中央研究院經濟所學術研討論文

IEAS Working Paper

Testing Over-Identifying Restrictions without Consistent Estimation of the Asymptotic Covariance Matrix

Wei-Ming Lee and Chung-Ming Kuan

IEAS Working Paper No. 06-A009

September , 2006

Institute of Economics
Academia Sinica
Taipei 115, TAIWAN
http://www.sinica.edu.tw/econ/

中央研究院　經濟研究所

INSTITUTE OF ECONOMICS, ACADEMIA SINICA

TAIWAN

Testing Over-Identifying Restrictions without Consistent Estimation of the Asymptotic Covariance Matrix

Wei-Ming Lee
Department of Economics
National Chung Cheng University

Chung-Ming Kuan
Institute of Economics
Academia Sinica

First version: September 15, 2006

† Author for correspondence: Chung-Ming Kuan, Institute of Economics, Academia Sinica, Taipei 115, Taiwan; ckuan@econ.sinica.edu.tw.

†† We thank Chang-Ching Lin for helpful comments. The research supports from the National Science Council of the Republic of China (NSC90-2415-H-001-016 for Kuan and NSC93-2415-H-194-010 for Lee) are gratefully acknowledged.

Abstract

We extend the KVB approach of Kiefer, Vogelsang, and Bunzel (2000, *Econometrica*) and Kiefer and Vogelsang (2002b, *Econometric Theory*) to construct a class of robust tests for over-identifying restrictions in the context of GMM. The proposed test does not require consistent estimation of the asymptotic covariance matrix but relies on kernel-based normalizing matrices to eliminate the nuisance parameters in the limit. Moreover, the proposed test is valid for any consistent GMM estimator, in contrast with the conventional test that requires the optimal GMM estimator, and hence is easy to implement. Our simulations show that the proposed test is properly sized and may even be more powerful than the conventional test computed with an inappropriate user-chosen parameter.

JEL classification: C12, C22

Keywords: generalized method of moments, kernel function, KVB approach, over-identifying restrictions, robust test

1 Introduction

The generalized method of moments (GMM) introduced in Hansen (1982) is a leading estimation technique in econometric applications. In the context of GMM, the validity of the moment conditions is tested using the over-identifying restrictions (OIR) test. Similar to many asymptotic tests, the OIR test requires consistent estimation of the asymptotic covariance matrix. When heteroskedasticity and serial correlations are present, the covariance matrix can be consistently estimated by the nonparametric kernel method; see den Haan and Levin (1997) for a review of this method. A well known problem with the kernel covariance-matrix estimator is that its behavior depends on the chosen kernel function and truncation lag (i.e., the number of autocovariances being estimated). The resulting OIR test is thus not robust because the choices of the kernel function and truncation lag are somewhat arbitrary in practice, even when some "automatic selection" methods for truncation lags are available (e.g., Andrews, 1991; Newey and West, 1994).

To circumvent the problems arising from nonparametric kernel estimation of the asymptotic covariance matrix, Kiefer, Vogelsang, and Bunzel (2000), hereafter KVB, proposed an alternative approach to constructing tests for parameters in linear regressions; see Bunzel, Kiefer, and Vogelsang (2001), Kiefer and Vogelsang (2002a, b), and Vogelsang (2003) for other applications and extensions of this approach. The main idea of the KVB approach is to obtain an asymptotically pivotal test by employing a normalizing matrix that can eliminate the nuisance parameters of the asymptotic covariance matrix. As for specification testing, Lobato (2001) obtained a test for serial correlations along the same line; Kuan and Lee (2006) proposed robust M tests for general moment conditions. However, the result of Kuan and Lee (2006) requires the asymptotic covariance matrix been nonsingular, yet this condition fails to hold in testing OIR. We are therefore motivated to find a robust OIR test in the spirit of KVB.

In this paper, we extend KVB and Kiefer and Vogelsang (2002b) to construct a class of robust OIR tests. The proposed OIR test does not require consistent estimation of the asymptotic covariance matrix but relies on kernel-based normalizing matrices to eliminate the nuisance parameters in the limit. This test is thus asymptotically pivotal, and its asymptotic critical values can be easily obtained via simulations (some critical values are already available in the literature). An important feature of the proposed OIR test is that it is valid for any consistent GMM estimator, in contrast with the conventional OIR test that requires the optimal GMM estimator. As for finite-sample performance, our simulations show that the proposed test is properly sized and may even be more powerful than the conventional test computed with an inappropriate user-chosen parameter.

This paper proceeds as follows. In Section 2, we review the GMM estimation and OIR test. A class of robust OIR tests and its asymptotic properties are presented in Section 3. Simulation results are reported in Section 4. Section 5 concludes the paper. All proofs are deferred to Appendix.

2 GMM and OIR Test

Consider the model characterized by a vector of q moment conditions:

$$\mathbb{E}[\boldsymbol{f}(\boldsymbol{\eta}_t;\boldsymbol{\theta}_o)] = \mathbf{0}, \text{ for a unique } \boldsymbol{\theta}_o \in \Theta \subset \mathbb{R}^p, \tag{1}$$

where $\boldsymbol{\eta}_t$ is a random data vector, $\boldsymbol{\theta}_o$ $(p \times 1)$ is the true parameter vector, and $\boldsymbol{f}$ $(q \times 1)$ is a vector of functions that are continuously differentiable in the neighborhood of $\boldsymbol{\theta}_o$. The parameter $\boldsymbol{\theta}_o$ is said to be over-identified (just-identified) if $q > (=)\ p$. Given a sample of T observations, the GMM estimator of $\boldsymbol{\theta}_o$ is

$$\hat{\boldsymbol{\theta}}_T = \operatorname{argmin}_{\boldsymbol{\theta}\in\Theta} Q_T(\boldsymbol{\theta}) = \boldsymbol{m}_T(\boldsymbol{\theta})'\boldsymbol{H}_T\boldsymbol{m}_T(\boldsymbol{\theta}),$$

where $\boldsymbol{H}_T$ is a symmetric, positive semi-definite weighting matrix and

$$\boldsymbol{m}_{[rT]}(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{[rT]} \boldsymbol{f}(\boldsymbol{\eta}_t;\boldsymbol{\theta}), \qquad 0 < r \leq 1,$$

with $\boldsymbol{m}_T(\boldsymbol{\theta})$ the full-sample average of $\boldsymbol{f}(\boldsymbol{\eta}_t;\boldsymbol{\theta})$.

In the analysis below, we shall consider the local departure from (1):

$$\mathbb{E}[\boldsymbol{f}(\boldsymbol{\eta}_t;\boldsymbol{\theta}_o)] = \boldsymbol{\delta}_o/\sqrt{T}, \tag{2}$$

where $\boldsymbol{\delta}_o$ is a non-zero vector. Clearly, (2) reduces to (1) when $\boldsymbol{\delta}_o = \mathbf{0}$. In what follows, we let $[c]$ denote the integer part of the real number c, $\Rightarrow$ weak convergence (of associated probability measures), $\xrightarrow{\mathbb{P}}$ convergence in probability, $\xrightarrow{D}$ convergence in distribution, $\stackrel{d}{=}$ equality in distribution, $\boldsymbol{W}_q$ a vector of q independent, standard Wiener processes, and $\boldsymbol{B}_q$ the Brownian bridge with $\boldsymbol{B}_q(r) = \boldsymbol{W}_q(r) - r\boldsymbol{W}_q(1)$ for $0 \leq r \leq 1$. Given a matrix $\boldsymbol{A}$ with full column rank, we write $\boldsymbol{M}_A = \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'$ and $\boldsymbol{V}_A = \boldsymbol{I} - \boldsymbol{M}_A$. We also write $\boldsymbol{B}^+$ as the Moore-Penrose generalized inverse of $\boldsymbol{B}$. We assume throughout the paper that $\boldsymbol{H}_T \xrightarrow{\mathbb{P}} \boldsymbol{H}_o$, where $\boldsymbol{H}_o$ is a $q \times q$ non-stochastic matrix that is symmetric and positive definite. This is a standard condition in the GMM literature; note that the class of optimal weighting matrices recommended by Hansen (1982) satisfies this condition.

To establish the properties of $\hat{\boldsymbol{\theta}}_T$, we impose the following "high-level" conditions, similar to those in Vogelsang (2003), Kiefer and Vogelsang (2005), and Kuan and Lee (2006). These conditions will also be used to analyze the proposed test in the next section.

[A1] Under the local alternative (2), $\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_o) = O_{\mathbb{P}}(1)$.

[A2] Under the local alternative (2),

$$\sqrt{T}\boldsymbol{m}_{[rT]}(\boldsymbol{\theta}_o) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} \boldsymbol{f}(\boldsymbol{\eta}_t; \boldsymbol{\theta}_o) \Rightarrow r\boldsymbol{\delta}_o + \boldsymbol{S}\boldsymbol{W}_q(r), \quad 0 < r \leq 1,$$

where $\boldsymbol{S}$ is the nonsingular, matrix square root of $\boldsymbol{\Sigma}_o$ (i.e., $\boldsymbol{\Sigma}_o = \boldsymbol{S}\boldsymbol{S}'$), and $\boldsymbol{\Sigma}_o = \lim_{T\to\infty} \operatorname{var}(T^{1/2}\boldsymbol{m}_T(\boldsymbol{\theta}_o))$.

