摂南経済研究
第1巻 第1･2号 (2011), 19－37ページ

研究論文

Stabilizing a GMM Bootstrap for Time Series: A Simulation Study

Masayuki Hirukawa

Abstract

Inoue and Shintani (2006) demonstrate that in order for their GMM bootstrap to achieve asymptotic refinements for symmetric two-sided confidence intervals and *J*-statistics of overidentifying restrictions, a kernel of order greater than two must be employed for HAC estimation. A well-known problem of employing a higher-order kernel for HAC estimation is that the resulting covariance estimate does not necessarily become positive semi-definite in finite samples, which often leads to unsatisfactory performance of the bootstrap. This paper proposes to stabilize the bootstrap through employing the nonparametric prewhitened HAC estimator by Xiao and Linton (2002) and Hirukawa (2006), which has the same bias property as a fourth-order kernel but always generates a positive semi-definite estimate in finite samples. Monte Carlo results indicate that the HAC estimator indeed stabilizes the GMM bootstrap.

Keywords: asymptotic refinements; block bootstrap; covariance estimation; generalized method of moments; Monte Carlo simulations; nonparametric prewhitening.

JEL classification numbers: C12; C22; C32.

1 Introduction

Since the seminal work by Hansen (1982), the generalized method of moments (“GMM”) has been a workhorse in the empirical analysis on macroeconomics and finance. However, a number of Monte Carlo experiments indicate rather poor finite sample performance of GMM: examples include Tauchen (1986), Kocherlakota (1990), Ferson and Foerster (1994), the articles in the special issue of the *Journal of Business & Economic Statistics* (vol.14, July 1996), and Smith (1999), among others. As regards finite sample performance of GMM-based inference, these articles find that empirical coverage frequencies of confidence intervals are often fewer than the nominal level, and that the chi-squared approximation to the *J*-statistic of overidentifying restrictions is poor.

There are two possible approaches to improve finite sample performance of the GMM-based inference. One is to provide better approximations to finite sample distributions of test statistics of interest via bootstrapping, and the other is to rely on alternatives to GMM, such as empirical likelihood (Kitamura, 1997) and exponential tilting (Kitamura and Stutzer, 1997). This paper adopts the former approach. More specifically, the paper aims at improving the bootstrap for two-step GMM estimation of overidentified linear models[1] proposed by Inoue and Shintani (2006; abbreviated as “IS” hereafter).

IS’s bootstrap is an extension of earlier work such as Hall and Horowitz (1996) and Andrews (2002a). While all these articles apply the technique of the moving block bootstrap (“MBB”) for weakly dependent data by Künsch (1989) to GMM estimation, IS differ from Hall and Horowitz (1996) and Andrews (2002a) in that IS leave the dependent structure of the moment function unspecified when establishing asymptotic refinements of their bootstrap. In order to deal with possibly infinite-order of dependence in the moment function, IS use the inverse of a heteroskedasticity and autocorrelation consistent (“HAC”) covariance estimator as the optimal weighting matrix for the efficient GMM estimation. IS further point out that the order of magnitude in the leading bias term of the HAC estimator is crucial to asymptotic refinements for the symmetric two-sided confidence interval and the *J*-statistic provided by their bootstrap. Finally, they demonstrate that a kernel of order greater than two (i.e. a higher-order kernel) must be employed in the HAC estimation for asymptotic refinements.

However, it is well-known that if we employ a higher-order kernel for HAC estimation, the resulting covariance estimate is not necessarily positive semi-definite (“psd”) in finite samples. In fact, IS report in their Monte Carlo simulations that even in the presence of positive serial dependence in the moment function, with 4-7% chances the HAC estimate with the bandwidth chosen by their automatic procedure fails to be psd. IS describe the problem of such non-psd HAC estimates as “a very troubling issue which

[1] While exclusive attention is paid to linear models, they are of particular interest in empirical macroeconomics and finance. Examples include consumption and labor demand (Rotemberg, 1984), the expectation hypothesis of the term structure (Campbell and Shiller, 1991), inventory models (Fuhrer, Moore and Schuh, 1995; West and Wilcox, 1996), investment (Oliner, Rudebusch and Sichel, 1996), the monetary policy reaction function (Clarida, Galí and Gertler, 2000), the permanent-income hypothesis (Runkle, 1991), and the present value model of stock prices (West, 1988), to name a few.

is difficult to resolve (Remark 4, p.539)".

The problem of non-psd HAC estimates may be more crucial for the GMM bootstrap as it appears.[2] While the theory refers to its large sample properties, the bootstrap is primarily intended to provide a better approximation to the *finite-sample* distribution of a test statistic of interest. Indeed an arbitrary correction of a non-psd HAC estimate is admissible in an asymptotic sense (see endnote 4 in Andrews, 2002b, for example); however, the correction tends to deteriorate the quality of the GMM estimate, which will in turn affect the performance of the bootstrap testing adversely. As we see in Section 3, the problem of non-psd HAC estimates is more pronounced in the presence of negative serial dependence in the moment function.

Generating a psd estimate in finite samples is a highly desirable property of a HAC estimator, and several efforts have been made to derive such an estimator. Leading examples are two HAC estimators, one by Newey and West (1987) for possibly infinite-order of dependence and the other by West (1997) for finite-order dependence. One of the main reasons why HAC estimators with the Bartlett, Parzen, and Quadratic Spectral ("QS") kernels are most popularly applied in empirical studies is that each of these kernels always generates a psd estimate in finite samples.[3] The nonparametric prewhitened ("NPW") HAC estimator by Xiao and Linton (2002) and Hirukawa (2006) has been developed in line with these achievements. NPW is a subclass of multiplicative bias corrections for spectral estimation, and equivalent bias correction techniques have been already applied in nonparametric regression (Linton and Nielsen, 1994), probability density estimation (Jones, Linton and Nielsen, 1995; Hirukawa, 2010a), and hazard estimation (Nielsen, 1998; Nielsen and Tanggaard, 2001). A remarkable feature of NPW is that when a second-order spectral window is employed, the resulting HAC estimator attains the same order of magnitude in the bias term as with fourth-order kernels while not inflating the order of the variance term. In this sense, NPW can be viewed as an operation that implicitly constructs a fourth-order spectral window from a second-order one. However, unlike many other operations of constructing a higher-order spectral window, NPW does not sacrifice positive semi-definiteness of the resulting spectral estimate for the rate improvement. Employing a second-order spectral window guarantees the NPW-HAC estimator to be psd in finite samples by construction.

Based on these attractive properties, the NPW-HAC estimator can be a remedy to what IS call the "troubling issue". Then, as an important application of the HAC estimator this paper proposes to apply its inverse as the optimal weighting matrix for the efficient GMM estimation. It is anticipated that because the NPW-HAC estimator automatically generates a psd covariance estimate, it can deliver the GMM estimate of better quality and stabilize the performance of IS's bootstrap. As the initial step of investigating IS's bootstrap with the NPW-HAC estimator employed (called "the NPW-HAC-GMM bootstrap" hereafter), this paper conducts a small Monte Carlo study to examine finite sample

[2] The last paragraph on p.703 in Newey and West (1987) also describes drawbacks of non-psd HAC estimates.

