Journal of Statistical Planning and Inference 178 (2016) 84–98

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

On the effect of weighting matrix in GMM specification test[☆]

Kazuhiko Hayakawa

Department of Economics, Hiroshima University, 1-2-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8525, Japan

ARTICLE INFO

Article history:
Received 26 April 2016
Received in revised form 3 May 2016
Accepted 9 June 2016
Available online 22 June 2016

Keywords:
GMM
Over-identifying restriction test
Size distortion
Dynamic panel data model
Covariance structure model

ABSTRACT

In this paper, we study the finite sample behavior of an over-identifying restriction test, the J test, in generalized method of moments(GMM). We consider two variants of the J test, one with centered weighting matrix and the other with uncentered weighting matrix. We demonstrate that the finite sample distribution of J test with centered weighting matrix is close to F distribution whereas that with uncentered weighting matrix is close to beta distribution and that both are far different from chi-square distribution. Using this, we demonstrate why competing simulation results relating to the size property is reported in the context of dynamic panel data and covariance structure models.

1. Introduction

Since the seminal work of Hansen (1982), the generalized method of moments (GMM) estimator has been widely used in applied econometric analysis. When using GMM, we need to first check the validity of the moment conditions via the over-identified restriction test, often called the J test, since the GMM estimator becomes inconsistent when the moment conditions are invalid. Unfortunately, the finite sample behavior of the J test is found to be poor: the J test reveals large size distortions in finite samples, especially when the number of moment conditions is large. In the context of GMM estimation of dynamic panel data models, Bowsher (2002) provided Monte Carlo evidence that the J test tends to become under-sized as the number of moment conditions gets larger. However, Hayakawa (2015) has recently provided exactly opposite results: the J test tends to be over-sized as the number of moment conditions becomes larger. Similar results are reported in the literature of covariance structure analysis.[1] For instance, Hu et al. (1992) conduct an extensive simulation study and show that GMM-based tests tend to become over-sized when the sample size is small.[2] Several approaches have been proposed to address this problem. Yuan and Bentler (1997) proposed an alternative form of weighting matrix, and Yuan and Bentler (1999) proposed an F test rather than the conventional chi-square test. However, with regard to the former test, Yuan and Bentler (1997) noted that the alternative approach tends to over-correct and results in an under-sized test when the sample size is small.

Thus, the literature reports contrasting results with regard to the size property of tests. However, the reason for this difference has not been well discussed in the literature. This paper tries to find the reason by investigating the finite sample distribution of tests with centered and uncentered weighting matrices. We show that finite sample distribution of a test

☆ The author is grateful to an associate editor, two referees and Yasunori Fujikoshi for helpful comments. He also acknowledges financial support from the Grant-in-Aid for Scientific Research (KAKENHI 25780153, 25285067, 16H03606) provided by the JSPS.

E-mail address: kazuhaya@hiroshima-u.ac.jp.

1 Although the term "GMM" is not used in the literature of covariance structure analysis, the methods used in this literature can be seen as a special case of GMM. For details, see Section 4.2

2 The asymptotically distribution-free (ADF) test corresponds to the GMM specification test in the context of covariance structure analysis.

http://dx.doi.org/10.1016/j.jspi.2016.06.003
0378-3758/

with centered weighting matrix can be approximated by F distribution whereas that with uncentered weighting matrix can be approximated by beta distribution. From these findings, we try to explain why a test with centered weighting matrix tends to become over-sized whereas that with uncentered weighting matrix tends to become under-sized. We also provide a detailed analysis in the context of dynamic panel data and covariance structure models.

In Section 2, we briefly review the GMM specification test. Section 3 explores distributions that approximate the finite sample distribution of GMM specification tests. Section 4 provides a detailed analysis of dynamic panel data and covariance structure models. Finally, we conclude the study in Section 5.

