Spectral Density Bandwidth Choice and Prewhitening in the Estimation of Heteroskedasticity and Autocorrelation Consistent Covariance Matrices In Panel Data Models

Min-Hsien Chiang
Institute of International Business
National Cheng-Kung University
Tainan, TAIWAN

Yongmiao Hong
Department of Economics &
Department of Statistical Science
Cornell University
and

Chihwa Kao
Department of Economics &
Center for Policy Research
Syracuse University

February 1, 2002

ABSTRACT

This paper investigates the performances kernel-based with and without prewhitening and the parametric heteroskedasticity and autocorrelation consistent (HAC) covariance matrices in panel data models. A Monte Carlo study is conducted to evaluate effects of kernel choice, data-based bandwidth selection, and prewhitening on the HAC covariance matrics in finite samples for panel data models.

Key Words: HAC covariance matrix, Kernel estimation, Panel data model

1. INTRODUCTION

This paper is the first step in understanding how to construct the heteroskedasticity and autocorrelation consistent (HAC) covariance matrices in panel data models. An important development in panel data models has been the modelling of serial correlation. To account for serial correlation, researchers often assume a particular form for serial correlation in the model, e.g., Hause (1977, 1980), Lillard and Willis (1978), Anderson and Hsiao (1982), Berry et al. (1988), Baltagi and Li (1991), Keane and Runkle (1998), Levitt (1998), Karlsson and Skoglund (2000), Park and Sickles (2000), and Blundell and Bond (1999). All these papers assumed a known form (e.g., AR(1) or MA(1)) for serial correlation to obtain consistent estimators and related covariance matrix estimators. Hause (1977, 1980), Lillard and Willis (1978), and Baltagi and Li (1991) used the generalized least squares estimator (GLS), whereas Anderson and Hsiao (1982) considered a maximum likelihood estimation (MLE) with normality. Karlsson and Skoglund (2000) and Skoglund and Karlsson (2001a) considered the MLE in a two-way error component model when the time-specific component follows an AR(1) or MA(1) process. Skoglund and Karlsson (2001b) introduced the serial correlation of general form in the time effects as well as the error terms. They proposed a maximum likelihood estimator and discussed a model selection procedure for determining the orders of serial correlation and the importance of time and individual effects.

Theses estimators are consistent and efficient when the assumed form of serial correlation is correct, and the related covariance matrix estimators are expected to perform well in finite samples. However, in practice, no prior information about serial correlation is available. In particular, the assumption of a common structure of serial correlation for all individuals are overly strong and are likely not to hold in practice. It is a rule rather than an exception that degrees of serial correlation will vary across individuals. Any misspecification of serial correlation may lead to misleading conclusions in estimation, inference and hypothesis testing. This is so even when estimators remain consistent (e.g., the OLS for the static panel models), because in general the covariance estimators that incorporate the assumed serial correlation will be not consistent. It is important to develop procedures that are robust to serial correlation of unknown form. This paper considers a general panel model for the error terms that may have heterogeneity and serial dependence of unknown form. More importantly, we propose a class of HAC estimation of covariance matrices used in panel data.

HAC estimators are not as often used in panel models as in time series, but attention to this important issue has been increased over the last few years. Over the long time in the panel literature, various strong assumptions imposed render it unnecessary to estimate HAC estimators. In the related literature, Kiefer (1980), and Arellano (1987) are not directly concerned with the HAC but do touch on a number of related issues on covariance estimation. Kiefer (1980) proposed a within-GLS estimator with arbitrary intertemporal covariance, whereas Arellano (1987) suggested a simple method for obtaining robust estimates of the standard errors that allow a general covariance matrix. In a dynamic panel model with small time dimension, T, and large cross-sectional dimension, n, Arellano and Bond (1991, p. 279) proposed a heteroskedascity consistent covariance matrix which is robust to general heteroskedasticity over individuals and over time but is not robust to serial correlation. Such estimators will not be valid when there exists serial correlation of unknown form. Kezdi (2001) examined the within estimator with short time series. Kezdi showed serial correlation in the error terms and regressors will induce severe bias in the conventional standard error estimates in panel data models.

Recently, HAC estimators have been used in nonstationary panels, e.g., Kao (1999), Kao and Chiang (2000), and Phillips and Moon (1999). These authors use the kernel estimators developed for time series models. Often, a data-driven bandwidth is determined as a function of T only. As will be seen below, such a procedure, although optimal for time series models, is not optimal in the panel context. In particular, it oversmoothes the HAC estimator because it does not take into account the additional smoothing provided by n. This is an undesirable feature because it is well-known that kernels often tend to underestimate the covariance matrixes and undersmoothing will lead to further downward bias. Our procedure will take this into account and the results are optimal for panel models.

In fact, HAC estimators are equally important in panel models as in time series models. For example, in static panel models, the within estimator is still consistent when there exists serial correlation of unknown form, but the standard covariance matrix estimators that assume no serial correlation will be invalid. In this case, the use of conventional t-tests and F-tests will lead to misleading conclusions, e.g., Bhargava et al. (1982, p.545).

Estimation of HAC covariance matrices is a long-standing problem in time series econometrics, e.g., Newey and West (1987, 1994), Andrews (1991), Andrews and Monahan (1992), den Hann and Levin (1998), and Hong and Lee (2000). Leading examples in panel contexts that may require using HAC estimators are estimation of asymptotic covariance matrices of least square estimators in linear, nonlinear and unit root/cointegration regression models, of two-stage least squares, three-stage least squares, quasi-maximum likelihood and generalized method of moment (GMM) estimators. Such covariance matrix estimation is important for confidence interval estimation, inference and hypothesis testing in dynamic contexts. To represent a covariance matrix by a spectral density matrix at frequency 0 and to estimate it by nonparametric kernel methods was suggested by Brillinger (1975, p.184; 1979), Hansen (1982, p.1047) and Phillips and Ouliaris (1988), among others. Various kernel-based covariance estimators have been proposed. These include Domowitz and White (1982), Levine (1983), White (1984), White and Domowitz (1984), Newey and West (1987, 1994), Gallant (1987), Gallant and White (1988), Kool (1988), Andrews (1991), Andrews and Monahan (1992), Hansen (1992), De Jong and Davidson (1999), and Xiao and Linton (2000). Andrews (1991) and Newey and West (1994) propose some data-driven bandwidth choices suitable for covariance matrix estimation, making the kernel methods operational in practice. Andrews (1991) derives the optimal kernel—the Quadratic-Spectral (QS) kernel over a class of kernels that generate positive semi-definite covariance estimators. den Haan and Levin (1998, 2000) also propose an autoregression-based covariance estimator.

The bulk of the problem is the difficulty in estimating a spectral density matrix at frequency 0 when it has a peak there, which can arise due to strong dependence. It is well-known that positive autocorrelation is apt to entail a mode in the spectral density at frequency $\mathbf{0}$, and strong autocorrelation yields a peak at frequency $\mathbf{0}$. Kernel estimators often tend to underestimate the peak, leading to overly narrow confidence intervals and liberal tests. In fact, Priestley (1981, pp.547-556) shows that the modes of the spectral densities of some low order AR and ARMA processes, whose autocorrelations decay to 0 at an exponential rate, are still underestimated even if some undersmoothing bandwidths are used. Spectral peaks often arise in economic time series, due to seasonalities, business cycle periodicities and strong dependence. Cochrane (1988), for example, argues that for economic data, low order ARMA procedures tend to yield poor estimates of infinite sums of autocorrelations (i.e., the long-run variance), because the autocorrelation function often is positive and decays slowly. Granger (1969) points out that the typical spectral shape of many economic time series is that it has a sharp peak at frequency $\mathbf{0}$ and decays to $\mathbf{0}$ as frequency increases. For such time series, kernel methods may not work well.

Because of the unsatisfactory finite sample performances of the kernel-based covariance estimators, it has been emphasized in the literature (e.g., Newey and West 1994, p.632) that extensions or refinements to the existing kernel methods should be a priority for further work. More reliable sampling distribution theory and better covariance estimators are required for the statistics used in economic and financial time series analysis. To our knowledge, however, few progress has been made so far. The most noticeable progress is Andrews and Monahan's (1992) prewhitening procedure. Prewhitening is a technique aimed to improve the accuracy of spectral density estimators by making certain transformations to the data before applying spectral estimation procedures. The idea is to "flatten" the spectral density by passing the original series through a filter so that its output has a relatively flat spectrum. A flat spectrum is much easier to estimate and the corresponding kernel estimator is less sensible to the choice of a bandwidth. Andrews and Monahan's (1992) prewhitening kernel estimator is effective in reducing the bias, and leads to considerably better sizes for related test statistics. In the meantime, it is also found that prewhitening inflates the variance and may lead to a larger mean squared error (MSE) than the kernel estimator without prewhitening (see Andrews and Monahan 1992, Newey and West 1994, p.634).