[A3] $\boldsymbol{F}_{[rT]}(\boldsymbol{\theta}) = [rT]^{-1} \sum_{t=1}^{[rT]} \nabla_{\boldsymbol{\theta}} \boldsymbol{f}\big(\boldsymbol{\eta}_t; \boldsymbol{\theta}\big) \xrightarrow{\mathbb{P}} \boldsymbol{F}(\boldsymbol{\theta})$ uniformly in $\boldsymbol{\theta}$ and $0 < r \leq 1$, and $\boldsymbol{F}_o := \boldsymbol{F}(\boldsymbol{\theta}_o)$ is a $q \times p$ matrix with full column rank. Further, $\nabla_{\boldsymbol{\theta}} \boldsymbol{F}_{[rT]}(\boldsymbol{\theta}_o)$ is bounded in probability.

[A1] requires consistency of the GMM estimator, and [A2] regulates $\{\boldsymbol{f}(\boldsymbol{\eta}_t; \boldsymbol{\theta}_o)\}$ to obey a functional central limit theorem. Both conditions are assumed to hold under the local alternative (2) and hence permit analysis under local mis-specification; see also Hall (1999, pp. 101–103). Note that [A2] is somewhat weaker than that for establishing consistency of the asymptotic covariance matrix estimator, as pointed out in Kiefer and Vogelsang (2005). [A3] is also standard in the literature and implies $\boldsymbol{F}_T(\boldsymbol{\theta}_o) = T^{-1} \sum_{t=1}^{T} \nabla_{\boldsymbol{\theta}} \boldsymbol{f}(\boldsymbol{\eta}_t; \boldsymbol{\theta}_o) \xrightarrow{\mathbb{P}} \boldsymbol{F}_o$.

It is easy to verify that the GMM estimator has the Bahadur representation:

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_o) = -(\boldsymbol{F}_o'\boldsymbol{H}_o\boldsymbol{F}_o)^{-1}\boldsymbol{F}_o'\boldsymbol{H}_o \left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{f}(\boldsymbol{\eta}_t; \boldsymbol{\theta}_o)\right] + o_{\mathbb{P}}(1), \tag{3}$$

and by [A2], its asymptotic covariance matrix is

$$\boldsymbol{\Omega}_o(\boldsymbol{H}_o) = (\boldsymbol{F}_o'\boldsymbol{H}_o\boldsymbol{F}_o)^{-1}\boldsymbol{F}_o'\boldsymbol{H}_o\boldsymbol{\Sigma}_o\boldsymbol{H}_o\boldsymbol{F}_o(\boldsymbol{F}_o'\boldsymbol{H}_o\boldsymbol{F}_o)^{-1}.$$

In particular, $\boldsymbol{\Omega}_o(\boldsymbol{\Sigma}_o^{-1}) = (\boldsymbol{F}_o'\boldsymbol{\Sigma}_o^{-1}\boldsymbol{F}_o)^{-1}$. It is easily shown that $\boldsymbol{\Omega}_o(\boldsymbol{H}_o) - \boldsymbol{\Omega}_o(\boldsymbol{\Sigma}_o^{-1})$ is a positive semi-definite matrix for any $\boldsymbol{H}_o \neq \boldsymbol{\Sigma}_o^{-1}$. This suggests that the optimal GMM estimator, $\hat{\boldsymbol{\theta}}_T^*$, can be obtained by minimizing $\boldsymbol{m}_T(\boldsymbol{\theta})'\widehat{\boldsymbol{\Sigma}}_T^{-1}\boldsymbol{m}_T(\boldsymbol{\theta})$. A preliminary GMM estimator of $\boldsymbol{\theta}_o$ is thus needed to compute $\widehat{\boldsymbol{\Sigma}}_T$ before conducting the optimal estimation.

As far as testing the validity of the model (1) is concerned, it is natural to base a specification test on $T^{1/2}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)$. Such a test is known as the OIR test because it is only possible to test (1) when $q > p$. Define the OIR test with the ($\boldsymbol{H}_T$-based) GMM estimator $\hat{\boldsymbol{\theta}}_T$ and the weighting matrix $\ddot{\boldsymbol{H}}_T$ as

$$\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \ddot{\boldsymbol{H}}_T) = T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)'\ddot{\boldsymbol{H}}_T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T).$$

It can be seen that the OIR test of Hansen (1982) is a special case with the optimal GMM estimator $\hat{\boldsymbol{\theta}}_T^*$ and the optimal weighting matrix $\ddot{\boldsymbol{H}}_T = \widehat{\boldsymbol{\Sigma}}_T^{-1}$, i.e.,

$$\mathcal{J}^* = \mathcal{J}(\hat{\boldsymbol{\theta}}_T^*, \widehat{\boldsymbol{\Sigma}}_T^{-1}) = T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T^*)'\widehat{\boldsymbol{\Sigma}}_T^{-1}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T^*).$$

To derive the limit of $\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \ddot{\boldsymbol{H}}_T)$, note that the first-order Taylor expansion of $T^{1/2}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)$ about $\boldsymbol{\theta}_o$ is

$$\sqrt{T}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T) = \sqrt{T}\boldsymbol{m}_T(\boldsymbol{\theta}_o) + \boldsymbol{F}_o\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_o) + o_{\mathbb{P}}(1). \tag{4}$$

Thus, $T^{1/2}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)$ and $T^{1/2}\boldsymbol{m}_T(\boldsymbol{\theta}_o)$ are not asymptotically equivalent due to the presence of estimation effect (i.e., the second term on the right-hand side of (4)). Letting $\boldsymbol{\Lambda}$ denote the matrix square root of $\boldsymbol{H}_o$ such that $\boldsymbol{\Lambda}\boldsymbol{\Lambda}' = \boldsymbol{H}_o$, we have from the Bahadur representation (3) that

$$\begin{aligned}\sqrt{T}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T) &= \left[\boldsymbol{I}_q - \boldsymbol{F}_o(\boldsymbol{F}_o'\boldsymbol{H}_o\boldsymbol{F}_o)^{-1}\boldsymbol{F}_o'\boldsymbol{H}_o\right]\sqrt{T}\boldsymbol{m}_T(\boldsymbol{\theta}_o) + o_{\mathbb{P}}(1)\\ &= \boldsymbol{\Lambda}'^{-1}\boldsymbol{V}_{\boldsymbol{\Lambda}'\boldsymbol{F}_o}\boldsymbol{\Lambda}'\sqrt{T}\boldsymbol{m}_T(\boldsymbol{\theta}_o) + o_{\mathbb{P}}(1),\end{aligned} \tag{5}$$

where $\boldsymbol{V}_{\boldsymbol{\Lambda}'\boldsymbol{F}_o} = \boldsymbol{I}_q - \boldsymbol{\Lambda}'\boldsymbol{F}_o(\boldsymbol{F}_o'\boldsymbol{H}_o\boldsymbol{F}_o)^{-1}\boldsymbol{F}_o'\boldsymbol{\Lambda}$ is symmetric and idempotent with rank $q-p$. Letting $\boldsymbol{U} := \boldsymbol{\Lambda}\boldsymbol{V}_{\boldsymbol{\Lambda}'\boldsymbol{F}_o}\boldsymbol{\Lambda}^{-1}$, [A2] and (5) ensure that

$$\sqrt{T}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T) \xrightarrow{D} \boldsymbol{U}'[\boldsymbol{\delta}_o + \boldsymbol{S}\boldsymbol{W}_q(1)]. \tag{6}$$

As $\boldsymbol{U}$ is singular with rank $q-p$, this is a Gaussian limit with the singular asymptotic covariance matrix: $\boldsymbol{U}'\boldsymbol{\Sigma}_o\boldsymbol{U}$. The result below gives the limits of the OIR tests.

Theorem 2.1 *Given [A1]–[A3], we have under the local alternative (2) that*

$$\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \ddot{\boldsymbol{H}}_T) \xrightarrow{D} [\boldsymbol{\delta}_o + \boldsymbol{S}\boldsymbol{W}_q(1)]'\boldsymbol{U}\ddot{\boldsymbol{H}}_o\boldsymbol{U}'[\boldsymbol{\delta}_o + \boldsymbol{S}\boldsymbol{W}_q(1)];$$

in particular, $\mathcal{J}^* \xrightarrow{D} \chi^2(q-p,\ \boldsymbol{\delta}_o'\boldsymbol{S}'^{-1}\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}\boldsymbol{S}^{-1}\boldsymbol{\delta}_o)$, *where* $\boldsymbol{\delta}_o'\boldsymbol{S}^{-1'}\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}\boldsymbol{S}^{-1}\boldsymbol{\delta}_o$ *is the non-centrality parameter.*

It is readily seen from Theorem 2.1 that, under the null that $\boldsymbol{\delta}_o = \boldsymbol{0}$, $\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \ddot{\boldsymbol{H}}_T)$ still depends on the nuisance parameters $\boldsymbol{S}$, $\boldsymbol{\Lambda}$ and $\boldsymbol{F}_o$ in the limit and hence is not asymptotically pivotal. By contrast, $\mathcal{J}^*$ is asymptotically pivotal with the limiting distribution $\chi^2(q-p)$ and has asymptotic local power against (2), yet it requires optimal GMM estimation and a consistent estimator of $\boldsymbol{\Sigma}_o$.

