Bootstrapping GMM Estimators for Time Series*

Atsushi Inoue† — North Carolina State University
Mototsugu Shintani‡ — Vanderbilt University

First Draft: October 2000

This Version: February 2001

Abstract

This paper establishes that the bootstrap provides asymptotic refinements for the generalized method of moments estimator of overidentified linear models when autocovariance structures of moment functions are unknown. Because the heteroskedasticity and autocorrelation consistent covariance matrix estimator cannot be written as a function of sample moments and converges at a rate slower than $T^{-1/2}$, the asymptotic refinement cannot be proved in the conventional way. As a result, we find that the bootstrap approximation error for the distribution of the t test and the test of overidentifying restrictions is of larger order than typically found in the literature. We also find that the choice of kernels plays a more important role in our second-order asymptotic theory than in the conventional first-order asymptotic theory. Nevertheless, the bootstrap approximation improves upon the first-order asymptotic approximation. A Monte Carlo experiment shows that the bootstrap improves the accuracy of inference on regression parameters in small samples. We apply our bootstrap method to inference about the parameters in the monetary policy reaction function.

KEYWORDS: asymptotic refinements, block bootstrap, HAC covariance matrix estimator, dependent data, Edgeworth expansions, instrumental variables, J test.

---

*We thank Jordi Galí for providing us with the data and program used in Clarida, Galí and Gertler (2000). We also thank Alastair Hall, Lutz Kilian and seminar participants at Brown University, University of Michigan and the 2000 Triangle Econometrics Conference for helpful comments.

†Department of Agricultural and Resource Economics, North Carolina State University, Box 8109, Raleigh, NC 27695-8109. E-mail: atsushi@unity.ncsu.edu.

‡Department of Economics, Vanderbilt University, Box 1819 Station B, Nashville, TN 37235. E-mail: mototsugu.shintani@vanderbilt.edu.

1. Introduction

In this paper we establish that the bootstrap provides asymptotic refinements for the generalized method of moments (GMM) estimator of possibly overidentified linear models. Our analysis differs from earlier work in that we allow for general autocovariance structures of moment functions. In typical empirical situations, the autocovariance structure of moment functions is unknown and the inverse of the heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator is used as a weighting matrix in GMM estimation. It is well known, however, that coverage probabilities based on the HAC covariance estimator are often too low, and that the t test tends to reject too frequently (see Andrews, 1991). In this paper, we propose a bootstrap method for the GMM estimator to improve the finite sample performance of the t test and the test of overidentifying restrictions (J test). We use the block bootstrap originally proposed by Künsch (1989) for weakly dependent data (see also Carlstein, 1986). When the block length increases at a suitable rate with the sample size, such block bootstrap procedures eventually will capture the unknown structure of dependence.

Our linear framework is of particular interest in applied time series analysis. GMM estimation of linear models has been applied to the expectation hypothesis of the term structure (Campbell and Shiller, 1991), 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). Since the GMM estimates often have policy implications in structural econometric models, it is important for researchers to obtain accurate confidence intervals. For example, the interpretation of the policy rule crucially depends on the value of the estimated parameters (see Clarida, Galí and Gertler, 2000).

Not surprisingly, given the poor performance of the conventional asymptotic approximation, the econometric literature on the bootstrap for GMM is growing rapidly. Hahn (1996) shows the first-order validity of the bootstrap for GMM with iid observations.[1] For dependent data, Hall and Horowitz (1996) show that the block bootstrap provides asymptotic refinements for GMM. However, Hall and Horowitz (1996) assume that the autocovariances of the moment function are zero after finite lags, and thus their framework does not cover the use of the HAC covariance matrix estimator for the general dependence structure. Economic theory often provides information about the specification of moment conditions, but not necessarily about the dependence structure of the moment conditions. Therefore, it is important for applied work to be able to allow for more general forms of autocorrelation. This extension is not straightforward because the HAC covariance matrix estimator cannot be written as a function of sample moments and converges at a rate slower than $T^{-1/2}$. Thus, the conventional arguments cannot be applied directly to prove the existence of Edgeworth expansions and to establish asymptotic refinements of the bootstrap.

Recently, Götze and Künsch (1996) and Lahiri (1996) show that the block bootstrap can provide asymptotic refinements for a smooth function of sample means and for parameters in a linear regression model, respectively, even when the HAC covariance estimator is used. They show that the bootstrap provides asymptotic refinements for approximating the distribution of the estimator and for the coverage probability of one-sided confidence intervals. However, they do not show asymptotic refinements for the two-sided symmetric t test nor do they provide any result for the overidentified case which is of great interest in empirical work. The purpose of this paper is to prove that

[1] Brown and Newey (1995) propose an alternative efficient bootstrap method based on the empirical likelihood.

the bootstrap provides asymptotic refinements for these statistics in overidentified linear models estimated by GMM. To our knowledge, the higher-order properties of the block bootstrap for GMM with unknown autocovariance structures have not been formally investigated.

Our results are nonstandard for two reasons. First, we show that the order of the bootstrap approximation error is larger than typically found in the literature on the bootstrap for parametric estimators. The intuition behind this result is as follows: The HAC covariance matrix estimator is (proportional to) a nonparametric estimator of the spectral density at frequency zero, and its convergence rate is slower than $T^{-1/2}$. For the first-order asymptotic theory, all that matters is the consistency of the HAC covariance matrix estimator. However, the nonparametric nature of the HAC covariance matrix estimator becomes important in the higher-order asymptotic theory and complicates the analysis of the two-sided symmetric t test and the J test statistic. Nevertheless, we are able to establish that the bootstrap approximation error is smaller than the conventional normal approximation error.

Second, we note that the choice of kernels plays a more important role in our second-order asymptotic theory than in the conventional first-order asymptotic theory because the order of the bootstrap approximation error depends on the bias of the HAC covariance estimator. For the bootstrap to provide asymptotic refinements, the bias must vanish sufficiently fast. For the one-sided t test, most of the commonly used kernels satisfy this condition. For two-sided symmetric t test and for the J test statistic, however, one must use kernels, such as the truncated kernel (White, 1984) and the trapezoidal kernel (Politis and Romano, 1995), whose bias vanishes even faster. The resulting HAC covariance matrix estimator based on these kernels, however, is not necessarily positive semidefinite. In this paper, we propose a modified HAC covariance matrix estimator that is always positive semidefinite.

In a Monte Carlo experiment, we find that our bootstrap method improves the accuracy of inference in small samples, especially for the two-sided symmetric t test. To illustrate the usefulness of the bootstrap approach, we apply our bootstrap procedure to the monetary policy reaction function of Clarida, Galí and Gertler (2000). We find that the data do not necessarily support some of their conclusions.

The rest of the paper is organized as follows. Section 2 introduces the model and describes the proposed bootstrap procedure. Section 3 presents the assumptions and theoretical results. Section 4 provides some Monte Carlo results. Section 5 presents an empirical illustration. Section 6 concludes the paper. All proofs are relegated to an appendix.