[3] Andrews (1991) categorizes these kernels as is "$\mathcal{K}_2$-class".

performance of the bootstrap.

The remainder of this paper is organized as follows. Section 2 describes the procedure of the NPW-HAC-GMM bootstrap. Section 3 provides the Monte Carlo results with discussions. Finally, Section 4 concludes the paper with possible research extensions.

This paper adopts the following notational conventions: $\lfloor x \rfloor$ denotes the integer part of x; $\circ$ signifies convolution; vec (A) denotes the column by column vectorization function of matrix A; $\otimes$ is used to represent the tensor (or Kronecker) product; $*$ signifies conjugate transpose of a complex-valued matrix, i.e. $A^* = \bar{A}'$; and I_p denotes the p-dimensional identity matrix. K_{dd} is the $d^2 \times d^2$ commutation matrix that transforms vec (A) into vec (A'), i.e. $K_{dd} = \sum_{i=1}^{d} \sum_{j=1}^{d} e_i e_j' \otimes e_j e_i'$, where e_i is the ith elementary d-vector. Lastly, the expression, '$X_T \sim Y_T$' is used whenever $X_T / Y_T \to 1$ as $T \to \infty$.

2 Combining the NPW-HAC Estimator with the GMM Bootstrap

2.1 Bootstrap Algorithm

For a stationary and ergodic time series $\{(x_t', y_t, z_t')'\}$, consider an overidentified linear regression model

$$E\,(z_t u_t) := E\,\{z_t\,(y_t - \beta_0' x_t)\} = 0, \tag{1}$$

where $\beta_0 \in \mathbb{R}^p$ is a parameter vector, $x_t \in \mathbb{R}^p$, and $z_t \in \mathbb{R}^m$ $(p < m)$ is the vector of instruments that may possibly include predetermined values of y_t. Suppose that given T_0 observations $\{(x_t', y_t, z_t')'\}_{t=1}^{T_0}$,we are interested in estimating β_0 based on the moment condition (1) by the two-step GMM. For some $m \times m$ psd weighting matrix V_T , the first-step GMM estimator is given by

$$\tilde{\beta}_T = \arg\min_{\beta} \left\{ \frac{1}{T_0} \sum_{t=1}^{T_0} z_t\,(y_t - \beta' x_t) \right\}' V_{\mathrm{T}} \left\{ \frac{1}{T_0} \sum_{t=1}^{T_0} z_t\,(y_t - \beta' x_t) \right\}.$$

Based on the cross product of instruments z_t and the first-step GMM residual $\hat{u}_t := y_t - \tilde{\beta}'_T x_t$, we obtain the NPW-HAC estimator $\hat{\Omega}_T$, which is described in the next section. Set the bandwidth M for the HAC estimator in the second-step GMM estimation equal to the block length ℓ for MBB so that $M = \ell = \lfloor T^{2/9} \rfloor$, and put $T = T_0 - \ell + 1$. We use T observations and employ $\hat{\Omega}_T^{-1}$as the optimal weighting matrix that is intended for asymptotic refinements for the symmetric confidence interval and the distribution of the J-statistic. Then, the second-step GMM estimator becomes

$$\hat{\beta}_T = \arg\min_{\beta} \left\{ \frac{1}{T} \sum_{t=1}^{T} z_t\,(y_t - \beta' x_t) \right\}' \hat{\Omega}_T^{-1} \left\{ \frac{1}{T} \sum_{t=1}^{T} z_t\,(y_t - \beta' x_t) \right\}.$$

The GMM bootstrap considers the distributions of the studentized statistic of a linear combination of the parameter

$$\sqrt{T}\left(c'\hat{\Sigma}_T c\right)^{-1/2} c'\left(\hat{\beta}_T - \beta_0\right)$$

for an arbitrary nonzero vector $c \in \mathbb{R}^p$ and

$$\hat{\Sigma}_T = \left\{\left(\frac{1}{T}\sum_{t=1}^{T} x_t z_t'\right) \hat{\Omega}_T^{-1} \left(\frac{1}{T}\sum_{t=1}^{T} z_t x_t'\right)\right\}^{-1},$$

and the J-statistic

$$J_T = \left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t (y_t - \hat{\beta}_T' x_t)\right\}' \hat{\Omega}_T^{-1} \left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t (y_t - \hat{\beta}_T' x_t)\right\}.$$

The bootstrap applies the overlapping MBB by Künsch (1989). Suppose that we may write $T = b\ell$ for some integer b. The algorithm for bootstrap resampling takes the following five steps.

Step 1: Let $N_1, N_2, \ldots, N_b$ be *iid* uniform random variables on $\{0, 1, \ldots, T-\ell\}$ and let

$$\left(x^{*\prime}_{(j-1)\ell+i}, y^{*}_{(j-1)\ell+i}, z^{*\prime}_{(j-1)\ell+i}\right)' = \left(x'_{N_j+i}, y_{N_j+i}, z'_{N_j+i}\right)'$$

for $1 \le i \le \ell$ and $1 \le j \le b$.

Step 2: Calculate the first-step bootstrap GMM estimator as

$$\tilde{\beta}^*_T = \arg\min_{\beta} \left\{\frac{1}{T}\sum_{t=1}^{T} z^*_t (y^*_t - \beta' x^*_t) - \mu^*_T\right\}' V_{\mathrm{T}} \left\{\frac{1}{T}\sum_{t=1}^{T} z^*_t (y^*_t - \beta' x^*_t) - \mu^*_T\right\},$$

where

$$\mu^*_T = \frac{1}{T-\ell+1}\sum_{t=0}^{T-\ell} \frac{1}{\ell}\sum_{i=1}^{\ell} z_{t+i}\,(y_{t+i} - \hat{\beta}_T' x_{t+i}) .$$

Step 3: Compute the second-step bootstrap GMM estimator as

$$\hat{\beta}^*_T = \arg\min_{\beta} \left\{\frac{1}{T}\sum_{t=1}^{T} z^*_t (y^*_t - \beta' x^*_t) - \mu^*_T\right\}' \hat{\Omega}_T^{*-1} \left\{\frac{1}{T}\sum_{t=1}^{T} z^*_t (y^*_t - \beta' x^*_t) - \mu^*_T\right\},$$

where

$$\hat{\Omega}^*_T = \frac{1}{T} \sum_{k=1}^{b} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} (z^*_{Nk+i}\,\hat{u}^*_{Nk+i} - \mu^*_T)\,(z^*_{Nk+j}\,\hat{u}^*_{Nk+j} - \mu^*_T)',$$

$$\hat{u}^*_t = y^*_t - \tilde{\beta}^{*\prime}_T x^*_t.$$

Step 4: Obtain the bootstrap version of the studentized statistic

$$\sqrt{T}\,(c'\hat{\Sigma}^*_T c)^{-1/2} c'(\hat{\beta}^*_T - \hat{\beta}_T)$$

for

$$\hat{\Sigma}^*_T = \left\{ \left(\frac{1}{T} \sum_{t=1}^{T} x^*_t z^{*\prime}_t \right) \hat{\Omega}^{*-1}_T \left(\frac{1}{T} \sum_{t=1}^{T} z^*_t x^{*\prime}_t \right) \right\}^{-1}$$

and the J-statistic

$$J^*_T = \left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} z^*_t (y^*_t - \hat{\beta}^{*\prime}_T x^*_t) - \mu^*_T \right\}' \hat{\Omega}^{*-1}_T \left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} z^*_t (y^*_t - \hat{\beta}^{*\prime}_T x^*_t) - \mu^*_T \right\}.$$

Step 5: Repeat Steps 1-4 sufficiently many times to approximate the finite-sample distributions of the studentized statistic and the J-statistic by the empirical distributions of their bootstrap version.