When $\boldsymbol{f}(\boldsymbol{\eta}_t;\boldsymbol{\theta}_o)$ are heteroskedastic and serially correlated, a leading consistent estimator of $\boldsymbol{\Sigma}_o$ is the following nonparametric kernel estimator:

$$\widehat{\boldsymbol{\Sigma}}^{\kappa}_{\ell(T)} = \frac{1}{T}\sum_{i=1}^{T}\sum_{j=1}^{T}\kappa\left(\frac{|i-j|}{\ell(T)}\right)\left[\boldsymbol{f}(\boldsymbol{\eta}_i;\hat{\boldsymbol{\theta}}_T)-\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)\right]\left[\boldsymbol{f}(\boldsymbol{\eta}_j;\hat{\boldsymbol{\theta}}_T)-\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)\right]', \qquad (7)$$

where $\hat{\boldsymbol{\theta}}_T$ is a preliminary consistent estimator for $\boldsymbol{\theta}_o$ and κ is a kernel function that vanishes when $|i-j| > \ell(T)$, and $\ell(T)$ grows with T at a slower rate and is known as the truncation lag. It should be mentioned that a "non-centered" version of $\widehat{\boldsymbol{\Sigma}}^{\kappa}_{\ell(T)}$, with $\boldsymbol{f}(\boldsymbol{\eta}_i;\hat{\boldsymbol{\theta}}_T)\boldsymbol{f}(\boldsymbol{\eta}_j;\hat{\boldsymbol{\theta}}_T)'$ as summand, is not consistent under non-local alternatives. While the tests based on these two versions of kernel estimators have the same weak limit under both the null and local alternatives, the test with a "centered" $\widehat{\boldsymbol{\Sigma}}^{\kappa}_{\ell(T)}$ is more powerful than that with a "non-centered" $\widehat{\boldsymbol{\Sigma}}^{\kappa}_{\ell(T)}$, because the former is $O_{\mathbb{P}}(T)$ but the latter is $O_{\mathbb{P}}(T/\ell(T))$ under non-local alternatives, as shown in Hall (2000) and Hall, Inoue, and Peixe (2003). Note also that the performance of the OIR test with the kernel covariance-matrix estimator depends on the chosen κ and $\ell(T)$. Even though $\ell(T)$ may be chosen using a data-dependent method (e.g., Andrews, 1991; Newey and West, 1994), the selected $\ell(T)$ may still be arbitrary because it requires additional user-chosen parameters.

3 The Proposed OIR Tests

The OIR test is a special case of the M test. Given a set of general moment conditions of the form (1), Kuan and Lee (2006) proposed robust M tests without consistent estimation of asymptotic covariance matrix, analogous to the robust tests of parameters considered by KVB. In particular, they suggested two normalizing matrices:[1] (i) $\widehat{\boldsymbol{C}}_T = T^{-1}\sum_{t=1}^{T}\boldsymbol{\varphi}_t(\hat{\boldsymbol{\theta}}_T)\boldsymbol{\varphi}_t(\hat{\boldsymbol{\theta}}_T)'$, with $\hat{\boldsymbol{\theta}}_T$ a full-sample estimator and

$$\boldsymbol{\varphi}_t(\hat{\boldsymbol{\theta}}_T) = \frac{1}{\sqrt{T}}\sum_{i=1}^{t}\left[\boldsymbol{f}(\boldsymbol{\eta}_i;\hat{\boldsymbol{\theta}}_T)-\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)\right],$$

and (ii) $\widetilde{\boldsymbol{C}}_T = T^{-1}\sum_{t=p+1}^{T}\tilde{\boldsymbol{\varphi}}_t\tilde{\boldsymbol{\varphi}}_t'$ with

$$\tilde{\boldsymbol{\varphi}}_t = \frac{1}{\sqrt{T}}\sum_{i=1}^{t}\left[\boldsymbol{f}(\boldsymbol{\eta}_i;\tilde{\boldsymbol{\theta}}_t)-\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)\right],$$

[1] What Kuan and Lee (2006) suggested are "centered" normalizing matrices. They demonstrated that the robust M test with a "non-centered" normalizing matrix virtually has no power.

where $\tilde{\boldsymbol{\theta}}_t$ is the recursive counterpart of $\hat{\boldsymbol{\theta}}_T$, computed from the subsample of first t observations so that $\tilde{\boldsymbol{\theta}}_T = \hat{\boldsymbol{\theta}}_T$. It was shown that $T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)'\widehat{\boldsymbol{C}}_T^{-1}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)$ is not asymptotically pivotal unless the estimation effect is absent (i.e., $\boldsymbol{F}_o$ in (4) is a zero matrix), and the M test $T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)'\widetilde{\boldsymbol{C}}_T^{-1}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)$ has the same weak limit regardless of the estimation effect and hence is asymptotically pivotal in general.

3.1 A Robust OIR Test

A crucial condition ensuring the validity of these two normalizing matrices is that the asymptotic covariance matrix of $T^{1/2}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)$ is nonsingular. In the context of OIR testing, this condition fails because the asymptotic covariance matrix, $\boldsymbol{U}'\boldsymbol{\Sigma}_o\boldsymbol{U}$, is singular. As such, none of the tests proposed by Kuan and Lee (2006) can serve as a robust OIR test. To be sure, we first derive the limits of $\widehat{\boldsymbol{C}}_T$ and $\widetilde{\boldsymbol{C}}_T$.

Lemma 3.1 *Given [A1]–[A3], we have under the local alternative (2) that*

$$\widehat{\boldsymbol{C}}_T \Rightarrow \boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}',$$
$$\widetilde{\boldsymbol{C}}_T \Rightarrow \boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U},$$

where $\boldsymbol{P}_q = \int_0^1 \boldsymbol{B}_q(r)\boldsymbol{B}_q(r)'\, dr$.

In view of (6), it is clear that the limit of $\widehat{\boldsymbol{C}}_T$ is unable to eliminate the nuisance parameters in the limit of $T^{1/2}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)$. On the other hand, the limit of $\widetilde{\boldsymbol{C}}_T$ is a singular matrix, and hence the convergence of $\widetilde{\boldsymbol{C}}_T$ in Lemma 3.1 need not carry over under generalized inverse. This stems from the fact that $\text{rank}(\widetilde{\boldsymbol{C}}_T)$ need not converge (with probability one) to $\text{rank}(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U}) = q-p$; see, e.g., Andrews (1987) and Scott (1997, pp. 188–190). As a consequence, it is not even easy to determine the weak limit of $T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)'\widetilde{\boldsymbol{C}}_T^{+}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)$.

Instead of normalizing by $\widetilde{\boldsymbol{C}}_T^{+}$, consider the following statistic:

$$T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)'\big(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U}\big)^{+}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T).$$

By (6), the weak limit of this statistic under the null hypothesis is

$$\boldsymbol{W}_q(1)'\boldsymbol{S}'\boldsymbol{U}\big(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U}\big)^{+}\boldsymbol{U}'\boldsymbol{S}\boldsymbol{W}_q(1) \stackrel{d}{=} \boldsymbol{W}_{q-p}(1)'\boldsymbol{P}_{q-p}^{-1}\boldsymbol{W}_{q-p}(1); \tag{8}$$

the proof (8) is given in Appendix; see also equation (9) of Kuan and Lee (2006). Thus, it is possible to have an asymptotically pivotal OIR test if we can find a suitable normalizing matrix such that its Moore-Penrose generalized inverse converges weakly to

$\big(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U}\big)^+$. As the limit of $\widehat{\boldsymbol{C}}_T$ is $\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'$ (Lemma 3.1), a candidate for the desired normalizing matrix is $\widehat{\boldsymbol{\Gamma}}_T = \widehat{\boldsymbol{U}}_T'\widehat{\boldsymbol{C}}_T\widehat{\boldsymbol{U}}_T$, where $\widehat{\boldsymbol{U}}_T = \widehat{\boldsymbol{\Lambda}}_T\widehat{\boldsymbol{V}}_T\widehat{\boldsymbol{\Lambda}}_T^{-1}$ is a consistent estimator of $\boldsymbol{U}$, with $\widehat{\boldsymbol{\Lambda}}_T$ the matrix square root of $\boldsymbol{H}_T$, $\widehat{\boldsymbol{F}}_T = T^{-1}\sum_{t=1}^T \nabla_{\boldsymbol{\theta}}\boldsymbol{f}(\boldsymbol{\eta}_t;\hat{\boldsymbol{\theta}}_T)$, and

$$\widehat{\boldsymbol{V}}_T = \boldsymbol{I}_q - \widehat{\boldsymbol{\Lambda}}_T'\widehat{\boldsymbol{F}}_T[\widehat{\boldsymbol{F}}_T'\boldsymbol{H}_T\widehat{\boldsymbol{F}}_T]^{-1}\widehat{\boldsymbol{F}}_T'\widehat{\boldsymbol{\Lambda}}_T.$$

Note that $\widehat{\boldsymbol{U}}_T$ is of rank $q-p$ for all T.

The normalizing matrix $\widehat{\boldsymbol{\Gamma}}_T$ then leads to the robust M test:

$$\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \widehat{\boldsymbol{\Gamma}}_T^+) = T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)'\widehat{\boldsymbol{\Gamma}}_T^+\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T).$$

The following limits are immediate once we show that $\widehat{\boldsymbol{\Gamma}}_T^+ \Rightarrow \big(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U}\big)^+$.