Our contributions in this paper is that we compare the finite sample performances of kernel-based estimators: nonparametric with and without prewhitening and parametric method of de Haan and Levin (2000) in the panel contexts.

In Section 2, we describe the framework in which estimation of heteroskedasticity and autocorrelation consistent covariance matrices in panel data of interest. In Section 3, we report the Monte Carlo results. Section 4 provides a concluding remark and directions for further research. Unless indicated, all limits are taken as the sample size $n \to \infty$ and $T \to \infty$.

2. MODEL

To motivate the problem, we first consider an two-way error component panel model with a possibly heteroskedastic and autocorrelated disturbance error

$$Y_{it} = \alpha_0 + X_{it}'\beta_0 + \mu_i + \lambda_t + v_{it}, \qquad t = 1, ..., T_i, i = 1, ..., n, \quad n, T_i \in \mathbb{Z}^+, \tag{2.1}$$

where Y_{it} is a scalar, X_{it} is a $p \times 1$ vector of explanatory variables that may contain lagged dependent variables Y_{it-h} $(p, h \in \mathbb{Z}^+)$, α_0 is an intercept, β_0 is a $p \times 1$ vector of the slope parameters, μ_i is the individual effect, λ_t is the time effect, and v_{it} is the error term. We allow fixed effects or random effects. Throughout, we assume $T_i = c_i T$ for some integer T and $c_i \in [c, C]$. Thus, we permit unbalanced panel data. Moreover, we allow Y_{it}, X_{it}, α_0 and β_0 to depend on both n and T. (For notational simplicity, we suppress such dependence.) Throughout, we assume the following conditions on (2.1):

Assumption A.1: $\{Y_{it}, X_{it}'\}'$ are stochastic processes such that (i) for each i, $\{v_{it}\}$ is covariance-stationary with $E(v_{it}) = 0$, $E(v_{it}^2) = \sigma_i^2 \in [c, C]$ and $E(v_{it}^8) \in [c, C]$; (ii) there is no spatial dependence in $\{v_{it}\}$, i.e., v_{it} and v_{js} are independent for all $i \neq j$ and all t, s; (iii) the individual and time effects, μ_i and λ_t, can be stochastic (random effects) or deterministic (fixed effects).

No dependence assumptions on $\{\mu_i\}$ and $\{\lambda_t\}$ are imposed, because they will be differenced out in the construction of our estimators. We thus allow $\{\lambda_t\}$ to be serially correlated if λ_t is random, and $\{\mu_i\}$ to be spatially correlated if μ_i is random. We also allow a certain degree of heterogeneity in panel data—$\{Y_{it}, X'_{it}\}'$ need not be stationary for each i, and the errors $\{v_{it}\}$ may have different variances across i. In particular, we allow some nonstationary processes. One example of nonstationary panel time series is the deterministic trend process (e.g., Kao and Emerson 1999)

$$Y_{it} = \alpha + \gamma_1 t + \gamma_2 t^2 + \cdots + \gamma_p t^p + \mu_i + \lambda_t + v_{it}.$$

This is covered by (2.1) with $X_{it} \equiv [t/T, (t/T)^2, ..., (t/T)^p]'$ and $\theta \equiv (T\gamma_1, ..., T^p\gamma_p)'$. Note that X_{it} and β_0 depend on T. Another example is the panel cointegration process (e.g., Phillips and Moon 1999, Kao and Chiang 2000):

$$Y_{it} = \alpha + \gamma Z_{it} + \mu_i + \lambda_t + v_{it},$$

where $Z_{it} = Z_{it-1} + \varepsilon_{it}$, $\{\varepsilon_{it}\}$ is $I(0)$ for each i, and $\{\varepsilon_{it}\}$ may or may not be correlated with $\{v_{it}\}$. This process is also covered by (2.1) with $X_{it} \equiv T^{-1} Z_{it}$ and $\beta_0 \equiv T\gamma$. We will provide regularity conditions on transformed variables $\{X_{it}\}$ and transformed parameters β.

The parameter vector β_0 in (2.1) can be estimated by the popular within estimator

$$\begin{aligned} \widehat{\beta} \equiv & \left[\sum_{i=1}^{n}\sum_{t=1}^{T_i}\left(X_{it} - \bar{X}_i - \bar{X}_t + \bar{X}\right)\left(X_{it} - \bar{X}_i - \bar{X}_t + \bar{X}\right)'\right]^{-1} \\ & \times \left[\sum_{i=1}^{n}\sum_{t=1}^{T_i}\left(X_{it} - \bar{X}_i - \bar{X}_t + \bar{X}\right)\left(Y_{it} - \bar{Y}_i - \bar{Y}_t + \bar{Y}\right)\right], \end{aligned} \tag{2.2}$$

where $\bar{X}_i \equiv T_i^{-1}\sum_{t=1}^{T_i} X_{it}$, $\bar{X}_t \equiv n^{-1}\sum_{i=1}^{n} X_{it}$ and $\bar{X} \equiv (nT_i)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T_i} X_{it}$. The variables $\bar{Y}_i, \bar{Y}_t$ and $\bar{Y}$ are defined in the same ways. Its asymptotic covariance matrix is

$$\text{AVAR}\left[\sqrt{nT}(\widehat{\beta} - \beta_0)\right] = \left(p\lim \hat{M}_{nT}\right)^{-1} p\lim \hat{\Omega}_{nT}\left(p\lim \hat{M}_{nT}\right)^{-1}, \tag{2.3}$$

where

$$\hat{M}_{nT} \equiv \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\tilde{X}_{it}\tilde{X}'_{it},$$

$$\hat{\Omega}_{nT} \equiv \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\sum_{s=1}^{T_i}\tilde{X}_{it}\tilde{v}_{it}(\tilde{v}_{is}\tilde{X}_{is})',$$

$\tilde{X}_{it} \equiv X_{it} - \bar{X}_i - \bar{X}_t + \bar{X}$, and $\tilde{v}_{it} \equiv v_{it} - \overline{v}_i - \overline{v}_t + \overline{v}$. To estimate (2.3), $p\lim \hat{\Omega}_{nT}$ is more challenging to estimate.

More generally, for many panel estimators $\widehat{\beta}$, we have

$$(M_{nT}\Omega_{nT}M'_{nT})^{-1}\sqrt{nT}(\widehat{\beta} - \beta_0) \xrightarrow{d} N(0, I_r), \qquad r \in \mathbb{Z}^+, \tag{2.4}$$

where M_{nT} is a nonstochastic $r \times p$ matrix, I_r is a $r \times r$ identity matrix, and

$$p \lim \hat{\Omega}_{nT} = \frac{1}{nT_i} \sum_{i=1}^{n} \sum_{t=1}^{T_i} \sum_{s=1}^{T_i} V_{it}(\beta_0) V_{is}(\beta_0)' \tag{2.5}$$

for some stochastic $p \times 1$ vector process $V_{it}(\beta_0)$. For example, the function $V_{it}(\beta_0)$ can be the moment function in GMM estimation, e.g., Arellano and Bond (1991), Arellano and Bover (1995), Ahn and Schmidt (1995, 1997), Hahn (1997), Blundell and Bond (1998), and Im et al. (1999). Also the HAC estimators proposed in this paper can be used in many other panel models. For example, the panel cointegration tests and panel cointegration estimation (e.g., Kao, 1999; Kao and Chiang, 2000, Phillips and Moon, 1999) require the HAC estimation for the long-run variance matrix of the error terms in the models. In an extensive simulation study, Kao and Chiang (2000) pointed out the panel fully modified (FM) estimator and t-statistic based on FM estimator are severely downward biased due to the failure of the kernel-based HAC estimation for the long-run variance covariance matrix. More seriously, Kao and Chiang also pointed out that the FM t-steatitic become more negatively biased as the cross-sectional dimension, n, increases. All these indicate there is much to be done on the HAC estimation in panel data models. For the panel cointegration test estimation (e.g., Kao and Chiang, 2000, p. 187; Phillips and Moon 1999, p. 1084), $V_{it}(\beta_0)$ usually takes the form:

$$V_{it}(\beta) = \begin{bmatrix} \tilde{Y}_{it} - \tilde{X}_{it}'\beta \\ X_{it} - X_{it-1} \end{bmatrix}.$$

