Improving Consistent Moment Selection Procedures for Generalized Method of Moments Estimation.

Jean-Bernard Chatelain*

November 3, 2006

Abstract

This paper proposes consistent moment selection procedures for generalized method of moments estimation based on the J test of over-identifying restrictions (Hansen [1982]) and on the Eichenbaum, Hansen and Singleton [1988] test of the validity of a subset of moment conditions.

JEL classification: **C10 C12**

Keywords: Generalized method of moments, test of over-identifying restrictions, test of subset of over-identifying restrictions, Consistent Moment Selection.

*CEPREMAP, PSE and Economix, University of Paris 10, c/o 10 rue de Sully, F-94220 Charenton le Pont. E-mail: jean-bernard.chatelain@u-paris10.fr

1. Introduction

Hall [2005, chapter 7] presents a detailed overview of moment selection procedures when using the generalized method of moments (GMM) estimation. In particular, Andrews [1999] proposed three methods based on the overidentifying restrictions tests for determining the valid moment conditions from a candidate set of moment restrictions. This paper proposes a modification to one of these methods, the upward testing strategy. The proposal is to use not only the overidentifying restriction test but also a statistic for testing the validity of a subset of moment conditions proposed by Eichenbaum, Hansen and Singleton (EHS) [1988]. The particular version of the EHS statistic is the difference between the overidentifying restrictions using the k moments and the overidentifying restrictions test using the $k-1$ moments that are believed to be correct on the basis of the outcomes of the tests up to this point in the upward testing procedure. The argument for this modification is that, by taking advantage of the information on the likely source of the misspecification on the basis of the outcomes of the tests up to this point in the upward testing procedure, the EHS test has *greater local power* than the overidentifying restrictions test (Hall [2005, Chapter 5]), and that a sequence of tests with greater local power improves the moment selection procedure with respect to a sequence of tests with lower local power.

The paper proceeds as follows. First, it is shown that this modification does not invalidate the consistency of the method (the differing local power properties are irrelevant for the proof). By construction, the method in this paper will select

fewer moment conditions than Andrews' [1999] method, when their selected moments conditions differ. Second, a large sample counter example illustrates how the greater local power of the EHS test can improve the reliability of the moment selection.

2. Upward Testing Procedures Based on J and EHS Tests

Moment selection is represented by a moment selection r–vector (r is a maximal finite number of correct or incorrect moment restrictions) denoted c such that, if the jth element of c (denoted c_j) is a one (resp. a zero), then the jth moment condition is included (resp. excluded). The number of selected moments is given by $|c| = \sum_{i=1}^{r} c_i = k \leq r$. Hall [2005, chapter 5] analyses into details the differences between J and EHS tests for a given moment selection vector c.

For the J test of overidentifying restrictions (Hansen [1982]), the null hypothesis, denoted H_0^O, is that the $k-p$ overidentifying restrictions on the moment condition, where p is the number of parameters to estimate and where k is the total number of moment conditions with $k > p$, are valid for the vector of estimated parameters, as it should be in the "true" model. The product of the minimized value of the GMM objective function and of the number of observations n is the statistic $J_n(c) = n \inf_{\theta \in \Theta} G_{nc}(\theta)' W_n(c) G_{nc}(\theta)$ where $W_n(c)$ is the $|c| \times |c|$ weight matrix employed with the random vector of moment conditions $G_{nc}(\theta)$ for a p-vector of parameters θ in the set $\Theta \subset \mathbb{R}^p$: it has an asymptotic chi-square distribution with $k-p$ degrees of freedom under the null hypothesis. The overidentifying restrictions are rejected if this

χ^2 exceeds a critical value $\gamma_{J,n,k} = \chi^2_{k-p}(\alpha_{J,n})$, where $\chi^2_{k-p}(\alpha_{J,n})$ denotes the $1-\alpha_{J,n}$ quantile of a chi-squared distribution with $k-p$ degrees of freedom.