The entire procedure basically follows IS's. In particular, Steps 1 and 2 play a key role in the MBB. To preserve the dependent structure of the moment function, the MBB resamples ℓ consecutive observations b times and construct a bootstrap sample by gluing them together. Each ℓ consecutive observations are chosen via the uniform weighting scheme. Because bootstrap resampling does not reproduce the overidentified moment condition that holds for the data, recentering the moment condition by demeaning is required for each bootstrap sample. Alternatively, applying the implied probabilities from the moment condition (e.g. Brown and Newey, 2002; Allen, Gregory and Shimotsu, 2010) can get rid of recentering the moment condition, but this approach is not considered in this paper.

There are two remarkable differences in the NPW-HAC-GMM bootstrap from IS's, which are underlined in the above description. First, the inverse of the NPW-HAC estimator is used as the optimal weighting matrix to obtain the GMM estimate. As discussed in the next section, the NPW-HAC estimator is by construction psd and has the same order of magnitude in the bias term as the HAC estimator with a fourth-order kernel. Because of its positive semi-definiteness the NPW-HAC estimator is expected to yield the GMM estimate in good quality, which is a key element for better performance of the bootstrap.

Second, while the bandwidth M is set equal to the block length ℓ as in Götze and Künsch (1996) and IS, the divergence rate of M and ℓ is $O(T^{2/9})$, which is slower than $O(T^{1/3})$ in IS's framework. It can be conjectured from IS's technical discussions and properties of the NPW-HAC estimator that setting

$M = O(T^{\delta})$ for some $\delta \in$ (2/9, 1/4) and $\ell = O(M)$ will deliver asymptotic refinements for the symmetric confidence interval and the distribution of the J-statistic.[4] A rigorous proof is left for future work.

2.2 The NPW-HAC Estimator

The NPW-HAC estimator is originally proposed by Xiao and Linton (2002) and modified by Hirukawa (2006). The idea behind this estimator is to apply a multiplicative bias correction in the frequency domain to a HAC estimator so that the bias can be accelerated by an order of magnitude while the order of the variance remains unchanged. As a result, the NPW-HAC estimator is guaranteed to be psd in finite samples when a second-order kernel is employed, and at the same time it has the same order of magnitude in the bias term as with fourth-order kernels.

The NPW-HAC estimator is formally defined as follows. Let $v_t := z_t u_t$ so that the moment condition (1) is rewritten as $E(v_t) = 0$. Also let fundamental frequencies be $\lambda_j = 2\pi j/T, j = 0, \pm 1, \ldots, \pm \lfloor T/2 \rfloor$, and the periodogram at the frequency λ_j be $I_{vv}(\lambda_j) = \zeta_v(\lambda_j)\, \zeta_v(\lambda_j)^*$, where $\zeta_v(\lambda_j) = (2\pi T)^{-1/2} \sum_{t=1}^{T} v_t e^{-it\lambda_j}$ is the finite Fourier transform of $\{v_t\}$ evaluated at the frequency λ_j. Given a kernel $K(\cdot)$, the bandwidth M, and the spectral window corresponding to this kernel $W(\theta) = (2\pi)^{-1} \int_{-\infty}^{\infty} K(u)\, e^{-iu\theta} du$, define the amplitude window (Parzen, 1963) as $W_M(\theta) = M \sum_{j=-\infty}^{\infty} W\{M(\theta + 2\pi j)\}$.[5] Then, the (bias-uncorrected) HAC estimator of $\{v_t\}$ can be represented in two ways as a weighted autocovariance and a smoothed periodogram

$$\tilde{S}_T = \sum_{l=-(T-1)}^{T-1} K\left(\frac{l}{M}\right) \tilde{\Gamma}_l \sim \frac{4\pi^2}{T} \sum_{\lambda_j \in B(0)} W_M(\lambda_j)\, I_{vv}(\lambda_j),$$

where $\tilde{\Gamma}_l = T^{-1} \sum_{t=\max\{1,1+l\}}^{\min\{T+l,T\}} v_t v_{t-l}'$ is the lth sample autocovariance of $\{v_t\}$ and $B(0) = \{\lambda_j \mid -\pi < \lambda_j < \pi\}$ is the frequency band with width 2π centered at zero frequency.

Similarly, for $B(\omega) = \{\lambda_j \mid \omega - \pi < \lambda_j < \omega + \pi\}$, the frequency band with width 2π centered at the frequency ω, the spectral matrix of $\{v_t\}$ evaluated at the frequency $\omega \in (-\pi, \pi)$ can be estimated as

$$\tilde{f}_{vv}(\omega) = \frac{1}{2\pi} \sum_{l=-(T-1)}^{T-1} K\left(\frac{l}{M}\right) \tilde{\Gamma}_l e^{-il\omega} \sim \frac{2\pi}{T} \sum_{\lambda_j \in B(\omega)} W_M(\lambda_j - \omega)\, I_{vv}(\lambda_j).$$

Also define the "square-root" of the spectral matrix $\tilde{f}_{vv}(\omega)$ as the one obtained by the unitary decomposition

---

[4] Strictly speaking, the divergence rate of M and ℓ needs to be slightly faster than $O(T^{2/9})$. However, it appears that using $M = O(T^{2/9})$ is not problematic in practice, because, for example, $M = \lfloor T^{2/9} \rfloor$ and $M = \lfloor T^{201/900} \rfloor$ would yield the same bandwidth for the sample sizes considered in the Monte Carlo study. Although deriving the data-based choice rules for M and ℓ that can establish asymptotic refinements is an interesting challenge, it is beyond the scope of this paper.

[5] As described in Parzen (1963), $W_M(\theta) \sim MW(M\theta)$. In particular, the approximation is replaced by the equality when the spectral window $W(\theta)$ is band-limited such that $W(\theta) = 0$ for $|\theta| \geq \pi$.

$$\tilde{f}_{vv}^{1/2}(\omega) = U\Lambda^{1/2}U^*,$$

where $\Lambda^{1/2} = \mathrm{diag}\,(\lambda_1^{1/2}, \ldots, \lambda_m^{1/2})$ is the diagonal matrix containing the square roots of the eigen-values of $\tilde{f}_{vv}(\omega)$, and U is the unitary matrix, i.e. $UU^* = I_m$.[6] Finally, the NPW-HAC estimator is defined as

$$\tilde{\Omega}_T = \tilde{S}_T^{1/2}\,\tilde{\alpha}(0)\,\tilde{S}_T^{1/2} = \tilde{S}_T^{1/2}\left\{\sum_{\lambda_j \in B(0)} W_M(\lambda_j)\frac{2\pi}{T}\tilde{f}_{vv}^{-1/2}(\lambda_j)\,I_{vv}(\lambda_j)\tilde{f}_{vv}^{-1/2}(\lambda_j)\right\}\tilde{S}_T^{1/2},$$

where $\tilde{\alpha}(0)$ serves as the bias correction term. Note that $\hat{\Omega}_T$ in the previous section is obtained by replacing v_t with $\hat{v}_t = z_t\,\hat{u}_t = z_t\,(y_t - \tilde{\beta}'_T x_t)$, where $\tilde{\beta}_T$ is the first-step GMM estimator.

