one proposed in Section 4 in DL. As mentioned in the introduction, the comparison between T_n and $T_n^{(2)}$ should be carried out under the null and under the alternative. Under the null, $T_n^{(2)}$ may not control the type I error. Even when both tests control properly the type I error, under the alternative, $T_n^{(2)}$ does not lead to a more powerful test. The reason is clear: efficiency of $\tilde{\theta}$ is a property derived under the null hypothesis, assuming that the specified model is correct, whereas power refers to the behavior of the statistic under the alternative hypothesis. The following example illustrates. Consider the model $y_i = g(x_i)/\sqrt{n} + u_i$, where u_i is $N(0, 1)$ and x_i takes three values $-1, 0$ and 1 , with probabilities p_1, p_2 and $1 - p_1 - p_2$. For the null hypothesis $g(x_i) = 0$, it can be shown that, in general, against local quadratic alternatives, T_n dominates $T_n^{(2)}$, whereas for local linear alternatives $T_n^{(2)}$ dominates T_n . In particular, in Figure 1, we have plotted in black the combinations of (p_1, p_2) where T_n has more power than $T_n^{(2)}$, for alternatives of the form $g(x_i) = ax_i^2$, with $a > 0$. From this plot it is clear that T_n dominates $T_n^{(2)}$ unless p_1 takes high values and p_2 is low. In practice, T_n has the additional advantage of being computationally simpler.

In order to derive the asymptotic theory, it is useful to rewrite the statistic in terms of the rescaled integrated regression function that can be seen as a marked empirical process with marks given by $h(Y_t, \theta)$. That is, introduce the following empirical process

$$R_n(\theta, x) = n^{-1/2} \sum_{t=1}^n h(Y_t, \theta) I(X_t \leq x).$$

Then, we can write

$$T_n = \frac{1}{n} \sum_{\ell=1}^n R_n(\widehat{\theta}, X_\ell)^2$$

so that our statistic can be seen as a Cramer von Mises statistic applied to the marked empirical process $R_n(\widehat{\theta}, x)$.

In the next section we state the asymptotic theory for this test statistic. Since the asymptotic distribution is case-dependent, in Section 4 we propose to employ the bootstrap to estimate the asymptotic critical values.

3 Asymptotic Theory

Let $|\cdot|$ denote the Euclidean norm in the corresponding Euclidean space, and assume that all the considered functions are Borel measurable. The following set of assumptions are referred to as *assumptions A*. These assumptions are slightly weaker than the assumptions in DL, but are sufficient for the results in DL to hold, see Escanciano (2006).

Assumption A1. $h(y, \cdot)$ is continuous in Θ for each y in R^k , $|h(Y_t, \theta)| < k(Y_t)$ with $E k(Y_t) < \infty$ and $E(h(Y_t, \theta) | X_t) = 0$ a.s. if and only if $\theta = \theta_0$.

Assumption A2. Z_t is ergodic and strictly stationary.

Assumption A3. $\Theta \subset R^m$ is compact.

Assumption A4. $h(y, \cdot)$ is once continuously differentiable in a neighborhood of θ_0 and satisfies that $E \left[\sup_{\theta \in \mathfrak{N}_0} \left| \dot{h}(Y_t, \theta) \right| \right] < \infty$ where \mathfrak{N}_0 denotes a neighborhood of θ_0 and $\dot{h}(Y_t, \theta) = \partial h(Y_t, \theta) / \partial \theta$.

Assumption A5. $h(Y_t, \theta_0)$ is a martingale difference sequence with respect to $\{Z_s, s \leq t\}$.

Assumption A6. $\theta_0 \in \text{int}(\Theta)$.

Assumption A7. $E(h^2(Y_t, \theta_0)) < \infty$.

DL established that under the null hypothesis

$$\sqrt{n}(\widehat{\theta} - \theta_0) \rightarrow_d \Sigma_{\dot{H}\dot{H}'}^{-1} \Sigma_{\dot{H}B_\Gamma}, \quad (3)$$

where $\Sigma_{\dot{H}\dot{H}'} = \int \dot{H}\dot{H}' dP_{X_1}$, $\Sigma_{\dot{H}B_\Gamma} = \int \dot{H}B_\Gamma dP_{X_1}$, with $\dot{H}(x) = E(\dot{h}(Y_t, \theta_0) I(X_t \leq x))$ and B_Γ denotes a centered Gaussian process in $D[\mathbb{R}]^d$ (where $D[\mathbb{R}]^d$ is the space of real functions that are continuous from above and with limits from below), with covariance structure given by $\Gamma(r, s) = E(h^2(Y_t, \theta_0) I(X_t \leq r \wedge s))$. Note that, when h is homoskedastic and $d = 1$, B_Γ particularizes to a scaled Brownian motion. In addition, notice that (3) reminds

similar properties satisfied by popular estimators such as nonlinear least squares or GMM estimators. The difference with them is that, in our case the involved variables ("regressors" and errors or generalized residuals) are partial sum processes instead of raw variables.