Once the estimates of $V_{it}(\beta_0)$, $V_{it}(\beta_0)$ were estimated, the HAC estimator of the long run covariance matrix was estimated by

$$\widehat{\Omega} = \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{1}{T} \sum_{t=1}^{T} V_{it}(\widehat{\beta}) V_{it}'(\widehat{\beta}) + \frac{1}{T} \sum_{\tau=1}^{l} \varpi_{\tau l} \sum_{t=\tau+1}^{T} \left(V_{it}(\widehat{\beta}) V_{it-\tau}'(\widehat{\beta}) + V_{it-\tau}(\widehat{\beta}) V_{it}'(\widehat{\beta}) \right) \right\},$$

where $\widehat{\beta}$ is the within estimator or the FM estimator and $\varpi_{\tau l}$ is a weight function or a kernel. The distribution results for the FM estimator in Kao and Chiang (2000) and Phillips and Moon (1999) require $\sqrt{n}\left(\widehat{\Omega} - \Omega\right)$ does not diverge as n grows large. However, $\widehat{\Omega} - \Omega$ may not be small when T is fixed. It follows that $\sqrt{n}\left(\widehat{\Omega} - \Omega\right)$ may be non-negligible in panel data with finite samples. It may be one of the reasons for the poor performance for panel FM estimator. For GMM estimator in panel data, $V_{it}(\beta_0)$ usually takes the form:

$$V_{it}(\beta_0) = Z_{it}\left(\tilde{Y}_{it} - \tilde{X}_{it}'\beta_0\right)$$

where Z_{it} is a vector of instruments with $E\left[Z_{it}\left(\tilde{Y}_{it} - \tilde{X}\beta_0\right)\right]$ and $Z_{it}\left(\tilde{Y}_{it} - \tilde{X}\beta_0\right)$ may have serial correlation and heteroskedasticity of unknown form. However, for the GMM is the dynamic panel models, one must be careful about how to choose the valid instruments when the serial correlation has the unknown form. For example, the consistency of the GMM in Arellano and Bond (1991) that use the lagged dependent variables be the instruments requires no second-order serial correlation in the first difference residuals It seems that if the error terms are serially correlated with the unknown form, then one cannot choose the lagged dependent variables to be the valid instruments in the dynamic panel models. In stead of using lagged dependent variables, we may have to use exogenous variables to be instruments.

Usually, M_{nT} is relatively simple to estimate, often by its sample analog. Our focus is estimation of Ω_{nT}. When $V_{it}(\beta_0)$ is a second order stationary process with mean $\mathbf{0}$, we have

$$\lim_{n,T\to\infty} \Omega_{nT} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \Omega_i = \Omega \equiv \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 2\pi f_i(0), \tag{2.6}$$

where

$$f_i(0) = (2\pi)^{-1} \sum_{l=-\infty}^{\infty} \Gamma_i(l)$$

is the $p\times p$ spectral density matrix of $V_{it}(\beta_0)$ at frequency $\mathbf{0}$, with $\Gamma_i(l) = E[V_{it}(\beta_0)V_{i,t-l}(\beta_0)']$. Thus, Ω_i can be consistently estimated by a nonparametric spectral density estimator at frequency $\mathbf{0}$, as suggested in Brillinger (1975), Hansen (1982) and Phillips and Ouliaris (1988) among others. Newey and West (1987) propose a convenient positive semi-definite kernel estimator for Ω_i :

$$\hat{\Omega}_{i\mathrm{NW}} = \sum_{l=-B_n}^{B_n} K(j/B_{iT})\hat{\Gamma}_{iT}(l), \tag{2.7}$$

where $K(x) = (1-|x|)\mathbf{1}(|x|\le 1)$ is the Bartlett kernel, $\mathbf{1}(\cdot)$ is the indicator function, B_{iT} is a lag truncation parameter depending on the sample size T,

$$\hat{\Gamma}_{iT}(l) = \begin{cases} T^{-1}\sum_{t=l+1}^{T} V_{it}(\widehat{\beta})V_{i,t-l}(\widehat{\beta})', & l = 0,1,...,T-1, \\ T^{-1}\sum_{t=1-l}^{n} V_{i,t+l}(\widehat{\beta})V_{it}(\widehat{\beta})', & l = -1,...,-(T-1), \end{cases} \tag{2.8}$$

is the sample autocovariance matrix of $V_{it}(\widehat{\beta})$, and $\widehat{\beta}$ is a consistent estimator of θ.

Andrews (1991) consider a general class of estimators

$$\hat{\Omega}_{i\mathrm{A}} = \sum_{l=1-T}^{T-1} K(j/B_{iT})\hat{\Gamma}_{iT}(l), \tag{2.9}$$

where $K : \mathbb{R} \to [-1,1]$ is a general kernel, and B_n a bandwidth. Examples of $K(\cdot)$ include Bartlett, Parzen, QS, Tukey-Hanning, and truncated kernels (e.g., Andrews 1991). When

$K(\cdot)$ has infinite support, B_{iT} is no longer a lag truncation parameter. Andrews derives the optimal kernel —the QS kernel, that minimizes an asymptotic MSE; he also proposes a parametric "plug-in" data-driven bandwidth choice for B_{iT}. Based on Andrews (1991), Phillips and Moon (1999, p. 1084) proposed a HAC estimator by taking the average of $\hat{\Omega}_{i\mathrm{A}}$ over i:

$$\hat{\Omega} = \frac{1}{n} \sum_{i=1}^{n} \hat{\Omega}_{i\mathrm{A}}.$$

Phillips and Moon (1999) verified the consistency of $\hat{\Omega}$ by assuming $B_{iT}, n, T \to \infty$ with $B_{iT}/T \to 0$. Newey and West (1994) propose a nonparametric "plug-in" data-driven choice of B_{iT} for their Bartlett kernel-based estimator $\hat{\Omega}_{i\mathrm{NW}}$. Andrews and Manahan (1992) further propose a prewhitening kernel estimator

$$\hat{\Omega}_{i\mathrm{AM}} = \left[I - \hat{A}_i(\widehat{\beta})\right]^{-1} \left[\sum_{l=1-T}^{T-1} K(j/B_{iT})\hat{\Gamma}_{iT}^{+}(l)\right] \left[I - \hat{A}_i(\widehat{\beta})\right]^{-1}, \qquad (2.10)$$

where $\hat{A}_i(\widehat{\beta})$ is a filter based on a Vector AutoRgression (VAR) approximation for $\{V_{it}(\widehat{\beta})\}$ and

$$\hat{\Gamma}_{iT}^{+}(l) = \begin{cases} T^{-1} \sum_{t=l+1}^{T} V_{it}^{+}(\widehat{\beta}) V_{i,t-l}^{+}(\widehat{\beta})', & l = 0, 1, ..., T-1, \\ T^{-1} \sum_{t=l+1}^{T} V_{i,t+l}^{+}(\widehat{\beta}) V_{it}^{+}(\widehat{\beta})', & l = -1, ..., -(T-1) \end{cases}$$

is the sample autocovariance function of the VAR residual $V_{it}^{+}(\widehat{\beta})$.

Extensive simulation experiments in the literature show that kernel estimators perform poorly in finite samples when there is strong autocorrelation. They often lead to strong overrejection in testing and too narrow confidence intervals in estimation. This is true even if the (infeasible) finite sample optimal bandwidth parameter is used. It appears that it is the very nature of the kernel method, rather than the choice of a bandwidth or a kernel, that attributes its poor performance in finite samples when the data display strong dependence.

In our opinion, the main reason for the poor performance of the kernel estimators is that the spectral density has a peak at frequency $\mathbf{0}$ when there exists strong autocorrelation, but the kernel method, as a local averaging method, tends to underestimate the peak. Andrews and Monahan's (1992) prewhitening procedure alleviates this downward bias substantially and thus gives better test sizes. However, it also inflates the variance and thus may not dominate the same procedure applied to the original series in terms of MSE criteria. Below we study the finite sample properties of the various HAC estimators in the panel context.

3. MONTE CARLO EVIDENCE

We now compare the finite sample performances of three HAC covariance matrix estimators: nonparametric kernel-based with and without prewhitening estimators and the parametric VARHAC estimator. We consider two sets of simulation experiments: the first basically follow the designs of Baltagi and Li (1995), Andrews (1991) and Andrews and Monahan (1992), and the second are based on an empirical study on the U.S. juvenile crime rates by Levitt (1998).