2. Model and Bootstrap Procedure

Consider a stationary time series $(x_t', y_t, z_t')'$ which satisfies

$$E[z_t u_t] \ = \ 0, \tag{2.1}$$

where $u_t = y_t - \beta_0' x_t$, β_0 is a p-dimensional parameter, x_t is a p-dimensional vector, z_t is a k-dimensional vector and $p < k$. Given a realization $\{(x_t', y_t, z_t')'\}_{t=1}^{T_0}$, we are interested in two-step GMM estimation of β_0 based on the moment condition (2.1). Let ℓ denote the lag truncation parameter used in HAC covariance matrix estimation and

$T = T_0 - \ell + 1$.[2] We first obtain the first-step GMM estimator $\tilde{\beta}_T$ by minimizing

$$\left[\frac{1}{T_0}\sum_{t=1}^{T_0} z_t(y_t - \beta' x_t)\right]' V_T \left[\frac{1}{T_0}\sum_{t=1}^{T_0} z_t(y_t - \beta' x_t)\right]$$

with respect to β, where V_T is some $k \times k$ positive semidefinite matrix. Then we obtain the second-step GMM estimator $\hat{\beta}_T$ by minimizing

$$\left[\frac{1}{T}\sum_{t=1}^{T} z_t(y_t - \beta' x_t)\right]' \hat{S}_T^{-1} \left[\frac{1}{T}\sum_{t=1}^{T} z_t(y_t - \beta' x_t)\right],$$

where

$$\begin{aligned} \hat{S}_T &= \frac{1}{T}\sum_{t=1}^{T}\left[z_t \tilde{u}_t^2 z_t' + \sum_{j=1}^{\ell}\omega\left(\frac{j}{\ell}\right)\left(z_{t+j}\tilde{u}_{t+j}\tilde{u}_t z_t' + z_t \tilde{u}_t \tilde{u}_{t+j} z_{t+j}'\right)\right] \\ \tilde{u}_t &= y_t - \tilde{\beta}_T' x_t. \end{aligned}$$

is the HAC covariance matrix estimator for the moment function (2.1), $\omega(\cdot)$ is a kernel. We are interested in the distribution of the studentized statistic $\hat{\Sigma}_T^{-1/2}(\hat{\beta}_T - \beta_0)$ where $\hat{\Sigma}_T = (\sum_{t=1}^{T} x_t z_t' \hat{S}_T^{-1} \sum_{t=1}^{T} z_t x_t')^{-1}$ and in the distribution of the J test statistic

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

We propose the following block bootstrap procedure. Suppose that $T = b\ell$ for some integer b.

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

$$(x_{(j-1)\ell+i}^{*\prime}, y_{(j-1)\ell+i}^{*}, z_{(j-1)\ell+i}^{*\prime})' = (x_{N_j+i}', y_{N_j+i}, z_{N_j+i}')',$$

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

[2] We use T observations and the modified HAC covariance matrix estimator $\hat{S}_T$ to obtain asymptotic refinements for the two-sided symmetric t test and the J test statistic. This modification is not necessary for obtaining asymptotic refinements for one-sided confidence intervals. See also Hall and Horowitz (1996, p.895).

Step 2. Calculate the first-step bootstrap GMM estimator $\tilde{\beta}_T^*$ by minimizing

$$\left[\frac{1}{T}\sum_{t=1}^{T} z_t^*(y_t^* - \beta' x_t^*) - \mu_T^*\right]' V_T \left[\frac{1}{T}\sum_{t=1}^{T} z_t^*(y_t^* - \beta' x_t^*) - \mu_T^*\right]$$

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 $\hat{\beta}_T^*$ by minimizing

$$\left[\frac{1}{T}\sum_{t=1}^{T} z_t^*(y_t^* - \beta' x_t^*) - \mu_T^*\right]' \hat{S}_T^{*-1} \left[\frac{1}{T}\sum_{t=1}^{T} z_t^*(y_t^* - \beta' x_t^*) - \mu_T^*\right],$$

where

$$\hat{S}_T^* = \frac{1}{T}\sum_{k=1}^{b}\sum_{i=1}^{\ell}\sum_{j=1}^{\ell}(z_{N_k+i}^*\tilde{u}_{N_k+i}^* - \mu_T^*)(z_{N_k+j}^*\tilde{u}_{N_k+j}^* - \mu_T^*)',$$

$$\tilde{u}_t^* = y_t^* - \tilde{\beta}_T^{*\prime} x_t^*.$$

Step 4. Obtain the bootstrap version of the studentized statistic $\hat{\Sigma}_T^{*-1/2}(\hat{\beta}_T^* - \hat{\beta}_T)$ where $\hat{\Sigma}_T^* = (\sum_{t=1}^{T} x_t^* z_t^{*\prime} \hat{S}_T^{*-1} \sum_{t=1}^{T} z_t^* x_t^{*\prime})^{-1}$ and the J test statistic

$$J_T^* = \left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[z_t^*(y_t^* - \hat{\beta}_T^{*\prime} x_t^*) - \mu_T^*]\right\}' \hat{S}_T^{*-1} \left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[z_t^*(y_t^* - \hat{\beta}_T^{*\prime} x_t^*) - \mu_T^*]\right\}.$$

By repeating Steps 1–4 sufficiently many times, one can approximate the finite-sample distributions of the studentized statistic and the J test statistic by the empirical distributions of their bootstrap version.

Remarks: 1. As in Hall and Horowitz (1996), we recenter the bootstrap version of the moment functions. Unlike the just identified case, the bootstrap version of the moment condition does not hold without recentering in the case of overidentified restrictions. The expression μ_T^* is the mean of the bootstrapped moment function with respect to the probability measure induced by the bootstrap algorithm.

2. Davison and Hall (1993) show that naïve applications of the block bootstrap do not provide asymptotic refinements for studentized statistics involving the long-run variance estimator. Specifically, they show that the error of the naïve bootstrap is of order $O(b^{-1}) + O(\ell^{-1})$ and thus is greater than or equal to the error of the first order asymptotic approximation. We therefore modify the bootstrap version of the HAC covariance matrix estimator (see Götze and Hipp, 1996, for the just-identified case). The expression $\hat{S}_T^*$ given in Step 3 is a consistent estimator for the variance of the bootstrapped moment function with the bootstrap probability measure.

3. Asymptotic Theory

In this section, we present our main theoretical results. Unless noted otherwise, we shall denote the Euclidean norm of a vector x by $\|x\|$. First, we provide the following set of assumptions.

Assumption 1:

(a) $\{(x_t', y_t, z_t')'\}$ is strictly stationary and strong mixing with mixing coefficients satisfying $\alpha_m \leq (1/d)\exp(-dm)$ for some $d > 0$.

(b) There is a unique $\beta_0 \in \Re^p$ such that $E[z_t u_t] \equiv E[z_t(y_t - \beta_0' x_t)] = 0$.