2. GMM specification test

First, let us consider the moment conditions $E[\mathbf{g}(\mathbf{x}_i,\boldsymbol{\theta}_0)] = E[\mathbf{g}_i(\boldsymbol{\theta}_0)] = \mathbf{0}$, where $\mathbf{g}(\cdot,\cdot)$ is an $m \times 1$ known function, $\{\mathbf{x}_i\}_{i=1}^N$ are independent observations, and $\boldsymbol{\theta}_0$ is the true value of a $p \times 1$ vector of unknown parameters.

We consider two variants of two-step GMM estimators,

$$\widehat{\boldsymbol{\theta}} = \underset{\theta}{\operatorname{argmin}}\ \bar{\mathbf{g}}_N(\boldsymbol{\theta})'\widehat{\boldsymbol{\Omega}}(\check{\boldsymbol{\theta}})^{-1}\bar{\mathbf{g}}_N(\boldsymbol{\theta}),$$
$$\widetilde{\boldsymbol{\theta}} = \underset{\theta}{\operatorname{argmin}}\ \bar{\mathbf{g}}_N(\boldsymbol{\theta})'\widetilde{\boldsymbol{\Omega}}(\check{\boldsymbol{\theta}})^{-1}\bar{\mathbf{g}}_N(\boldsymbol{\theta}),$$

where $\check{\boldsymbol{\theta}}$ is a preliminary consistent estimator of $\boldsymbol{\theta}$, $\bar{\mathbf{g}}_N(\boldsymbol{\theta}) = N^{-1}\sum_{i=1}^N \mathbf{g}_i(\boldsymbol{\theta})$, and

$$\widehat{\boldsymbol{\Omega}}(\boldsymbol{\theta}) = \frac{1}{N-1}\sum_{i=1}^N [\mathbf{g}_i(\boldsymbol{\theta}) - \bar{\mathbf{g}}_N(\boldsymbol{\theta})][\mathbf{g}_i(\boldsymbol{\theta}) - \bar{\mathbf{g}}_N(\boldsymbol{\theta})]', \qquad \widetilde{\boldsymbol{\Omega}}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^N \mathbf{g}_i(\boldsymbol{\theta})\mathbf{g}_i(\boldsymbol{\theta})'.$$

Note that both $\widehat{\boldsymbol{\Omega}}(\boldsymbol{\theta}_0)$ and $\widetilde{\boldsymbol{\Omega}}(\boldsymbol{\theta}_0)$ are consistent estimators of $\boldsymbol{\Omega} = E[\mathbf{g}_i(\boldsymbol{\theta}_0)\mathbf{g}_i(\boldsymbol{\theta}_0)']$ when $E[\mathbf{g}_i(\boldsymbol{\theta}_0)] = \mathbf{0}$ holds. Also, $\widetilde{\boldsymbol{\Omega}}(\boldsymbol{\theta}_0)$ can be interpreted such that it uses the information that $E[\mathbf{g}_i(\boldsymbol{\theta}_0)] = \mathbf{0}$ when it holds, since $\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0) \xrightarrow{p} E[\mathbf{g}_i(\boldsymbol{\theta}_0)] = \mathbf{0}$.

In empirical studies using GMM, we first check the validity of moment conditions through the J test prior to inference of the parameters of interest. The J test statistics are typically computed as

$$\widehat{J} = N\cdot\bar{\mathbf{g}}_N(\widehat{\boldsymbol{\theta}})'\widehat{\boldsymbol{\Omega}}(\widehat{\boldsymbol{\theta}})^{-1}\bar{\mathbf{g}}_N(\widehat{\boldsymbol{\theta}}), \tag{1}$$
$$\widetilde{J} = N\cdot\bar{\mathbf{g}}_N(\widetilde{\boldsymbol{\theta}})'\widetilde{\boldsymbol{\Omega}}(\widetilde{\boldsymbol{\theta}})^{-1}\bar{\mathbf{g}}_N(\widetilde{\boldsymbol{\theta}}). \tag{2}$$