For the first set of experiments, We consider the following DGP for the panel data model:

$$\begin{aligned} Y_{it} &= \alpha + \beta X_{it} + \mu_i + v_{it}, \\ X_{it} &= 0.5 X_{it-1} + \eta_{it}; \end{aligned}$$

where $\eta_{it} \overset{iid}{\sim} U[-0.5, 0.5]$, $\mu_i \overset{iid}{\sim} N(0, \sigma_\mu^2)$, $\alpha = 5$ and $\beta = 0.5$. The initial values X_{i0} were chosen as in Baltagi *et al.* (1992). We let $\sigma^2 \equiv \sigma_\mu^2 + \sigma_v^2 = 20$ and $\tau \equiv \frac{\sigma_\mu^2}{\sigma^2}$ take five different values, $(0, 0.4, 0.8)$. The value of τ measures the relative strength of random effects (when $\tau = 0$, there is no random effect). We note that a similar DGP has been used in Baltagi and Li (1995) and Bera *et al.* (2000). We consider three sample size combinations: $(n, T) = (25, 32), (50, 64), (100, 128)$. To examine the performance of kernel-based HAC estimators, we consider the following processes for $\{v_{it}\}$, where $\{v_{it}\}$ follows an ARMA(p_0, q_0) process, i.e.,

$$v_{it} = \sum_{l=1}^{p_0} \rho_l v_{it-1} + \sum_{l=1}^{q_0} \eta_l \varepsilon_{t-l} + \varepsilon_t, \ \{\varepsilon_t\} \sim \text{IID}.N(0, \sigma^2).$$

In our simulation, we consider ARMA(1,1), ARMA(2,2) and ARMA(4,4). For ARMA(1,1), we set the AR parameter $\rho_l = 0.6$ and 0.9. For ARMA(2,2) and ARMA(4,4), we set the values of $\{\rho_l\}$ such that their sum equals to 0.6 and 0.9. For each model, we set each MA parameter $\eta_l = 0.6$ or 0.9. We also set $\sigma^2 = 1$ in each model.

We compare the following covariance estimators: Newey and West's (NW, 1994) Bartlett kernel-based estimator, Andrews' (1991) QS estimator, and the VARHAC estimator of den Haan. For NW, we select the data-driven bandwidth using Newey and West's (1994, pp.637) nonparametric plug-in method. For QS, we select the data-driven bandwidth using Andrews' (1991) parametric plug-in method based on univariate $\{\text{AR}(m_a)\}_{a=1}^p$ models, where the order m_a is selected by AIC or BIC. The resulting estimator is denoted as QS(A) or QS(B). We also use AIC or BIC, and obtain the estimator VARHAC, i.e., VARHAC(A) and VARHAC(B).

Alternatively, we also apply a prewhitening (PW) procedure to NW and QS: we first fit a prehwitening VAR(m) model for $\{V_t(\widehat{\beta}_n)\}$ with the order m determined by AIC or BIC again, use the resulting residual vector series $\{V_t^+(\widehat{\beta}_n)\}$ to construct NW and QS estimators, and then recolor them. For NW, we only need to use AIC or BIC to select the order of the prewhitening VAR(m) model. We obtain PW-NW(A) or PW-NW(B). For QS we apply the same order selection criterion to both the VAR(m) model for prewhitening and the univariate $\{\text{AR}(m_a)\}_{a=1}^{p}$ models for choosing the plug-in bandwidth or the finest scale. These are denoted as PW-QS(A) and PW-QS(B) respectively.

In order to study the potential heterogeneity in the serial correlation across the cross-sectional units of the serial correlation in panel data for each case we study the following five subcases:

ARMA(4,4) Alternatives:

$$\begin{cases} \text{ARMA(4,4)}^a : & v_{it} = -\rho v_{it-4} + \varepsilon_{it} + \eta\varepsilon_{it-4}, \quad i = 1, ..., n, \\ \text{ARMA(4,4)}^b : & v_{it} = \rho v_{it-4} + \varepsilon_{it} - \eta\varepsilon_{it-4}, \quad i = 1, ..., n, \\ \text{ARMA(4,4)}^c : & v_{it} = \begin{cases} -\rho v_{it-4} + \varepsilon_{it} + \eta\varepsilon_{it-4}, & i = 1, ..., \frac{n}{2}, \\ \varepsilon_{it}, & i = \frac{n}{2} + 1, ..., n, \end{cases} \\ \text{ARMA(4,4)}^d : & v_{it} = \begin{cases} \rho v_{it-4} + \varepsilon_{it} - \eta\varepsilon_{it-4}, & i = 1, ..., \frac{n}{2}, \\ \varepsilon_{it}, & i = \frac{n}{2} + 1, ..., n, \end{cases} \\ \text{ARMA(4,4)}^e : & v_{it} = \begin{cases} -\rho v_{it-4} + \varepsilon_{it} + \eta\varepsilon_{it-4}, & i = 1, ..., \frac{n}{2}, \\ \rho v_{it-1} + \varepsilon_{it} - \eta\varepsilon_{it-4}, & i = \frac{n}{2} + 1, ..., n. \end{cases} \end{cases} \tag{7.6}$$

We estimate β by the within estimator $\widehat{\beta}$. We examine the various estimators for the asymptotic variance of $\widehat{\beta}$. We examine their biases, variances and MSEs.

We also examine the size of the t-test of

$$H_0: \ \beta = 0 \qquad \text{v.s.} \qquad H_A: \ \beta = \delta,$$

The test is constructed using the within estimator $\widehat{\beta}$ and the various covariance estimators.

Table 1 reports the bias, variance, MSE, and the size of the t test under ARMA (1,1) for $(n, T) = (25, 32)$ with $\tau = (0.0, 0.4, 0.8)$. The VARHAC(A) has the smallest downward bias, then followed by QS(B), QS(A), and NW. As found in previous studies in time series, the prewhitening estimators of PW-NW(A), PW-NW(B), PW-QS(A), and PW-QS(B) have larger MSEs than non-prewithening methods. In general, QS methods have smaller biases than their NW counterparts. For the size of tests, VARHAC(A) gains a better size than rests of methods but has a largest variance. Also, QS methods have good sizes in contrast to NW counterparts. The prewhitening methods, however, have relatively worse sizes than non-prewhitening methods. On the other hand, biases of all methods reduce when the random effect increases. This can be observed that MSEs for all methods tend to be smaller in correspondence to decrease of τ. The sizes of tests show no signs of improvements even though MSEs decrease. Tables 2 and 3 share the similar observations as found in Table 1. The biases and MSEs, however, tend to become larger when the dimension of combination of N and T increases for all methods. Meanwhile, the sizes of tests improves when sample size increases for both N and T.

The similar observations can be found in ARMA (2,2) and ARMA(4,4) cases from Tables 4 to 9. Furthermore, the biases and MSEs are relatively larger than those in ARMA (1,1) cases. This indicates that estimated variances deviate more when the errors become much more correlated.

In summary, we observe the following:

1) The VARHAC(A) estimator has a smaller bias and a larger variance than other estimators. The MSEs of the prewhitening estimators are larger than those of non-prewhitening methods.

2) For the size of the tests, the VARHAC(A) estimator outperforms other methods.

3) The prewhitening procedures enlarge MSEs.

4. CONCLUSION

As is well-known, a HAC covariance matrix is proportional to a spectral density matrix at frequency 0, and can be consistently estimated by the popular kernel methods of Andrews-Newey-West. When data displays strong dependence, the spectral density has a peak at frequency 0. Kernels, as a local averaging method, tend to underestimate the peak. This often leads to overrejection in testing and too narrow confidence intervals in estimation. In this paper, the issues of bandwidth selection and prewhitening of estimating the HAC covariance matrix in panel data models are investigated using Monte Carlo.

REFERENCES

Ahn, S. C. and Schmidt, P. (1995), "Efficient Estimation of Models for Dynamic Panel Data," *Journal of Econometrics*, 68, 5-27.

Ahn, S. C. and Schmidt, P. (1997), "Efficient Estimation of Models for Dynamic Panel Data: Alternative Assumptions and Simplified Estimation," *Journal of Econometrics*, 76, 309-321.

Anderson, T. W. and Hsiao, C. (1982), "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 18, 47-82.

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

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

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

Arellano, M. (1987), "Computing Robust Standard Errors for Within-Groups Estimators," *Oxford Bulletin of Economics and Statistic*s, 49, 431-434.

Arellano, M. and Bover, O. (1995), "Another Look at the Instrumental Variables Estimation of Error-Component Models," *Journal of Econometrics*, 68, 29-51.

