# A SIMPLE TEST FOR IDENTIFICATION IN GMM UNDER CONDITIONAL MOMENT RESTRICTIONS

**By**

**Francesco Bravo, Juan Carlos Escanciano and Taisuke Otsu**

**March 2011**

**COWLES FOUNDATION DISCUSSION PAPER NO. 1790**

**COWLES FOUNDATION FOR RESEARCH IN ECONOMICS**
**YALE UNIVERSITY**
**Box 208281**
**New Haven, Connecticut 06520-8281**

**http://cowles.econ.yale.edu/**

# A Simple Test for Identification in GMM under Conditional Moment Restrictions[1]

Francesco Bravo[2] Juan Carlos Escanciano[3] Taisuke Otsu[4]

This version: April 10, 2010.

**Abstract:** This paper proposes a simple, fairly general, test for global identification of unconditional moment restrictions implied from point-identified conditional moment restrictions. The test is based on the Hausdorff distance between an estimator that is consistent even under global identification failure of the unconditional moment restrictions, and an estimator of the identified set of the unconditional moment restrictions. The proposed test has a chi-squared limiting distribution and is also able to detect weak identification alternatives. Some Monte Carlo experiments show that the proposed test has competitive finite sample properties already for moderate sample sizes.

*Key words and phrases:* Conditional moment restrictions; Generalized method of moments; Global identification; Hausman test; Asset pricing.

*JEL Classification Number:* C12; C13; C32.

[1]We thank Lin Zhu for excellent research assistance. We thank Barbara Rossi and Atsushi Inoue for providing us with their Matlab codes, and to Jonathan Wright for providing us with his Matlab codes and for drawing our attention to his paper "Testing the Adequacy of Conventional Asymptotics in GMM". We also thank comments from seminar and conference participants at the Bergen University, Queen's University, ESRC Econometric Study Group 2009, Far Eastern Meeting of the Econometric Society 2009, and Midwest Econometrics Group 2009. All errors are our own.

[2]Department of Economics, The University of York.

[3]Corresponding author. Department of Economics, Indiana University, email: jescanci@indiana.edu. Research funded by the Spanish Plan Nacional de I+D+I, reference number SEJ2007-62908.

[4]Cowles Foundation and Department of Economics, Yale University, email: taisuke.otsu@yale.edu. Research funded by the National Science Foundation (SES-0720961)

# 1 Introduction

Economic models can often be characterized by conditional moment restrictions in the underlying economic variables. For example rational expectations and dynamic asset pricing models used in macroeconomics and international finance give rise to conditional moment restrictions in the form of stochastic Euler conditions. Alternatively, exogeneity or other statistical assumptions can also lead to conditional moment restrictions.

The typical approach to estimate these models is to find a set of unconditional moment restrictions implied from the conditional ones, and use Hansen's (1982) generalized method of moments (GMM) estimator, as for example Hansen and Singleton (1982) did in their seminal paper on estimating a consumption based capital asset pricing model.[5] The key assumption in this unconditional GMM-based approach is that the parameters identified in the conditional moment restrictions can be globally identified by the implied unconditional moment restrictions. However, as recently emphasized by Dominguez and Lobato (2004, henceforth DL), this needs not be the case. These authors showed that the global identification condition of GMM can fail in nonlinear models, regardless on whether the instruments are optimally chosen or not. Moreover, since the seminal theoretical works by Staiger and Stock (1997) and Stock and Wright (2000), there is growing empirical evidence indicating potential (weak) identification problems in commonly used macroeconomic models: Canova and Sala (2006) in dynamic stochastic general equilibrium models, Nason and Smith (2008) in the new Keynesian Phillips curve, Yogo (2004) in consumption Euler equations to name just a few. Given the popularity of unconditional GMM-based approach and its potential identification problem, it seems natural to investigate the possibility of testing for global identification of the unconditional moment restrictions.[6]

This paper proposes a simple test for the hypothesis that the unconditional moment restrictions globally identify the true parameter identified in the conditional moment restrictions. The test is a Hausman-type test based on the Hausdorff distance between an estimator that is always consistent for the true parameter, namely DL's consistent estimator, and a GMM-based estimator of the identified set of the un-

[5]Further examples of applications of GMM abound in the economics and financial literature; see e.g. the monograph by Hall (2005) and the anniversary issue on GMM of the *Journal of Business and Economics Statistics* in 2002.

[6]It should be noted that our methodology can be applied to other estimators different from GMM such as the continuous updating estimator (Hansen, Heaton and Yaron (1996)) or any of the members of the generalized empirical likelihood family (Newey and Smith (2004)).

conditional moments.[7] Under the null hypothesis of global identification any member of this estimated identified set will be consistent and asymptotically normal under some regularity conditions. Thus, the proposed test has a simple chi-squared limiting distribution. Under the alternative of identification failure, the Hausdorff distance is expected to be non-zero, leading to a powerful test.

The proposed test can also be useful in the context of weakly identified unconditional moment restriction models (Stock and Wright (2000)), where the unconditional moments may be close to zero in finite samples for possibly a large set of parameter values. Under weak identification GMM estimators have nonstandard limiting distributions and traditional approaches to inference are not valid; see Stock, Wright and Yogo (2005) for a review. The proposed test is able to detect weak identification alternatives with probability tending to one as the sample size increases, and with high probability in finite samples for the data generating processes used in the simulations.

Most of the existing tests for identification have been confined to linear models; see earlier contributions by Koopmans and Hood (1953) and Sargan (1958), and more recent ones by Cragg and Donald (1993), Hahn and Hausman (2002), and Stock and Yogo (2005). In linear models, global identification reduces to a rank condition. Although this rank condition can also be applied to nonlinear models, as in Wright (2003), it is in general neither necessary nor sufficient for global identification, see Sargan (1983) and an example below illustrating this point. In addition to Wright's (2003) test, the only available tests for identification in nonlinear models we are aware of are those of Arellano, Hansen and Sentana (2009), and Wright (2009).[8] Arellano, Hansen and Sentana's (2009) test is an overidentification test (similar to that of Carrasco and Florens (2000)) in which the parameter of interest is a continuous reparametrization of the original (not identified) one. Wright's (2009) test compares the volume of a Wald confidence set with that of Stock and Wright's (2000) $S$-confidence set in overidentified unconditional moment restrictions. The proposed test is a useful complement to these existing tests because it is does not require the specification of the identified set, it can be applied to both just-identified and overidentified models, and is computationally very simple.

In practice the proposed test can be used as follows: if the null hypothesis of global

[7] We stress that our methodology can be based on any consistent estimator for the conditional moment restrictions parameters, as for example those proposed by Carrasco and Florens (2000), Donald, Imbens and Newey (2003), Kitamura, Tripathi and Ahn (2004). We chose DL's estimator because of its simple implementation; in particular no tuning parameters, such as the bandwidth or the rate of growth of the approximating functions, are necessary.