The name of "NPW" comes from the fact that the bias correction technique reminds us of the well-known prewhitening procedure (see Andrews and Monahan, 1992, for example), because the transformed periodograms $\tilde{f}_{vv}^{-1/2}(\lambda_j)\,I_{vv}(\lambda_j)\tilde{f}_{vv}^{-1/2}(\lambda_j)$ are approximately constant. Also observe that when $W(\theta) \geq 0, \forall\theta \in \mathbb{R}$, the NPW-HAC estimator $\tilde{\Omega}_T$ is psd in finite samples because of its "sandwich form" structure.

Hirukawa (2006) further establishes that under the standard regularity conditions $\tilde{\Omega}_T$ has an asymptotic expansion

$$\tilde{\Omega}_T = \Omega_0 + \mathcal{B} + \mathcal{V} + o_p\left(M^{-4} + \sqrt{\frac{M}{T}}\right). \tag{2}$$

where

$$\Omega_0 = \lim_{T\to\infty}\frac{1}{T}E\left\{\left(\sum_{t=1}^{T} v_t\right)\left(\sum_{t=1}^{T} v_t\right)'\right\}$$

is the long-run variance of $\{v_t\}$, and $\mathcal{B}$ and $\mathcal{V}$ constitute the leading bias and variance terms of $\tilde{\Omega}_T$. The leading bias term $\mathcal{B}$ is approximated by

$$\mathcal{B} \sim -K_2^2\,\Psi M^{-4},$$

where $K_2 = \lim_{x\to 0}\{1 - K(x)\}/|x|^2 \in (0, \infty)$, $\Psi = \Omega_0^{1/2}\,\Phi''(0)\,\Omega_0^{1/2}$, $\Phi''(0) = d^2\Phi(\omega)/d\omega^2|_{\omega=0}$, $\Phi(\omega) = f_{vv}^{-1/2}(\omega)\,f''_{vv}(\omega) f_{vv}^{-1/2}(\omega)$, and $f''_{vv}(\omega) = d^2 f_{vv}(\omega)/d\omega^2$. Moreover, the leading variance term $\mathcal{V}$ is approximated by

$$Var\{\mathrm{vec}(\mathcal{V})\} \sim \frac{M}{T}\int_{-\infty}^{\infty} t_k^2(x)dx\,(I_{m^2} + K_{mm})\,\Omega_0 \otimes \Omega_0,$$

where $t_K(x) = \int_{-\infty}^{\infty} T_W(\theta)\,e^{ix\theta}d\theta$ and $T_W(\theta) = 2W(\theta) - W \circ W(\theta)$ is the fourth-order spectral window obtained by twicing (Stuetzle and Mittal, 1979). Observe that the bias term is of order M^{-4}, which is

[6] Indeed $\tilde{f}_{vv}^{1/2}(\omega)$ is well-defined. The $I_{vv}(\lambda_j)$ are Hermitian and psd by construction, so is $\tilde{f}_{vv}(\omega)$ when $W(\theta) \geq 0$, $\forall\theta \in \mathbb{R}$. Hence, $\tilde{f}_{vv}(\omega)$ has positive eigenvalues $\lambda_1, \ldots, \lambda_m$. For the reference to the related argument, see Section 3.7 in Brillinger (1975).

usually achieved under the fourth-order spectral window, whereas the order of magnitude of the variance term is not inflated, i.e. usual $O(M/T)$ is maintained.

3 Monte Carlo Experiment

3.1 Setup

This section conducts a small Monte Carlo simulation study to investigate the reliability of the NPW-HAC-GMM bootstrap in finite samples. Following IS, consider the two-step GMM estimation of a simple linear regression

$$y_t = \beta_1 + \beta_2 x_t + u_t .$$

Without loss of generality the true parameter value is set equal to $\beta = (\beta_1, \beta_2)' = (0, 0)'$. The regressor $\{x_t\}$ and the error term $\{u_t\}$ are generated independently by AR(1) and either AR(1) or MA(1) processes.

$$\begin{aligned} x_t &= \rho x_{t-1} + \varepsilon_{1t}, \rho \in \{0.5, 0.8\}. \\ u_t &= \begin{cases} \phi u_{t-1} + \varepsilon_{2t}, \phi = \pm\rho & : \text{AR(1)} \\ \varepsilon_{2t} + \psi\varepsilon_{2t-1}, \psi = \pm\rho & : \text{MA(1)} \end{cases} . \\ \varepsilon_t &= (\varepsilon_{1t}, \varepsilon_{2t})' \overset{iid}{\sim} N(0, I_2). \end{aligned}$$

Observe that the AR or MA coefficient in the error term has the same magnitude (although the sign may be opposite) as the AR coefficient in the regressor. Because predetermined regressors are frequently used as instruments in the presence of endogeneity, this experiment adopts $z_t = (1, x_{t-1}, x_{t-2})'$ as the vector of instruments. Hence, the moment condition model is overidentified with 1 degree of freedom for the *J*-test. In the simulations, 5000 random samples are generated for two sample sizes (T), namely, 64 and 128.

Our primary interest is to evaluate finite sample null rejection frequencies of the *t*-ratio of β_2 (that is intended for the two-sided test) and the *J*-statistic. Because there is parameter uncertainty due to the first-step estimation and we apply a univariate testing procedure to each element of the moment function, it is hard to control the size of this procedure. Therefore, this Monte Carlo experiment concentrates on the critical values for the nominal 10% level to be conservative. Furthermore, the bootstrap critical values are calculated based on 499 replications for each sample.

The first-step GMM estimation uses the 3×3 identity matrix as the initial weighting matrix. In the second-step GMM estimation, the following HAC estimators and bandwidths (and block lengths, whenever applicable) are employed. First, the NPW-HAC-GMM bootstrap applies NPW-HAC estimators with Parzen and Bohman (Bohman, 1961) kernels.[7] The bandwidth M and the block length ℓ are chosen

[7] The Bohman kernel is expressed as

$$K(x) = \begin{cases} (1-|x|)\cos(\pi x) + \sin(\pi|x|)/\pi & \text{if } |x| \le 1 \\ 0 & \text{if } |x| > 1 \end{cases} .$$

as $M = \ell = \lfloor T^{2/9} \rfloor$. Observe that $M = \ell = 2$ for each sample size.

Second, for IS's bootstrap, HAC estimators are computed using the truncated kernel, the trapezoidal kernel (Politis and Romano, 1995) with parameter 1/2, and the Parzen (b) kernel (Parzen, 1957) of order 3.[8] Again M is set equal to ℓ, and ℓ is chosen by IS's automatic procedure similar to the general-to-specific modelling strategy: see p.540 in IS for details. Each of the three kernels does not necessarily generate psd estimates in finite samples. When the HAC estimate is not psd, the block length chosen automatically is shortened until the resulting covariance estimate becomes psd.