Eichenbaum, Hansen and Singleton [1988] tests subsets of Δk orthogonality conditions. The null hypothesis, denoted H_0^S, is that the $k-p$ overidentifying restrictions on the moment condition are valid. The alternative hypothesis, denoted H_A^S, is that $k-p$ overidentifying restrictions on the moment conditions are *not* valid *and* that $k-\Delta k-p$ overidentifying restrictions are valid. The difference between the $J_n(k-p)$ statistic of the model estimated with k moment conditions and the $J_n(k-\Delta k-p)$ statistic of the model estimated with $k-\Delta k$ moment conditions has an asymptotic chi-square distribution with Δk degrees of freedom under the null hypothesis. The subset of Δk overidentifying restrictions is rejected if this χ^2 exceeds a specified value $\gamma_{\Delta J,n,k} = \chi^2_{\Delta k}(\alpha_{\Delta J,n})$, where $\chi^2_{\Delta k}(\alpha_{\Delta J,n})$ denotes the $1-\alpha_{\Delta J,n}$ quantile of a chi-squared distribution with Δk degrees of freedom. Eichenbaum, Hansen and Singleton [1988] and Hall [2005, p.156]) demonstrates that the EHS test is a consistent test of H_0^S against H_A^S and that the EHS test has greater local power than the J test.

Andrews [1999] specifies moment selection criteria based on the J_n statistic and on a "bonus term" that rewards moment selection vectors c that utilize more overidentifying restrictions. For example, the GMM-BIC moment selection criterion equals $J_n(|c|-p,c)-(|c|-p)\ln(n)$. The moment selection criterion estimator $\widehat{c}_{MSC}$ is the moment selection vector that minimizes the moment selection criterion over the parameter space for the moment selection vectors $\mathcal{C}$, a smaller set than the set of moment

selection vectors C that should incorporate the information that certain blocks of moment conditions are either correct or incorrect block by block rather than moment conditions by moment conditions. Under regularity conditions for the J_n statistic and for the bonus term, $\widehat{c}_{MSC}$ is consistent.

For example, Andrews [1999] considers upward testing procedures based on the statistic $J_n(c)$ and critical values $\{\gamma_{J,n,k} : k = 1, ..., r\}$. Starting with moment selection vectors $c \in \mathcal{C}$ which have the smallest positive values of $|c|$, we carry out tests with progressively larger $|c|$ until we find that all tests with the same value of $|c|$ reject the null hypothesis that the moment conditions considered are all valid. Define $\widehat{k}_{J,UT}$ to be the largest integer (number of moments) in $\mathcal{K} = \{|c| : c \in \mathcal{C}\}$ for which $\forall k \in \mathcal{K}$ with $k \leq \widehat{k}_{J,UT}$, $\min_{c \in \mathcal{C}, |c|=k} J_n(k-p, c) \leq \gamma_{J,n,k}$. Define the upward testing selection vector estimator $\widehat{c}_{J,UT}$ of c^0 to be any vector in $\mathcal{C}$ for which $|\widehat{c}_{J,UT}| = \widehat{k}_{J,UT}$ and $J_n(\widehat{c}_{J,UT}) = \min_{c \in \mathcal{C}, |c|=\widehat{k}_{J,UT}} J_n\left(\widehat{k}_{J,UT} - p, c\right)$.

Consider "EHS upward testing moment selection procedures" based on the $J_n(c)$ test *and* on the EHS test and on critical values $\{\gamma_{J,\Delta J,n,k,\Delta k} : k = 1, ..., r\}$. Starting with the moment selection vector $c \in \mathcal{C}$ which have the smallest positive values of $|c|$ and which *minimizes* the moment selection criterion, we carry out tests with progressively larger $|c|$ until we find that all tests with the same value of $|c|$ reject the null hypothesis that the moment conditions considered are valid using both J and EHS tests, while keeping the moment selection vector *minimizing* the moment selection criterion at each step $|c|$ and not only at the final step as in upward testing