Baltagi, B. H., and Li, Q. (1995), "Testing AR(1) Against MA(1) Disturbances in an Error Component Model," *Journal of Econometrics*, 68, 133 - 151.

Beltrao, K. and P. Bloomfield (1987), "Determining the Bandwidth of a Kernel Spectrum Estimate", *Journal of Time Series Analysis* 8, 21-38.

Berry, S., Gottschalk, P., and Wissoker, D. (1988), "An Error Components Model of the Impact of Plant Closing on Earnings," *Review of Economics and Statistics*, 70, 701 - 707.

Bhargava, A., Franzini, L., and Narendranathan, W. (1982), "Serial Correlation and the Fixed Effects Model," *Review of Economic Studies*, 49, 533-549.

Blundell, R. W., and Bond, S. (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115-143.

Blundell, R. W., and Bond, S. (1999), "GMM Estimation with Persistent Panel Data: An Application to Production Functions," Working Paper Series No. W99/4, The Institute for Fisical Studies.

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

Christiano and W. den Haan (1995), "Small Sample Properties of GMM for Business Cycle Analysis", *Journal of Business and Economic Statistics.*

Cochrane, J. (1988), "How Big is the Random Walk in GNP?" *Journal of Political Economy* 96, 893-920.

den Haan, W. and A. Levin (1997), "A Practitioner's Guide to Robust Covariance Matrix Estimation", in *Handbook of Statistics: Robust Inference*, Vol. 15, eds. G.S. Maddala and C.R. Rao. Elsevier; New York.

den Haan, W. and A. Levin (1998), "Vector Autoregressive Covariance Matrix Estimation," Department of Economics Working Paper, University of California, San Diego.

den Haan, W. and A. Levin (2000), "Robust Covariance Matrix Estimation with Data-Dependent VAR Prewhitening Order," Department of Economics Working Paper, University of California, San Diego.

De Jong, R. and J. Davidson (1999), "Consistent Estimation of Heteroskedasticity and Autocorrelation Consistent Covariance Matrices," *Econometrica*.

Gallant, A.R. (1987), Nonlinear Statistical Models. Wiley: New York.

Gallant, A.R. and H. White (1988), A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models. Basil Blackwell: New York.

Granger, C. (1969), "Investigating causal Relations by Econometric Models and Cross-Spectral Methods", *Econometrica* 37, 315-340.

Hahn, J. (1997), "Efficient Estimation of Panel Data Models with Sequential Moment Restrictions," *Journal of Econometrics*, 79, 1-21.

Hahn, J. (1999), "How Informative is the Initial Condition in the Dynamic Panel Model with Fixed Effects?" *Journal of Econometrics*, 93, 309-326.

Hannan, E. (1970), *Multiple Time Series*. Wiley: New York.

Hansen, B. (1992), "Consistent Covariance Matrix Estimation for Dependent Heterogenous Processes", *Econometrica* 60, 967-972.

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

Hause, J. (1977), "The Covariance Structure of Earnings and the On-the-Job Training Hypothesis," *Annals of Economic and Social Measurement*, 6, 335-365.

Hause, J. (1980), "The Fine Structure of Earnings and the On-the-Job Training Hypothesis," *Econometrica*, 48, 1013-1029.

Hong, Y., and Lee, J. (2000), "Wavelet-Based Estimation of Heterokedasticity and Autocorrelation Consistent Covariance Matrices," Manuscript.

Im, K. S., Ahn, S. C., Schmidt, P., and Wooldridge, J. M. (1999), "Efficient Estimation of Panel Models with Strictly Exogenous Explanatory Variables," *Journal of Econometrics*, 93, 177-201.

Kiefer, N. M. (1980), "Estimation of Fixed Effect Models for Time Series of Cross-Sections with Arbitrary Intertemporal Covariance," *Journal of Econometrics*, 14, 195-202.

Kao, C. (1999), "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data," *Journal of Econometrics*, 90, 1-44.

Kao, C., and Chiang, M.-H. (2000), "On the Estimation and Inference of a Cointegrated Regression in Panel Data." *Advances in Econometrics*, Volume 15, forthcoming.

Karlsson, S., and Skoglund, J. (2000), "Maximum-Likelihood Based Inference in the Two-way Random Effects Model with Serially Correlated Time Effects," No 383, SSE/EFI Working Paper Series in Economics and Finance, Stockholm School of Economics.

Keane, M. P., and Runkle, D. E. (1992), "On the Estimation of Panel-Data Models with Serial Correlation when Instruments are not Strictly Exogenous," *Journal of Business and Economics Statistics*, 10, 1- 9.

Keener, R. W., J. Kmenta and N.C. Weber (1991), "Estimation of the Covariance Matrix of the Least Squares Regression Coefficients When the Disturbance Covariance Matrix is of Unknown Form", *Econometric Theory* 7, 22-45.

Kezdi, G. (2001), "Robust Standard Error Estimation in Fixed-Effects Panel Models," Manuscript, University of Michigan.

King, R., C. Plosser, J. Stock and M. Watson (1991), "Stochastic Trends and Economic Fluctuations," *American Economic Review* 81, 819-840.

Kool, H. (1988), "A Note on Consistent Estimation of Heteroskedastic and Autocorrelated Covariance Matrices", Research Memorandum 21, Department of Econometrics, Free University, Amsterdam..

Levine, D. (1983), "A Remark on Serial Correlation in Maximum Likelihood", *Journal of Econometrics* 23, 337-342.

Levitt, S. D. (1998), "Juvenile Crime and Punishment," *Journal of Political Economy*, 106, 1156 - 1185.

Lillard, L. A., and Willis, R. J. (1978), "Dynamic Aspects of Earnings Mobility," *Econometrica*, 46, 985-1012.

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

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

Park, B., and Sickles, R. C. "Semiparametric Efficient Estimation of AR(1) Panel Data Models," Manuscript, Rice University.

Parzen, E. (1957), "On Consistent Estimates of the Spectrum of a Stationary Time Series",

Annals of Mathematical Statistics 28, 329-348.
Phillips, P. and B. Hansen (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes," *Review of Economic Studies* 57, 99-125.
Phillips, P. and S. Ouliaris (1988), "Testing for Cointegration Using Principal Components methods", *Journal of Economic Dynamics and Control* 12, 205-230.
Press, H. and T. J. Tukey (1956), "Power Spectral Methods of Analysis and Their Applications to Problems in Airplane Dynamics", Bell Systems Monograph No. 2606.
Priestley, M.B. (1981), Spectral Analysis and Time Series. Academic Press: New York.
Schwert, G. (1987), "Test for Unit Roots: A Monte Carlo Investigation", National Bureau of Economic Research Working Paper No. 73.
Skoglund, J. and S. Karlsson (2001a), "Asymptotic Properties of the Maximum Likelihood Estimator of Random Effects Models with Serial Correlation," No 432, SSE/EFI Working Paper Series in Economics and Finance, Stockholm School of Economics.
Skoglund, J. and S. Karlsson (2001a), "Specification and Estimation of Random Effects Models with Serial Correlation of general form," No 433, SSE/EFI Working Paper Series in Economics and Finance, Stockholm School of Economics.
White, H. (1984), Asymptotic Theory for Econometricians. Academic Press: New York.
White, H. and I. Domowitz (1984), "Nonlinear regression with Dependent Observations", *Econometrica* 52, 143-161.
Xiao, Z. and O. Linton (2000), "A Nonparametric Prewhitened Covariance Estimator," Forthcoming, *Journal of Time Series Analysis*.