Using the previous results, we can derive the following properties of T_n . These theorems are straightforward, given the results in DL and in Escanciano (2006), and so, their proofs are omitted.

Theorem 1. Under *assumptions A* and the null hypothesis

$$T_n \rightarrow_d \int \left(B_\Gamma + \dot{H}' \Sigma_{HH'}^{-1} \Sigma_{HB_\Gamma} \right)^2 dP_{X_1}.$$

Remark 1. Note that the covariance structure of the process $B_\Gamma + \dot{H}' \Sigma_{HH'}^{-1} \Sigma_{HB_\Gamma}$ is given by

$$\begin{aligned} \Phi(t, s) &= \Gamma(t, s) + \dot{H}'(t) \Sigma_{HH'}^{-1} \Sigma_{H\Gamma}'(s) + \Sigma_{H\Gamma}(t) \Sigma_{HH'}^{-1} \dot{H}'(s) \\ &\quad + \dot{H}'(t) \Sigma_{HH'}^{-1} \left[\int \dot{H}'(u) \Gamma(u, v) \dot{H}(v) P_{X_1}(du) P_{X_1}(dv) \right] \Sigma_{HH'}^{-1} \dot{H}(s), \end{aligned}$$

where $\Sigma_{H\Gamma}'(s) = \int \dot{H}'(u) \Gamma(u, s) dP_{X_1}(u)$. Therefore, the critical values of the test statistic T_n depend on the data generating process (DGP), complicating statistical inference.

Concerning the behavior under the alternative, it is straightforward to show that the test statistic T_n diverges under fixed alternatives. Under a sequence of local alternatives, such as

$$H_{A,n} : E(h(Y_t, \theta_0) | X_t) = \frac{g(X_t)}{\sqrt{n}} \quad a.s.$$

note that θ_0 still minimizes (2). Therefore, we have the following result.

Theorem 2. Under *assumptions A* and under a sequence of local alternatives

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightarrow_d \Sigma_{HH'}^{-1} \Sigma_{H(B_\Gamma+G)},$$

where $\Sigma_{H(B_\Gamma+G)} = \int \dot{H}(B_\Gamma + G) dP_{X_1}$, $G(x) = E(g(X)I(X \leq x))$, and

$$T_n \rightarrow_d \int \left(B_\Gamma + G + \dot{H}' \Sigma_{HH'}^{-1} \Sigma_{H(B_\Gamma+G)} \right)^2 dP_{X_1}.$$

Notice that the asymptotic distribution of T_n under $H_{A,n}$ is a Gaussian process centered at the function $G + \dot{H}' \Sigma_{HH'}^{-1} \Sigma_{HG}$, where $\Sigma_{HG} = \int \dot{H} G dP_{X_1}$. Note that the structure of the asymptotic distribution is essentially equivalent to the structure of the tests proposed

in Bierens and Ploberger (1997) and Stute (1997). These papers prove that the existence of this bias, G , is all that is needed to show that the probability of rejecting under $H_{A,n}$ is larger than α . In particular, denote $T_0 = \int \left(B_\Gamma + \dot{H}' \Sigma_{\dot{H}\dot{H}'}^{-1} \Sigma_{\dot{H}B_\Gamma} \right)^2 dP_{X_1}$ and $T = \int \left(B_\Gamma + G + \dot{H}' \Sigma_{\dot{H}\dot{H}'}^{-1} \Sigma_{\dot{H}(B_\Gamma+G)} \right)^2 dP_{X_1}$. Then, it is straightforward to show that for any t , we have that $P(T_0 > t) < P(T > t)$, so the next theorem follows.

Theorem 3. Under *assumptions A*, T_n test has nontrivial power against a sequence of local alternatives $H_{A,n}$.

