E C O N O M I C S B U L L E T I N

The asymptotic global power comparisons of the GMM overidentifying restrictions tests

Sheng-Kai Chang
Wayne State University

Abstract

In this paper, the asymptotic power comparisons of two versions of GMM overidentifying restrictions tests are conducted globally through the concept of approximate slopes. It is found that the GMM overidentifying restrictions test with the consistent mean deviation variance-covariance matrix estimator is more powerful than the test with the conventional non-mean deviation one. The results shed new light on the findings of Chang (2005) and Hall (2000).

I would like to thank Professor Bruce Hansen and Professor Yuichi Kitamura for their valuable comments and helpful discussions. All errors which remain are mine.
Citation: Chang, Sheng-Kai, (2007) "The asymptotic global power comparisons of the GMM overidentifying restrictions tests." *Economics Bulletin,* Vol. 3, No. 44 pp. 1-6
Submitted: July 19, 2007. **Accepted:** September 18, 2007.
URL: http://economicsbulletin.vanderbilt.edu/2007/volume3/EB-07C10003A.pdf

1 Introduction

This paper investigates the power properties of the Generalized Method of Moments (GMM) overidentifying restrictions test using the concept of approximate slopes. Under the martingale difference assumption, the power properties of the overidentifying restrictions test is investigated in Chang (2005) through the concept of approximate slopes. However, as many empirical applications indicated, it is important to allow the weakly dependent structure in GMM estimation. For example, Hansen and Singleton (1983) found that the moment conditions which contain stock return and difference in the logarithms of consumption are serially correlated when the monthly data set is used.

Without the restrictions of the martingale difference assumption, Hall (2000) also studies the power of the overidentifying restrictions test and it is demonstrated that the overidentifying restrictions test with a consistent mean deviation variance-covariance matrix estimator is more powerful than that with a traditional non-mean deviation one if the model is misspecified.

In this paper, the same results of Hall (2000) are found but with less limitations imposed on the estimators.[1] Most importantly, the asymptotic power comparisons of the GMM overidentifying restrictions tests are conducted globally through the concept of approximate slopes.[2] In other words, by performing the asymptotic power comparisons globally, it is shown that the GMM overidentifying restrictions test with the mean deviation variance-covariance matrix estimator is more powerful than the test using the conventional non-mean deviation one under suitable regularity conditions, when the model is misspecified.

The paper is organized as follows. Section 2 provides the basic framework for two versions of the GMM overidentifying restrictions test. The main theorem of the paper is presented in Section 3. Finally, the conclusion is stated in Section 4.

2 The Model

The two-step GMM estimator in the correctly specified model is considered first. Let x_t be a set of observed variables, Θ be a parameter space, and θ_0 be the $p \times 1$ unknown parameter vectors. Thus, the $q \times 1$ population moment conditions are assumed satisfied. That is,

$$E[f(x_t, \theta_0)] = 0 \tag{1}$$

[1] For example, the Assumption 5 (iii) in Hall (2000) is not required in this paper for global power comparisons of the GMM overidentifying restrictions tests.

[2] See Geweke (1981) for more details.

Since the weighting matrix of GMM estimation plays an important role only when $q > p$, I assume that $q > p$ throughout the entirety of the paper.

First, the condition imposed on x_t and $f(x_t, \theta)$ is

Assumption 1 *$\{x_t \subseteq \Re^s, t = 1, 2, \cdots\}$ is a sequence of strictly stationary and ergodic random vectors. In addition, $f : x_t \times \Theta \longrightarrow \Re^q$, where Θ is a compact set, $f(., \theta)$ is measurable for each $\theta \in \Theta$ and $f(x_t, .)$ is continuous on Θ for all x_t.*

Let $g_T(\theta) = T^{-1} \sum_{t=1}^{T} f(x_t, \theta)$ and W_T be a $q \times q$ positive semidefinite weighting matrix. Then the GMM estimator for θ_0 can be written as

$$\widehat{\theta}_T = argmin_{\theta \in \Theta} \quad g_T(\theta)' W_T g_T(\theta) \tag{2}$$

Let $\widehat{\theta}_T(1)$ be the first step GMM estimator obtained by using the suboptimal choice of W_T, and let $\widehat{S}_T$ be a consistent positive semidefinite non-mean deviation variance-covariance matrix estimator of S, where $S = lim_{T \longrightarrow \infty} var[T^{1/2} g_T(\theta_0)]$ and $\widehat{S}_T$ is constructed by using the $\widehat{\theta}_T(1)$.[3] Moreover, the optimal choice of W_T is to set $W_T = \widehat{S}_T^{-1}$ as demonstrated in Hansen (1982).

Let $\widehat{\theta}_T(2)$ be the second step GMM estimator by using $\widehat{S}_T^{-1}$ as the weighting matrix. Thus, the overidentifying restrictions test can be written as

$$J_T^{nc} = T g_T(\widehat{\theta}_T(2))' \widehat{S}_T^{-1} g_T(\widehat{\theta}_T(2)) \tag{3}$$

where "*nc*" indicates that J_T^{nc} is obtained by using the weighting matrix constructed by non-centering sample moments. Hansen (1982) also indicated that J_T^{nc} converges to χ^2_{q-p} in distribution when the model (1) is correctly specified.