Table 1: Bias, Variance, MSE, and Size of t test under ARMA(1,1):(n,T)=(25,32)

τ	Estimator	Bias	MSE	Variance	10%	5%	1%
	NW	-4.6589	21.7081	0.0022	0.2960	0.1760	0.0540
	QS(A)	-4.6350	21.4864	0.0028	0.2820	0.1560	0.0440
	QS(B)	-4.6334	21.4712	0.0028	0.2780	0.1510	0.0430
	VARHAC(A)	-4.3683	20.4186	1.3381	0.2200	0.1150	0.0310
0.0	VARVAR(B)	-4.5212	20.5881	0.1465	0.2250	0.1280	0.0340
	PW-NW(A)	-4.7531	22.6346	0.0427	0.4450	0.3350	0.1470
	PW-NW(B)	-4.7615	22.6949	0.0224	0.4550	0.3420	0.1530
	PW-QS(A)	-4.7386	22.6031	0.1494	0.4310	0.3190	0.1460
	PW-QS(B)	-4.7472	22.5467	0.0111	0.4310	0.3100	0.1370
	NW	-2.7982	7.8308	0.0007	0.3100	0.1860	0.0630
	QS(A)	-2.7816	7.7380	0.0009	0.2810	0.1770	0.0480
	QS(B)	-2.7808	7.7339	0.0009	0.2810	0.1770	0.0430
	VARHAC(A)	-2.4941	7.7783	1.5593	0.2240	0.1330	0.0350
0.4	VARHAC(B)	-2.7090	7.3973	0.0584	0.2300	0.1350	0.0430
	PW-NW(A)	-2.8542	8.1565	0.0096	0.4470	0.3260	0.1660
	PW-NW(B)	-2.8605	8.1859	0.0031	0.4440	0.3290	0.1780
	PW-QS(A)	-2.8432	8.1218	0.0382	0.4250	0.2970	0.1470
	PW-QS(B)	-2.8460	8.1062	0.0063	0.4180	0.2980	0.1400
	NW	-0.9324	0.8695	0.0000	0.3110	0.1920	0.0610
	QS(A)	-0.9271	0.8597	0.0001	0.2490	0.1490	0.0430
	QS(B)	-0.9270	0.8594	0.0001	0.2460	0.1460	0.0430
	VARHAC(A)	-0.8331	0.9842	0.2904	0.2200	0.1310	0.0430
0.8	VARHAC(B)	-0.9001	0.8174	0.0073	0.2340	0.1310	0.0360
	PW-NW(A)	-0.9528	0.9082	0.0004	0.4370	0.3340	0.1660
	PW-NW(B)	-0.9542	0.9112	0.0001	0.4200	0.3070	0.1570
	PW-QS(A)	-0.9475	0.9005	0.0027	0.3750	0.2670	0.1250
	PW-QS(B)	-0.9542	0.9112	0.0001	0.4200	0.3070	0.1570

Note:

The number of simulations is 1000.

The sums of coefficients of AR and MA are 0.9, respecitvely

Table 2: Bias, Variance, MSE, and Size of t test under ARMA(1,1):(n,T)=(50,64)

τ	Estimator	Bias	MSE	Variance	10%	5%	1%
	NW	-4.8123	23.1584	0.0001	0.2770	0.1680	0.0550
	QS(A)	-4.8034	23.0725	0.0001	0.2700	0.1600	0.0530
	QS(B)	-4.8039	23.0780	0.0001	0.2690	0.1620	0.0500
	VARHAC(A)	-4.7777	22.8507	0.0243	0.2270	0.1280	0.0360
0.0	VARHAC(B)	-4.7835	22.8828	0.0013	0.2160	0.1290	0.0260
	PW-NW(A)	-4.8532	23.5669	0.0134	0.4450	0.3350	0.1760
	PW-NW(B)	-4.8564	23.5969	0.0123	0.4830	0.3680	0.1950
	PW-QS(A)	-4.8581	23.6012	0.0000	0.4720	0.3510	0.1880
	PW-QS(B)	-4.8567	23.5878	0.0000	0.4660	0.3430	0.1770
	NW	-2.8877	8.3389	0.0000	0.2790	0.1580	0.0500
	QS(A)	-2.8823	8.3078	0.0001	0.2610	0.1490	0.0340
	QS(B)	-2.8825	8.3089	0.0000	0.2600	0.1500	0.0350
	VARHAC(A)	-2.8482	8.2487	0.1367	0.2180	0.1210	0.0180
0.4	VARHAC(B)	-2.8674	8.2299	0.0078	0.2220	0.1060	0.0230
	PW-NW(A)	-2.9148	8.4961	0.0001	0.4560	0.3390	0.1670
	PW-NW(B)	-2.9161	8.5034	0.0000	0.4860	0.3510	0.1890
	PW-QS(A)	-2.9146	8.4947	0.0000	0.4750	0.3440	0.1760
	PW-QS(B)	-2.9137	8.4898	0.0000	0.4680	0.3360	0.1660
	NW	-0.9627	0.9269	0.0000	0.2900	0.1680	0.0470
	QS(A)	-0.9608	0.9232	0.0000	0.2490	0.1530	0.0410
	QS(B)	-0.9609	0.9233	0.0000	0.2500	0.1570	0.0410
	VARHAC(A)	-0.9444	0.9188	0.0270	0.2160	0.1240	0.0350
0.8	VARHAC(B)	-0.9558	0.9142	0.0006	0.2130	0.1060	0.0260
	PW-NW(A)	-0.9718	0.9444	0.0000	0.4910	0.3520	0.1780
	PW-NW(B)	-0.9721	0.9450	0.0000	0.4510	0.3290	0.1860
	PW-QS(A)	-0.9665	0.9565	0.0223	0.4220	0.3130	0.1700
	PW-QS(B)	-0.9710	0.9429	0.0000	0.4150	0.3040	0.1690

Note:

The number of simulations is 1000.

The sums of coefficients of AR and MA are 0.9, respecitvely

Table 3: Bias, Variance, MSE, and Size of t test under ARMA(1,1):(n,T)=(100,128)

τ	Estimator	Bias	MSE	Variance	10%	5%	1%
	NW	-4.8733	23.7492	0.0000	0.2490	0.1310	0.0420
	QS(A)	-4.8716	23.7321	0.0000	0.2300	0.1270	0.0410
	QS(B)	-4.8718	23.7340	0.0000	0.2340	0.1290	0.0400
	VARHAC(A)	-4.8669	23.6876	0.0007	0.2030	0.1110	0.0310
0.0	VARHAC(B)	-4.8668	23.6854	0.0000	0.2020	0.1030	0.0230
	PW-NW(A)	-4.8883	23.8951	0.0000	0.4230	0.3250	0.1540
	PW-NW(B)	-4.8893	23.9048	0.0000	0.4600	0.3430	0.1760
	PW-QS(A)	-4.8880	23.8924	0.0000	0.4380	0.3170	0.1580
	PW-QS(B)	-4.8877	23.8898	0.0000	0.4330	0.3070	0.1550
	NW	-2.9241	8.5504	0.0000	0.2370	0.1440	0.0440
	QS(A)	-2.9229	8.5434	0.0000	0.2260	0.1280	0.0350
	QS(B)	-2.9230	8.5440	0.0000	0.2270	0.1270	0.0360
	VARHAC(A)	-2.9206	8.5299	0.0000	0.1940	0.1050	0.0280
0.4	VARHAC(B)	-2.9200	8.5265	0.0000	0.1970	0.1020	0.0240
	PW-NW(A)	-2.9330	8.6025	0.0000	0.4430	0.3260	0.1650
	PW-NW(B)	-2.9334	8.6051	0.0000	0.4500	0.3340	0.1700
	PW-QS(A)	-2.9328	8.6011	0.0000	0.4200	0.3080	0.1520
	PW-QS(B)	-2.9326	8.6001	0.0000	0.4160	0.3040	0.1480
	NW	-0.9748	0.9502	0.0000	0.2770	0.1470	0.0380
	QS(A)	-0.9743	0.9493	0.0000	0.2100	0.1160	0.0230
	QS(B)	-0.9743	0.9493	0.0000	0.2110	0.1160	0.0240
	VARHAC(A)	-0.9735	0.9478	0.0000	0.1850	0.0970	0.0110
0.8	VARHAC(B)	-0.9734	0.9474	0.0000	0.2180	0.1010	0.0230
	PW-NW(A)	-0.9777	0.9559	0.0000	0.4750	0.3530	0.1770
	PW-NW(B)	-0.9778	0.9562	0.0000	0.4470	0.3180	0.1570
	PW-QS(A)	-0.9776	0.9556	0.0000	0.4180	0.2820	0.1380
	PW-QS(B)	-0.9775	0.9555	0.0000	0.4080	0.2780	0.1330

Note:

The number of simulations is 1000.