Finally, Bierens and Ploberger (1997) show that this kind of test enjoys optimality properties: a sequence of local alternatives, which depend on the spectral decomposition of the bilinear operator Φ defined in Remark 1, can be defined such that T_n is asymptotically equivalent to a likelihood ratio under H_0 . On the other hand, under regularity conditions it can be shown that $\sqrt{n}(\hat{\theta} - \theta_0)$ is asymptotically equivalent to $\Sigma_{\dot{H}\dot{H}'}^{-1} \left(n^{-1/2} \sum_t \Sigma_{\dot{H}'\Delta}(X_t) h(Y_t, \theta_0) \right)$, where $\Sigma_{\dot{H}'\Delta}(x) = \int \dot{H}'(u) 1(x \leq u) dP_{X_1}(u)$. Therefore, the optimality results in Stute (1997) also apply to T_n .

4 Bootstrap test

Since Φ depends on the DGP, the asymptotic distributions of both $R_n(\theta, x)$ and T_n generally also do. Hence, the theory established in the previous section cannot be automatically applied for statistical inference because there are not generally valid critical values. There are two approaches to constructing feasible tests: to estimate the critical values using the bootstrap or to obtain an asymptotically distribution free test statistic via a martingalization. Koul and Sakhanenko (2005) report that in finite samples, tests based on the bootstrap control worse the type I error, although they have more empirical power. We prefer to follow the bootstrap approach for three reasons. First, the bootstrap test preserves the optimality properties of the original unfeasible test, see Domínguez (2004). Second, the bootstrap test is valid under heteroskedasticity of any form and it is not a case specific procedure. Finally, it is unclear whether the martingalization approach would lead to abandon the unifying inference approach advocated in this article.

Next, we explain and justify the proposed bootstrap-based test procedure. Recall

$$R_n(\hat{\theta}, x) = n^{-1/2} \sum_{t=1}^n h(Y_t, \hat{\theta}) I(X_t \leq x),$$

so that,

$$R_n(\widehat{\theta}, x) = R_{1n}(\theta_0, x) + R_{2n}(\widehat{\theta}, x),$$

where

$$R_{1n}(\theta, x) = \frac{1}{\sqrt{n}} \sum_{t=1}^n h(Y_t, \theta) I(X_t \leq x), \quad R_{2n}(\theta, x) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \dot{h}(Y_t, \theta^*)(\theta - \theta_0) I(X_t \leq x)$$

and θ^* is intermediate between $\widehat{\theta}$ and θ_0 . R_{1n} is the process that a test would use for model checking when the parameters are known, while R_{2n} corrects R_{1n} for the effect of the estimation of the model parameters. Then, using (3), we define

$$R_n^*(\widehat{\theta}, x) = R_{1n}^*(\widehat{\theta}, x) + R_{2n}^*(\widehat{\theta}, x)$$

where

$$R_{1n}^*(\widehat{\theta}, x) = \frac{1}{\sqrt{n}} \sum_{t=1}^n h(Y_t, \widehat{\theta}) I(X_t \leq x) W_t, \quad \text{and} \quad R_{2n}^*(\widehat{\theta}, x) = \widehat{H}(x)' \widehat{\Sigma}_{HH'}^{-1} \widehat{\Sigma}_{HB_\Gamma},$$

where

$$\widehat{H}(x) = \frac{1}{n} \sum_{s=1}^n \dot{h}(Y_s, \widehat{\theta}) I(X_s \leq x),$$

$$\widehat{\Sigma}_{HH'} = \frac{1}{n} \sum_{k=1}^n \widehat{H}(X_k) \widehat{H}'(X_k), \quad \widehat{\Sigma}_{HB_\Gamma} = \frac{1}{n} \sum_{k=1}^n \widehat{H}(X_k) R_{1n}^*(\widehat{\theta}, X_k),$$

and where $\{W_t\}$ is a sequence of independent random variables with zero mean, unit variance and bounded support. The main idea is to estimate the distribution of $\sqrt{n}R_n(\widehat{\theta}, x)$ by the distribution of $\sqrt{n}R_n^*(\widehat{\theta}, x)$, and hence to estimate the distribution of T_n by the distribution of T_n^* , defined by

$$T_n^* = \frac{1}{n} \sum_{\ell=1}^n R_n^*(\widehat{\theta}, X_\ell)^2. \quad (4)$$

This procedure has been called a wild or external bootstrap, see Wu (1986), Mammen (1993) and Delgado and Fiteni (2002) for applications in econometrics.