procedures based on the J test. *When a priori information is not available, this minimizes the risk of potential misspecification on the subset of k moment conditions which would violate the null and alternative hypothesis of the EHS test for the next step with $k+1$ moment conditions.* Define $\widehat{k}_{J,\Delta J,UT}$ to be the largest integer (number of moments) in $\mathcal{K}$ for which $\forall k \in \mathcal{K}$ with $k \leq \widehat{k}_{J,\Delta J,UT}$: $\min_{c\in\mathcal{C},|c|=k} J_n(k-p,c) \leq \gamma_{J,n,k}$ and $\min_{c\in\mathcal{C},|c|=k}\left(J_n(k-p,c) - \min_{c\in\mathcal{C},|c|=k-1} J_n(k-1-p,c)\right) \leq \gamma_{\Delta J,n,\Delta k}$ or:

$$\min_{c\in\mathcal{C},|c|=k} J_n(k-p,c) \leq \underbrace{\min\left(\gamma_{J,n,k}\ ;\ \gamma_{\Delta J,n,\Delta k} + \min_{c\in\mathcal{C},|c|=k-1} J_n(k-1-p,c)\right)}_{=\gamma_{J,\Delta J,n,k,\Delta k}} \leq \gamma_{J,n,k}\ . \tag{2.1}$$

The critical values $\gamma_{J,\Delta J,n,k,\Delta k}$ are recursively defined according to the above equation. Define the EHS upward testing selection vector estimator $\widehat{c}_{J,\Delta J,UT}$ of c^0 to be any vector in $\mathcal{C}$ for which $|\widehat{c}_{J,\Delta J,UT}| = \widehat{k}_{J,\Delta J,UT}$ and $J_n(\widehat{c}_{J,\Delta J,UT}) = \min_{c\in\mathcal{C},|c|=\widehat{k}_{J,\Delta J,UT}} J_n\left(\widehat{k}_{J,\Delta J,UT}-p,c\right)$.

Upward testing procedures based on the J test only and upward testing procedures based on both the J test and the EHS test differ only with respect to more restrictive critical values for EHS upward testing procedures: $\gamma_{J,\Delta J,n,k,\Delta k} \leq \gamma_{J,n,k}$. Consistency of $\widehat{c}_{J,UT}$ is established by Andrews' [1999] theorem 3. The only assumption of this theorem related to critical values is **Assumption** T : $\gamma_{J,n,k} \to +\infty$ and $\gamma_{J,n,k} = o(n)$, $\forall k = p+1, ..., r$.[1] Then assumption T holds with the significance level α_n satisfying

[1]Andrews [1999] theorem 1 requires two other assumptions labeled 1 (standard condition used to obtain consistency of GMM estimators) and UT (the parameter space for the moment selection vectors $\mathcal{C}$ can always be chosen such as assumption UT holds).

$\alpha_n \to 0$ and $\ln(\alpha_n) = o(n)$ (see Theorem 5.8 of Pötscher [1983] and the discussion in Hall [2005, p.259]). Theorem 1 establishes that the consistency of $\widehat{c}_{J,UT}$ implies the consistency of $\widehat{c}_{J,\Delta J,UT}$:

Theorem 1. *If assumption T holds for critical values $\gamma_{J,n,k}$, then assumption T holds for critical values $\gamma_{J,\Delta J,n,k,\Delta k}$.*

Proof: The critical values $\gamma_{J,n,k}$ are assumed to satisfy assumption T: $\gamma_{J,n,k} \to +\infty$ and $\gamma_{J,n,k} = o(n)$, $\forall k = p+1, ..., r$. The critical values for the J_n test are taken as: $\gamma_{J,n,k} = \chi^2_{k-p}(\alpha_{J,n})$. As $\gamma_{J,\Delta J,n,k,\Delta k} \leq \gamma_{J,n,k}$, $\forall k = p+1, ..., r$, $\forall n \in \mathbb{N}$, then $\gamma_{J,\Delta J,n,k,\Delta k} = o(n)$, $\forall k = p+1, ..., r$. The critical values for the ΔJ_n test are taken as $\gamma_{\Delta J,n,\Delta k} = \chi^2_{\Delta k=1}(\alpha_{\Delta J,n}) > 0$. The $J_n(k-1-p, c)$ test statistic being the number of observation n times a statistic converging to a finite positive number (Andrews assumption 1) implies that $\gamma_{J,\Delta J,n,k,,\Delta k} \to +\infty$ according to equation (2.1). Then assumption T holds for critical values $\gamma_{J,\Delta J,n,k,\Delta k}$. Q.E.D.