(c) Let $R_t = ((z_t u_t)', \mathrm{vec}(z_t x_t)')'$. Then $E\|R_t\|^{r+\eta} < \infty$ for some $r \geq 12$ and $\eta > 0$.

(d) Let $\mathcal{F}_a^b$ denote the sigma-algebra generated by $R_a, R_{a+1}, ..., R_b$. For all $m, s, t = 1, 2, ...$ and $A \in \mathcal{F}_{t-s}^{t+s}$,

$$E|P(A|\mathcal{F}_{-\infty}^{t-1} \cup \mathcal{F}_{t+1}^{\infty}) - P(A|\mathcal{F}_{t-s-m}^{t-1} \cup \mathcal{F}_{t+1}^{t+s+m})| \leq (1/d)\exp(-dm).$$

(e) For all $m, t = 1, 2, ...$ and $\theta \in \Re^{p+k+1}$ such that $1/d < m < t$ and $|\theta| \geq d$,

$$E\left|E\left\{\exp\left[i\theta' \sum_{s=t-m}^{t+m} (R_s - E(R_s))\right]\middle| \mathcal{F}_{-\infty}^{t-1} \cup \mathcal{F}_{t+1}^{\infty}\right\}\right| \leq \exp(-d).$$

(f) $\omega : \Re \to [-1, 1]$ satisfies (i) $\omega(0) = 1$, (ii) $\omega(x) = \omega(-x)$ $\forall x \in \Re$, (iii) $\omega(x) = 0$ $\forall |x| \geq 1$, (iv) $\omega(\cdot)$ is continuous at 0 and at all but a finite number of other points.

(g) $\ell \to \infty$ as $T \to \infty$ such that $\ell \neq O(T^{1/6})$ and $\ell = o(T^{1/4})$.

(h) $\hat{S}_T = \sum_{j=-T+1}^{T-1} \omega(j/\ell)\hat{\Gamma}_j$ is a positive semidefinite matrix that converges in probability to a positive definite matrix $S_0 \equiv \sum_{j=-\infty}^{\infty} E(z_0 u_0 u_j z_j')$.

(i) The first-step estimator $\tilde{\beta}_T$ satisfies $E|T^{1/2}(\tilde{\beta}_T - \beta_0)|^r = O(1)$, and V_T is a positive semidefinite matrix that converges to a positive definite matrix V at rate $O(\ell^{1/2}T^{-1/2})$.

Remarks: Assumption 1(c) requires that at least the 12th moment of the moment function be finite, and we will later require that at least the 36th moment be finite. Although this condition is strong, it is not atypical in the literature on higher-order asymptotic theory. For example, a sufficient (but not necessary) condition for Assumptions 3(f) and 4 of Hall and Horowitz (1996) is the finiteness of the 33rd moment of the moment functions and of their derivatives. Assumptions 1(d) and 1(e) are from Götze and Künsch (1996). Hall and Horowitz (1996, Assumptions 1 and 6) impose similar assumptions. Assumption 1(f) is a subset of Andrews' (1991) class of kernels $\mathcal{K}_1$. For example, the truncated kernel (White, 1984), Bartlett kernel (Newey and West, 1989) and Parzen kernel (Gallant, 1987) satisfy Assumption 1(f).[3] The range of divergence

[3]Our proofs depend on the assumption that lags of order greater than or equal to ℓ receive zero weight. We do not know whether the bootstrap provides asymptotic refinements for one-sided confidence intervals when the quadratic spectral kernel (Andrews, 1991) is used. The bootstrap does not provide asymptotic refinements for the two-sided symmetric t test and the J test statistic when this kernel is used as its characteristic exponent is two.

rates of ℓ allowed in Assumption 1(g) is narrower than the one typically assumed in the literature on HAC covariance matrix estimation (e.g., Theorem 1 of Andrews, 1991) but is wider than the one Hall and Horowitz (1996) assumed for the divergence rate of the block length. While the $\sqrt{T}$-consistency of the first-step estimator is sufficient for the first-order asymptotic theory (e.g., Assumption B(i) of Andrews, 1991), Assumption 1(i) requires further conditions.

Next, we present our three main theorems. Let q denote the characteristic exponent of the kernel ω. That is, q is the largest real number such that $\lim_{x\to 0}(1-\omega(x))/|x|^q \in [0,\infty)$.

Theorem 1: Suppose that Assumption 1 holds. Let

$$\begin{aligned}\Psi_T(x) &= \Phi(x) + T^{-1/2}p_1(x)\phi(x) + \ell T^{-1}p_2(x)\phi(x)\\ \Psi_{J,T}(x) &= F_{\chi^2_{k-p}}(x) + \ell T^{-1}p_J(x) f_{\chi^2_{k-p}}(x)\end{aligned}$$

denote the Edgeworth expansions of $P(\hat{\Sigma}_T^{-1/2}(\hat{\beta}_T-\beta_0)\le x)$ and $P(J_T\le x)$, respectively, where $\Phi(x)$ denote the p-dimensional standard normal distribution, $F_{\chi^2_{k-p}}$ and $f_{\chi^2_{k-p}}$ are the distribution and density functions of a χ^2 random variable with degree of freedom $k-p$, p_1 is even, and p_2 and p_J are odd. Then

$$\sup_{x\in\Re^p}|P(\hat{\Sigma}_T^{-1/2}(\hat{\beta}_T-\beta_0)\le x) - \Psi_T(x)| = o(\ell T^{-1}) + O(\ell^{-q}), \quad (3.2)$$

$$\sup_{x\ge 0}|P(J_T\le x) - \Psi_{J,T}(x)| = o(\ell T^{-1}) + O(\ell^{-q}). \quad (3.3)$$

Theorem 2: Suppose that Assumption 1 holds with $r\ge 12$ replaced by $r\ge 36$. Let

$$\begin{aligned}\Psi^*_T(x) &= \Phi(x) + T^{-1/2}p^*_1(x)\phi(x) + \ell T^{-1}p^*_2(x)\phi(x)\\ \Psi^*_{J,T}(x) &= F_{\chi^2_{q-p}}(x) + \ell T^{-1}p^*_J(x) f_{\chi^2_{k-p}}(x)\end{aligned}$$

denote the Edgeworth expansions of $P(\hat{\Sigma}_T^{*-1/2}(\hat{\beta}_T^* - \hat{\beta}_T) \leq x)$ and $P(J_T^* \leq x)$, respectively, where p_1^* is even, and p_2^* and p_J^* are odd. Then

$$\sup_{x \in \Re^p} |P^*(\hat{\Sigma}_T^{*-1/2}(\hat{\beta}_T^* - \hat{\beta}_T) \leq x) - \Psi_T^*(x)| = o_p(\ell T^{-1}), \tag{3.4}$$

$$\sup_{x \geq 0} |P^*(J_T^* \leq x) - \Psi_{J,T}^*(x)| = o_p(\ell T^{-1}) \tag{3.5}$$

where P^* is the probability measure induced by the bootstrap conditional on the data.