[8] Inoue and Rossi (2008) propose a test for a related but different hypothesis to ours. They test for the equality of two parameter values identified by two different GMM objective functions.

identification is rejected one should use DL. If the null hypothesis is accepted then the standard GMM estimator can be used. Note that DL's estimator is not efficient whereas its one-step efficient version (Section 4 of DL) requires the computation of conditional expectations and it is therefore subject to the curse of dimensionality. Thus from a theoretical and practical point of view the use of the standard GMM seems preferable when the null hypothesis is accepted.

The rest of this paper is organized as follows. The next section briefly reviews standard GMM estimation and the associated potential identification problem. Section 3 makes explicit the null and alternative hypotheses and introduces the Hausman-type test statistic for global identification. Section 4 develops the asymptotic theory, whereas Section 5 reports a Monte Carlo experiment showing that the new test possesses satisfactory finite-sample properties. Section 6 concludes. Finally an Appendix contains formulae for the variances and the proof of the main theorem.

# 2 GMM and Global Identification Failure

The model we consider is defined by a set of conditional moment restrictions

$$E[h(Y_t,\theta_0) \mid X_t] = 0 \text{ almost surely (a.s.), at some unique } \theta_0 \in \Theta \subset \mathbb{R}^p, \quad (1)$$

for a measurable moment function $h : \mathbb{R}^{d_y} \times \Theta \longrightarrow \mathbb{R}^{d_h}$ that is assumed to be known up-to the finite dimensional parameter $\theta_0$. For the sake of exposition we only consider the case $d_h = 1$, the extension of our methods to the case $d_h > 1$ being straightforward. The vector-valued stochastic process $\{Z_t \equiv (Y_t', X_t')'\}_{t\in\mathbb{Z}}$ is a strictly stationary and ergodic time series. Henceforth, $|A|$ and $A'$ denote the Euclidean norm $|A| \equiv (tr(A'A))^{1/2}$ and the transpose of a matrix $A$, respectively. The conditioning variable $X_t$ takes values in $\mathbb{R}^{d_x}$ and can contain lagged values of $Y_t$ and other exogenous variables. Throughout the paper we shall assume that the conditional moment (1) uniquely identifies the parameter $\theta_0$.

The standard unconditional GMM estimator for $\theta_0$ is constructed as follows. Given an $r \times 1$ vector of "instruments" $a(X_t)$ with $r \geq p$, possibly depending on $\theta_0$, the unconditional GMM estimator $\hat{\theta}_{GMM}$ is defined as any solution of the optimization problem

$$\min_{\theta\in\Theta} Q_n(\theta) \equiv \left(\frac{1}{n}\sum_{t=1}^{n} f(Z_t,\theta)\right)' W_n \left(\frac{1}{n}\sum_{t=1}^{n} f(Z_t,\theta)\right), \quad (2)$$

where $f(Z_t,\theta) \equiv a(X_t)h(Y_t,\theta)$ and $W_n$ is a possibly stochastic matrix satisfying some mild conditions; see Assumption A3 below. The critical assumption in the unconditional GMM is that the identified set $\Theta_I \equiv \{\theta \in \Theta : E[f(Z_t, \theta)] = 0\}$ by the unconditional moment restrictions is a singleton $\Theta_I = \{\theta_0\}$, i.e.,

$$E[f(Z_t, \theta)] = 0 \Longrightarrow \theta = \theta_0. \tag{3}$$

This is the global identification assumption of GMM. This assumption was questioned by DL. They provided some examples where (3) is not satisfied. The following example generalizes in an important way one of the examples of DL.

EXAMPLE 1 (DL'S EXAMPLE 2): Assume that $(Y, X)$ satisfies $E[Y|X] = \theta_0^2 X + \theta_0 X^2$ at $\theta_0 = 5/4$. For the unconditional moment restriction $E[a(X)(Y - \theta^2 X - \theta X^2)] = 0$ with a scalar instrument $a(X)$, it can be shown that $\Theta_I = \{\theta_0, \theta_1\}$, where $\theta_1 = (-E[X^2 a(X)]/E[X a(X)]) - \theta_0$. So, global identification of $\theta_0$ holds if and only if $E[a(X)a^*(X)] = 0$, where $a^*(X) = 2\theta_0 X + X^2$. This example contains three important features. (i) Note that $a^*(X)$ is the optimal instrument of (1) provided $Var[Y|X] = 1$. Hence, the use of the optimal instrument (i.e. $a(X) = a^*(X)$) leads to global identification failure, *regardless of the distribution of* $X$. (ii) The identification failure for the instrument $a(X) = a^*(X)$ occurs even if the rank condition $E[a^*(X)^2] > 0$ is satisfied. The full rank condition in GMM does not imply nor is implied by global identification. (iii) When the feasible optimal instrument $a(X) = 2\theta X + X^2$ is employed, the moment conditions becomes $E[(2\theta X + X^2)(Y - \theta^2 X - \theta X^2)] = 0$. In this case, it can be proved that the parameter is not identified when $X \sim N(1, \sigma^2)$, provided $\sigma^2 < 2.0163$. In particular, if $\sigma^2 = 1$, the identified set is $\Theta_I = \{5/4, -5/4, -3\}$. $\square$

The previous example illustrates the identification problem of the unconditional GMM, but is this a pathological example or is the rule rather than the exception? The following theorem complements the examples given in DL and confirms our intuition that the problem is quite general.

THEOREM 1: *A necessary and sufficient condition of* $E[f(Z_t, \theta_1)] = 0$ *with* $\theta_1 \neq \theta_0$ *for some distribution of* $Z_t$ *with support* $\mathcal{Z}$ *in a fixed measurable space is*

$$0 \in \mathcal{R} \equiv \left\{ \sum_{i=1}^{n} p_i f(z_i, \theta_1), z_i \in \mathcal{Z}, \sum_{i=1}^{n} p_i = 1, p_i \geq 0, n \in \mathbb{N} \right\}.$$

**Proof of Theorem 1:** By Lemma 3 in Chamberlain (1987) we can assume without loss of generality that the distribution of $Z_t$ has a finite support on $\mathcal{Z}$. Thus, the proof follows from the definition of $\mathcal{R}$. ■

Related results to Theorem 1 are available in the mathematical literature. Most notably, Riesz (1911) first proved Theorem 1 for univariate and bounded $Z_t$; see also Theorem 3.4 in Krein and Nudelman (1977). Recently, in his Theorem 1, Georgiou (2005) proved Theorem 1 here under the assumption that $\mathcal{Z}$ is a closed interval in $\mathbb{R}^{d_y} \times \mathbb{R}^{d_x}$. The proof in Georgiou (2005) clearly depends on the latter assumption and is much more cumbersome than the one given here.