The following large sample counter-example illustrates that the greater local power of the EHS test can improve the reliability of the moment selection. An autoregressive distributed lag model of capital demand is estimated using Arrelano and Bond GMM dynamic panel program DPD using a large sample which consists of balance sheet data for $N = 4025$ French firms estimated over 1993-1996 ($T = 4$ and $NT = 16100$) (see a detailed description of the model, data and results in Chatelain et Al. [2003]):

$$\begin{aligned}\Delta k_{it} &= (\gamma_1 - 1)\,\Delta k_{it} + \beta_0 \Delta s_{it} + (\beta_0 + \beta_1)\,\Delta s_{i,t-1} + (\gamma_2 + \gamma_1 - 1)\,(k_{i,t-2} - s_{i,t-2}) + \\ & (\beta_0 + \beta_1 + \beta_2 + \gamma_2 + \gamma_1 - 1)\,s_{i,t-2} - \sigma_0 c_{it} - \sigma_1 c_{i,t-1} - \sigma_2 c_{i,t-2} + \\ & \theta_0 \Pi_{it} + \theta_1 \Pi_{i,t-1} + \theta_2 \Pi_{i,t-2} + f_i + f_t + \varepsilon_{it} \qquad (2.2)\end{aligned}$$

where k is the logarithm of capital stock, s is the logarithm of sales, c is the logarithm of the user cost, Π is cash flow divided by capital, f_i is for firms fixed effects, ε is a disturbance, i represents a firm, t a year, γ_j, β_j, σ_j, θ_j are parameters to estimate (for $j = 0, 1, 2$). The number of estimated parameters including the 4 year dummies f_t is $p = 15$. The approximation of the growth of capital $\Delta k_{it} = \frac{I_{it}}{K_{i,t-1}} - \delta$ introduces the depreciation rate δ assumed to be constant and the investment ratio.

In Arrelano and Bond [1991] method, when the above equation is estimated in *first differences*, the lags of levels of explanatory variables could be valid instruments and the number of instruments is $k = T \cdot k_t$ where T is the number of periods and k_t is the number of instruments for a given year. Available candidates are lags of order 2, 3 and 4 of the variables $y_t, c_t, \Pi_t, k_t, I_t/K_{t-1}$. A block of moment conditions includes four moment conditions for a given lagged variable and for each year of estimation. In the first step, the blocks of moments conditions related to $y_{t-4}, c_{t-4}, \Pi_{t-4}$ and k_{t-4} and year dummies minimizes the J_n statistic at the low value 2.75 with a corresponding high p-value of $73,8\%$: they are kept as the initial set of moment conditions. The

second step performs J and EHS tests for the validity of blocks of moment conditions related to each of the remaining 11 instruments. The block of moment conditions related to c_{t-2} passes both tests and minimizes the J_n statistic: it is added to the initial set of valid instrument. In the following steps, the blocks of moment conditions are accepted in this order: I_{t-4}/K_{t-5}, c_{t-3}, s_{t-2}, $\Pi_{t\ -3}$, s_{t-3}. Each of the five remaining instruments $\Pi_{t\ -2}$, I_{t-2}/K_{t-3}, k_{t-3}, I_{t-3}/K_{t-4}, k_{t-2} are rejected by the EHS test at the $\alpha_{\Delta J} = 6\%$ level but none of them is rejected by the J test (this is explained by *differing local power* properties of the two tests). These results suggest that the correlation between the residuals ε_{it} and $\Pi_{i,t\ -2}$ is higher than the correlation between ε_{it} and $\Pi_{i,t\ -3}$ using this particular data set. Therefore, the EHS upward testing procedures stops at $k_{\Delta J,UT} = 44$ for a p-value of the J test equal to $79,1\%$, whereas upward testing procedures based on the J test stops at $k_{J,UT} = 64$, including all blocks of moment conditions, for a p-value of the J test of 59.1%. If one sets α_J up to 70% instead of the usual 5%, then the two upward testing procedures provide the same results ($k_{J,UT} = k_{\Delta J,UT} = 44$). The two sets of moments conditions yield J_n statistics such that the test of overidentifying restrictions does not reject the null hypothesis that all moment conditions are correct for standard critical values, whereas they yield parameter estimates for the user cost $c_{i,t-1}$ and for cash flow $\Pi_{i,t-1}$ that differ noticeably from each other: they are not significantly different from zero with the upward testing procedure based on the J test but significantly different from zero with the EHS upward testing procedure. **Insert table 1.**