Theorem 3: Suppose that Assumption 1 holds with $r \geq 12$ replaced by $r \geq 36$. Let τ_T denote the t-statistic for the kth element of β. Let $\tau_{1,\alpha}^*$, $\tau_{2,\alpha}^*$ and χ_α^* denote the 100α level critical values for the one-sided t test, the two-sided symmetric-t test and the J test statistic, respectively. Then

$$P(\tau_T \leq \tau_{1,\alpha}^*) = 1 - \alpha + O(\ell T^{-1}) + O(\ell^{-q}), \tag{3.6}$$

$$P(|\tau_T| \leq \tau_{2,\alpha}^*) = 1 - \alpha + o(\ell T^{-1}) + O(\ell^{-q}), \tag{3.7}$$

$$P(J_T > \chi_\alpha^*) = \alpha + o(\ell T^{-1}) + O(\ell^{-q}). \tag{3.8}$$

Remarks: Theorems 1 and 2 show that the distributions of the studentized statistic and the J test statistic and their bootstrap versions can be approximated by their Edgeworth expansions. Theorem 3 shows the order of the bootstrap approximation error. For the one-sided t test, the two-sided symmetric t test and the J test statistic, the approximation errors made by the first-order asymptotic theory are of order

$$O(T^{-1/2}) + O(\ell^{-q}),\ O(\ell T^{-1}) + O(\ell^{-q}) \text{ and } O(\ell T^{-1}) + O(\ell^{-q}), \tag{3.9}$$

respectively, whereas the bootstrap approximation errors are of order

$$O(\ell T^{-1}) + O(\ell^{-q}),\ o(\ell T^{-1}) + O(\ell^{-q}) \text{ and } o(\ell T^{-1}) + O(\ell^{-q}). \tag{3.10}$$

Thus the bootstrap provides asymptotic refinements if the bias of the HAC covariance matrix estimator vanishes fast enough, i.e.,

$$O(\ell^{-q}) = o(T^{-1/2}),\ O(\ell^{-q}) = o(\ell T^{-1}) \text{ and } O(\ell^{-q}) = o(\ell T^{-1}). \tag{3.11}$$

for the three statistics, respectively.

For the one-sided t test, the bootstrap provides asymptotic refinements for a wide class of kernels that satisfy $O(\ell^{-q}) = o(T^{-1/2})$, such as the Parzen kernel. However, the bootstrap does not provide asymptotic refinements for the Bartlett kernel as it does not satisfy (3.11), because its characteristic exponent is one. For the two-sided symmetric t test and the J test statistic, the bootstrap can provide asymptotic refinements only for kernels whose characteristic exponent is greater than 2, such as the truncated kernel,

$$\omega(x) = \begin{cases} 1 & \text{for } |x| < 1 \\ 0 & \text{otherwise} \end{cases},$$

the trapezoidal kernel (Politis and Romano, 1995)

$$\omega(x) = \begin{cases} 1 & \text{for } |x| \leq \alpha \\ 1 - \frac{|x|-\alpha}{1-\alpha} & \text{for } \alpha < |x| \leq 1 \\ 0 & \text{otherwise} \end{cases},$$

where $0 < \alpha < 1$, and the Parzen (b) kernel (Parzen, 1957)

$$\omega(x) = \begin{cases} 1 - |x|^q & \text{for } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where $q > 2$. Under the assumption of exponentially decaying mixing coefficients, the truncated and trapezoidal kernels have no asymptotic bias and thus satisfy (3.11). If $q > 2$ and $\ell \neq O(T^{1/(q+1)})$, the Parzen (b) kernel also satisfies (3.11). A potential problem with these kernels is that the resulting weighting matrix is not necessarily positive semidefinite. To eliminate this problem, the weighting matrix can be modified as follows: By Schur's decomposition theorem (e.g., Theorem 13 of Magnus and Neudecker, 1999, p.16), there exist an orthogonal $k \times k$ matrix E whose columns are eigenvalues of $W_T = \hat{S}_T^{-1}$ and a diagonal matrix $\Lambda = \text{diag}(\lambda_1, ..., \lambda_k)$, whose elements are the eigenvalues of W_T, such that

$$W_T = E'^{-1} \Lambda E^{-1}. \tag{3.12}$$

Define a modified HAC covariance matrix estimator by

$$W_T^+ = E'^{-1} \Lambda^+ E^{-1}, \tag{3.13}$$

where $\Lambda^+ = \text{diag}(\max(\lambda_1, 0), ..., \max(\lambda_k, 0))$. Then W_T^+ is positive semidefinite, asymptotically equivalent to (3.12) and thus is consistent. Politis and Romano (1995, equation 12) uses a similar modification in the context of univariate spectral density estimation. For the trapezoidal kernel, the frequency of positive semidefinite corrections can be reduced by choosing small α. However, Politis and Romano (1995) recommends $\alpha = 1/2$.

4. Monte Carlo Results

In this section, we conduct a small simulation study to examine the accuracy of the proposed bootstrap procedure. We consider the following stylized linear regression model with an intercept and a regressor, x_t:

$$y_t = \beta_1 + \beta_2 x_t + u_t, \qquad \text{for } t = 1, \ldots, T. \tag{4.14}$$

The disturbance and the regressors are generated from the following AR(1) processes with common ρ,

$$u_t = \rho u_{t-1} + \varepsilon_{1t}, \tag{4.15}$$

$$x_t = \rho x_{t-1} + \varepsilon_{2t}, \tag{4.16}$$

where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})' \sim N(0, I_2)$. In the simulation, we use $\beta = (\beta_1, \beta_2)' = (0, 0)'$ for the regression parameter and $\rho \in \{0.5, 0.9, 0.95\}$ for the AR parameters. For instruments, we use x_t, x_{t-1} and x_{t-2} in addition to an intercept. This choice of instruments implies an over-identified model with 2 degrees of freedom for the J test. Two values for the sample size T, 64 and 128, are considered. The kernel functions employed are the trapezoidal, Parzen (b) and truncated kernels. In all experiments, the number of Monte Carlo trials is 1000.

The choice of the block length is important in practice. Ideally, one would choose a longer block length for more persistent processes and a shorter block length for less persistent processes. In the literature, this is typically accomplished by selecting the lag truncation parameter that minimizes the mean squared error of the HAC covariance matrix estimator (see Andrews, 1991; and Newey and West, 1994). Because the trapezoidal and truncated kernels have no asymptotic bias, however, one cannot take advantage of the usual bias-variance trade-off and thus no optimal block length can be defined for these kernels. Thus, we propose the following procedure that is similar to the general-to-specific modeling strategy for selecting the lag order of autoregressions in the literature on unit root testing (see Hall, 1994; Ng and Perron, 1995). By the Wold representation theorem, the moment function has a moving average (MA) representation of possibly infinite order. The idea is to approximate this MA representation by a sequence of finite-order MA processes. Because the block bootstrap is originally designed to capture the dependence of m-dependent-type processes when ℓ is fixed, it makes sense to approximate the process by an MA process that is m-dependent.

The proposed procedure takes the following steps.