In order to investigate the GMM estimator under a misspecified model, I assume that there is no value of θ at which the population moment condition (1) is satisfied. Then, following Hall (2000), the misspecified model can be captured by the following assumption:

Assumption 2 *Let $E[f(x_t, \theta)] = \mu(\theta)$. Then, $\mu : \Theta \longrightarrow \Re^q$ such that $\|\mu(\theta)\| > 0$ for all $\theta \in \Theta$.*

where $\| \, . \, \|$ denotes any norm.

The following assumption about the weighting matrix W_T and the conditions for identification are assumed to be satisfied.

Assumption 3 *W_T is a positive semidefinite matrix which converges in probability to the positive semidefinite matrix of constants W. Also, there exists $\theta^* \in \Theta$ such that $Q_0(\theta^*) < Q_0(\theta)$ for all $\theta \in \Theta \setminus \theta^*$, where $Q_0(\theta) = E[f(x_t, \theta)]' W E[f(x_t, \theta)]$.*

[3] In other words, the $j - th$ order autocovariance matrix is estimated by using $\widehat{\Gamma}_j = \frac{1}{T} \sum_{t=j+1}^{T} f_t f'_{t-j}$, where $f_t = f(x_t, \widehat{\theta}_T(1))$.

Following Newey and McFadden's (1994) Theorem 2.1, Wooldridge's (1994) Theorem 7.1 and Hall's (2000) Lemma 1, $\widehat{\theta}_T(1) \longrightarrow \theta^*$ in probability under Assumptions 1 to 3. Moreover, $\mu^* = E[f(x_t, \theta^*)]$ and $\mu^* \neq 0$ under Assumption 3.

Let $\widehat{V}_T$ be a consistent positive semidefinite mean deviation variance-covariance matrix estimator of V, where $V = lim_{T \longrightarrow \infty} var[T^{1/2}(f(x_t, \widehat{\theta}_T(1)) - \mu^*)]$ and $\widehat{V}_T$ is constructed by using the $\widehat{\theta}_T(1)$.[4]

Let $\widetilde{\theta}_T(2)$ be the second step GMM estimator by using $\widehat{V}_T^{-1}$ as the weighting matrix. Thus, the overidentifying restrictions test can be written as

$$J_T^c = T g_T(\widetilde{\theta}_T(2))' \widehat{V}_T^{-1} g_T(\widetilde{\theta}_T(2)) \qquad (4)$$

where "c" denotes that J_T^c is obtained using the weighting matrix constructed by centering sample moments. Furthermore, J_T^c also converges to χ^2_{q-p} in distribution when the model is correctly specified.

The following assumption for the matrix V and the conditions for identification are also assumed to be satisfied.

Assumption 4 *V is a positive semidefinite matrix of constants. Also, there exists $\theta^{**} \in \Theta$ such that $Q_{00}(\theta^{**}) < Q_{00}(\theta)$ for all $\theta \in \Theta \backslash \theta^{**}$, where $Q_{00}(\theta) = E[f(x_t, \theta)]' V^{-1} E[f(x_t, \theta)]$.*

3 Results

The approximate slope is introduced by Bahadur in order to facilitate a global power comparison of statistical tests (Bahadur (1960); Serfling (1980)). The overidentifying restrictions test under J_T^{nc} and J_T^c is the problem associated with testing

$$\begin{aligned} H_0 &: E[f(x_t, \theta)] = 0 \quad for \quad some \quad \theta \in \Theta \\ H_A &: E[f(x_t, \theta)] \neq 0 \quad for \quad any \quad \theta \in \Theta \end{aligned}$$

The main theorem of this paper is presented as follows:

Theorem 1 *If Assumptions 1-4 hold, then J_T^c is a more powerful test than J_T^{nc} in large samples for any $\theta^*, \theta^{**} \in \Theta$ when the model is misspecified, where θ^* and θ^{**} are defined in Assumptions 3 and 4.*

Proof of Theorem 1:

[4] In other words, the $j-th$ order autocovariance matrix is estimated by using $\widehat{\Psi}_j = \frac{1}{T}\sum_{t=j+1}^{T}(f_t - g_T(\widehat{\theta}_T(1)))(f_{t-j} - g_T(\widehat{\theta}_T(1)))'$, where $f_t = f(x_t, \widehat{\theta}_T(1))$.