Theorem 3.2 *Given [A1]-[A3], we have under the local alternative (2) that*

$$\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \widehat{\boldsymbol{\Gamma}}_T^+) \xrightarrow{D} [\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{W}_{q-p}(1)]'\boldsymbol{P}_{q-p}^{-1}[\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{W}_{q-p}(1)],$$

where $\boldsymbol{A}$ is a $q\times(q-p)$ matrix with $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I}_{q-p}$ and $\boldsymbol{\Delta}$ is a $(q-p)\times(q-p)$ diagonal matrix with positive diagonal elements such that $\boldsymbol{U}'\boldsymbol{S}\boldsymbol{S}'\boldsymbol{U} = \boldsymbol{A}\boldsymbol{\Delta}^2\boldsymbol{A}'$. Under the null hypothesis that $\boldsymbol{\delta}_o = \boldsymbol{0}$, $\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \widehat{\boldsymbol{\Gamma}}_T^+) \xrightarrow{D} \boldsymbol{W}_{q-p}(1)'\boldsymbol{P}_{q-p}^{-1}\boldsymbol{W}_{q-p}(1)$.

While the robust M tests of Kuan and Lee (2006) converge weakly to $\boldsymbol{W}_q(1)'\boldsymbol{P}_q^{-1}\boldsymbol{W}_q(1)$ under the null, the null limit of $\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \widehat{\boldsymbol{\Gamma}}_T^+)$ is of a similar form but depends on the number of over-identifying restrictions, $q-p$. It is worth mentioning that $\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \widehat{\boldsymbol{\Gamma}}_T^+)$ does not require recursive estimation and hence is computationally simpler than the robust M test based on $\widetilde{\boldsymbol{C}}_T$ proposed by Kuan and Lee (2006).

3.2 Extension to Kernel-Based OIR Tests

Analogous to Kiefer and Vogelsang (2002a), it can be easily shown that $\widehat{\boldsymbol{C}}_T$ is algebraically equivalent to one half of the Bartlett-kernel-based covariance matrix estimator without truncation, i.e., $\widehat{\boldsymbol{\Sigma}}_{\ell(T)}^B = 2\widehat{\boldsymbol{C}}_T$ with $\ell(T) = T$. Moreover, other kernel-based covariance matrix estimators without truncation ($\widehat{\boldsymbol{\Sigma}}_{\ell(T)}^\kappa$ in (7) with $\ell(T) = T$) can be expressed as

$$\widehat{\boldsymbol{\Sigma}}_T^\kappa = \sum_{i=1}^{T-1}\sum_{j=1}^{T-1}\Big[(\kappa_{i,j} - \kappa_{i,j+1}) - (\kappa_{i+1,j} - \kappa_{i+1,j+1})\Big]\boldsymbol{\varphi}_i(\hat{\boldsymbol{\theta}}_T)\boldsymbol{\varphi}_j(\hat{\boldsymbol{\theta}}_T)', \tag{9}$$

where $\kappa_{i,j} = \kappa(|i-j|/T)$, as in Kiefer and Vogelsang (2002b) and Lee (2006). These results suggest that $\widehat{\boldsymbol{\Sigma}}_T^\kappa$ is closely related to the normalizing matrix in the KVB approach.

To be specific, we consider the kernel function κ that satisfies the following conditions.

[A4] The kernel function κ is such that:

(a) $\kappa(0) = 1$, $\kappa(z) = \kappa(-z)$ and $|\kappa(z)| \leq 1$ for all $z \in \mathbb{R}$, $\int_{-\infty}^{\infty} \kappa(z)^2 \, \mathrm{d}z < \infty$, κ is continuous at the origin;

(b) $\int_{-\infty}^{\infty} \kappa(z) \exp(-iz\lambda) \geq 0$ for all $\lambda \in \mathbb{R}$, where i is the imaginary unit;

(c) κ is twice continuously differentiable with the second order derivative κ''.

The conditions [A4](a) and (b) are standard in the literature and admit the Bartlett, Parzen, and quadratic spectral kernels; see, e.g., Andrews (1991). In particular, [A4](b) ensures that the kernel-based covariance matrix estimator is positive semi-definite for all samples, but it rules out the Tukey-Hanning kernel. Kiefer and Vogelsang (2002b) and Vogelsang (2003) also require [A4](c) for kernel-based normalizing matrices, yet this condition rules out the Bartlett kernel. The lemma below suggests that $\widehat{\boldsymbol{\Sigma}}_T^{\kappa}$ may play the role as $\widehat{\boldsymbol{C}}_T$; cf. Lemma 3.1.

Lemma 3.3 *Given [A1]–[A4], we have under the local alternative (2) that*

$$\widehat{\boldsymbol{\Sigma}}_T^{\kappa} \Rightarrow \boldsymbol{S}\boldsymbol{P}_{\kappa,q}\boldsymbol{S}',$$

where $\boldsymbol{P}_{\kappa,q} = \int_0^1 \int_0^1 \kappa''(r-s)\boldsymbol{B}_q(r)\boldsymbol{B}_q(s)' \, dr \, ds = \int_0^1 \int_0^1 \kappa(r-s) \, d\boldsymbol{B}_q(r) \, d\boldsymbol{B}_q(s)'$.

Remark: Although the Bartlett kernel is excluded by [A4](c), Lemma 3.3 remains valid for $\widehat{\boldsymbol{\Sigma}}_T^{B}$ because $\widehat{\boldsymbol{\Sigma}}_T^{B} = 2\widehat{\boldsymbol{C}}_T \Rightarrow \boldsymbol{S}\boldsymbol{P}_{B,q}\boldsymbol{S}'$ with $\boldsymbol{P}_{B,q-p} = 2\int_0^1 \boldsymbol{B}_{q-p}(r)\boldsymbol{B}_{q-p}(r)' \, \mathrm{d}r$.

Analogous to $\widehat{\boldsymbol{\Gamma}}_T$ in the preceding subsection, we can consider the kernel-based normalizing matrix: $\widehat{\boldsymbol{\Gamma}}_T^{\kappa} = \widehat{\boldsymbol{U}}_T'\widehat{\boldsymbol{\Sigma}}_T^{\kappa}\widehat{\boldsymbol{U}}_T$. While depending on the selected kernel function, this normalizing matrix avoids choosing a truncation lag. With $\widehat{\boldsymbol{\Gamma}}_T^{\kappa}$, we obtain a class of robust OIR tests:

$$\mathcal{J}\big(\hat{\boldsymbol{\theta}}_T, (\widehat{\boldsymbol{\Gamma}}_T^{\kappa})^+\big) = T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)'(\widehat{\boldsymbol{\Gamma}}_T^{\kappa})^+\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T);$$

note that for the Bartlett kernel B, $\mathcal{J}\big(\hat{\boldsymbol{\theta}}_T, (\widehat{\boldsymbol{\Gamma}}_T^{B})^+\big) = 2\mathcal{J}\big(\hat{\boldsymbol{\theta}}_T, \widehat{\boldsymbol{\Gamma}}_T^{+}\big)$. The limiting behavior of this class of tests is established in the result below. In view of the remark after Lemma 3.3, this result in fact includes Theorem 3.2 as a special case.

Theorem 3.4 *Given [A1]-[A4], we have under the local alternative (2) that*

$$\mathcal{J}\big(\hat{\boldsymbol{\theta}}_T, (\widehat{\boldsymbol{\Gamma}}_T^{\kappa})^+\big) \xrightarrow{D} [\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{W}_{q-p}(1)]'\boldsymbol{P}_{\kappa,q-p}^{-1}[\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{W}_{q-p}(1)],$$

where $\boldsymbol{A}$ and $\boldsymbol{\Delta}$ are defined in Theorem 3.2. Under the null that $\boldsymbol{\delta}_o = \boldsymbol{0}$, $\mathcal{J}\big(\hat{\boldsymbol{\theta}}_T, (\widehat{\boldsymbol{\Gamma}}_T^{\kappa})^+\big) \xrightarrow{D}$ $\boldsymbol{W}_{q-p}(1)'\boldsymbol{P}_{\kappa,q-p}^{-1}\boldsymbol{W}_{q-p}(1)$.

Remarks:

1. Note that the proposed robust OIR test requires only a consistent GMM estimator, in contrast with the conventional OIR test of Hansen (1982) which requires the optimal GMM estimator. As such, the proposed test is easy to implement and can serve to evaluate GMM models and determine if the optimal estimation is worthwhile under present specification.

2. When κ is the Bartlett kernel, the (asymptotic) critical values of $\mathcal{J}\big(\hat{\boldsymbol{\theta}}_T, (\widehat{\boldsymbol{\Gamma}}_T^{\kappa})^+\big)$ with different $q-p$ are available from Lobato (2001, Table 1). As for other κ and different $q-p$, the critical values can be obtained via simulations; for $q-p=1$, the square root of the critical values can be found in Kiefer and Vogelsang (2002b, Table 1) and Phillips, Sun, and Jin (2006, Table 6).

4 Monte Carlo Simulations

In this section, the finite sample performance of the proposed $\mathcal{J}\big(\hat{\boldsymbol{\theta}}_T, (\widehat{\boldsymbol{\Gamma}}_T^{\kappa})^+\big)$ test is evaluated via Monte Carlo simulations. We consider two nominal sizes: 5% and 10%, the samples $T = 50$, 100, and 500 for size simulations, and $T = 50$ and 100 for power simulations. The number of replications is 10,000 for all simulations. As the results under different nominal sizes are qualitatively similar, we report only the results for 5% nominal size; the results for 10% nominal size are available from the authors upon request.