To circumvent the identification problem of the unconditional GMM, DL proposed a consistent estimator based on a continuum number of unconditional moment restrictions that are equivalent to the original conditional moment restrictions (1), and hence preserve the identification of $\theta_0$. This DL estimator plays a crucial role in our arguments and is introduced in the next section.

# 3 Test for Identification

This section formally introduces our null and alternative hypotheses and the test statistic for global identification of the unconditional moment restrictions implied from the conditional ones. Let $P_0$ be the true unknown (joint) probability measure of $Z_t$, and $\mathcal{M}$ be the set of all possible measures for $Z_t$ consistent with Assumptions A2 and A3 below. Define the subset of measures which are compatible with the conditional moment restrictions (1) as

$$\mathcal{P}_c = \left\{ P \in \mathcal{M} : \int h(y, \theta_0) dP_{Y|X=x} = 0 \text{ a.s at some unique } \theta_0 \in \Theta \right\},$$

where $P_{Y|X=x}$ is the conditional probability measure of $Y_t$ given $X_t = x$. On the other hand, the subset of measures where the parameter value $\theta_0$, identified from (1), is also globally identified from the unconditional moments is defined as

$$\mathcal{P}_u = \left\{ P \in \mathcal{P}_c : \int f(z, \theta) dP = 0 \text{ only at } \theta = \theta_0 \right\}.$$

Based on the above notation, our testing problem is written as

$$H_0 : P_0 \in \mathcal{P}_u, \quad H_A : P_0 \in \mathcal{P}_c \setminus \mathcal{P}_u. \tag{4}$$

Note that the correct specification of the conditional moment restriction (1) is always maintained in our testing problem.[9] Several test statistics are available in

[9]If the conditional model is misspecified but DL's population objective function is uniquely minimized, say at $\theta_*$, then it can be shown that all our theory goes through provided we replace $\theta_0$ with $\theta_*$ and use some adequate asymptotic variance estimator. Details are omitted to save space.

the literature for the correct specification hypothesis, i.e., $P_0 \in \mathcal{P}_c$, such as Bierens (1982).

We now introduce our identification test based on the following intuition. Under the null hypothesis $H_0$, it holds that $\Theta_I = \{\theta_0\}$ and the unconditional GMM estimator $\hat{\theta}_{GMM}$ is consistent for $\theta_0$ under some regularity conditions; under the alternative hypothesis $H_1$, the identified set $\Theta_I$ contains elements different from $\theta_0$ and $\hat{\theta}_{GMM}$ is typically inconsistent. On the other hand there are estimators that are consistent for $\theta_0$ under both the null and alternative hypotheses, which can provide a basis for a Hausman-type test statistic. To construct such an estimator there are at least two available possibilities: one based on unknown instruments that lead to unique identification, for instance $a(X) = E[h(Y,\theta)|X]$, and another based on the characterization of the conditional moment (1) by an infinite number of unconditional moments,[10]

$$E[h(Y_t,\theta_0) \mid X_t] = 0 \text{ a.s.} \Longleftrightarrow H(x,\theta_0) \equiv E[h(Y_t,\theta_0)1(X_t \leq x)] = 0 \text{ a.s. } x \in \mathbb{R}^{d_x},$$

see Billingsley (1995, Theorem 16.10iii). In other words, $\theta_0$ is the unique solution of the minimization problem

$$\min_{\theta \in \Theta} \int_{\mathbb{R}^{d_x}} |H(x,\theta)|^2 \, dF_X(x),$$

where $F_X$ is the cumulative distribution function of $X_t$. This suggests the following minimum-distance estimator proposed in DL

$$\hat{\theta}_C = \arg\min_{\theta \in \Theta} \int_{\mathbb{R}^{d_x}} |H_n(x,\theta)|^2 \, dF_{n,X}(x),$$

where $H_n(x,\theta) = n^{-1/2} \sum_{t=1}^{n} h(Y_t,\theta) 1(X_t \leq x)$, and $F_{n,X}$ is the empirical distribution function of $\{X_t\}_{t=1}^{n}$. For computational purposes it is better to write the

[10] This characterization generally holds for the unconditional moments in the form of $E[h(Y_t,\theta_0)w(X_t,x)] = 0$ a.s. $x \in \Pi \subseteq \mathbb{R}^q$, where $w(\cdot,x)$ is a suitable parametric family of functions; see Stinchcombe and White (1998), Bierens and Ploberger (1997) and Escanciano (2006b) for examples of $\Pi$ and $w$. To simplify the exposition, we follow DL and choose $w(X_t,x) = 1(X_t \leq x)$, but it must be stressed that all the theory that follows holds with other choices of $w(X_t,x)$ as well. For practical reasons, when $d_x$ is moderate or large (say $> 3$ for commonly used sample sizes) it is better to use other weighting functions different from $w(X_t,x) = 1(X_t \leq x)$. The reason is that in a given sample it could be the case that most of the indicators $1(X_t \leq x)$ are zero when $x$ is evaluated at the sample. Alternative weighting functions such as those proposed in Bierens (1982) and Escanciano (2006a) solve this practical deficiency.

above objective function by a quadratic form of "errors". If we denote $h(\theta) \equiv (h(Y_1,\theta),...,h(Y_n,\theta))'$ and $P \equiv (1(X_t \leq X_s)_{t,s})$, then DL's estimator can be simply computed as

$$\hat{\theta}_C = \arg\min_{\theta\in\Theta} h(\theta)'(P'P)h(\theta).$$

Under certain regularity conditions DL's estimator $\hat{\theta}_C$ is strongly consistent and asymptotically normal under both the null and alternative hypotheses in (4). Since the unconditional GMM estimator $\hat{\theta}_{GMM}$ is consistent and asymptotically normal only under the null $H_0$, this suggests to construct a Hausman-type test statistic for $H_0$ based on the contrast between $\hat{\theta}_C$ and $\hat{\theta}_{GMM}$, i.e.,

$$T_n(\hat{\theta}_{GMM}) \equiv n\left(\hat{\theta}_C - \hat{\theta}_{GMM}\right)' \Sigma_n^{-1}\left(\hat{\theta}_C - \hat{\theta}_{GMM}\right), \tag{5}$$

where $\Sigma_n$ is a consistent estimator for the asymptotic variance-covariance matrix $\Sigma$ of $\sqrt{n}(\hat{\theta}_C - \hat{\theta}_{GMM})$, and $\Sigma_n$ and $\Sigma$ are both defined in Appendix A. However, since under the alternative $H_A$ we may expect many solutions to the optimization problem (2), the test statistic $T_n(\hat{\theta}_{GMM})$ may not be uniquely defined even in finite samples. Example 1 illustrates this issue.