Remark 2. Note that the standard bootstrap approach, based on constructing a bootstrap sample (Y_t^*, X_t) from resampling the residuals, cannot be followed. The reason is that Y_t^* would be defined as the implicit solution of the equation $W_t h(y, \widehat{\theta}) = 0$. However, this solution may not exist or may not be unique.

Remark 3. The wild bootstrap proposed in (4) is original in specification testing. Different authors have proposed wild bootstrap procedures in similar contexts, see for instance,

Stute, González-Manteiga and Presedo-Quindimil (1998) or Dominguez (2004). In these references, the bootstrap procedure is asymptotically equivalent to resampling a complicated process, which can not be regarded as a marked empirical process because of the particular form of the corresponding term R_{2n} . On the contrary, in our case both the estimator and the test statistic are defined in terms of the same process $R_{1n}(\theta, x)$. Consequently, the effects of the errors in T_n are fully summarized in $R_{1n}(\theta, x)$, and hence, only this simple marked empirical process has to be resampled. As a result, in order to bootstrap T_n , the wild bootstrap only involves $R_{1n}^*(\hat{\theta}, x)$, which is just the marked empirical process that one would consider in case the parameters were known.

The next theorem establishes the consistency of the bootstrapped process $\sqrt{n}R_n^*(\hat{\theta}, x)$. This means that asymptotically the probability law of $\sqrt{n}R_n^*(\hat{\theta}, x)$ given the data \mathcal{X}_n is the null asymptotic distribution of $\sqrt{n}R_n(\hat{\theta}, x)$ for almost all samples.

Theorem 4. *Under assumptions A,*

$$\sqrt{n}R_n^*(\hat{\theta}, x) \Rightarrow_* B_\Gamma + \dot{H}'\Sigma_{HH'}^{-1} \cdot \Sigma_{HB_\Gamma} \quad a.s.,$$

where \Rightarrow_* a.s. denotes weak convergence almost surely under the bootstrap law, that is,

$$P(\sqrt{n}R_n^*(\hat{\theta}, x) \leq s \mid \mathcal{X}_n) \rightarrow_{a.s.} P(B_\Gamma + \dot{H}'\Sigma_{HH'}^{-1} \cdot \Sigma_{HB_\Gamma} \leq s) \text{ as } n \rightarrow \infty$$

plus tightness a.s.

Therefore, the asymptotic distribution of $\sqrt{n}R_n(\hat{\theta}, x)$ can be estimated with that of $\sqrt{n}R_n^*(\hat{\theta}, x)$. Similarly, the asymptotic distribution of T_n can be estimated with that of T_n^* . In fact, a straightforward application of the Continuous Mapping yields the following corollary.

Corollary. *Under assumptions A,*

$$T_n^* \Rightarrow_* \int \left(B_\Gamma + \dot{H}'\Sigma_{HH'}^{-1} \cdot \Sigma_{HB_\Gamma} \right)^2 dP_{X_1} \quad a.s..$$

This corollary justifies the estimation of the asymptotic critical values of T_n by those of T_n^* . In practice, the critical values of T_n^* are approximated by simulations. Hence, the proposed general bootstrap test consists in the following steps:

- a) Calculate the test statistic T_n .
- b) Generate $\{W_t\}$ a sequence of n bounded independent random variables with zero mean and unit variance. This sequence is serially independent and is also independent of the original sample \mathcal{X}_n .

c) Compute $\sqrt{n}R_n^*(\hat{\theta}, x)$ and T_n^* .

d) Repeat steps b) and c) B times where in step b) each sequence $\{W_t\}$ is independent of each other. This produces a set of B independent (conditionally in the sample) values of T_n^* that share the asymptotic distribution of T_n .

e) Let $T_{[1-\alpha]}^*$ be the $1 - \alpha$ -quantile of the empirical distribution of the B values of T_n^* . The proposed test of nominal level α rejects the null hypothesis if $T_n > T_{[1-\alpha]}^*$.