The sums of coefficients of AR and MA are 0.9, respecitvely

Table 4: Bias, Variance, MSE, and Size of t test under ARMA(2,2):(n,T)=(25,32)

τ	Estimator	Bias	MSE	Variance	10%	5%	1%
	NW	-4.8273	23.3048	0.0015	0.3190	0.1760	0.0550
	QS(A)	-4.8086	23.1250	0.0019	0.2950	0.1570	0.0410
	QS(B)	-4.8073	23.1118	0.0020	0.2910	0.1560	0.0410
	VARHAC(A)	-4.5029	21.7122	1.4374	0.2240	0.1220	0.0360
0.0	VARHAC(B)	-4.6820	22.1005	0.1796	0.2460	0.1280	0.0430
	PW-NW(A)	-4.9011	24.0499	0.0290	0.4640	0.3400	0.1520
	PW-NW(B)	-4.9089	24.1140	0.0166	0.4610	0.3430	0.1560
	PW-QS(A)	-4.8947	24.0050	0.0471	0.4580	0.3290	0.1410
	PW-QS(B)	-4.8950	23.9913	0.0329	0.4510	0.3210	0.1420
	NW	-2.8987	8.4030	0.0005	0.3060	0.1860	0.0580
	QS(A)	-2.8858	8.3287	0.0006	0.2770	0.1780	0.0470
	QS(B)	-2.8854	8.3261	0.0007	0.2770	0.1780	0.0460
	VARHAC(A)	-2.6642	7.9949	0.8978	0.2290	0.1330	0.0290
0.4	VARHAC(B)	-2.8113	7.9795	0.0765	0.2460	0.1390	0.0390
	PW-NW(A)	-2.9462	8.6665	0.0041	0.4370	0.3310	0.1620
	PW-NW(B)	-2.9440	8.6791	0.0120	0.4300	0.3180	0.1490
	PW-QS(A)	-2.9381	8.6427	0.0105	0.4150	0.3110	0.1640
	PW-QS(B)	-2.9362	8.6354	0.0139	0.4120	0.3050	0.1620
	NW	-0.9658	0.9329	0.0000	0.3020	0.1870	0.0540
	QS(A)	-0.9619	0.9254	0.0001	0.2490	0.1420	0.0500
	QS(B)	-0.9619	0.92532	0.0001	0.2480	0.1440	0.0500
	VARHAC(A)	-0.8072	1.4377	0.7868	0.2010	0.1160	0.0400
0.8	VARHAC(B)	-0.9371	0.88521	0.0072	0.2370	0.1290	0.0350
	PW-NW(A)	-0.9809	0.9634	0.0014	0.4310	0.3170	0.1710
	PW-NW(B)	-0.9831	0.9666	0.0001	0.4240	0.3020	0.1490
	PW-QS(A)	-0.9796	0.9600	0.0004	0.3850	0.2620	0.1300
	PW-QS(B)	-0.9793	0.9595	0.0004	0.3800	0.2620	0.1250

Note:

The number of simulations is 1000.

The sums of coefficients of AR and MA are 0.9, respecitvely

Table 5: Bias, Variance, MSE, and Size of t test under ARMA(2,2):(n,T)=(50,64)

τ	Estimator	Bias	MSE	Variance	10%	5%	1%
	NW	-4.9428	24.4319	0.0001	0.2860	0.1710	0.0590
	QS(A)	-4.9345	24.3495	0.0001	0.2740	0.1720	0.0450
	QS(B)	-4.9355	24.3596	0.0001	0.2750	0.1710	0.0460
	VARHAC(A)	-4.9025	24.1024	0.0682	0.2400	0.1380	0.0280
0.0	VARHAC(B)	-4.9135	24.1559	0.0130	0.2250	0.1290	0.0290
	PW-NW(A)	-4.9796	24.8149	0.0181	0.4410	0.3420	0.1810
	PW-NW(B)	-4.9835	24.8361	0.0012	0.4920	0.3690	0.2050
	PW-QS(A)	-4.9829	24.8298	0.0003	0.4800	0.3650	0.1990
	PW-QS(B)	-4.9825	24.8253	0.0002	0.4820	0.3580	0.1920
	NW	-2.9659	8.7970	0.0000	0.2880	0.1630	0.0500
	QS(A)	-2.9608	8.7665	0.0000	0.2600	0.1540	0.0450
	QS(B)	-2.9613	8.7693	0.0000	0.2620	0.1570	0.0420
	VARHAC(A)	-2.9352	8.6727	0.0574	0.2200	0.1290	0.0250
0.4	VARHAC(B)	-2.9503	8.7054	0.0011	0.2220	0.1220	0.0240
	PW-NW(A)	-2.9885	8.9345	0.0037	0.4640	0.3520	0.1790
	PW-NW(B)	-2.9883	8.9358	0.0055	0.4910	0.3620	0.1930
	PW-QS(A)	-2.9845	8.9376	0.0304	0.4820	0.3510	0.1860
	PW-QS(B)	-2.9839	8.9342	0.0304	0.4760	0.3430	0.1790
	NW	-0.9887	0.9776	0.0000	0.2900	0.1760	0.0480
	QS(A)	-0.9871	0.9744	0.0000	0.2580	0.1460	0.0470
	QS(B)	-0.9872	0.9745	0.0000	0.2570	0.1460	0.0470
	VARHAC(A)	-0.9775	0.9646	0.0091	0.2220	0.1220	0.0360
0.8	VARHAC(B)	-0.9834	0.9672	0.0001	0.2280	0.1110	0.0240
	PW-NW(A)	-0.9966	0.9932	0.0000	0.4830	0.3630	0.1930
	PW-NW(B)	-0.9970	0.9939	0.0000	0.4620	0.3380	0.1810
	PW-QS(A)	-0.9966	0.9931	0.0000	0.4480	0.3310	0.1700
	PW-QS(B)	-0.9964	0.9928	0.0000	0.4380	0.3280	0.1610

Note:

The number of simulations is 1000.

The sums of coefficients of AR and MA are 0.9, respecitvely

Table 6: Bias, Variance, MSE, and Size of t test under ARMA(2,2):(n,T)=(100,128)

τ	Estimator	Bias	MSE	Variance	10%	5%	1%
	NW	-4.9947	24.9473	0.0000	0.2420	0.1380	0.0430
	QS(A)	-4.9929	24.9286	0.0000	0.2250	0.1410	0.0410
	QS(B)	-4.9931	24.9315	0.0000	0.2300	0.1420	0.0440
	VARHAC(A)	-4.9893	24.8931	0.0000	0.2070	0.1190	0.0280
0.0	VARHAC(B)	-4.9888	24.8885	0.0000	0.1910	0.0990	0.0260
	PW-NW(A)	-5.0088	25.0879	0.0000	0.4290	0.3140	0.1580
	PW-NW(B)	-5.0095	25.0951	0.0000	0.4690	0.3470	0.1830
	PW-QS(A)	-5.0086	25.0861	0.0000	0.4460	0.3270	0.1690
	PW-QS(B)	-5.0085	25.0852	0.0000	0.4460	0.3220	0.1670
	NW	-2.9969	8.9814	0.0000	0.2480	0.1420	0.0150
	QS(A)	-2.9957	8.9743	0.0000	0.2290	0.1260	0.0320
	QS(B)	-2.9959	8.9752	0.0000	0.2280	0.1270	0.0340
	VARHAC(A)	-2.9932	8.9593	0.0000	0.1970	0.1010	0.0290
0.4	VARHAC(B)	-2.9931	8.9589	0.0000	0.1940	0.1080	0.0250
	PW-NW(A)	-3.0047	9.0285	0.0004	0.4460	0.3190	0.1630
	PW-NW(B)	-3.0056	9.0338	0.0000	0.4480	0.3230	0.1710
	PW-QS(A)	-3.0051	9.0309	0.0000	0.4290	0.3110	0.1470
	PW-QS(B)	-3.0051	9.0305	0.0000	0.4270	0.3120	0.1460
	NW	-0.9990	0.9981	0.0000	0.2700	0.1630	0.0430
	QS(A)	-0.9986	0.9972	0.0000	0.2350	0.1190	0.0220
	QS(B)	-0.9986	0.9973	0.0000	0.2340	0.1230	0.0220
	VARHAC(A)	-0.9977	0.9954	0.0000	0.1970	0.0960	0.0130
0.8	VARHAC(B)	-0.9977	0.9954	0.0000	0.2110	0.0970	0.0230
	PW-NW(A)	-1.0018	1.0036	0.0000	0.4720	0.3460	0.1790
	PW-NW(B)	-1.0019	1.0038	0.0000	0.4500	0.3220	0.1720
	PW-QS(A)	-1.0017	1.0034	0.0000	0.4220	0.3080	0.1440
	PW-QS(B)	-1.0017	1.0034	0.0000	0.4200	0.3020	0.1420

Note:

The number of simulations is 1000.