Example 1 (Cont.): There are two solutions to the estimating function $n^{-1}\sum_{t=1}^n a(X_t)(Y_t - \theta^2 X_t - \theta X_t^2) = 0$, namely

$$\frac{-n^{-1}\sum_{t=1}^n X_t^2 a(X_t) \pm \sqrt{\left(-n^{-1}\sum_{t=1}^n X_t^2 a(X_t)\right)^2 + 4\left(n^{-1}\sum_{t=1}^n X_t a(X_t)\right)\left(n^{-1}\sum_{t=1}^n Y_t a(X_t)\right)}}{2n^{-1}\sum_{t=1}^n X_t a(X_t)}.$$

Denote by $\hat{\theta}_{GMM,+}$ and $\hat{\theta}_{GMM,-}$ these two solutions. Associated to these two solutions we have two possible values for the test statistic (5), i.e., $T_n(\hat{\theta}_{GMM,+})$ and $T_n(\hat{\theta}_{GMM,-})$.

To solve the difficulty of defining the test statistic when there is more than one global minimum in (2), we propose a modified test statistic based on the Hausdorff distance

$$T_n \equiv \max_{\theta\in\hat{\Theta}_{GMM}} T_n(\theta) = \max_{\theta\in\hat{\Theta}_{GMM}} n\left(\hat{\theta}_C - \theta\right)' \Sigma_n^{-1}\left(\hat{\theta}_C - \theta\right),$$

where $\hat{\Theta}_{GMM}$ is a suitable estimator of the identified set $\Theta_I$. In particular cases such as Example 1, there is a natural choice for $\hat{\Theta}_{GMM}$, namely $\hat{\Theta}_{GMM} = \{\hat{\theta}_{GMM,+}, \hat{\theta}_{GMM,+}\}$. In general, we suggest to construct $\hat{\Theta}_{GMM}$ using the following algorithm; see Veall

(1990) for a related method. First we generate $m \in \mathbb{N}$ random independent initial conditions from a random variable taking values in $\Theta$ and compute the resulting $m$ minima $Q_n(\hat{\theta}_{GMM}^{(j)})$ for $j = 1, \ldots, m$, where $Q_n(\theta)$ is the GMM objective function in (2). We define $\hat{\Theta}_{GMM}$ as the set of values $\hat{\theta}_{GMM}^{(j)}$ satisfying

$$\hat{\Theta}_{GMM} = \{\hat{\theta}_{GMM}^{(j)} : Q_n(\hat{\theta}_{GMM}^{(j)}) \leq \min_{1 \leq j \leq m} Q_n(\hat{\theta}_{GMM}^{(j)}) + a_n\}, \tag{6}$$

where $a_n = 1/(n \log(n))$. For this choice of $a_n$ (or indeed any other one satisfying $a_n = o(1/n)$) any member of $\hat{\Theta}_{GMM}$ is characterized by the same pointwise asymptotic distribution under the null hypothesis $H_0$; see Theorem 5.23 in van der Vaart (2000). This is the key point of the paper because it implies that under $H_0$ the asymptotic behavior of the statistic $T_n$ can be studied using standard methods.

On the other hand under the alternative hypothesis $H_A$, no general pointwise asymptotic theory is available; $\hat{\Theta}_{GMM}$ is not a consistent estimator for $\Theta_I$ and standard methods cannot be applied (see for example Chernozhukov, Hong and Tamer (2007)). As a result we are unable to fully characterize the asymptotic power properties of our test for general alternatives. We can however show consistency of the test under a general class of unidentified models that includes weak identification alternatives. This is because in this case the asymptotic properties of the GMM estimators (and hence those of $\hat{\Theta}_{GMM}$) are well-known; see Stock and Wright (2000). The simulations of Section 4 below suggest that the finite sample power of the proposed test is high for small and moderate sample sizes.

# 4 Asymptotic Theory

## 4.1 Asymptotic Null Distribution

This section investigates the asymptotic null distribution of the test statistic $T_n$. Let $\mathcal{F}_t \equiv \sigma(Y_{t-1}, X_t, ...)$ be the $\sigma$-field generated by the information set obtained up to time $t$, and $G_t(x) \equiv E[E[h(Y_t, \theta_0)^2 \mid X_t] 1(X_t \leq x) | \mathcal{F}_{t-1}]$. We impose the following assumptions.

Assumption A1: There exists a unique $\theta_0 \in \Theta$ such that $E[h(Y_t, \theta_0)|X_t] = 0$ a.s.

Assumption A2: (i) $\{(Y_t', X_t')'\}_{t \in \mathbb{Z}}$ is a strictly stationary and ergodic process with an absolutely continuous distribution function $F_X$ for $X_t$; (ii) $E[h(Y_t, \theta_0)|\mathcal{F}_t] = E[h(Y_t, \theta_0)|X_t]$ a.s.; (iii) $h : \mathbb{R}^{d_y} \times \Theta \longrightarrow \mathbb{R}$ is continuous in $\Theta$ a.s. and is continuously differentiable in a neighborhood $\mathcal{N}$ of $\theta_0$, $E[\sup_{\theta \in \Theta} |h(Y_t, \theta)|] < \infty$, and

$E[\sup_{\theta\in\mathcal{N}} |(\partial/\partial\theta')h(Y_t,\theta)|] < \infty$; (iv) $\Theta$ is compact and $\theta_0$ belongs to the interior of $\Theta$; (v) the matrix $\Sigma_{GG'}$, given in Appendix A, is nonsingular; (vi) there exist a constant $s > 0$ and stationary sequence $\{C_t\}_{t\in\mathbb{Z}}$ with $E|C_t| < \infty$ such that $|G_t(x_1) - G_t(x_2)| \leq C_t |x_1 - x_2|^s$ for each $x_1, x_2 \in \mathbb{R}^{d_x}$.

Assumption A3: $f : \mathbb{R}^{d_y+d_x} \times \Theta \longrightarrow \mathbb{R}^r$ is continuous in $\Theta$ a.s. and is continuously differentiable in $\mathcal{N}$, $E[\sup_{\theta\in\Theta} |f(Z_t,\theta)|] < \infty$, and $E[\sup_{\theta\in\mathcal{N}} |(\partial/\partial\theta')f(Z_t,\theta)|] < \infty$. $W_n \to_P W$, with a symmetric positive definite matrix $W$. $\Xi' W \Xi$, given in Appendix A, is nonsingular.

Assumption A4: The matrix $\Sigma$ is positive definite and the estimator $\Sigma_n \to_P \Sigma$.

Assumption A1 imposes identification of $\theta_0$ in the conditional moment restrictions. Assumption A2 guarantees consistency and asymptotic normality of the DL estimator $\hat{\theta}_C$. Our Assumption A2 is in general weaker than conditions assumed in DL and related literature. For example, DL assumed $E[|h(Y_t,\theta_0)|^4 |X|^{1+\delta}] < \infty$, whereas we only require the second-order moments. Also, DL required the density of $X_t$ conditioning on $\mathcal{F}_{t-1}$ to be bounded and continuous. By Hölder's inequality, this latter assumption implies our Assumption A2(c) with $s_1 = \delta/(1+\delta)$ for some $\delta > 0$, provided $E[|h(Y_t,\theta_0)|^{2(1+\delta)}] < \infty$. Finally the smoothness condition for $h$ can be relaxed at the cost of longer proofs; see Escanciano (2009). Assumption A3 is standard in the literature of GMM. Assumption A4 involves the variance matrix $\Sigma$ used in the construction of the test. Its expression and a consistent estimator can be found in Appendix A.
Under these assumptions, we obtain the asymptotic null distribution of our global identification test statistic $T_n$.