Third, besides the bootstrap results, the first-order asymptotic results based on widely applied HAC estimators are also obtained. For this purpose, two most popular automatic bandwidth choice rules by Andrews (1991) and Newey and West (1994) are employed for the QS and Bartlett kernels, respectively. For the former, AR(1) reference is used. For the latter, the bandwidth for the *normalized curvature* (Hirukawa, 2010b, p.713) estimator using the truncated kernel is set equal to $\lfloor 4(T/100)^{2/9} \rfloor$. In addition, recently Hirukawa (2010b) has proposed yet another bandwidth choice rule called the solve-the-equation plug-in rule. Bartlett and Parzen HAC estimators using this rule are also calculated. Moreover, because prewhitening (Andrews and Monahan, 1992) is popularly applied in empirical studies, each of these HAC estimators is computed for the prewhitened series as well as the original (= non-prewhitened) series. The procedure of prewhitening follows Andrews and Monahan (1992) with the eigenvalues of the fitted VAR(1) coefficient matrix adjusted to being less than 0.97 in magnitude.

The spectral windows corresponding to the Parzen and Bohman kernels are $W(\theta) = 24(1 - \cos(\theta/2))^2 / (\pi\theta^4)$ and $W(\theta) = 2\pi(1 + \cos(\theta)) / (\pi^2 - \theta^2)^2$, respectively.

[8] The trapezoidal kernel and the Parzen (b) kernel of order 3 are expressed as

$$K(x) = \begin{cases} 1 & \text{if } |x| \le 1/2 \\ 2(1-|x|) & \text{if } 1/2 < |x| \le 1 \\ 0 & \text{if } |x| > 1 \end{cases},$$

and

$$K(x) = \begin{cases} 1-|x|^3 & \text{if } |x| \le 1 \\ 0 & \text{if } |x| > 1 \end{cases},$$

respectively.

Table 1: Finite Sample Null Rejection Frequencies against Nominal 10% Level for AR(1) Error

Panel (a): *t*-ratio

(%)

Kernel	$(\rho, \phi) =$ / $T_0 + 1 =$	(0.8, -0.8) 64	(0.8, -0.8) 128	(0.5, -0.5) 64	(0.5, -0.5) 128	(0.5, 0.5) 64	(0.5, 0.5) 128	(0.8, 0.8) 64	(0.8, 0.8) 128
(1) Bootstrap:									
Truncated		16.4	23.8	21.0	25.9	15.1	13.9	18.2	9.2
		(60.5)	(51.3)	(18.0)	(8.1)	(4.1)	(3.7)	(4.1)	(3.1)
Trapezoidal		27.7	34.1	21.3	24.2	15.0	13.4	17.7	8.4
		(16.3)	(5.6)	(15.6)	(5.0)	(2.7)	(1.7)	(1.9)	(1.0)
Parzen (b)		25.3	36.6	19.7	21.4	14.8	13.5	17.9	8.6
		(40.0)	(23.1)	(7.8)	(2.0)	(2.5)	(1.4)	(1.6)	(0.9)
Parzen-NPW		8.3	8.8	7.3	8.0	13.1	13.2	26.2	24.3
Bohman-NPW		12.0	13.7	8.5	8.9	13.0	12.6	24.8	22.5
(2) First-Order Asymptotic:									
Non-Prewhitened:									
QS-A		7.5	8.5	6.4	6.9	17.4	14.7	29.6	21.8
Bartlett-NW		19.4	9.4	13.8	12.0	19.6	16.5	31.5	23.1
Bartlett-H		13.3	12.2	9.2	9.1	18.0	15.8	31.2	23.0
Parzen-H		14.5	12.4	10.6	10.3	18.2	15.3	33.9	24.1
Prewhitened:									
QS-A		13.7	12.6	10.3	9.9	17.0	13.7	25.0	18.3
Bartlett-NW		18.6	15.0	14.7	12.8	20.0	15.8	26.1	19.1
Bartlett-H		13.4	12.1	10.7	9.9	17.0	13.7	25.2	18.6
Parzen-H		13.6	11.7	10.1	9.6	17.3	13.8	25.1	18.2

Panel (b): *J*-statistic

(%)

Kernel	$(\rho, \phi) =$ / $T_0 + 1 =$	(0.8, -0.8) 64	(0.8, -0.8) 128	(0.5, -0.5) 64	(0.5, -0.5) 128	(0.5, 0.5) 64	(0.5, 0.5) 128	(0.8, 0.8) 64	(0.8, 0.8) 128
(1) Bootstrap:									
Truncated		10.5	11.0	10.8	11.8	9.6	10.8	11.6	12.1
		(60.5)	(51.3)	(18.0)	(8.1)	(4.1)	(3.7)	(4.1)	(3.1)
Trapezoidal		12.7	11.7	10.8	11.9	9.5	10.6	11.1	11.7
		(16.3)	(5.6)	(15.6)	(5.0)	(2.7)	(1.7)	(1.9)	(1.0)
Parzen (b)		11.4	11.2	10.6	11.8	9.3	10.5	11.0	11.3
		(40.0)	(23.1)	(7.8)	(2.0)	(2.5)	(1.4)	(1.6)	(0.9)
Parzen-NPW		9.0	10.2	9.1	10.3	9.1	9.9	10.9	11.7
Bohman-NPW		8.9	10.2	9.1	10.6	9.0	9.7	10.5	11.8
(2) First-Order Asymptotic:									
Non-Prewhitened:									
QS-A		11.5	10.7	10.5	10.9	16.0	16.9	23.7	23.4
Bartlett-NW		5.5	9.4	8.4	10.9	14.1	15.9	18.6	19.5
Bartlett-H		8.0	9.7	9.9	11.0	13.8	15.2	18.3	19.3
Parzen-H		11.9	10.8	10.7	11.2	15.2	16.7	19.9	21.5
Prewhitened:									
QS-A		12.7	11.0	11.4	11.2	19.3	19.6	27.3	25.6
Bartlett-NW		12.7	11.2	11.5	11.3	19.3	19.5	27.5	25.8
Bartlett-H		12.6	11.3	11.5	11.0	19.6	19.9	27.4	25.7
Parzen-H		12.7	11.2	11.4	10.9	19.4	19.7	27.2	25.8

Note: *Numbers in parentheses are frequencies of psd corrections.*

Table 2: Finite Sample Null Rejection Frequencies against Nominal 10% Level for MA(1) Error

Panel (a): *t* -ratio

(%)