Alternatively, we can consider

$$\dot{J} = N\cdot\bar{\mathbf{g}}_N(\widehat{\boldsymbol{\theta}})'\widetilde{\boldsymbol{\Omega}}(\widehat{\boldsymbol{\theta}})^{-1}\bar{\mathbf{g}}_N(\widehat{\boldsymbol{\theta}}), \tag{3}$$

where $\widehat{J}$ and $\dot{J}$ are related such that

$$\dot{J} = \frac{N\widehat{J}}{(N-1)+\widehat{J}}. \tag{4}$$

From this, it follows that $\dot{J} < \widehat{J}$. Then, under $H_0 : E[\mathbf{g}_i(\boldsymbol{\theta}_0)] = \mathbf{0}$ and the regularity conditions, it follows that $\widehat{J}, \widetilde{J}, \dot{J} \xrightarrow{d} \chi^2_{m-p}$ as $N \to \infty$, where m is fixed.

For later use, we reformulate statistics $\widetilde{J}, \widehat{J}$, and $\dot{J}$ as follows:

$$\widehat{J} = N\cdot\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0)'\left(\widehat{\boldsymbol{\Omega}}_0^{-1} - \widehat{\boldsymbol{\Omega}}_0^{-1}\mathbf{G}(\mathbf{G}'\widehat{\boldsymbol{\Omega}}_0^{-1}\mathbf{G})^{-1}\mathbf{G}'\widehat{\boldsymbol{\Omega}}_0^{-1}\right)^{-1}\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0) + o_p(1),$$
$$\widetilde{J} = N\cdot\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0)'\left(\widetilde{\boldsymbol{\Omega}}_0^{-1} - \widetilde{\boldsymbol{\Omega}}_0^{-1}\mathbf{G}(\mathbf{G}'\widetilde{\boldsymbol{\Omega}}_0^{-1}\mathbf{G})^{-1}\mathbf{G}'\widetilde{\boldsymbol{\Omega}}_0^{-1}\right)^{-1}\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0) + o_p(1) = \widetilde{J}_0 + o_p(1),$$
$$\dot{J} = \widetilde{J}_0 + o_p(1)$$

where $\widehat{\boldsymbol{\Omega}}_0 = \widehat{\boldsymbol{\Omega}}(\boldsymbol{\theta}_0)$, $\widetilde{\boldsymbol{\Omega}}_0 = \widetilde{\boldsymbol{\Omega}}(\boldsymbol{\theta}_0)$, $\mathbf{G} = \text{plim}_{N\to\infty}\,\partial\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}'$ and $\widetilde{J}_0 = N\cdot\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0)'\left(\widetilde{\boldsymbol{\Omega}}_0^{-1} - \widetilde{\boldsymbol{\Omega}}_0^{-1}\mathbf{G}(\mathbf{G}'\widetilde{\boldsymbol{\Omega}}_0^{-1}\mathbf{G})^{-1}\mathbf{G}'\widetilde{\boldsymbol{\Omega}}_0^{-1}\right)^{-1}\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0)$. We then reformulate these statistics. For this, we assume that $\mathbf{G}_c$ is an $m \times (m-p)$ matrix whose columns are orthogonal to those of $\mathbf{G}$; that is, $\mathbf{G}_c'\mathbf{G} = \mathbf{0}$. Then, using Lemma 1 of Khatri (1966), we rewrite the above expressions as

$$\widehat{J} = N\cdot\left[\mathbf{G}_c'\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0)\right]'\left(\mathbf{G}_c'\widehat{\boldsymbol{\Omega}}_0^{-1}\mathbf{G}_c\right)^{-1}\left[\mathbf{G}_c'\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0)\right] + o_p(1), \tag{5}$$
$$\widetilde{J} = N\cdot\left[\mathbf{G}_c'\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0)\right]'\left(\mathbf{G}_c'\widetilde{\boldsymbol{\Omega}}_0^{-1}\mathbf{G}_c\right)^{-1}\left[\mathbf{G}_c'\bar{\mathbf{g}}_N(\boldsymbol{\theta}_0)\right] + o_p(1) = \widetilde{J}_{c0} + o_p(1), \tag{6}$$
$$\dot{J} = \widetilde{J}_{c0} + o_p(1), \tag{7}$$