Theorem 2: *Suppose Assumptions A1-A4 hold. Under $H_0$ , the test statistic $T_n$ with $\hat{\Theta}_{GMM}$ in (6) satisfies*

$$T_n \to_d \chi^2_p.$$

The proof of Theorem 2 can be found in Appendix B.

## 4.2 Power Properties

This section shows that the proposed test is consistent under a high-level condition that is satisfied for a general class of unidentified or weakly identified unconditional moment restriction models, including weak identified alternatives as defined by Stock and Wright (2000).

Assumption A5: For $\hat{\Theta}_{GMM}$ defined in (6), it holds $\lim_{n\to\infty} \max_{\theta\in\hat{\Theta}_{GMM}} \sqrt{n}\,|\theta-\theta_0| = \infty$ a.s.

Theorem 3: *Suppose Assumptions A1-A5 hold. Then* $T_n \to_P \infty$.

The proof of Theorem 3 is straightforward, and hence it is omitted. We now verify that Assumption A5 holds for weakly identified unconditional moment restriction models in the sense of Stock and Wright (2000). Assume that there exists a measurable function $m$ satisfying $m(\theta_0) = 0$ and

$$E\left[\frac{1}{n}\sum_{t=1}^{n} f(Z_t,\theta)\right] = \frac{m(\theta)}{\sqrt{n}},$$

for each $n \in \mathbb{N}$ and $\theta \in \Theta$. In this case the asymptotic distribution of the DL estimator $\hat{\theta}_C$ does not change under Assumptions A1 and A2, whereas a simple extension of the results in Stock and Wright (2000, Theorem 1) yields that any member of $\hat{\Theta}_{GMM}$ converges in distribution to a stochastic limit that is different from $\theta_0$ with probability one. Hence, under the conditions in Stock and Wright (2000) we conclude that Assumption A5 holds, and our test is consistent against weak identification alternatives.

For general alternatives the power of our test has to be analyzed on a case-by-case basis, given that there is no general pointwise limit distribution theory for estimation under lack of identification. For instance, in Example 1 we can prove Assumption A5 by simple arguments.

Example 1 (Cont.): Suppose that $a(X)$ is such that $E[a(X)a^*(X)] \neq 0$, so the unconditional moment restriction $E[a(X)(Y - \theta^2 X - \theta X^2)] = 0$ does not identify $\theta_0$. It can be shown that $\hat{\theta}_{GMM,+}$ and $\hat{\theta}_{GMM,-}$ are consistent and asymptotically normal for $\theta_1$ and $\theta_0$, respectively. If we define $\hat{\Theta}_{GMM} \equiv \{\hat{\theta}_{GMM,-}, \hat{\theta}_{GMM,+}\}$, then by $\theta_1 \neq \theta_0$, we obtain

$$\lim_{n\to\infty} \max_{\theta\in\hat{\Theta}_{GMM}} \sqrt{n}\,|\theta-\theta_0| = \lim_{n\to\infty} \max\left\{\sqrt{n}\left|\hat{\theta}_{GMM,-}-\theta_0\right|, \sqrt{n}\left|\hat{\theta}_{GMM,+}-\theta_0\right|\right\} = \infty \text{ a.s.}$$

Thus, our test is consistent for this example. Note that in this example, there exists a consistent and asymptotically normal estimator for $\theta_0$ even under identification failure, namely $\hat{\theta}_{GMM,-}$. Hence, this example nicely illustrates that taking the Hausdorff distance is crucial to make our test consistent.

## 5 Monte Carlo Simulations

In this section we investigate the finite-sample performance of the proposed test using some Monte Carlo experiments. First we consider Example 1 using feasible instruments and several marginal distributions for the conditioning variable $X$. This example is useful because it illustrates a simple situation in which existing tests for identification either cannot be used or could potentially lead to misleading conclusions. The model is

$$Y = \theta_0^2 X + \theta_0 X^2 + \varepsilon,$$

where $\varepsilon$ is distributed as $N(0,1)$ independently of $X$, and $\theta_0 = 5/4$. For the unconditional GMM, we consider the unconditional moment restriction using the optimal instrument $a^*(X,\theta) = 2\theta X + X^2$, i.e.,

$$E[(2\theta X + X^2)(Y - \theta^2 X - \theta X^2)] = 0. \quad (7)$$

As in DL we consider four possible distributions for $X$, namely $X \sim N(0,1)$, $X \sim N(1,3)$, $X \sim N(1,1)$, and $X \sim N(1,2)$. In the first two cases, the parameter $\theta_0$ is identified and GMM provides consistent estimates. In the third and fourth cases there are, respectively, three and two solutions to the unconditional moments so that (7) does not globally identify $\theta_0$ and GMM is inconsistent.

To approximate $\hat{\Theta}_{GMM}$ as given in (6) we generate $m = 20$ random independent initial conditions from $N(0,1)$. Table I reports the rejection probabilities (RP) for the sample sizes $n = 50$, 100, 200, and 500 at the 1, 5 and 10 percent nominal level using 1000 Monte Carlo replications.

Table I: RP for DL's Example 2

| | | Size | | | | Power | | |
|---|---|---|---|---|---|---|---|---|
| $X$ | $n$ | 1% | 5% | 10% | $X$ | 1% | 5% | 10% |
| $N(0,1)$ | 50 | 0.032 | 0.049 | 0.081 | $N(1,2)$ | 0.642 | 0.644 | 0.662 |
| | 100 | 0.016 | 0.038 | 0.077 | | 0.631 | 0.637 | 0.662 |
| | 200 | 0.013 | 0.043 | 0.087 | | 0.573 | 0.589 | 0.614 |
| | 500 | 0.006 | 0.045 | 0.105 | | 0.577 | 0.591 | 0.609 |
| $N(1,3)$ | 50 | 0.178 | 0.194 | 0.221 | $N(1,1)$ | 0.978 | 0.979 | 0.980 |
| | 100 | 0.088 | 0.105 | 0.148 | | 0.978 | 0.998 | 0.998 |
| | 200 | 0.039 | 0.071 | 0.121 | | 0.999 | 0.999 | 0.999 |
| | 500 | 0.009 | 0.048 | 0.101 | | 1.000 | 1.000 | 1.000 |

The size performance of our test is satisfactory for both DGPs in the null hypothesis, although there are some size distortions when $X \sim N(1,3)$ for $n = 50$.