For the proposed test, we adopt six different kernel functions: Bartlett (B), quadratic spectral (QS), Daniel (D), Parzen (P), and exponentiated Parzen (EP) with power $\rho = 8$ and 32 (EP08 and EP32). The EP kernel, proposed by Phillips et al. (2006), is obtained by exponentiating the Parzen kernel with power $\rho \geq 1$ and satisfies [A4].[2] We consider the EP kernel because, as shown by Phillips et al. (2006), it delivers faster rate of convergence of the covariance matrix estimator and yields test power close to the power envelope when $\rho = 32$. For these tests, we employ the GMM estimator based on the identity weighting matrix. For comparison, we simulate the conventional $\mathcal{J}^*$ test of Hansen (1982) which

[2]The EP kernel with power ρ is

$$\mathrm{EP}_\rho(x) = \begin{cases} (1-6x^2+6|x|^3)^\rho, & 0 \leq |x| \leq 1/2, \\ (2(1-|x|)^3)^\rho, & 1/2 \leq |x| \leq 1, \\ 0, & \text{otherwise}, \end{cases}$$

and it reduces to the Parzen kernel when $\rho = 1$.

is based on the consistent estimator $\widetilde{\boldsymbol{\Sigma}}^{B}_{\ell(T)}$. This is a "centered" version of the Bartlett-kernel-based covariance matrix estimator, as recommended by Hall (2000). To compute $\widetilde{\boldsymbol{\Sigma}}^{B}_{\ell(T)}$, the identity-matrix based GMM estimator is used as a preliminary GMM estimator for θ_o. As for the selection of $\ell(T)$, we follow Hall (2000) and adopt the nonparametric method of Newey and West (1994) with the weighting vector: $[1\ -1]'$ and two preliminary truncation lags: $[c(T/100)^{2/9}]$, where $c = 4$ and 12; see Hall (2000, pp. 1522–1523) for more detail. The resulting OIR tests are denoted as $\mathcal{J}^*_{nw,4}$ and $\mathcal{J}^*_{nw,12}$.

Similar to Hall (2000), we consider the GMM specification:

$$\mathbb{E}[\boldsymbol{z}_t(y_t - \theta x_t)] = \mathbf{0}, \tag{10}$$

where y_t and x_t are random variables, $\boldsymbol{z}_t = [z_{1,t}\ \ z_{2,t}]'$ is a 2×1 random vector, and θ is an unknown parameter. The data generating processes (DGPs) for y_t and x_t are

$$y_t = x_t + \gamma z_{1,t} + e_t,$$
$$x_t = z_{1,t} + z_{2,t} + u_t,$$

where e_t and u_t are random errors. Let $\boldsymbol{\xi}_t = [z_{1,t}\ \ z_{2,t}\ \ e_t\ \ u_t]'$. Then the data for $\boldsymbol{\xi}_t$ are generated according to the VAR(1) model: $\boldsymbol{\xi}_t = a\boldsymbol{\xi}_{t-1} + \boldsymbol{v}_t$, where $a \in (-1, 1)$ is a scalar parameter and $\boldsymbol{v}_t \sim$ i.i.d. $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{v}})$ with the diagonal elements of $\boldsymbol{\Sigma}_{\boldsymbol{v}}$ being $1 - a^2$ and nonzero off-diagonal elements being such that $\mathrm{corr}(z_{1,t}, z_{2,t}) = \mathrm{corr}(e_t, u_t) = 0.5$. Note that this DGP reduces to that of Hall (2000) when $a = 0$. In this study, we set $a = 0$, 0.5, 0.8, 0.9, -0.5 so that we can examine the effect of serial correlations on the performance of the proposed tests. When $\gamma = 0$, it is easy to see that the model (10) is correctly specified in the sense that there exists a unique $\theta_o = 1$ such that $\mathbb{E}[\boldsymbol{z}_t(y_t - \theta_o x_t)] = \mathbf{0}$. For $\gamma \neq 0$, the model (10) is misspecified. We thus consider $\gamma = 0$ for size simulations and various γ in $(0, 2]$ for power simulations.

The empirical sizes of these tests are reported in Table 1. It is found that when the data possess no or moderate serial correlation (i.e., $a = 0$ or 0.5), the empirical sizes of all $\mathcal{J}\big(\hat{\boldsymbol{\theta}}_T, (\widehat{\boldsymbol{\Gamma}}^{\kappa}_T)^+\big)$ tests are quite close to the nominal size 5% even when $T = 50$. When a gets larger such that data become more persistent, the proposed tests with QS, P, and D remain properly sized, but those with B, EP08 and EP32 tend to be over-sized when T is small; the size distortions of the latter tests diminish quickly with T. For example, when $a = 0.8$ (0.9), $T = 100$ (500) is enough for these tests to be properly sized. Note that the proposed test based on the EP kernel seems to be robust to ρ under moderate serial correlation, but it has more size distortion for $\rho = 32$ when the data are highly persistent and T is small. By contrast, $\mathcal{J}^*_{nw,c}$ are over-sized for all samples, and the distortions do

Table 1: Empirical sizes of the OIR tests.

DGP	T	Proposed test with κ						Conventional test	
		B	QS	D	P	$EP08$	$EP32$	$\mathcal{J}^*_{nw,4}$	$\mathcal{J}^*_{nw,12}$
$a = 0.0$	50	4.65	4.47	4.50	4.44	4.91	5.00	9.81	19.66
	100	4.89	4.65	4.67	4.67	4.66	4.98	7.25	13.51
	500	5.08	4.86	4.89	5.00	5.43	5.47	5.76	7.14
$a = 0.5$	50	5.28	4.88	4.92	4.63	4.69	5.51	12.11	20.69
	100	4.99	4.90	4.90	4.76	4.59	5.02	9.87	15.15
	500	5.13	5.09	5.08	4.60	4.89	4.69	7.30	7.84
$a = 0.8$	50	6.37	4.79	4.78	4.76	5.09	7.89	19.05	24.83
	100	5.78	4.79	4.77	4.43	4.61	5.61	14.25	17.96
	500	5.24	4.74	4.76	4.96	5.05	5.13	8.70	10.07
$a = 0.9$	50	8.83	5.20	5.17	5.70	7.30	14.40	28.39	31.41
	100	7.14	4.47	4.47	4.52	5.59	8.63	21.19	22.53
	500	5.82	5.14	5.19	5.23	4.75	5.00	11.19	10.81
$a = -0.5$	50	5.62	5.04	5.05	4.80	5.12	5.52	12.28	20.65
	100	5.16	4.70	4.71	5.07	4.73	5.05	9.55	15.12
	500	5.17	5.25	5.29	5.33	5.17	5.29	7.33	7.77

Note: The entries are rejection frequencies in percentage; the nominal size is 5%.

not disappear even when $T = 500$. For a given sample, the size distortions deteriorate rapidly when a increases. This shows that $\mathcal{J}^*_{nw,c}$ is not quite robust to serial correlation of unknown form. Also, the size distortions depend on the user-chosen parameter c and are much larger for $c = 12$.

To provide a proper comparison between the power performance of different tests, we simulate the size-adjusted powers. We first examine the effect of kernel function on the power of the proposed test and plot the power curves in Figure 1, with γ on the horizontal axis. We consider the cases that $a = 0.5, 0.9$ and $T = 50, 100$. It can be seen that in all cases, the tests with the EP kernel dominate the tests with other kernels (the one with EP32 has the highest power), and the test with the Bartlett kernel dominates those with other conventional kernels. As for the effect of persistence in data, we observe that the test powers are lower for a larger a, yet the ranking of power performance is not altered.

Given that the proposed test with EP32 enjoys power advantage over the tests with other kernel functions, we further evaluate its power performance relative to the conventional OIR tests: $\mathcal{J}^*_{nw,4}$ and $\mathcal{J}^*_{nw,12}$. As $\mathcal{J}^*_{nw}$ tests are based on the Bartlett kernel, we also consider the proposed test with the same kernel. The power curves of these tests are plotted in Figure 2. It can be seen that the performance of $\mathcal{J}^*_{nw,c}$ depends on the user-chosen parameter c, and the test with $c = 4$ is more powerful than the test with $c = 12$. It has been documented in the literature that the tests with a KVB-type normalizing matrix typically suffer from power loss, even though they are properly sized. Nonetheless, it can be seen from Figure 2 that, for moderately correlated data (e.g., $a = 0.5$), the test with EP32 performs similarly to $\mathcal{J}^*_{nw,4}$ (with very minor power loss) and hence is more powerful than $\mathcal{J}^*_{nw,12}$ in both samples. The test with B, on the other hand, has more power loss relative to $\mathcal{J}^*_{nw,4}$, but it still outperforms $\mathcal{J}^*_{nw,12}$ in a small sample (it has the lowest power in a larger sample). When the data are highly persistent (e.g., $a = 0.9$), the tests with EP32 and B in a small sample are even more powerful than $\mathcal{J}^*_{nw,4}$ for large γ (so that the DGP is far away from the null). In a larger sample, the test with EP32 performs similarly to $\mathcal{J}^*_{nw,4}$; the test with B is less powerful but performs similarly to $\mathcal{J}^*_{nw,12}$. These results together suggest that the proposed tests, especially the one with EP32, can serve as practically useful diagnostic tools for testing OIR.

5 Conclusions

In this paper, the KVB approach is modified to construct a class of robust specification tests for OIR in the context of GMM. The proposed test does not require consistent estimation of the asymptotic covariance matrix so that it avoids the problems arising from nonparametric kernel estimation. Moreover, the proposed test is computationally simple. First, it does not require optimal GMM estimation, in contrast with the conventional OIR test. Second, it does not require recursive estimation, in contrast with the robust M test of Kuan and Lee (2006). The proposed test is thus practically useful because it can serve as a preliminary check of GMM models without going through the more complex process of optimal estimation. The simulation results also confirm that the proposed test is properly sized and may have power advantage relative to the conventional OIR test with an inappropriate user-chosen parameter. In particular, with a properly selected kernel function (e.g., EP32), our test may even outperform the conventional OIR test in terms of finite sample power.