Step 1. Let $\ell_1 < \ell_2 < \cdots < \ell_{\max}$ be candidate block lengths that satisfy Assumption 1(g) and set $k = \max - 1$.

Step 2. Test the null that every element of the moment function is $MA(\ell_k)$ against the alternative that at least one of the elements is $MA(\ell_{k+1})$.

Step 3. If the null is accepted and if $k > 1$, then let $k = k - 1$ and go to Step 2. If the null is accepted and if $k = 1$, then let $\ell = \ell_1$. If the null is rejected, then set $\ell = \ell_{k+1}$.

Because there is parameter uncertainty due to first-step estimation and because we apply a univariate testing procedure to each element of the moment function, it is difficult to control the size of this procedure. In this Monte Carlo experiment, therefore, we use the 99% level critical value to be conservative.

Our primary interest is to compare the size properties of tests based on asymptotic and bootstrap critical values. For each experiment, the empirical size for the t test for the regression slope parameter β_2 as well as for the J test is obtained using the 10% nominal significance level. Each bootstrap critical value is constructed from 999 replications of the bootstrap sampling process. In addition to the results based on the asymptotic and bootstrap critical values using our proposed procedure, we report the asymptotic results based on the Bartlett and QS kernels, with Andrews' (1991) data-dependent bandwidth estimator and Andrews and Monahan's (1992) prewhitening procedure.

Table 1 summarizes the result of the simulation study. In all cases, the size properties of the bootstrap t test are better than those of the asymptotic t test. The choice of kernel function does not make much of a difference for the performance. Indeed the empirical sizes of the bootstrap test are very close to the nominal size when T is 128. The degree of the reduction in the size distortion depends on the value of the AR parameters as well as the sample size. The bootstrap works quite well with persistent processes.

Because the moment functions have an AR(1) autocovariance structure, the prewhitening procedure has a considerable advantage in our simulation design. However, the bootstrap outperforms the conventional prewhitened HAC procedure with asymptotic critical values. In contrast, the advantage of the bootstrap for the J test is not clear because the J test performs quite well even with asymptotic critical values.[4] Based on this experiment, we recommend our bootstrap procedure especially for the t test for regression parameters.

5. Empirical Illustration

To illustrate the usefulness of the proposed bootstrap approach, we conduct bootstrap inference about the parameters in the monetary policy reaction function of Clarida, Galí and Gertler (2000, hereafter CGG). CGG model the target for the federal funds rate r_t^* by

$$r_t^* = r^* + \beta(E[\pi_{t+1}|\Omega_t] - \pi^*) + \gamma E[x_t|\Omega_t] \tag{5.17}$$

where π_t is the inflation rate, π^* is the target for inflation, Ω_t is the information set at time t, x_t is the output gap, and r^* is the target with zero inflation and output gap. Policy rules (5.17) with $\beta > 1$ and $\gamma > 0$ are stabilizing and those with $\beta \leq 1$ and $\gamma \leq 0$ are destabilizing. CGG obtain the GMM estimates of β and γ based on the set of unconditional moment conditions

$$E\{[r_t - (1-\rho_1-\rho_2)[rr^* - (\beta-1)\pi^* + \beta\pi_{t+1} + \gamma x_t] + \rho_1 r_{t-1} + \rho_2 r_{t-2}]z_t\} = 0, \tag{5.18}$$

where r_t is the actual federal fund rate, rr^* is the equilibrium real rate and z_t is a vector of instruments. They find that the GMM estimate of β is significantly less than unity

[4] See Tauchen (1986) and Hall and Horowitz (1996) for similar findings.

during the pre-Volcker era, while the estimate is significantly greater than unity during the Volcker-Greenspan era.

We reexamine these findings by applying our bootstrap procedure as well as the bootstrap procedure of Hall and Horowitz (1991) and the standard HAC asymptotics. We obtain GMM estimates of β and γ based on the linear moment conditions

$$E\{[r_t - c - \theta_1\pi_{t+1} - \theta_2 x_t - \rho_1 r_{t-1} - \rho_2 r_{t-2}]z_t\} \;=\; 0, \tag{5.19}$$

where $c = (1 - \rho_1 - \rho_2)[rr^* - (\beta - 1)\pi^*]$. Then $\hat{\beta}_T = \hat{\theta}_{1T}/(1 - \hat{\rho}_{1T} - \hat{\rho}_{2T})$ and $\hat{\gamma}_T = \hat{\theta}_{2T}/(1 - \hat{\rho}_{1T} - \hat{\rho}_{2T})$, where $\hat{\theta}_{1T}, \hat{\theta}_{2T}, \hat{\rho}_{1T}$ and $\hat{\rho}_{2T}$ are the GMM estimates of $\theta_1, \theta_2, \rho_1$ and ρ_2, respectively. We use CGG's baseline dataset and two sample periods, the pre-Volcker period (1960:1-1979:2) and the Volcker-Greenspan period (1979:3-1996:3) (see CGG for the description of the data source). In addition to their baseline specification, we construct the optimal weighting matrix using the inverse of the HAC covariance matrix estimator to allow for more general dynamic specifications in the determination of the actual funds rate. For the asymptotic confidence intervals, we use the conventional prewhitened and recolored estimates based on the Bartlett and QS kernels with the automatic bandwidth selection method (Andrews 1991, Andrews and Monahan 1992). For the confidence intervals constructed from our bootstrap, we use the trapezoidal, Parzen (b) and truncated kernels. We use the data-dependent procedure described in the previous section to select the block length for the bootstrap. The number of bootstrap replications is set to 999.

Table 2 presents GMM estimates of these parameters. Asymptotic standard errors are reported in parentheses. The first two rows of each of Tables 2(a) and (b) replicate CGG's results. These findings are robust to whether or not the HAC covariance matrix estimator is used.

Table 3 shows 90% two-sided confidence intervals of these parameters. Consistent with CGG's findings, the upper bound of the asymptotic confidence interval for β is less than unity during the pre-Volcker period, and the lower bound is far greater than unity during the Volcker-Greenspan period. Based on these estimates, CGG suggest that the Fed was accommodating inflation before 1979, but not after 1979. The bootstrap confidence interval, however, indicates that β may be greater than unity even during the pre-Volcker period, consistent with the view that the Fed has always been combating inflation. Moreover, unlike the asymptotic confidence interval, the bootstrap confidence interval does not rule out that γ is negative during the Volcker-Greenspan period.

6. Concluding Remarks

In this paper we establish that the bootstrap provides asymptotic refinements for the GMM estimator of possibly overidentified linear models when the autocovariance structure of the moment function is unknown. Because the HAC covariance matrix estimator cannot be written as a function of sample moments and converges at a rate slower than $T^{-1/2}$, the conventional techniques cannot be used directly to prove the existence of the Edgeworth expansions. Because of the nonparametric nature of the HAC covariance matrix estimator, the order of the bootstrap approximation error is larger than the typical order of the bootstrap approximation error for parametric estimators. Nevertheless, the bootstrap provides improved approximations relative to the first-order approximation. We also find that the choice of kernels plays a more important role in our second-order asymptotic theory than in the conventional first-order asymptotic theory because the order of the bootstrap approximation error depends on the bias of the HAC covariance estimator. We note that an extension of the present results to nonlinear dynamic models as well as further investigation of data-dependent methods for selecting the optimal block length would be useful.