3. Conclusion

Consistent moment selection procedures for GMM estimation based on the J test and the EHS test differ from the larger set of consistent moment selection procedures based on the J test only with respect to more restrictive critical values $\gamma_{J,\Delta J,n,k,\Delta k}$ for the J test, related to the fact that the moment selection vector which minimizes the J_n statistic is kept at each step of k moment restrictions if it does not reject the null hypotheses of the J test and of the EHS test. Due to their greater local power, EHS upward testing procedures may select a smaller set of moment conditions than upward testing procedures based on the J test, that yield parameter estimates that differ noticeably from each other.

References

[1] Andrews D. [1999]. "Consistent Moment Selection Procedures for Generalized Method of Moments Estimation". *Econometrica* 67(3). pp.543-564.

[2] Arellano M. and Bond S. [1991]. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies.* (58). 277-297.

[3] Chatelain J.B., Ehrmann M., Generale A., Martínez-Pagés J., Vermeulen P., Worms A. [2003]. "Monetary policy transmission in the euro area - Evidence from

micro data on Banks and Firms", *Journal of the European Economic Association*, 1(3), pp. 731-742.

[4] Eichenbaum M.S., Hansen L.P., and Singleton K.J. [1988]. "A Time Series Analysis of Representative Agents Models of Consumption and Leisure Choice Under Uncertainty". *Quarterly Journal of Economics.* 103. pp.51-78.

[5] Hall A.R. [2005]. *Generalized Method of Moments,* Oxford University Press.

[6] Hansen L.P. [1982]. Large sample properties of generalized method of moments estimators. *Econometrica.* 50. pp.1029-1054.

[7] Pötscher B.M. [1983]. "Order Estimation in ARMA-models by Lagrangian Multiplier Tests". *Annals of Statistics,* 11, pp.872-885.

Table 1: GMM estimators of the growth of capital.

Variables	Regression 1 EHS and J tests	Student 1	Regression 2 J tests	Student 2
$I_{i,t-1}/K_{i,t-2}$	-0.59	(-4.15)	-0.073	(-3.43)
Δs_{it}	-0.026	(-0.72)	-0.046	(-1.39)
$\Delta s_{i,t-1}$	-0.014	(-0.37)	-0.014	(-0.47)
$k_{i,t-2} - s_{i,t-2}$	-0.28	(-5.24)	-0.162	(-6.86)
$s_{i,t-2}$	-0.23	(-4.26)	-0.163	(-6.20)
c_{it}	0.022	(0.64)	0.011	(0.34)
$c_{i,t-1}$	-0.044	(-2.74)	-0.006	(-0.43)
$c_{i,t-2}$	0.00009	(0.01)	0.001	(0.22)
Π_{it}	-0.017	(-0.26)	0.014	(0.27)
$\Pi_{i,t-1}$	0.11	(2.45)	0.027	(0.89)
$\Pi_{i,t-2}$	-0.036	(-1.28)	0.002	(0.24)
J_n statistic p-value	22.54 79.1%	$k-p$ $=29$	46.11 59.1%	$k-p$ $=49$