Appendix

Proof of Theorem 2.1: As $\ddot{\boldsymbol{H}}_T \xrightarrow{\mathbb{P}} \ddot{\boldsymbol{H}}_o$, the limit of $\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \ddot{\boldsymbol{H}}_T)$ follows immediately from (6) and the continuous mapping theorem. For the second assertion, note that when $\boldsymbol{H}_o = \ddot{\boldsymbol{H}}_o = \boldsymbol{\Sigma}_o^{-1}$, $\boldsymbol{\Lambda} = \boldsymbol{S}'^{-1}$, $\boldsymbol{V}_{\boldsymbol{\Lambda}'\boldsymbol{F}_o} = \boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}$, and $\boldsymbol{U} = \boldsymbol{S}'^{-1}\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}\boldsymbol{S}'$. Then by (6), $T^{1/2}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T^*) \Rightarrow \boldsymbol{S}\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}\boldsymbol{S}^{-1}[\boldsymbol{\delta}_o + \boldsymbol{S}\boldsymbol{W}_q(1)]$ and

$$
\begin{aligned}
\mathcal{J}^* \xrightarrow{D} & \big[\boldsymbol{S}^{-1}\boldsymbol{\delta}_o + \boldsymbol{W}_q(1)\big]'\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}\boldsymbol{S}'\boldsymbol{\Sigma}_o^{-1}\boldsymbol{S}\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}\big[\boldsymbol{S}^{-1}\boldsymbol{\delta}_o + \boldsymbol{W}_q(1)\big] \\
= & \big[\boldsymbol{S}^{-1}\boldsymbol{\delta}_o + \boldsymbol{W}_q(1)\big]'\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}\big[\boldsymbol{S}^{-1}\boldsymbol{\delta}_o + \boldsymbol{W}_q(1)\big],
\end{aligned}
$$

because $\boldsymbol{S}'\boldsymbol{\Sigma}_o^{-1}\boldsymbol{S} = \boldsymbol{I}$ and $\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}$ is symmetric and idempotent. Given $\operatorname{rank}(\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}) = q-p$, we have from Theorem 9.11 of Scott (1997, p. 381) that the quadratic form on the right-hand side above is distributed as a non-central χ^2 distribution with $q-p$ degrees of freedom and the non-centrality parameter: $\boldsymbol{\delta}_o'\boldsymbol{S}'^{-1\prime}\boldsymbol{V}_{\boldsymbol{S}^{-1}\boldsymbol{F}_o}\boldsymbol{S}^{-1}\boldsymbol{\delta}_o$. $\square$

Proof of Lemma 3.1: Setting $t = [rT]$, $0 < r \le 1$, the first-order Taylor expansion of $T^{1/2}\boldsymbol{m}_{[rT]}(\hat{\boldsymbol{\theta}}_T)$ about $\boldsymbol{\theta}_o$ is

$$
\begin{aligned}
\sqrt{T}\boldsymbol{m}_{[rT]}(\hat{\boldsymbol{\theta}}_T) &= \sqrt{T}\boldsymbol{m}_{[rT]}(\boldsymbol{\theta}_o) + \frac{[rT]}{T}\boldsymbol{F}_{[rT]}(\boldsymbol{\theta}_o)\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_o) + o_{\mathbb{P}}(1) \\
&= \sqrt{T}\boldsymbol{m}_{[rT]}(\boldsymbol{\theta}_o) + \frac{[rT]}{T}\boldsymbol{F}_o\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_o) + o_{\mathbb{P}}(1),
\end{aligned}
$$

where $\boldsymbol{F}_{[rT]}(\boldsymbol{\theta}_o)$ and $\boldsymbol{F}_o$ are defined in [A3]. It follows from [A2] that

$$
\begin{aligned}
\boldsymbol{\varphi}_{[rT]}(\hat{\boldsymbol{\theta}}_T) &= \sqrt{T}\boldsymbol{m}_{[rT]}(\hat{\boldsymbol{\theta}}_T) - \frac{[rT]}{T}\sqrt{T}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T) \\
&= \sqrt{T}\boldsymbol{m}_{[rT]}(\boldsymbol{\theta}_o) - \frac{[rT]}{T}\sqrt{T}\boldsymbol{m}_T(\boldsymbol{\theta}_o) + o_{\mathbb{P}}(1) \\
&\Rightarrow \boldsymbol{S}\boldsymbol{B}_q(r),
\end{aligned}
$$

regardless of the values of $\boldsymbol{F}_o$ and $\boldsymbol{\delta}_o$. The first assertion on $\widehat{\boldsymbol{C}}_T$ now follows from the continuous mapping theorem.

As for $\widetilde{\boldsymbol{C}}_T$, it can be shown that, analogous to (5),

$$
\sqrt{T}\boldsymbol{m}_{[rT]}(\tilde{\boldsymbol{\theta}}_{[rT]}) = \boldsymbol{U}'\sqrt{T}\boldsymbol{m}_{[rT]}(\boldsymbol{\theta}_o) + o_{\mathbb{P}}(1).
$$

By [A2], $T^{1/2}\boldsymbol{m}_{[rT]}(\tilde{\boldsymbol{\theta}}_{[rT]}) \Rightarrow \boldsymbol{U}'\big[r\boldsymbol{\delta}_o + \boldsymbol{S}\boldsymbol{W}_q(r)\big]$, and hence

$$
\tilde{\boldsymbol{\varphi}}_{[rT]} = \sqrt{T}\boldsymbol{m}_{[rT]}(\tilde{\boldsymbol{\theta}}_{[rT]}) - \frac{[rT]}{T}\sqrt{T}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T) \Rightarrow \boldsymbol{U}'\boldsymbol{S}[\boldsymbol{W}_q(r) - r\boldsymbol{W}_q(1)] = \boldsymbol{U}'\boldsymbol{S}\boldsymbol{B}_q(r).
$$

The weak limit of $\widetilde{\boldsymbol{C}}_T$ also follows from the continuous mapping theorem. $\square$

Proof of Equation (8): By the singular value decomposition, we have $\boldsymbol{U}'\boldsymbol{S} = \boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{B}'$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are $q \times (q-p)$ matrices such that $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{B}'\boldsymbol{B} = \boldsymbol{I}_{q-p}$, and $\boldsymbol{\Delta}$ is a $(q-p)\times(q-p)$ diagonal matrix with positive diagonal elements. As $\boldsymbol{U}'\boldsymbol{S}\boldsymbol{S}'\boldsymbol{U} = \boldsymbol{A}\boldsymbol{\Delta}^2\boldsymbol{A}'$, it is easily seen that $\boldsymbol{U}'\boldsymbol{S}\boldsymbol{W}_q(r) \stackrel{d}{=} \boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{W}_{q-p}(r)$ and $\boldsymbol{U}'\boldsymbol{S}\boldsymbol{B}_q(r) \stackrel{d}{=} \boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{B}_{q-p}(r)$. Hence,

$$\begin{aligned}
\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U} &= \int_0^1\int_0^1 \boldsymbol{U}'\boldsymbol{S}\boldsymbol{B}_q(r)\boldsymbol{B}_q(s)'\boldsymbol{S}'\boldsymbol{U}\,\mathrm{d}r\,\mathrm{d}s \\
&\stackrel{d}{=} \left[\int_0^1\int_0^1 \boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{B}_{q-p}(r)\boldsymbol{B}_{q-p}(s)'\boldsymbol{\Delta}\boldsymbol{A}'\,\mathrm{d}r\,\mathrm{d}s\right] \\
&= \boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{P}_{q-p}\boldsymbol{\Delta}\boldsymbol{A}'.
\end{aligned}$$

As shown in Kuan and Lee (2006), taking the generalized inverse of both sides of the equation above yields

$$(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U})^+ \stackrel{d}{=} \boldsymbol{A}\boldsymbol{\Delta}^{-1}\boldsymbol{P}_{q-p}^{-1}\boldsymbol{\Delta}^{-1}\boldsymbol{A}'.$$

Under the null, $T^{1/2}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T) \Rightarrow \boldsymbol{U}'\boldsymbol{S}\boldsymbol{W}_q(1)$ by (6), so that

$$\begin{aligned}
&T\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T)'(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U})^+\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T) \\
&\qquad \Rightarrow \boldsymbol{W}_q(1)'\boldsymbol{S}'\boldsymbol{U}(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U})^+\boldsymbol{U}'\boldsymbol{S}\boldsymbol{W}_q(1) \\
&\qquad \stackrel{d}{=} \boldsymbol{W}_{q-p}(1)\boldsymbol{\Delta}\boldsymbol{A}'\boldsymbol{A}\boldsymbol{\Delta}^{-1}\boldsymbol{P}_{q-p}^{-1}\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{W}_{q-p}(1) \\
&\qquad = \boldsymbol{W}_{q-p}(1)'\boldsymbol{P}_{q-p}^{-1}\boldsymbol{W}_{q-p}(1). \quad \square
\end{aligned}$$