Appendix

Notation

To simplify the notation, we will assume $p = 1$ throughout the appendix. In the proof for the case $p > 1$, the scalar β in the current proof is replaced by an arbitrary linear combination of β.

denotes the Kronecker product operator. If α is an n-dimensional nonnegative integral, $|\alpha|$ denotes its length, i.e., $|\alpha| = \sum_{i=1}^{n} |\alpha_i|$. $\|\cdot\|$ denotes the Euclidean norm, i.e., $\|x\| = (\sum_{i=1}^{n} x_i^2)^{1/2}$, where x is an n-dimensional vector. We will write $\omega(j/\ell)$ as ω_j for notational simplicity. $\kappa_j(x)$ denotes the jth cumulant of a random variable x. $\mathrm{vec}(\cdot)$ is the column-by-column vectorization function. $\mathrm{vech}(\cdot)$ denotes the column stacking operator that stacks the elements on and below the leading diagonal. For a nonnegative integral vector $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$, let

$$D^{\alpha} \;=\; \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}.$$

ℓ and l are treated differently: ℓ denotes the lag truncation parameter and l denotes an integer. Let $u_t = y_t - \beta_0' x_t$, $\hat{u}_t = y_t - \hat{\beta}_T' x_t$, $\tilde{u}_t = y_t - \tilde{\beta}_T' x_t$, $v_t = z_t u_t$, $\hat{v}_t = z_t \hat{u}_t$, $\tilde{v}_t = z_t \tilde{u}_t$, $w_t = z_t x_t'$,

$$\hat{\Gamma}_j = \begin{cases} (1/T)\sum_{t=1}^{T} \tilde{v}_{t+j}\tilde{v}_t' & j \geq 0 \\ (1/T)\sum_{t=1}^{T} \tilde{v}_t \tilde{v}_{t-j}' & j < 0 \end{cases}, \quad \nabla\tilde{\Gamma}_j = \begin{cases} (1/T)\sum_{t=1}^{T} v_{t+j}w_t' + w_{t+j}v_t' & j \geq 0 \\ (1/T)\sum_{t=1}^{T} v_t w_{t-j}' + v_t w_{t-j}' & j < 0 \end{cases},$$

$$\tilde{\Gamma}_j = \begin{cases} (1/T)\sum_{t=1}^{T} v_{t+j}v_t' & j \geq 0 \\ (1/T)\sum_{t=1}^{T} v_t v_{t-j}' & j < 0 \end{cases}, \quad \nabla\Gamma_j = \begin{cases} E(v_{t+j}w_t' + w_{t+j}v_t') & j \geq 0 \\ E(v_t w_{t-j}' + v_t w_{t-j}') & j < 0 \end{cases},$$

$$\Gamma_j = \begin{cases} E(v_{t+j}v_t') & j \geq 0 \\ E(v_t v_{t-j}') & j < 0 \end{cases}, \quad \nabla^2\Gamma_j = \begin{cases} (1/T)\sum_{t=1}^{T} w_{t+j}w_t' & j \geq 0 \\ (1/T)\sum_{t=1}^{T} w_t w_{t-j}' & j < 0 \end{cases},$$

$$\begin{aligned} \hat{S}_T &= \textstyle\sum_{j=-\ell}^{\ell} \omega_j \hat{\Gamma}_j, & \tilde{S}_T &= \textstyle\sum_{j=-\ell}^{\ell} \omega_j \tilde{\Gamma}_j, & \bar{S}_T &= \textstyle\sum_{j=-\ell}^{\ell} \omega_j \Gamma_j, \\ S_T &= \textstyle\sum_{j=-T+1}^{T-1} (1 - \frac{|j|}{T})\Gamma_j, & \nabla\tilde{S}_T &= \textstyle\sum_{j=-\ell}^{\ell} \omega_j \nabla\tilde{\Gamma}_j, & \nabla\bar{S}_T &= \textstyle\sum_{j=-\ell}^{\ell} \omega_j \nabla\Gamma_j, \\ \nabla S &= \textstyle\sum_{j=-\infty}^{\infty} \nabla\Gamma_j, & \nabla^2\tilde{S}_T &= \textstyle\sum_{j=-\ell}^{\ell} \omega_j \nabla^2\tilde{\Gamma}_j. \end{aligned}$$

Let $G_T = (1/T)\sum_{t=1}^{T} w_t$ and $m_T = T^{-1/2}\sum_{t=1}^{T} v_t$. Then the studentized statistic can be written as

$$f_T \;=\; \sqrt{T}\hat{\Sigma}^{-1/2}(\hat{\beta}_T - \beta_0) \;=\; (G_T'\hat{S}_T^{-1}G_T)^{-\frac{1}{2}} G_T'\hat{S}_T^{-1} m_T.$$

We use the following notation for the bootstrap. Let

$$\begin{aligned} m_T^* &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T} (z_t^* u_t^* - \mu_T^*) \;=\; \frac{1}{\sqrt{b}}\sum_{k=1}^{b} B_{N_k}, \\ B_{N_k} &= \frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell} (z_{N_k+i}\hat{u}_{N_k+i} - \mu_T^*) \;=\; \frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell} (\hat{v}_{N_k+i} - \mu_T^*), \\ \widehat{B}_{N_k} &= \frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell} \left(z_{Nk+i}^* \widehat{u}_{Nk+i}^* - \mu_T^*\right) \;, \widehat{u}_i^* \;=\; y_i^* - \widetilde{\beta}^{*\prime} x_i^*, \\ G_T^* &= \frac{1}{T}\sum_{t=1}^{T} z_t^* x_t^{*\prime} \;=\; \frac{1}{b}\sum_{k=1}^{b} F_{N_k}, \\ F_{N_k} &= \frac{1}{\ell}\sum_{i=1}^{\ell} z_{N_k+i} x_{N_k+i}' \;=\; \frac{1}{\ell}\sum_{i=1}^{\ell} w_{N_k+i}. \end{aligned}$$

$$\hat{S}_T^* \;=\; \frac{1}{b}\sum_{k=1}^{b} \widehat{B}_{N_k}\widehat{B}_{N_k}', \quad \widetilde{S}_T^* \;=\; \frac{1}{b}\sum_{k=1}^{b} B_{N_k}B_{N_k}', \quad S_T^* \;=\; \mathrm{Var}^*(m_T^*).$$