$(\rho, \psi) =$	(0.8, -0.8)		(0.5, -0.5)		(0.5, 0.5)		(0.8, 0.8)	
Kernel $T_0 + 1 =$	64	128	64	128	64	128	64	128
(1) Bootstrap:								
Truncated	15.4	18.2	17.3	19.4	15.1	13.9	14.2	11.4
	(30.2)	(26.6)	(10.6)	(5.8)	(4.6)	(3.9)	(4.4)	(4.0)
Trapezoidal	15.9	19.5	17.6	19.1	15.4	13.6	13.8	11.3
	(29.8)	(24.7)	(9.3)	(3.2)	(3.0)	(1.6)	(2.5)	(1.7)
Parzen (b)	15.1	17.4	15.3	15.9	14.9	13.5	13.5	11.4
	(15.7)	(9.1)	(4.7)	(1.7)	(2.7)	(1.4)	(2.2)	(1.5)
Parzen-NPW	2.7	2.1	5.7	6.2	11.7	11.8	12.7	11.1
Bohman-NPW	3.1	2.6	6.4	6.7	11.6	11.3	11.8	10.2
(2) First-Order Asymptotic:								
Non-Prewhitened:								
QS-A	2.5	2.2	5.4	5.2	15.5	13.6	19.3	15.7
Bartlett-NW	14.4	13.2	13.1	11.1	18.2	15.7	22.9	18.4
Bartlett-H	7.5	6.8	7.4	7.6	15.4	14.3	19.6	15.6
Parzen-H	9.2	8.8	8.2	8.0	16.5	14.5	24.3	17.6
Prewhitened:								
QS-A	3.7	2.5	8.0	6.6	15.4	12.9	15.2	11.5
Bartlett-NW	14.6	13.1	13.4	11.1	18.9	15.4	21.6	16.4
Bartlett-H	6.3	5.7	8.2	6.9	15.6	13.1	16.0	11.8
Parzen-H	4.3	2.3	7.8	6.4	16.1	13.5	18.7	14.0

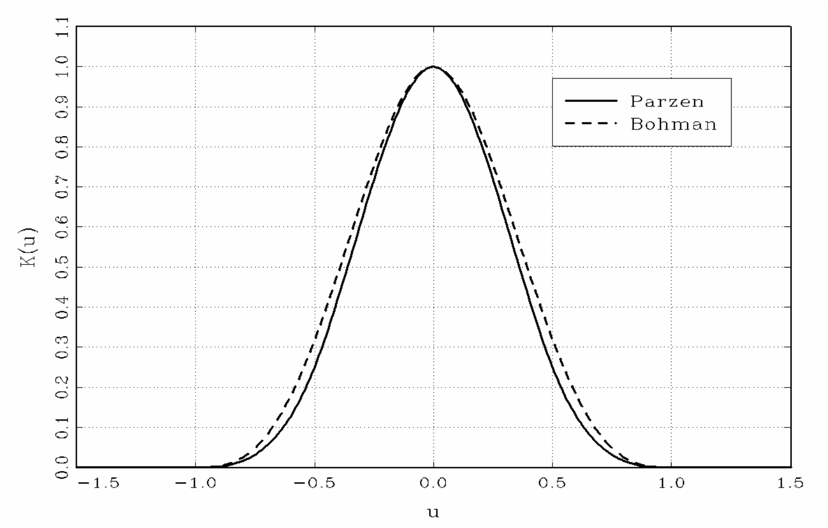
Panel (b): *J* -statistic

(%)

$(\rho, \psi) =$	(0.8, -0.8)		(0.5, -0.5)		(0.5, 0.5)		(0.8, 0.8)	
Kernel $T_0 + 1 =$	64	128	64	128	64	128	64	128
(1) Bootstrap:								
Truncated	9.8	11.0	10.4	11.8	9.7	10.7	10.4	11.2
	(30.2)	(26.6)	(10.6)	(5.8)	(4.6)	(3.9)	(4.4)	(4.0)
Trapezoidal	10.0	11.0	10.5	11.9	9.7	10.9	10.4	10.9
	(29.8)	(24.7)	(9.3)	(3.2)	(3.0)	(1.6)	(2.5)	(1.7)
Parzen (b)	9.7	11.0	10.5	11.9	9.7	10.7	10.3	10.8
	(15.7)	(9.1)	(4.7)	(1.7)	(2.7)	(1.4)	(2.2)	(1.5)
Parzen-NPW	8.7	10.5	9.1	10.4	8.9	10.1	9.8	10.4
Bohman-NPW	8.9	10.7	9.3	10.3	8.6	9.9	9.9	10.5
(2) First-Order Asymptotic:								
Non-Prewhitened:								
QS-A	11.0	11.4	10.5	11.0	15.0	15.2	16.8	15.2
Bartlett-NW	6.6	8.0	9.2	10.9	14.2	14.4	14.6	14.2
Bartlett-H	9.5	10.6	10.4	11.3	13.7	13.7	15.0	14.2
Parzen-H	12.1	12.1	10.9	11.4	14.6	14.7	15.6	14.8
Prewhitened:								
QS-A	11.7	11.5	11.6	11.2	17.1	16.9	18.4	16.2
Bartlett-NW	10.9	11.1	11.6	11.6	16.8	16.6	17.0	15.2
Bartlett-H	11.3	11.3	11.6	11.1	18.0	17.4	19.7	17.2
Parzen-H	11.4	11.3	11.6	11.0	17.9	17.2	18.7	16.0

Note: *Numbers in parentheses are frequencies of psd corrections.*

Figure 1: Shapes of the Parzen and Bohman Kernels



3.2 Results

Tables 1 and 2 report finite sample null rejection frequencies of the *t*-ratio and the *J*-statistic for the cases with AR(1) and MA(1) errors, respectively. Numbers in parentheses for the truncated, trapezoidal and Parzen (b) kernels are frequencies of psd corrections. In addition, the automatic bandwidth choice procedures by Andrews (1991), Newey and West (1994) and Hirukawa (2010b) are denoted as A, NW and H, respectively.

We start looking into the results for AR(1) errors. Panel (a) in Table 1 presents null rejection frequencies of the *t*-ratio. At a first glance we can recognize that there is little difference in performance between two kernel choices for the NPW-HAC estimation. This is not surprising, because their shapes look alike (see Figure 1) and exactly the same block length (and thus exactly the same bandwidth) is used. We now closely look at the results for each parameter setting. In the presence of negative serial dependence in the moment function (i.e. when $\phi = -0.8, -0.5$), all three versions of IS's bootstrap suffer from considerable frequencies of psd corrections. The non-psd covariance estimates deteriorate the quality of the GMM estimates and lead to large null rejection frequencies. Moreover, contrary to what the large sample theory predicts, the null rejection frequencies deviate from the nominal level as the sample size increases. Since the NPW-HAC estimator is free of the non-psd problem, both versions of the NPW-HAC-GMM bootstrap exhibit better size properties than IS's. When $\phi=0.5$, the NPW-HAC-GMM bootstrap still maintains better performance, but the margin gets thinner due to infrequent psd corrections in IS's bootstrap. When $\phi=0.8$, IS's bootstrap at last outperforms the NPW-HAC-GMM bootstrap in a decisive manner. However, the unsatisfactory performance of the latter is attributed mainly to a short block length:

see below for additional analysis. In addition, first-order asymptotic-based inference when $\phi=-0.8$, -0.5 is in general satisfactory. Often null rejection frequencies are quite close to the nominal level, regardless of whether prewhitened or non-prewhitened series is chosen. On the other hand, the results based on first-order asymptotics get worse as ϕ becomes positive and larger. Even prewhitening does not resolve the issue of size distortion substantially.