Proof of Theorem 3.2: Clearly, $\widehat{\boldsymbol{\Gamma}}_T \Rightarrow \boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U}$. By Lemma 3.1, $\boldsymbol{P}_q$ is positive definite with probability one, hence so is $\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'$. Then, $\mathrm{rank}(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U}) = q-p$. As $\widehat{\boldsymbol{C}}_T$ converges to $\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'$ by Lemma 3.1, it must have rank q for all T large. Given that $\widehat{\boldsymbol{U}}$ is of rank $q-p$ for all T, $\widehat{\boldsymbol{\Gamma}}_T$ is thus of rank $q-p$ and satisfies the rank condition for the continuity of the Moore-Penrose generalized inverse; see, e.g., Scott (1997, 188–190). It follows that $(\widehat{\boldsymbol{\Gamma}}_T)^+ \Rightarrow (\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U})^+$. In view of the proof of equation (8), we know there exist $\boldsymbol{A}$ and $\boldsymbol{\Delta}$ such that

$$\sqrt{T}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T) \Rightarrow \boldsymbol{U}'[\boldsymbol{\delta} + \boldsymbol{S}\boldsymbol{W}_q(1)] \stackrel{d}{=} \boldsymbol{U}'\boldsymbol{\delta} + \boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{W}_{q-p}(1),$$

and $(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U}')^+ \stackrel{d}{=} \boldsymbol{A}\boldsymbol{\Delta}^{-1}\boldsymbol{P}_{q-p}^{-1}\boldsymbol{\Delta}^{-1}\boldsymbol{A}'$. It follows from the continuous mapping theorem that

$$\mathcal{J}(\hat{\boldsymbol{\theta}}_T, \widehat{\boldsymbol{\Gamma}}_T^{+}) \xrightarrow{D} [\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{U}'\boldsymbol{S}\boldsymbol{W}_q(1)]'\big(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_q\boldsymbol{S}'\boldsymbol{U}\big)^{+}[\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{U}'\boldsymbol{S}\boldsymbol{W}_q(1)]$$
$$\stackrel{d}{=} [\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{W}_{q-p}(1)]'\boldsymbol{A}\boldsymbol{\Delta}^{-1}\boldsymbol{P}_{q-p}^{-1}\boldsymbol{\Delta}^{-1}\boldsymbol{A}'[\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{W}_{q-p}(1)]$$
$$= [\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{W}_{q-p}(1)]'\boldsymbol{P}_{q-p}^{-1}[\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{W}_{q-p}(1)].$$

The null limit follows immediately by setting $\boldsymbol{\delta} = \boldsymbol{0}$. $\square$

Proof of Lemma 3.3: As in the proof of Lemma 3.1, $\boldsymbol{\varphi}_{[rT]}(\hat{\boldsymbol{\theta}}_T) \Rightarrow \boldsymbol{S}\boldsymbol{B}_q(r)$. When the kernel function κ is twice continuously differentiable with the second order derivative κ'', it is easily shown that

$$T^2\Big[(\kappa_{[rT],[sT]} - \kappa_{[rT],[sT]+1}) - (\kappa_{[rT]+1,[sT]} - \kappa_{[rT]+1,[sT]+1})\Big] \to -\kappa''(r-s),$$

uniformly in r and s; see Kiefer and Vogelsang (2002b, pp. 1364–1365). It then follows from (9) and the continuous mapping theorem that

$$\widehat{\boldsymbol{\Sigma}}_T^{\kappa} \Rightarrow \boldsymbol{S}\bigg[\int_0^1\int_0^1 -\kappa''(r-s)\boldsymbol{B}_q(r)\boldsymbol{B}_q(s)'\,\mathrm{d}r\,\mathrm{d}s\bigg]\boldsymbol{S}'.$$

The second equality of $\boldsymbol{P}_{\kappa,q}$ follows from integration by parts, as in Phillips et al. (2006, pp. 890–891). $\square$

Proof of Theorem 3.4: The proof is analogous to that of Theorem 3.2. Given Lemma 3.3, it can be shown that $\widehat{\boldsymbol{\Gamma}}_T^{\kappa} \Rightarrow \boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_{\kappa,q}\boldsymbol{S}'\boldsymbol{U}$ and $(\widehat{\boldsymbol{\Gamma}}_T^{\kappa})^{+} \Rightarrow (\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_{\kappa,q}\boldsymbol{S}'\boldsymbol{U})^{+}$. Moreover, equation (8) carries over to the present case:

$$\boldsymbol{W}_q(1)'\boldsymbol{S}'\boldsymbol{U}\big(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_{\kappa,q}\boldsymbol{S}'\boldsymbol{U}\big)^{+}\boldsymbol{U}'\boldsymbol{S}\boldsymbol{W}_q(1) \stackrel{d}{=} \boldsymbol{W}_{q-p}(1)'\boldsymbol{P}_{\kappa,q-p}^{-1}\boldsymbol{W}_{q-p}(1).$$

We also have

$$\sqrt{T}\boldsymbol{m}_T(\hat{\boldsymbol{\theta}}_T) \Rightarrow \boldsymbol{U}'\boldsymbol{\delta} + \boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{W}_{q-p}(1),$$

and $(\boldsymbol{U}'\boldsymbol{S}\boldsymbol{P}_{\kappa,q}\boldsymbol{S}'\boldsymbol{U}')^{+} \stackrel{d}{=} \boldsymbol{A}\boldsymbol{\Delta}^{-1}\boldsymbol{P}_{\kappa,q-p}^{-1}\boldsymbol{\Delta}^{-1}\boldsymbol{A}'$. Therefore,

$$\mathcal{J}\big(\hat{\boldsymbol{\theta}}_T, (\widehat{\boldsymbol{\Gamma}}_T^{\kappa})^{+}\big) \xrightarrow{D} [\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{W}_{q-p}(1)]'\boldsymbol{P}_{\kappa,q-p}^{-1}[\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{U}'\boldsymbol{\delta}_o + \boldsymbol{W}_{q-p}(1)],$$

and the null limit follows by setting $\boldsymbol{\delta} = \boldsymbol{0}$. $\square$

References

Andrews, D. W. K. (1987). Asymptotic results for generalized Wald tests, *Econometric Theory*, **3**, 348–358.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica*, **59**, 817–858.

Bunzel, H., N. M. Kiefer and T. J. Vogelsang (2001). Simple robust testing of hypotheses in nonlinear models, *Journal of the American Statistical Association*, **96**, 1088–1096.

den Haan, W. J. and A. T. Levin (1997). A practitioners's guide to robust covariance matrix estimation, in *Handbook of Statistics: Robust Inference*, edited by G. S. Maddala and C. R. Rao, Amsterdam: Elsevier, 299–342.

Hall, A. R. (1999). Hypothesis testing in models estimated by GMM, *Generalized Method of Moments Estimation*, ed. by L. Mátyás, Cambridge: Cambridge University Press.

Hall, A. R. (2000). Covariance matrix estimation and the power of the overidentifying restrictions test, *Econometrica*, **68**, 1517–1527.

Hall, A. R., A. Inoue, and F. P. M. Peixe (2003). Covariance matrix estimation and the limiting behavior of the overidentifying restrictions test in the presence of neglected structural instability, *Econometric Theory*, **19**, 962–983.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators, *Econometrica*, **50**, 1029–1054.

Kiefer, N. M. and T. J. Vogelsang (2002a). Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation, *Econometrica*, **70**, 2093–2095.

Kiefer, N. M. and T. J. Vogelsang (2002b). Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size, *Econometric Theory*, **18**, 1350–1366.

Kiefer, N. M. and T. J. Vogelsang (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests, *Econometric Theory*, **21**, 1130–1164

Kiefer, N. M., T. J. Vogelsang and H. Bunzel (2000). Simple robust testing of regression hypothesis, *Econometrica*, **68**, 695–714.

Kuan, C.-M. and W.-M. Lee (2006). Robust M tests without consistent estimation of asymptotic covariance matrix, *Journal of the American Statistical Association*, forthcoming.

Lee, W.-M. (2006). Robust M tests using kernel-based estimators with bandwidth equal to sample size, Working Paper, Department of Economics, National Chung Cheng University.

Lobato, I. N. (2001). Testing that a dependent process is uncorrelated, *Journal of the American Statistical Association*, **96**, 1066–1076.

Newey, W. K. and K. D. West (1994). Automatic lag selection in covariance matrix estimation, *Review of Economic Studies*, **61**, 631–653.

Phillips, P. C. B., Y. Sun, and S. Jin (2006). Spectral density estimation and robust hypothesis testing using steep origin kernels without truncation, *International Economic Review*, **47**, 837–894.

Scott, J. R. (1997). *Matrix Analysis for Statistics*, New York: Wiley.

Vogelsang, T. J. (2003). Testing in GMM models without truncation, *Advances in Econometrics*, **17**, 199–233.