Table I also shows that the power performance of our test is satisfactory. The finite sample power against the alternative with $X \sim N(1,1)$ is high already for $n = 50$. For the alternative with $X \sim N(1,2)$ the test has a satisfactory power, but it is lower than for $X \sim N(1,1)$, and it does not increase with the sample size. This result is consistent with the fact that the alternative corresponding to $X \sim N(1,2)$ is very close to the identification region $\sigma^2 < 2.0163$.

Before we consider the second example it should be noted that neither of Wright's (2003) and Wright's (2009) tests can be used in this example because the full rank condition holds regardless of the identification and because the original model is not overidentified. Also Inoue and Rossi's (2008) test could lead to a misleading conclusion as the following example suggests. Suppose that $X \sim N(1,3)$ and that a researcher applying their test chooses the GMM estimator with the optimal instrument assuming homoskedasticity and the GMM estimator with a constant instrument. The former identifies the true parameter, however the latter does not by the results of the example. Indeed unreported simulations show that in this case their test rejects the null hypothesis of identification with probability one when $n = 200$. Interestingly the proposed test would also be useful in this context because it would allow to find out which estimator is causing the rejection of Inoue and Rossi's null hypothesis.

In the second experiment, we investigate the finite-sample performance of our test in a consumption capital asset pricing model (CCAPM) with a constant relative risk aversion preferences model that has been used in much of the literature on GMM identification; see e.g. Stock and Wright (2000), Wright (2003), Inoue and Rossi (2008), and Wright (2009). The CCAPM data are generated using Tauchen and Hussey's (1991) method which involves fitting a 16-state Markov chain to consumption and dividend growth calibrated so as to approximate the first-order vector autoregression

$$\begin{bmatrix} \log\left(\frac{C_t}{C_{t-1}}\right) \\ \log\left(\frac{D_t}{D_{t-1}}\right) \end{bmatrix} = \mu + \Phi \begin{bmatrix} \log\left(\frac{C_{t-1}}{C_{t-2}}\right) \\ \log\left(\frac{D_{t-1}}{D_{t-2}}\right) \end{bmatrix} + \begin{bmatrix} u_{ct} \\ u_{dt} \end{bmatrix},$$

where $C_t$ is the consumption, $D_t$ is the dividend $\mu$ is a $2 \times 1$ vector, $\Phi$ is a $2 \times 2$ matrix of constants, and $(u_{ct}, u_{dt})' \sim N(0, \Lambda)$. Assets prices are then generated so that they satisfy the stochastic Euler equation

$$E[\delta R_{t+1} \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} - 1 \quad | \quad \mathcal{F}_t] = 0 \ a.s.,$$

where $\delta$ is the discount factor, $R_t$ is the gross stock return, and $\gamma$ is the coefficient of relative risk aversion. Following Inoue and Rossi (2008), we use the instruments

$X_t = (1, R_t, C_t/C_{t-1})'$, and consider five different combinations of the parameters $\theta = (\delta, \gamma)'$, $\mu$, $\Phi$, and $\Lambda$, as listed in Table II.

Table II: parameter values for CCAPM

| Model | $\mu$ | $\Phi$ | $\Lambda$ | $\delta$ | $\gamma$ |
|---|---|---|---|---|---|
| SI | $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} -0.5 & 0.1 \\ 0.1 & -0.5 \end{bmatrix}$ | $\begin{bmatrix} 0.01 & 0.005 \\ 0.005 & 0.01 \end{bmatrix}$ | 0.97 | 1.3 |
| PI1 | $\begin{bmatrix} 0.018 \\ 0.013 \end{bmatrix}$ | $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ | $\begin{bmatrix} 0.0012 & 0.0017 \\ 0.0017 & 0.0146 \end{bmatrix}$ | 0.97 | 1.3 |
| PI2 | $\begin{bmatrix} 0.018 \\ 0.013 \end{bmatrix}$ | $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ | $\begin{bmatrix} 0.0012 & 0.0017 \\ 0.0017 & 0.0146 \end{bmatrix}$ | 1.139 | 13.7 |
| WI1 | $\begin{bmatrix} 0.021 \\ 0.004 \end{bmatrix}$ | $\sqrt{\frac{90}{n}} \begin{bmatrix} -0.161 & 0.017 \\ 0.414 & 0.117 \end{bmatrix}$ | $\begin{bmatrix} 0.0012 & 0.0017 \\ 0.0017 & 0.0146 \end{bmatrix}$ | 0.97 | 1.3 |
| WI2 | $\begin{bmatrix} 0.021 \\ 0.004 \end{bmatrix}$ | $\sqrt{\frac{90}{n}} \begin{bmatrix} -0.161 & 0.017 \\ 0.414 & 0.117 \end{bmatrix}$ | $\begin{bmatrix} 0.0012 & 0.0017 \\ 0.0017 & 0.0146 \end{bmatrix}$ | 1.139 | 13.7 |

The first one (SI) is for the strongly identified case; the second and third ones (PI1 and PI2) are for two partially identified cases where the instruments are independent of $R_{t+1}, C_{t+1}/C_t$; the last two ones (WI1 and WI2) are for two weakly identified cases; see Inoue and Rossi (2008) for further details. Table III reports the rejection probabilities of the proposed test and those of Wright's (2003) test for his null of lack of identification.

Table III: RP of the proposed test and Wright's (2003)

| | | $\alpha = 0.01$ | | $\alpha = 0.05$ | | $\alpha = 0.10$ | |
|---|---|---|---|---|---|---|---|
| Model | $n$ | Wright | Proposed | Wright | Proposed | Wright | Proposed |
| SI | 50 | 0.660 | 0.051 | 0.748 | 0.101 | 0.795 | 0.171 |
| | 200 | 0.983 | 0.016 | 0.987 | 0.061 | 0.988 | 0.116 |
| PI1 | 50 | 0.075 | 0.186 | 0.142 | 0.224 | 0.202 | 0.284 |
| | 200 | 0.036 | 0.292 | 0.097 | 0.334 | 0.141 | 0.374 |
| PI2 | 50 | 0.702 | 0.203 | 0.751 | 0.240 | 0.778 | 0.282 |
| | 200 | 0.530 | 0.303 | 0.606 | 0.352 | 0.657 | 0.394 |
| WI1 | 50 | 0.220 | 0.538 | 0.330 | 0.729 | 0.389 | 0.810 |
| | 200 | 0.219 | 0.781 | 0.337 | 0.869 | 0.407 | 0.914 |
| WI2 | 50 | 0.220 | 0.756 | 0.511 | 0.804 | 0.533 | 0.829 |
| | 200 | 0.219 | 0.826 | 0.281 | 0.874 | 0.322 | 0.892 |

In the SI case both tests perform well. The proposed test is slightly oversized for $n = 50$ but its accuracy improves as the sample size increases. In the two PI cases the proposed test has some power for $n = 50$, but it increases with the sample size. On the other hand Wright's (2003) test performs rather differently for the two cases: it is slightly oversized for PI1 but it is very oversized for the second case. Finally the proposed test has good power in both WI cases even for $n = 50$.