Panel (b) in Table 1 displays null rejection frequencies of the *J*-statistic. We can find a clear advantage of bootstraps over first-order asymptotics. The null rejection frequencies vary very little across bootstraps, and they are quite close to the nominal level. While first-order asymptotics perform equally well for negative ϕ's, they tend to reject the null too often for positive ϕ's. This tendency is pronounced when the moment function becomes persistent.

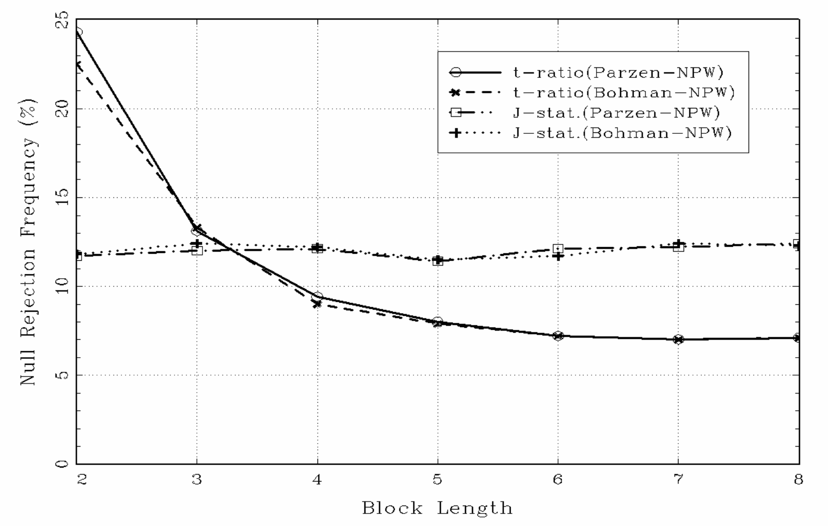
We now examine the results for MA(1) errors given in Table 2. The general tendencies shown in Panels (a) and (b) are basically the same as those for AR(1) errors. A notable difference is that when $\phi=-0.8$, -0.5, *t*-ratios from both versions of the NPW-HAC-GMM bootstrap (as well as those based on two versions of the QS-HAC estimator) tend to reject the null too infrequently. In other words, the bootstrap quantiles tend to inflate in this situation, but it seems that this issue is not an outcome from the quantile breakdown by outliers (see Camponovo, Scaillet and Trojani, 2010, for example) but again due to a short block length: see below for details.

Table 3: Finite Sample Null Rejection Frequencies of the NPW-HAC-GMM Bootstrap for MA(1) Error with $(\rho,\psi)=(0.8,-0.8)$, $(0.5,-0.5)$ When $M=\ell=3$

(%)

$(\rho,\psi)=$	(0.8, -0.8)		(0.5, -0.5)	
Kernel $T_\phi+1=$	64	128	64	128
***t*-ratio**				
Bootstrap:				
Parzen-NPW	9.6	8.4	11.6	11.1
Bohman-NPW	11.2	10.5	12.5	12.5
***J*-statistic**				
Bootstrap:				
Parzen-NPW	10.1	11.0	10.4	11.2
Bohman-NPW	9.9	11.3	10.3	11.8

Figure 2: A Plot of Null Rejection Frequencies against Block Lengths
($T_0 + 1 = 128$; AR(1) Error with (ρ, ϕ) = (0.8, 0.8))



Tables 1 and 2 reveal that the NPW-HAC-GMM bootstrap performs poorly in the following two scenarios: (i) for the MA(1) error with (ρ,ψ) = (0.8, −0.8), (0.8, −0.5); and (ii) for the AR(1) error with (ρ, ϕ) = (0.8, 0.8). In each scenario only the *t*-ratio is affected adversely. Invoke that the block length is chosen in a non-data-dependent manner; to be more precise, it is fixed at 2 regardless of the sample size. Therefore, we reasonably anticipate that the issue of poor performance can be resolved once a longer block length is chosen. As regards the scenario (i), the numbers in Table 3 are obtained by using the same bootstrap samples but setting $M = \ell = 3$. Indeed the block length makes null rejection frequencies of the *t*-ratio close to the nominal level, while maintaining the highly satisfactory performance of the *J*-statistic.

Next, the scenario (ii) corresponds to the case of a persistent moment function. For such a persistent process, a large value of the block length is expected to be a remedy. Figure 2 plots null rejection frequencies of the *t*-ratio and the *J*-statistic based on the NPW-HAC-GMM bootstrap for the sample size 128. Again the numbers are obtained by using the same bootstrap samples and changing only the value of M and ℓ. We can see that the performance of the *t*-ratio is highly sensitive to the block length, whereas that of the *J*-statistic is insensitive. Judging from the performance of the *t*-ratio, setting $M = \ell =$ 4 appears to be the best for sample size 128.

Overall, the NPW-HAC-GMM bootstrap performs well, and it often outperforms IS's bootstrap and first-order asymptotics; in other words, the NPW-HAC estimator contributes to stabilizing IS's bootstrap. In particular, the good performance for MA-type dependence is encouraging, because linear rational expectations models often exhibit MA-type dependence in the moment function. On the other hand, there is a pressing need for a data-based choice rule of the block length. Because the performance of the *J*-statistic is insensitive to the block length, a possible strategy of deriving the choice rule would be based on the coverage accuracy of the symmetric two-sided confidence interval.

4 Concluding Remarks

To stabilize IS's bootstrap, this paper has proposed to employ the inverse of the NPW-HAC estimator as the optimal weighting matrix for the second-step GMM estimation. Monte Carlo results indicate that the HAC estimator can indeed stabilize IS's bootstrap and improve its finite sample performance, in particular, in the presence of either negative AR-type dependence or MA-type dependence in the moment function.

There are several important research extensions. First, the asymptotic theory for the NPW-HAC-GMM bootstrap, including the condition for divergence rates of the bandwidth and the block length that can provide asymptotic refinements, should be established. While IS's asymptotic results can be carried through in principle, an important modification should be made. IS's theory is built on HAC estimators of a weighted autocovariance form. Since the NPW-HAC estimator takes a more complicated form, relevant parts of IS's proofs should be suitably overwritten. The Edgeworth expansion of the *t*-ratio in the frequency domain by Velasco and Robinson (2001) may be a nice hint for the modification.

Second, an equal priority is given to a data-dependent choice rule of the bandwidth and the block length in the framework of the NPW-HAC-GMM bootstrap. Although there are a number of block length choice rules proposed, it is important to see whether one of these rules is readily applicable to the bootstrap or a new rule must be tailor-made.

Third, it is highly motivated to investigate finite sample performance of the bootstrap in a more realistic framework. In particular, since West and Wilcox (1996) report that test statistics implied by their inventory model are often poorly sized, it is of interest to see how much the NPW-HAC-GMM bootstrap can contribute to size corrections. Furthermore, as an empirical example, the bootstrap procedure can be applied to the monetary policy reaction function by Clarida, Galí and Gertler (2000), as well as the inventory model. Because the GMM estimate of a structural econometric model often has a policy implication, it is essential to construct more reliable confidence intervals for key structural parameters. All these extensions are currently under the author's investigation, and they will be addressed in a separate paper.