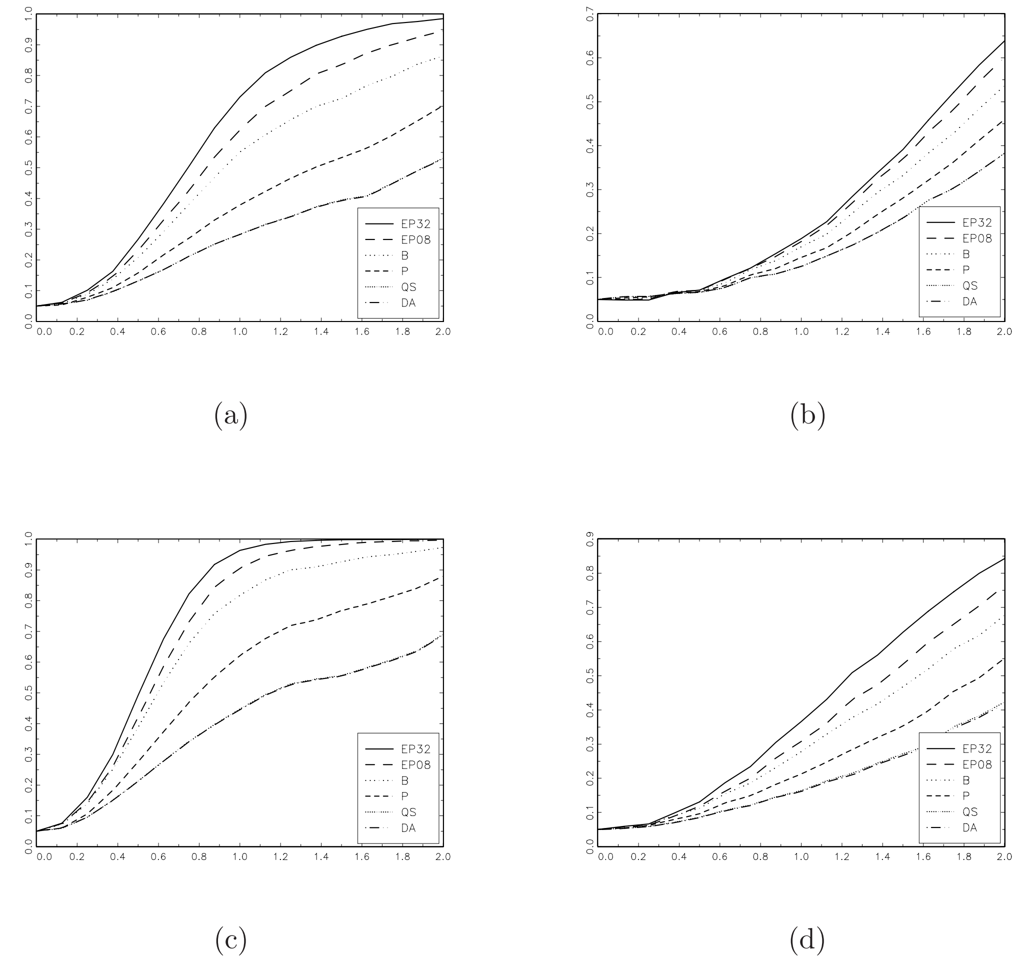
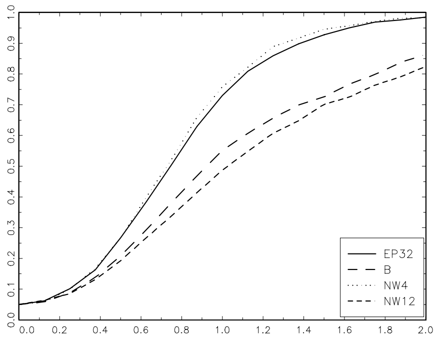
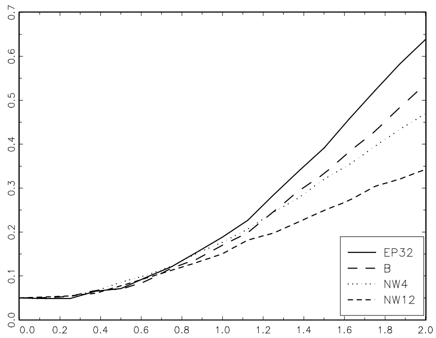


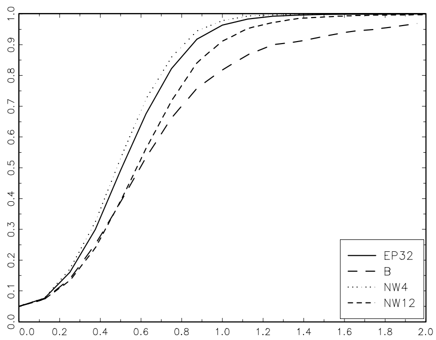
Figure 1: The size-adjusted powers of the proposed test with different kernels: (a) $a = 0.5$ and $T = 50$, (b) $a = 0.9$ and $T = 50$, (c) $a = 0.5$ and $T = 100$, and (d) $a = 0.9$ and $T = 100$.



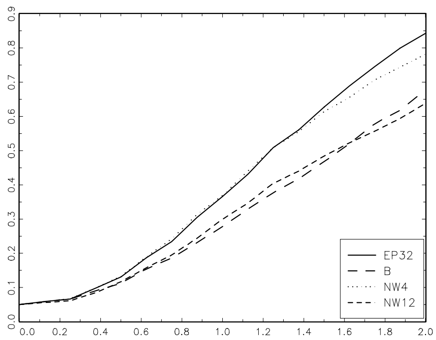
(a)



(b)



(c)



(d)

Figure 2: The size-adjusted powers of the proposed test with EP32 and B, $\mathcal{J}^*_{nw,4}$, and $\mathcal{J}^*_{nw,12}$: (a) $a = 0.5$ and $T = 50$, (b) $a = 0.9$ and $T = 50$, (c) $a = 0.5$ and $T = 100$, and (d) $a = 0.9$ and $T = 100$.

Number	Author(s)	Title	Date
06-A009	Wei-Ming Lee Chung-Ming Kuan	Testing Over-Identifying Restrictions without Consistent Estimation of the Asymptotic Covariance Matrix	09/06
06-A008	Chung-Ming Kuan. Yu-Wei Hsieh	Improved HAC Covariance Matrix Estimation Based on Forecast Errors	09/06
06-A007	Yu-Chin Hsu Chung-Ming Kuan	Change-Point Estimation of Nonstationary *I(d)* Processes	09/06
06-A006	Yuko Kinishita Chia-Hui Lu	On the Role of Absorptive Capacity: FDI Matters to Growth	08/06
06-A005	Kamhon Kan	Residential Mobility and Social Capital	07/06
06-A004	Kamhon Kan	Cigarette Smoking and Self –Control	07/06
06-A003	林忠正	懲罰怠惰、流失人才	06/06
06-A002	Shin-Kun Peng Takatoshi Tabuchi	Spatial Competition in Variety and Number of Stores	02/06
06-A001	Mamoru Kaneko J. Jude Kline	Inductive Game Theory: A Basic Scenario	01/06
05-A011	Amy R. Hwang	An Ecological-Economic Integrated General Equilibrium Model	12/05
05-A010	Juin-jen Chang Ming-fu Shaw Ching-chong Lai	A “Managerial” Trade Union and Economic Growth	12/05
05-A009	Lin-Ti Tan	Spatial Economic Theory of Pollution Control under Stochastic Emissions	10/05
05-A008	Kamhon KAN Wei-Der TSAI	Entrepreneurship and Risk Aversion	10/05
05-A007	江豐富	台灣縣市失業率的長期追蹤研究—1987- 2001	08/05
05-A006	Shin-Kun Peng, Jacques-Francois Thisse Ping Wang	Economic Integration and Agglomeration in a Middle Product Economy	06/05
05-A005	譚令蒂 洪乙禎 謝啓瑞	論藥價差	05/05
05-A004	Lin-Ti Tan Yan-Shu Lin	Spatial Monopoly Pricing in a Stochastic Environment	05/05
05-A003	Been-Lon Chen Shun-Fa Lee	Congestible Public Goods and Indeterminacy in a Two-sector Endogenous Growth Model*	03/05
05-A002	C. Y. Cyrus Chu Hung-Ken Chien	The Optimal Decoupled Liabilities: A General Analysis	02/05
05-A001	Cyrus C.Y. Chu Hung-Ken Chien	Durable-Goods Monopolists, Network Effects and Penetration Pricing	02/05

04-A015	Been-Lon Chen	Multiple Equilibria in a Growth Model with Habit Persistence	11/04
04-A014	C. Y. Cyrus Chu R. R. Yu Ruey S. Tsay	A New Model for Family Resource Allocation Among Siblings: Competition, Forbearance, and Support	05/04
04-A013	C. Y. Cyrus Chu Ruey S. Tsay Huoying Wu	Transmission of Sex Preferences Across Generations: The Allocation of Educational Resources Among Siblings	05/04
04-A012	C. Y. Cyrus Chu	Children as Refrigerators: When Would Backward Altruism Appear?	05/04
04-A011	Marcus Berliant Shin-Kun Peng Ping Wang	Welfare Analysis of Number and Locations of Local Public Facilities	05/04
04-A010	Daigee Shaw Yue-Mi Tsai	Assessing Alternative Policies for Reducing Household Waste in Taiwan	03/04
04-A009	Daigee Shaw Chin-Hsiung Loh Chin-Hsun Yeh Wen-Yu Jean Yen-lien Kuo	A Probabilistic Risk Analysis for Taipei Seismic Hazards: An Application of HAZ-Taiwan with its Pre-processor and Post-processor	03/04
04-A008	Yu-Lan Chien Cliff J. Huang Daigee Shaw	A General Model of Starting Point Bias in Double-Bounded Dichotomous Contingent Valuation Surveys	03/04
04-A007	鍾經樊 詹維玲 張光亮	財富在不同時期對台灣消費行爲的影響：多變量馬可夫結構轉換模型的應用	02/04
04-A006	Chun-chieh Huang Ching-Chong Lai Juin-Jen Chang	Working Hours Reduction and Endogenous Growth	02/04
04-A005	Juin-Jen Chang Ching-Chong Lai Ping Wang	On the Public Economics of Casino Gambling	02/04
04-A004	Ming-Fu Shaw Shu-Hua Chen Ching-Chong Lai Juin-Jen Chang	Interest Rate Rules, Target Policies, and Endogenous Economic Growth in an Open Economy	02/04
04-A003	Po-Hsuan Hsu Chung-Ming Kuan	Re-Examining the Profitability of Technical Analysis with White’s Reality Check	02/04