The sums of coefficients of AR and MA are 0.9, respecitvely

Table 7: Bias, Variance, MSE, and Size of t test under ARMA(4,4):(n,T)=(25,32)

τ	Estimator	Bias	MSE	Variance	10%	5%	1%
	NW	-4.8564	23.5859	0.0008	0.3230	0.1840	0.0540
	QS(A)	-4.8842	23.4678	0.0010	0.2990	0.1600	0.0430
	QS(B)	-4.8435	23.4609	0.0010	0.2970	0.1590	0.0380
	VARHACA)	-4.5626	22.5929	1.7770	0.2410	0.1400	0.0430
0.0	VARHAC(B)	-4.7836	22.9872	0.1042	0.2740	0.1750	0.0510
	PW-NW(A)	-4.9042	24.0722	0.0207	0.4520	0.3190	0.1310
	PW-NW(B)	-4.9060	24.0892	0.0203	0.4520	0.3260	0.1390
	PW-QS(A)	-4.8964	24.0161	0.0413	0.4380	0.3040	0.1270
	PW-QS(B)	-4.8981	24.0034	0.0119	0.4360	0.3080	0.1250
	NW	-2.9156	8.5011	0.0003	0.2970	0.1720	0.0520
	QS(A)	-2.9075	8.4537	0.0003	0.2920	0.1620	0.0470
	QS(B)	-2.9072	8.4522	0.0003	0.2900	0.1580	0.0490
	VARHAC(A)	-2.7460	8.3621	0.8223	0.2520	0.1470	0.0460
0.4	VARHAC(B)	-2.8745	8.2976	0.0350	0.2770	0.1670	0.0510
	PW-NW(A)	-2.9412	8.6566	0.0058	0.4210	0.2970	0.1300
	PW-NW(B)	-2.9424	8.6657	0.0082	0.4260	0.3060	0.1470
	PW-QS(A)	-2.9246	8.6500	0.0000	0.4100	0.2970	0.1370
	PW-QS(B)	-2.9399	8.6490	0.0059	0.4140	0.2930	0.1320
	NW	-0.9714	0.9437	0.0000	0.2960	0.1800	0.0510
	QS(A)	-0.9691	0.9393	0.0000	0.2410	0.1390	0.0430
	QS(B)	-0.9691	0.9392	0.0000	0.2400	0.1390	0.0440
	VAR(HACA)	-0.8691	1.1746	0.4196	0.2100	0.1250	0.0480
0.8	VARHAC(B)	-0.9562	0.9184	0.0040	0.2560	0.1540	0.0460
	PW-NW(A)	-0.9803	0.9622	0.0012	0.4250	0.3010	0.1430
	PW-NW(B)	-0.9826	0.9656	0.0000	0.3940	0.2800	0.0340
	PW-QS(A)	-0.9804	0.9614	0.0002	0.3680	0.2490	0.1180
	PW-QS(B)	-0.9808	0.9621	0.0001	0.3670	0.2570	0.1210

Note:

The number of simulations is 1000.

The sums of coefficients of AR and MA are 0.9, respecitvely

Table 8: Bias, Variance, MSE, and Size of t test under ARMA(4,4):(n,T)=(50,64)

τ	Estimator	Bias	MSE	Variance	10%	5%	1%
	NW	-4.9334	24.3389	0.0001	0.2920	0.1620	0.0580
	QS(A)	-4.9271	24.2763	0.0001	0.2690	0.1690	0.0450
	QS(B)	-4.9286	24.2908	0.0001	0.2750	0.1740	0.0490
	VARHAC(A)	-4.8925	24.5929	0.0781	0.2450	0.1440	0.0270
0.0	VARHAC(B)	-4.9167	24.1781	0.0039	0.2430	0.1290	0.0410
	PW-NW(A)	-4.9602	24.6154	0.0120	0.4500	0.3400	0.1710
	PW-NW(B)	-4.9641	24.6424	0.0000	0.4460	0.3400	0.1910
	PW-QS(A)	-4.9639	24.6407	0.0000	0.4520	0.3450	0.1950
	PW-QS(B)	-4.9636	24.6375	0.0000	0.4440	0.3410	0.1870
	NW	-2.9602	8.7627	0.0000	0.2840	0.1780	0.0470
	QS(A)	-2.9565	8.7407	0.0000	0.2710	0.1580	0.0410
	QS(B)	-2.9571	8.7445	0.0000	0.2770	0.1620	0.0460
	VARHAC(A)	-2.9247	8.7036	0.1499	0.2390	0.1360	0.0310
0.4	VARHAC(B)	-2.9510	8.7089	0.0007	0.2430	0.1470	0.0330
	PW-NW(A)	-2.9766	8.8616	0.0016	0.4500	0.3260	0.1860
	PW-NW(B)	-2.9783	8.8702	0.0001	0.4840	0.3630	0.1840
	PW-QS(A)	-2.9780	8.8688	0.0001	0.4730	0.3580	0.1850
	PW-QS(B)	-2.9781	8.8690	0.0000	0.4710	0.3560	0.1780
	NW	-0.9868	0.9737	0.0000	0.2940	0.1750	0.0440
	QS(A)	-0.9857	0.9716	0.0000	0.2580	0.1560	0.0480
	QS(B)	-0.9858	0.9718	0.0000	0.2630	0.1550	0.0510
	VARHAC(A)	-0.9656	1.0039	0.0715	0.2260	0.1310	0.0400
0.8	VARHAC(B)	-0.9835	0.9675	0.0001	0.2450	0.1210	0.0290
	PW-NW(A)	-0.9928	0.9856	0.0000	0.4820	0.3340	0.1740
	PW-NW(B)	-0.9928	0.9857	0.0000	0.4560	0.3370	0.1730
	PW-QS(A)	-0.9926	0.9853	0.0000	0.4540	0.3310	0.1690
	PW-QS(B)	-0.9925	0.9851	0.0000	0.4490	0.3290	0.1650

Note:

The number of simulations is 1000.

The sums of coefficients of AR and MA are 0.9, respecitvely

Table 9: Bias, Variance, MSE, and Size of t test under ARMA(4,4):(n,T)=(100,128)

τ	Estimator	Bias	MSE	Variance	10%	5%	1%
	NW	-4.9720	24.7206	0.0000	0.2460	0.1350	0.0430
	QS(A)	-4.9703	24.7035	0.0000	0.2490	0.1510	0.0340
	QS(B)	-4.9706	24.7073	0.0000	0.2520	0.1530	0.0350
	VARHAC(A)	-4.9643	24.6519	0.0075	0.2270	0.1290	0.0260
0.0	VARHAC(B)	-4.9678	24.6786	0.0000	0.2070	0.1090	0.0270
	PW-NW(A)	-4.9837	24.8371	0.0000	0.4180	0.3040	0.1480
	PW-NW(B)	-4.9840	24.8460	0.0000	0.4540	0.3360	0.1930
	PW-QS(A)	-4.9836	24.8367	0.0000	0.4480	0.3290	0.1760
	PW-QS(B)	-4.9837	24.8371	0.0000	0.4520	0.3280	0.1770
	NW	-2.9832	8.8995	0.0000	0.2530	0.1520	0.0450
	QS(A)	-2.9822	8.8935	0.0000	0.2390	0.1220	0.0340
	QS(B)	-2.9824	8.8948	0.0000	0.2410	0.1220	0.0360
	VARHAC(A)	-2.9803	8.8822	0.0000	0.2080	0.1080	0.0240
0.4	VARHAC(B)	-2.9803	8.8824	0.0000	0.2170	0.1180	0.0270
	PW-NW(A)	-2.9902	8.9413	0.0000	0.4330	0.3110	0.1630
	PW-NW(B)	-2.9904	8.9425	0.0000	0.4390	0.3150	0.1560
	PW-QS(A)	-2.9902	8.9412	0.0000	0.4290	0.3090	0.1500
	PW-QS(B)	-2.9902	8.9414	0.0000	0.4290	0.3100	0.1530
	NW	-0.9944	0.9889	0.0000	0.2800	0.1500	0.0320
	QS(A)	-0.9941	0.9883	0.0000	0.2380	0.1270	0.0200
	QS(B)	-0.9942	0.9884	0.0000	0.2400	0.1270	0.0200
	VARHAC(A)	-0.9934	0.9868	0.0000	0.2010	0.1010	0.0120
0.8	VARHAC(B)	-0.9935	0.9871	0.0000	0.2250	0.1140	0.0220
	PW-NW(A)	-0.9968	0.9935	0.0000	0.4510	0.3450	0.1650
	PW-NW(B)	-0.9968	0.9936	0.0000	0.4450	0.3260	0.1460
	PW-QS(A)	-0.9967	0.9935	0.0000	4340	0.3200	0.1400
	PW-QS(B)	-0.9967	0.9935	0.0000	0.4330	0.3220	0.1410

Note:

The number of simulations is 1000.

The sums of coefficients of AR and MA are 0.9, respecitvely