Then the bootstrap version of the first-step and the second-step GMM estimators can be written as

$$\begin{aligned}
\widetilde{\beta}^{*} &= \hat{\beta} + \left[\frac{1}{b}\sum_{k=1}^{b} F_{N_k}' V_T \frac{1}{b}\sum_{k=1}^{b} F_{N_k}\right]^{-1} \frac{1}{b}\sum_{k=1}^{b} F_{N_k}' V_T \frac{1}{\sqrt{Tb}}\sum_{k=1}^{b} B_{N_k} \\
&= \hat{\beta} + \left[G_T^{*\prime} V_T G_T^{*}\right]^{-1} G_T^{*\prime} V_T \frac{1}{\sqrt{T}} m_T^{*}, \\
\hat{\beta}^{*} &= \hat{\beta} + \left[\frac{1}{b}\sum_{k=1}^{b} F_{N_k}' \hat{S}_T^{*-1} \frac{1}{b}\sum_{k=1}^{b} F_{N_k}\right]^{-1} \frac{1}{b}\sum_{k=1}^{b} F_{N_k}' \hat{S}_T^{*-1} \frac{1}{\sqrt{Tb}}\sum_{k=1}^{b} B_{N_k} \\
&= \hat{\beta} + \left[G_T^{*\prime} \hat{S}_T^{*-1} G_T^{*}\right]^{-1} G_T^{*\prime} \hat{S}_T^{*-1} \frac{1}{\sqrt{T}} m_T^{*},
\end{aligned}$$

respectively.

Proofs of Lemmas

Next, we will present the lemmas used in the proofs of the theorems. Lemma A.1 produces a Taylor series expansion of the studentized statistic f_T. Lemma A.2 provides bounds on the moments and will be used in the proofs of Lemmas A.3–A.6. Lemma A.3 shows the limits and the convergence rates of the first three cumulants of g_T in (A.1), that will be used to derive the formal Edgeworth expansion. Lemmas A.5 and A.6 provide bounds on the approximation error. For convenience, we present Lemma B.1 that will be used in the proofs of Lemmas B.2 and B.3. Lemma B.2 shows the consistency and convergence rate of the bootstrap version of the moments. Lemma B.3 shows the limits and the convergence rates of the first three cumulants of the bootstrap version.

Lemma A.1:

$$\begin{aligned}
& f_T \\
= \; & \mathbf{a}' m_T + \mathbf{b}'[(G_T - G_0) \quad m_T] + \mathbf{c}'[\mathrm{vech}(\hat{S}_T - S_0) \quad m_T] \\
& + \mathbf{d}'[(G_T - G_0) \quad \mathrm{vech}(\hat{S}_T - S_0) \quad m_T] + \mathbf{e}'[\mathrm{vech}(\hat{S}_T - S_0) \quad \mathrm{vech}(\hat{S}_T - S_0) \quad m_T] \\
& + O_p((\ell/T)^{3/2}) \\
= \; & \mathbf{a}' m_T + \mathbf{b}'[(G_T - G_0) \quad m_T] + \mathbf{c}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T] + \mathbf{c}'[\mathrm{vech}(\bar{S}_T - S_0) \quad m_T] \\
& + \mathbf{d}'[(G_T - G_0) \quad \mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T] + \mathbf{e}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad \mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T] \\
& + \mathbf{d}'[(G_T - G_0) \quad \mathrm{vech}(\bar{S}_T - S_0) \quad m_T] + \mathbf{e}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad \mathrm{vech}(\bar{S}_T - S_0) \quad m_T] \\
& + \mathbf{e}'[\mathrm{vech}(\bar{S}_T - S_0) \quad \mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T] + \mathbf{e}'[\mathrm{vech}(\bar{S}_T - S_0) \quad \mathrm{vech}(\bar{S}_T - S_0) \quad m_T] \\
& + O_p((\ell/T)^{3/2}) \\
\equiv \; & g_T + \mathbf{c}'[\mathrm{vech}(\bar{S}_T - S_0) \quad m_T] + \mathbf{d}'[(G_T - G_0) \quad \mathrm{vech}(\bar{S}_T - S_0) \quad m_T] \\
& + \mathbf{e}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad \mathrm{vech}(\bar{S}_T - S_0) \quad m_T] + \mathbf{e}'[\mathrm{vech}(\bar{S}_T - S_0) \quad \mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T] \\
& + \mathbf{e}'[\mathrm{vech}(\bar{S}_T - S_0) \quad \mathrm{vech}(\bar{S}_T - S_0) \quad m_T] + O_p((\ell/T)^{3/2}), \qquad \text{(A.1)}
\end{aligned}$$

where a, b, c, d and e are q, q^2, $q(q^2+q)$, $q(q^2+q)/2$, $q^2(q^2+q)/2$ and $q((q^2+q)/2)^2$-dimensional vectors of smooth functions of G_0 and S_0, respectively.

Proof of Lemma A.1: (A.1) immediately follows from a Taylor series expansion of f_T around

$$(m_T', G_T', \mathrm{vech}(\hat{S}_T)')' \; = \; (0_{1\times q}, G_0', \mathrm{vech}(S_0)')'$$

and from Theorem 1 of Andrews (1991). *Q.E.D.*

Lemma A.2:

$$E\|m_T\|^{r+\eta} \;=\; O(1), \qquad \text{(A.2)}$$

$$
\begin{aligned}
E\|T^{1/2}(G_T - G_0)\|^{r+\eta} &= O(1), &\text{(A.3)}\\
E\|(T/\ell)^{1/2}\text{vech}(\tilde{S}_T - \bar{S}_T)\|^{r/2} &= O(1), &\text{(A.4)}\\
E\|(T/\ell)^{1/2}\text{vech}(\nabla\tilde{S}_T - \nabla\bar{S}_T)\|^{r/2} &= O(1), &\text{(A.5)}\\
E\|T^{1/2}\text{vech}(\hat{S}_T - \tilde{S}_T)\|^{r/2} &= O(1). &\text{(A.6)}
\end{aligned}
$$

Proof of Lemma A.2: First, (A.2) and (A.3) immediately follow from the moment inequality of Yokoyama (1980). Second, we will show (A.4). Note that

$$
\begin{aligned}
(T/\ell)^{1/2}(\tilde{S}_T - \bar{S}_T) &= (T/\ell)^{1/2}\sum_{j=-\ell}^{\ell}\omega_j(\tilde{\Gamma}_j - \Gamma_j) = (\ell/T)^{1/2}\sum_{i=1}^{[T/\ell]} W_i \\
&= (\ell/T)^{1/2}\Big(\sum_{i=0 mod 3} W_i + \sum_{i=1 mod 3} W_i + \sum_{i=2 mod 3} W_i\Big), \qquad \text{(A.7)}
\end{aligned}
$$

where

$$
W_i = \frac{1}{\ell}\sum_{t=(i-1)\ell+1}^{i\ell}\{v_t v_t' - E(v_t v_t') + \sum_{j=1}^{\ell}\omega_j[v_{t+j}v_t' - E(v_{t+j}v_t') + v_t v_{t+j}' - E(v_t v_{t+j}')]\}.
$$