References

Allen, J., A. W. Gregory, and K. Shimotsu (2010): "Empirical Likelihood Block Bootstrapping," *Journal of Econometrics*, forthcoming.

Andrews, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817 - 858.

Andrews, D. W. K. (2002a): "Higher-Order Improvements of a Computationally Attractive *k*-Step Bootstrap for Extremum Estimators," *Econometrica*, 70, 119 - 162.

Andrews, D. W. K. (2002b): "Equivalence of the Higher Order Asymptotic Efficiency of *k*-Step and Extremum Statistics," *Econometric Theory*, 18, 1040 - 1085.

Andrews, D. W. K., and J. C. Monahan (1992): "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60, 953 - 966.

Bohman, H. (1961): "Approximate Fourier Analysis of Distribution Functions," *Arkiv för Matematik*, 4, 99 - 157.

Brillinger, D. R. (1975): *Time Series: Data Analysis and Theory*. New York: Holt, Rinehart & Winston.

Brown, B. W., and W. K. Newey (2002): "Generalized Method of Moments, Efficient Bootstrapping, and Improved Inference," *Journal of Business & Economic Statistics*, 20, 507 - 517.

Campbell, J. Y., and R. J. Shiller (1991): "Yields Spread and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58, 495 - 514.

Camponovo, L., O. Scaillet, and F. Trojani (2010): "Robust Resampling Methods for Time Series," Working Paper, HEC Genève.

Clarida, R., J. Galí, and M. Gertler (2000): "Monetary Policy Rule and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147 - 180.

Götze, F., and H. R. Künsch (1996): "Second-Order Correctness of the Blockwise Bootstrap for Stationary Observations," *Annals of Statistics*, 24, 1914 - 1933.

Ferson, W. E., and S. R. Foerster (1994): "Finite Sample Properties of the Generalized Method of Moments in Tests of Conditional Asset Pricing Models," *Journal of Financial Economics*, 36, 29 - 55.

Fuhrer, J. C., G. R. Moore, and S. D. Schuh (1995): "Estimating the Linear-Quadratic Inventory Model: Maximum Likelihood versus Generalized Method of Moments," *Journal of Monetary Economics*, 35, 115 - 157.

Hall, P., and J. L. Horowitz (1996): "Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators," *Econometrica*, 64, 891 - 916.

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

Hirukawa, M. (2006): "A Modified Nonparametric Prewhitened Covariance Estimator," *Journal of Time Series Analysis*, 27, 441 - 476.

Hirukawa, M. (2010a): "Nonparametric Multiplicative Bias Correction for Kernel-Type Density Estimation on the Unit Interval," *Computational Statistics & Data Analysis*, 54, 473 - 495.

Hirukawa, M. (2010b): "A Two-Stage Plug-In Bandwidth Selection and Its Implementation for Covariance Estimation," *Econometric Theory*, 26, 710 - 743.

Inoue, A., and M. Shintani (2006): "Bootstrapping GMM Estimators for Time Series," *Journal of Econometrics*, 133, 531 - 555.

Jones, M. C., O. Linton, and J. P. Nielsen (1995): "A Simple Bias Reduction Method for Density Estimation," *Biometrika*, 82, 327 - 338.

Kitamura, Y. (1997): "Empirical Likelihood Methods with Weakly Dependent Processes," *Annals of Statistics*, 25, 2084 - 2102.

Kitamura, Y., and M. Stutzer (1997): "An Information-Theoretic Alternative to Generalized Method of Moments Estimation," *Econometrica*, 65, 861 - 874.

Kocherlakota, N. R. (1990): "On Tests of Representative Consumer Asset Pricing Models," *Journal of Monetary Economics*, 26, 285 - 304.

Künsch, H. R. (1989): "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, 17, 1217 - 1241.

Linton, O., and J. P. Nielsen (1994): "A Multiplicative Bias Reduction Method for Nonparametric Regression," *Statistics & Probability Letters*, 19, 181 - 187.

Nielsen, J. P. (1998): "Multiplicative Bias Correction in Kernel Hazard Estimation," *Scandinavian Journal of Statistics*, 25, 541 - 553.

Nielsen, J. P., and C. Tanggaard (2001): "Boundary and Bias Correction in Kernel Hazard Estimation," *Scandinavian Journal of Statistics*, 28, 675 - 698.

Newey, W. K., and K. D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703 - 708.

Newey, W. K., and K. D. West (1994): "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631 - 653.

Oliner, S. D., G. D. Rudebusch, and D. Sichel (1996): "The Lucas Critique Revisited: Assessing the Stability of Empirical Euler Equations for Investment," *Journal of Econometrics*, 70, 291 - 316.

Parzen, E. (1957): "On Consistent Estimates of the Spectrum of a Stationary Time Series," *Annals of Mathematical Statistics*, 28, 329 - 348.

Parzen, E. (1963): "Notes on Fourier Analysis and Spectrum Windows," Technical Report No.48, Department of Statistics, Stanford University. (Published in Parzen, E. (1967): *Time Series Analysis Papers*. San Francisco: Holden-Day.)

Politis, D. N., and J. P. Romano (1995): "Bias-Corrected Nonparametric Spectral Estimation," *Journal of Time Series Analysis*, 16, 67 - 103.

Rotemberg, J. J. (1984): "Interpreting the Statistical Failures of Some Rational Expectations Macroeconomic Models," *American Economic Review Papers and Proceedings*, 74, 188 - 193.

Runkle, D. E. (1991): "Liquidity Constraints and the Permanent-Income Hypothesis," *Journal of Monetary Economics*, 27, 73 - 98.

Smith, D. C. (1999): "Finite Sample Properties of Tests of the Epstein-Zin Asset Pricing Model," *Journal of Econometrics*, 97, 113 - 148.

Stuetzle, W., and Y. Mittal (1979): "Some Comments on the Asymptotic Behavior of Robust Smoothers," in T. Gasser and M. Rosenblatt (eds.), *Smoothing Techniques for Curve Estimation: Proceedings of a Workshop Held in Heidelberg, April 2 - 4, 1979*. Berlin: Springer-Verlag, 191 - 195.

Tauchen, G. (1986): "Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained from Financial Market Data," *Journal of Business & Economic Statistics*, 4, 397 - 425.

Velasco, C., and P. M. Robinson (2001): "Edgeworth Expansions for Spectral Density Estimates and Studentized Sample Mean," *Econometric Theory*, 17, 497 - 539.

West, K. D. (1988): "Dividend Innovations and Stock Price Volatility," *Econometrica*, 56, 37 - 61.

West, K. D. (1997): "Another Heteroskedasticity- and Autocorrelation-Consistent Covariance Matrix Estimator," *Journal of Econometrics*, 76, 171 - 191.

West, K. D., and D. W. Wilcox (1996): "A Comparison of Alternative Instrumental Variables Estimators of a Dynamic Linear Model," *Journal of Business & Economic Statistics*, 14, 281 - 293.

Xiao, Z., and O. Linton (2002): "A Nonparametric Prewhitened Covariance Estimator," *Journal of Time Series Analysis*, 23, 215 - 250.