Let $\mu(\theta) = E[f(x_t, \theta)]$. $\widehat{S}_T$ is a consistent positive semidefinite variance-covariance matrix estimator of S, and

$$S = lim_{T\to\infty}\widehat{S}_T = \Gamma_0 + \sum_{j=1}^{\infty}(\Gamma_j + \Gamma_j')$$

where $\Gamma_j = E[f(x_t, \theta^*)f'(x_{t-j}, \theta^*)]$.[5] Moreover, $\widehat{V}_T$ is a consistent positive semidefinite variance-covariance matrix estimator of V, and

$$V = lim_{T\to\infty}\widehat{V}_T = \Psi_0 + \sum_{j=1}^{\infty}(\Psi_j + \Psi_j')$$

where $\Psi_j = E[(f(x_t, \theta^*) - \mu^*)(f(x_{t-j}, \theta^*) - \mu^*)']$, and $\mu^* = E[f(x_t, \theta^*)]$.

Then, $\widehat{S}_T$ can be rewritten as $\widehat{S}_T = V + B_T\mu^*\mu^{*'} + op(1)$, where $B_T = Op(T)$. Based on the Theorem 1 in Hall (2000),

$$lim_{T\longrightarrow\infty}\widehat{S}_T^{-1} = S_* = V^{-1/2'}[I_q - h(h'h)^{-1}h']V^{-1/2}$$

where $h = V^{-1/2}\mu^*$ and S_* is a positive semidefinite matrix.

Let $J^{nc}(\theta^*) = \mu^{*'}S_*\mu^*$ and $J^c(\theta^{**}) = \mu^{**'}V^{-1}\mu^{**}$ where $\mu^{**} = E[f(x_t, \theta^{**})]$. Thus, $\frac{1}{T}J_T^{nc} \longrightarrow J^{nc}(\theta^*)$ *a.s.* and $\frac{1}{T}J_T^c \longrightarrow J^c(\theta^{**})$ *a.s.* when the H_A is true, where J_T^{nc} and J_T^c are defined in (3) and (4). Since both J_T^{nc} and J_T^c have asymptotic χ^2_{q-p} distributions under the null hypothesis, by Geweke's (1981) Theorem 1, the approximate slope of the J_T^{nc} is $J^{nc}(\theta^*)$, and the approximate slope of the J_T^c is $J^c(\theta^{**})$.
Claim: $J^{nc}(\theta^*) < J^c(\theta^{**})$ for any θ^* and $\theta^{**} \in \Theta$ when the H_A is true.
Let $J^{nc}(\theta^{**}) = \mu^{**'}S_*\mu^{**}$, then,

$$J^c(\theta^{**}) - J^{nc}(\theta^{**}) = \mu^{**'}[V^{-1} - S_*]\mu^{**} \tag{5}$$

where

$$V^{-1} - S_* = V^{-1/2'}V^{-1/2} - V^{-1/2'}[I_q - h(h'h)^{-1}h']V^{-1/2} = V^{-1/2'}[h(h'h)^{-1}h']V^{-1/2}$$

Thus, $J^c(\theta^{**}) \geq J^{nc}(\theta^{**})$ since $V^{-1} - S_*$ is a positive semidefinite matrix. Moreover $J^{nc}(\theta^*) < J^{nc}(\theta^{**})$ by Assumption 3. Therefore, $J^{nc}(\theta^*) < J^c(\theta^{**})$ for any θ^* and $\theta^{**} \in \Theta$ when the model is misspecified.

Thus, the approximate slope of J_T^c is greater than that of J_T^{nc}. In other words, J_T^c is more powerful than J_T^{nc} in terms of approximate slopes in large samples. *Q.E.D.*

[5] θ^* is defined in the Assumption 3, and $\widehat{\theta}_T(1) \longrightarrow \theta^*$ in probability.

4 Conclusion

By conducting the asymptotic power comparisons globally through the concept of approximate slopes, it is demonstrated that the overidentifying restrictions test with the consistent mean deviation version variance-covariance matrix estimator is more powerful than the test with the traditional non-mean deviation one. Therefore, it is recommended that the mean deviation version of the consistent variance-covariance matrix estimator should be used for the GMM overidentifying restrictions test.

References

[1] Bahadur, R. (1960) Stochastic comparison of tests. *Annals of Mathematical Statistics* 31, 276-295.

[2] Chang, S. (2005) The approximate slopes and the power of the GMM overidentifying restrictions test. *Applied Economics Letters*, 12, 845-848.

[3] Geweke, J. (1981) The approximate slopes of econometric tests. *Econometrica* 49, 1427-1442.

[4] Hall, A. (2000) Covariance matrix estimation and the power of the overidentifying restrictions test. *Econometrica* 68, 1517-1527.

[5] Hansen, L. (1982) Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029-1054.

[6] Hansen, L. and Singleton, K. (1983) Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy* 91, 249-265.

[7] Newey, W.K. and McFadden, D. (1994) Large sample estimation and hypothesis testing. In: *Engel, R. and McFadden, D. (Eds.), Handbook of Econometrics,* Vol. IV, 2111-2245.

[8] Serfling, R. (1980) Approximation theorems of mathematical statistics. *John Wiley & Sons, Inc.*, New York, NY, U.S.A..

[9] Wooldridge, J. (1994) Estimation and inference for dependent processes. In: *Engel, R. and McFadden, D. (Eds.), Handbook of Econometrics,* Vol. IV, 2641-2739.