The corollary establishes that under the null hypothesis, T_n and T_n^* share the same asymptotic distribution for almost all samples. Hence, under the null, the rejection probability of the bootstrap test converges to α (the theoretical level). In addition, using arguments similar to Domínguez (2004), it can be shown that the proposed bootstrap does not alter the critical region. Therefore,

$$P(T_n > T_{[1-\alpha]}^*) \rightarrow \begin{cases} \alpha & \text{under the null,} \\ 1 & \text{under the alternative,} \\ C & \text{under the sequence of local alternatives,} \end{cases}$$

where $\alpha < C < 1$. Hence, the proposed bootstrap test has an α asymptotic level, it is consistent, it is able to detect alternatives tending to the null at the $n^{-1/2}$ rate, and preserves admissibility.

5 Conclusions

In this article we have proposed a consistent specification test for models defined by CMR. Together with DL, this article provides a simple unified methodology for performing consistent statistical inference for CMR models. Consistency derives from the use of a compatibility index that takes into account an infinite number of unconditional restrictions, which fully impose the definition of the model. Our approach highlights the importance of the estimation stage in the model checking stage, an issue that has been overlooked in the previous literature.

Compared to closely related existing tests, there are three main differences. The first is that our test is a part of a unified approach, as mentioned above. The second is that either the model or the assumptions of the rival tests are different from ours. Finally, the unified approach allows us to define a new bootstrap procedure for estimating the critical values.

References

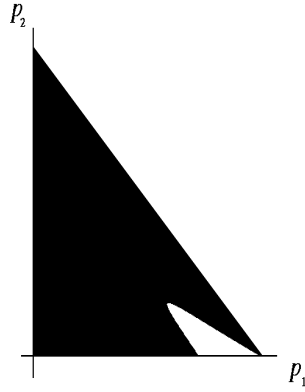


Figure 1: Local power comparison of T_n vs $T_n^{(2)}$. In black, the region where T_n is more powerful than $T_n^{(2)}$.

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

_____ (1990), "A consistent conditional moment test of functional form", *Econometrica*, 58, 1443-1458.

_____ and Ploberger, W. (1997), "Asymptotic theory of integrated conditional moment test", *Econometrica*, 65, 1129-1151.

Carrasco, M. and Florens J.-P. (2000), "Generalization of GMM to a continuum of moment conditions", *Econometric Theory*, 16, 797-834.

Davidson, J. (2000), *Econometric Theory*, Oxford University Press.

Delgado, M. A. and Fiteni, I. (2002), "External bootstrap tests for parameter stability", *Journal of Econometrics*, 109, 275-303.

Domínguez, M.A. (2004), "On the power of bootstrapped specification tests", *Econometric Reviews*, 23, 215-228.

Domínguez, M.A. and I.N. Lobato (2003), "Testing the martingale difference hypothesis", *Econometric Reviews*, 22, 351-377.

_____ (2004), "Consistent Estimation of Models Defined by Conditional Moment Restrictions",

Econometrica, 1601-1615.

Escanciano, J.C. (2006), "Goodness-of-fit Tests for Linear and Nonlinear Time Series Models", *Journal of the American Statistical Association* 101, 531 - 541.

Khamaladze, E. V. (1981), "Martingale approach in the theory of goodness-of-fit tests", *Theory of Probability and Applications*, 26, 240-257.

Koul, H. and Sakhanenko, L. (2005), "Goodness-of-fit testing in regression: A finite sample comparison of bootstrap methodology and Khamaladze transformation", *Statistics and Probability Letters*, 74 (3), 290-302.

Koul, H. L. and W. Stute (1999), "Nonparametric model checks for time series", *Annals of Statistics*, 27, 204-236.

Mammen, E. (1993), "Bootstrap and Wild Bootstrap for High Dimensional Linear Models", *Annals of Statistics*, 21, 255-285.

Newey, W. (1985), "Maximum Likelihood Specification Testing and Conditional Moment Tests", *Econometrica* 53, no. 5, 1047-1070.

Stute, W. (1997), "Nonparametric model checks for regression", *Annals of Statistics*, 25, 613-641.

_____, González-Manteiga, W.G., and Presedo-Quindimil, M. (1998), "Bootstrap approximations in model checks for regression", *Journal of the American Statistical Association*, 83, 141-149.

Tauchén, G. (1985), "Diagnostic testing and evaluation of maximum likelihood models", *Journal of Econometrics* 30, 415-443.

Wooldridge, J.M. (1990): "A Unified Approach to Robust, Regression-Based Specification Tests", *Econometric Theory*, 6, 17-43.

Wu, C. F. J. (1986), "Jackknife, bootstrap and other resampling methods in regression analysis (with discussion)", *Annals of Statistics*, 14, 1261-1350.