# 6 Conclusions

There are growing evidences that many unconditional moment restriction models used in empirical economics are potentially not (or weakly) identified. In this paper we propose a Hausman-type test statistic that can be used to test the null hypothesis of identification for the unconditional moment restrictions. Our test has a number of appealing properties that we summarize as follows: it is computationally simple, does not require any choice of tuning parameters and has a simple chi-squared limiting distribution. Furthermore it can be applied to dependent data and to just identified models as well. The test is consistent against weak identification alternatives, although we are not able to fully characterize its asymptotic properties under a general alternative of no identification. Monte Carlo simulations suggest that the proposed test has good finite sample size and power properties under both the alternative of lack of identification and weak identification. These results illustrate the general applicability and usefulness of the proposed test.

# 7 Appendix

## 7.1 Appendix A: Asymptotic Variance and Its Estimator

Define $h_\theta(Y_t,\theta) \equiv (\partial/\partial\theta')h(Y_t,\theta)$, $f_\theta(Z_t,\theta) \equiv (\partial/\partial\theta')f(Z_t,\theta)$, $\Xi \equiv E[f_\theta(Z_t,\theta_0)]$, $S \equiv E[f(Z_t,\theta_0)f(Z_t,\theta_0)']$, $G(x,\theta) \equiv E[h_\theta(Y_t,\theta)1(X_t \leq x)]$, $J(x) := E[G(X_t,\theta)1(x \leq X_t)]$, and $\Sigma_{GG'} \equiv \int G(x,\theta_0)G(x,\theta_0)'dF_X(x)$. Let $I_p$ be the $p \times p$ identity matrix and $\tilde{I} \equiv [I, -I]$ be a $p \times 2p$ matrix. Then, the variance $\Sigma$ is defined as

$$\Sigma \equiv \tilde{I}VAV'\tilde{I}',$$

where $V \equiv \begin{bmatrix} \Sigma_{GG'}^{-1} & 0 \\ 0 & (\Xi'W\Xi)^{-1} \end{bmatrix}$ and

$$A \equiv \begin{bmatrix} E[h^2(Y_t,\theta_0)\, J(X_t)J(X_t)'] & E[J(X_t)h(Y_t,\theta_0)f(Z_t,\theta_0)']W\Xi \\ (E[J(X_t)h(Y_t,\theta_0)f(Z_t,\theta_0)']W\Xi)' & \Xi'WSW\Xi \end{bmatrix}.$$

Let $\Xi_n \equiv n^{-1}\sum_{t=1}^n f_\theta(Z_t,\hat{\theta}_C)$, $S_n \equiv n^{-1}\sum_{t=1}^n f(Z_t,\hat{\theta}_C)f(Z_t,\hat{\theta}_C)'$, $G_n(x,\hat{\theta}_C) \equiv n^{-1}\sum_{t=1}^n h_\theta(Y_t,\hat{\theta}_C)1(X_t \leq x)$, $J_n(x) \equiv n^{-1}\sum_{t=1}^n G_n(X_t,\hat{\theta}_C)1(x \leq X_t)$, and $\Sigma_{GG',n} = n^{-1}\sum_{t=1}^n G_n(X_t,\hat{\theta}_C)G_n(X_t,\hat{\theta}_C)'$. A consistent estimator of $\Sigma$ is defined as

$$\Sigma_n \equiv \tilde{I}V_n A_n V_n' \tilde{I}', \tag{8}$$

where $V_n$ and $A_n$ replace population expectations with sample counterparts and make use of the previous quantities in a routine fashion. The consistency of $\Sigma_n$ follows easily under Assumptions A1-A3.

## 7.2 Appendix B: Proof of Theorem 2

By standard arguments, see e.g. Theorem 1 in DL, Assumptions A1 and A2 imply that $\hat{\theta}_C \to_P \theta_0$, and likewise Assumption A3 implies $\hat{\theta}_{GMM} \to_P \theta_0$ under $H_0$. By the definition of $\hat{\Theta}_{GMM}$, we can apply Theorem 5.23 of van der Vaart (2000) to $\hat{\theta}_{GMM}$, and hence the first-order conditions of the estimators imply that for sufficiently large $n$,

$$\begin{bmatrix} \sqrt{n}(\hat{\theta}_C - \theta_0) \\ \sqrt{n}(\hat{\theta}_{GMM} - \theta_0) \end{bmatrix} = -\begin{bmatrix} \Sigma_{1n}^{-1} & 0 \\ 0 & \Sigma_{2n}^{-1} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{n}}\sum\limits_{t=1}^n G_n(X_t,\hat{\theta}_C)H_n(X_t,\theta_0) \\ \tilde{\Xi}_n' W_n \frac{1}{\sqrt{n}}\sum\limits_{t=1}^n f(Z_t,\theta_0) \end{bmatrix},$$

where $\Sigma_{1n} \equiv \frac{1}{\sqrt{n}}\sum\limits_{t=1}^n G_n(X_t,\hat{\theta}_C)G_n(X_t,\bar{\theta}_c)$, $\Sigma_{2n} \equiv \tilde{\Xi}_n' W_n \bar{\Xi}_n$, $\tilde{\Xi}_n \equiv \frac{1}{n}\sum\limits_{t=1}^n f_\theta(Z_t,\hat{\theta}_{GMM})$, $\bar{\Xi}_n \equiv \frac{1}{n}\sum\limits_{t=1}^n f_\theta(Z_t,\bar{\theta}_{GMM})$, and $\bar{\theta}_c$ and $\bar{\theta}_{GMM}$ are mean values. The uniform law of large numbers of Ranga Rao (1962) and standard arguments imply that $\tilde{\Xi}_n, \bar{\Xi}_n \to_P \Xi$,

$$G_n(x,\theta) \to_{a.s.} G(x,\theta), \qquad \text{uniformly in } (x,\theta),$$

and $V_n \to_P V$.

By Theorem A1 in Delgado and Escanciano (2007), see also Theorem 1 in Escanciano (2007), it follows that $H_n(x,\theta_0)$ is asymptotically stochastic equicontinuous in $x$ with respect to the pseudo-metric $d(x_1,x_2) \equiv |F_X(x_1) - F_X(x_2)|$. Hence, from Lemma 3.1 in Chang (1990) we conclude that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^n G_n(X_t,\hat{\theta}_C)H_n(X_t,\theta_0) = \frac{1}{\sqrt{n}}\sum_{t=1}^n h(Y_t,\theta_0)J(X_t) + o_P(1).$$

Hence, the theorem follows from a straightforward application of a multivariate central limit theorem of strictly stationary and ergodic martingales, see e.g. Billingsley (1961), applied to the vector

$$\begin{bmatrix} \frac{1}{\sqrt{n}}\sum_{t=1}^{n} h(Y_t,\theta_0)J(X_t) \\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n} f(Z_t,\theta_0) \end{bmatrix}.$$

The conclusion follows by the continuous mapping theorem. ■

# References


[1] Arellano, M., Hansen, L. P. and Sentana, E. (2009): "Underidentification?" unpublished manuscript.

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

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

[4] Billingsley, P. (1961): "The Lindeberg-Levy Theorem for Martingales," *Proceedings of the American Mathematical Society,* 12, 788-792.

[5] Billingsley, P. (1995): *Probability and Measure.* Wiley and Sons. New York.

[6] Carrasco, M. and Florens, J. P. (2000): "Generalization of GMM to a Continuum of Moment Conditions," *Econometric Theory,* 16, 797-834.

[7] Canova, F. and Sala, L. (2006): "Back to Square One: Identification Issues in DSGE models," unpublished manuscript.

[8] Chamberlain, G. (1987): "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions," *Journal of Econometrics*, 34, 305-334.

[9] Chang, N. M. (1990): "Weak Convergence of a Self-Consistent Estimator of a Survival Function with Doubly Censored Data," *Annals of Statistics*, 18**,** 391-404.

[10] Chernozhukov, V. Hong, H. and Tamer, E. (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243-1284.

[11] Delgado, M. A. and Escanciano, J. C. (2007): "Nonparametric Tests for Conditional Symmetry in Dynamic Models," *Journal of Econometrics*, 141, 652-682.

[12] Dominguez, M. A. and Lobato, I. N. (2004): "Consistent Estimation of Models Defined by Conditional Moment Restrictions," *Econometrica*, 72, 1601-1615.

[13] Donald, S. G., Imbens, G. W. and Newey, W. K. (2003) "Empirical Likelihood Estimation and Consistent Tests with Conditional Moment Restrictions," *Journal of Econometrics*, 117, 55-93.

[14] Escanciano, J. C. (2006a): "Consistent Diagnostic Test for Regression Models Using Projections," *Econometric Theory*, 22, 1030-1051.

[15] Escanciano, J. C. (2006b): "Goodness-of-fit Tests for Linear and Non-linear Time Series Models," *Journal of the American Statistical Association,* 101, 531-541.

[16] Escanciano, J. C. (2007): "Weak Convergence of Non-stationary Multivariate Marked Processes with Applications to Martingale Testing," *Journal of Multivariate Analysis,* 98, 1321-1336.

[17] Escanciano, J. C. (2009): "Simple and Powerful Bootstrap Tests for Conditional Moment Restrictions," unpublished manuscript.

[18] Georgiou, T. T. (2005): "Solution of The General Moment Problem Via a One-Parameter Imbedding," *IEEE Transactions on Automatic Control*, 50, 811-826.

[19] Gragg, J. G. and Donald, S. G. (1993): "Testing Identifiability and Specification in Instrumental Variables Models," *Econometric Theory,* 9, 222-240.

[20] Hall, A. R. (2005): *Generalized Method of Moments*, Oxford University Press.

[21] Hahn, J. and Hausman, J. (2002): "A New Specification Test for the Validity of Instrumental Variables," *Econometrica,* 70,163–89.

[22] Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moment Estimators," *Econometrica*, 50, 1029-1054.

[23] Hansen, L. P. and Singleton, K. J. (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269-1286.

[24] Hansen, L.P., Heaton, J. and Yaron, A. (1996): "Finite Sample Properties of some Alternative GMM Estimators," *Journal of Business and Economic Statistics,* 14, 262-280.

[25] Inoue, A. and Rossi, B. (2008): "Testing for Identifcation in Possibly Nonlinear Models," unpublished manuscript.

[26] Kitamura, Y., Tripathi, G. and Ahn, H. (2004): "Empirical Likelihood-Based Inference for Conditional Moment Restriction Models," *Econometrica*, 72, 1667-1714.

[27] Koopmans, T. C. and Hood, W. C. (1953): "The Estimation of Simultaneous Linear Economic Relationships," in Hood, W. C. and Koopmans, T. C. (eds.): *Studies in Econometric Methods*, Cowles Commission Monograph No. 14, Chapter 6, Wiley.

[28] Krein, M. G. and Nudelman, A. A. (1977): "The Markov Moment Problem and Extremal Problems. Ideas and problems of P. L. Chebyshev and A. A. Markov and their further development," Translated from the Russian by D. Louvish. *Translations of Mathematical Monographs*, Vol. 50. American Mathematical Society, Providence, R.I.

[29] Nason, J. M. and Smith, G. W. (2008): "Identifying the New Keynesian Curve," *Journal of Applied Econometrics*, 23, 525-551.

[30] Newey, W. K. and Smith, R. J. (2004): "Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators," *Econometrica*, 72, 219-255.

[31] Ranga Rao, R. (1962): "Relations Between Weak and Uniform Convergence of Measures with Applications," *Annals of Mathematical Statistics*, 33, 659–680.

[32] Riesz, F. (1911): "Sur certains systemes singuliers d'equations integrales," *Ann. Sci. Ecole Norm. Sup.* (3), 15, 69-81.

[33] Sargan, J. D. (1958): "The Estimation of Economic Relationships Using Instrumental Variables," *Econometrica*, 26, 393-415.

[34] Sargan, J. D. (1983): "Identification and Lack of Identification," *Econometrica*, 51, 1605-1634.

[35] Staiger, D. and Stock, J. H. (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica,* 65, 557-586.

[36] Stinchcombe, M. and White, H. (1998): "Consistent Specification Testing with Nuisance Parameters Present Only under the Alternative," *Econometric Theory,* 14, 295-325.

[37] Stock, J. H. and Wright, J. H. (2000): "GMM with Weak Identification," *Econometrica,* 68, 1055-1096.

[38] Stock, J. H., Wright, J. H. and Yogo, M. (2005): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 20, 518-529.

[39] Stock, J. H. and Yogo, M. (2005): "Testing for Weak Instruments in Linear IV Regression." In: *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Cambridge University Press.

[40] Tauchen, G. and Hussey, R. (1991): "Quadrature-based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica,* 59, 371-396.

[41] Veall, M. R. (1990): "Testing for a Global Maximum in an Econometric Context," *Econometrica*, 58, 1459-1465.

[42] van der Vaart, A. W. (2000): *Asymptotic Statistics.* Cambridge University Press.

[43] Yogo, M. (2004): "Estimating the Elasticity of of Intertemporal Substitution when Instruments are Weak," *Review of Economics and Statistics*, 86, 797-810.

[44] Wright, J. H. (2003): "Detecting Lack of Identification in GMM," *Econometric Theory*, 19, 322-330.

[45] Wright, J. H. (2009): "Testing the Adequacy of Conventional Asymptotics in GMM," unpublished manuscript.