Note that the summands in each sum on the RHS of (A.7) are asymptotically independent by construction. Thus,

$$
E\left\|(T/\ell)^{1/2}\text{vech}(\tilde{S}_T - \bar{S}_T)\right\|^{\frac{r}{2}} = O(E\|\text{vech}(W_2)\|^{\frac{r}{2}}) = \sum_{i=1}^{3} O(E\|\text{vech}(W_2(i))\|^{\frac{r}{2}}) \quad \text{(A.8)}
$$

where

$$
W_2(1) = \ell^{-1}\sum_{t=\ell+1}^{2\ell}\sum_{j=0}^{\ell-1}\omega_j v_{t+j}v_t',\ W_2(2) = \ell^{-1}\sum_{t=\ell+1}^{2\ell}\sum_{j=-\ell+1}^{-1}\omega_j v_t v_{t-j}',\ W_2(3) = \sum_{j=-\ell+1}^{\ell-1}E(v_0 v_{-j}').
$$

Thus it suffices to show that, for $i, j = 1, 2, ..., q$,

$$
\begin{aligned}
E|W_2(1)^{(i,j)}|^{\frac{r}{2}} &= O(1), &\text{(A.9)}\\
E|W_2(2)^{(i,j)}|^{\frac{r}{2}} &= O(1), &\text{(A.10)}\\
E|W_2(3)^{(i,j)}|^{\frac{r}{2}} &= O(1), &\text{(A.11)}
\end{aligned}
$$

where $W_2(\cdot)^{(i,j)}$ denotes the (i,j)th element of $W_2(\cdot)$. By Assumptions 1(a) and 1(f), it follows that

$$
E|W_2(1)^{(i,j)}|^{r/2} = O(\ell^{r/2}\sum_{t_1\le t_2\le\cdots\le t_r} E|v_{t_1}^{(k_1)}v_{t_2}^{(k_2)}\cdots v_{t_r}^{(k_r)}|), \qquad \text{(A.12)}
$$

where $0 \le t_l \le 2\ell$ and $k_l = i, j$ for $l = 1, 2, ..., r$. Then the standard arguments used in proofs of the moment inequality complete the proof of (A.9). The proof of (A.10) is analogous to that of (A.9) and thus is omitted. By the mixing inequality of Hall and Heyde (1980, Corollary A.2), it follows that for some $d' > 0$

$$
E|W_2(3)^{(i,j)}|^{\frac{r}{2}} = (\sum_{j=-\ell+1}^{\ell-1}E(v_0 v_{-j}'))^{\frac{r}{2}} = (\sum_{j=-\ell+1}^{\ell-1}\alpha_j^{d'})^{\frac{r}{2}} = O(1), \qquad \text{(A.13)}
$$

and thus (A.11) holds. Therefore, (A.4) immediately follows from (A.7)–(A.11). The proof of (A.5) is analogous to that of (A.4) and thus is omitted.

Lastly, we will prove (A.6). Note that

$$
T^{1/2}(\hat{S}_T - \tilde{S}_T) = \nabla\tilde{S}_T T^{1/2}(\tilde{\beta}_T - \beta_0) + \nabla^2\tilde{S}_T T^{1/2}(\tilde{\beta}_T - \beta_0)^2. \qquad \text{(A.14)}
$$

Thus it follows from (A.5) and Minkowski's inequality that

$$[E\|\nabla\tilde{S}_T\|^r]^{1/r} \;\leq\; [E\|\nabla\tilde{S}_T - \nabla\bar{S}_T\|^r]^{1/r} + [E\|\nabla\bar{S}_T\|^r]^{1/r} \;=\; O(\ell^{1/2}T^{-1/2}) + O(1), \quad \text{(A.15)}$$

$$\begin{aligned}
[E\|\nabla^2\tilde{S}_T\|^r]^{1/r} &\leq [E\|\sum_{j=-\ell}^{\ell}\omega_j(\nabla^2\Gamma_j - E(\nabla^2\Gamma_j))\|^r]^{1/r} + [E\|\sum_{j=-\ell}^{\ell}\omega_j E(\nabla^2\Gamma_j)\|^r]^{1/r} \\
&= O(\ell T^{-1/2}) + O(\ell). \qquad \text{(A.16)}
\end{aligned}$$

Therefore (A.6) follows from (A.14), (A.15), (A.16), Assumption 1(i) and Hölder's inequality. *Q.E.D.*

Lemma A.3:

$$\begin{aligned}
T^{1/2}\kappa_1(g_T) &= \alpha_\infty + O(\ell^{-q}) + o(\ell T^{-1/2}), & \text{(A.17)} \\
(T/\ell)(\kappa_2(g_T) - 1) &= \gamma_\infty + O(\ell^{-1/2}), & \text{(A.18)} \\
T^{1/2}\kappa_3(g_T) &= \kappa_\infty - 3\alpha_\infty + O(\ell^{-q}) + o(\ell T^{-1/2}), & \text{(A.19)} \\
(T/\ell)(\kappa_4(g_T) - 3) &= \zeta_\infty + O(\ell^{-1/2}), & \text{(A.20)}
\end{aligned}$$

where

$$\begin{aligned}
\alpha_\infty &= \mathbf{b}'\sum_{i=-\infty}^{\infty}E[w_0 \quad v_i] + \mathbf{c}'\sum_{i,j=-\infty}^{\infty}E[\text{vech}(v_0v_i') \quad v_j] \\
&\quad + \mathbf{c}'\sum_{i=-\infty}^{\infty}E\{\text{vech}[\nabla\bar{S}(E(w_0)'VE(w_0))^{-1}E(w_0)'Vv_0] \quad v_i\} \\
\gamma_\infty &= 2\lim_{T\to\infty}\frac{1}{\ell}\sum_{j=-\ell}^{\ell}\sum_{i,k=-T}^{T}E\{\mathbf{a}'v_0\mathbf{c}'[\text{vech}(v_iv_{i-j}' - \Gamma_j) \quad v_k]\} \\
&\quad + 2\lim_{T\to\infty}\frac{1}{\ell T}\sum_{i,l=-\ell}^{\ell}\sum_{i,k,m=-T}^{T}E\{\mathbf{a}'v_0\mathbf{e}'[\text{vech}(v_iv_{i-j}' - \Gamma_j) \quad \text{vech}(v_kv_{k-l}' - \Gamma_l) \quad v_m]\} \\
&\quad + \lim_{T\to\infty}\frac{1}{\ell T}\sum_{j,k,m=-T}^{T}\sum_{i,l=-\ell}^{\ell}E\{\mathbf{c}'[\text{vech}(v_0v_{-i}' - \Gamma_i) \quad v_j]\mathbf{c}'[\text{vech}(v_kv_{k-l}' - \Gamma_k) \quad v_m]\}, \\
\kappa_\infty &= \sum_{i,j=-\infty}^{\infty}E(\mathbf{a}'v_0\mathbf{a}'v_i\mathbf{a}'v_j) + 3\lim_{T\to\infty}\frac{1}{T}\sum_{i,j,k=-T+1}^{T-1}E\{\mathbf{a}'v_0\mathbf{a}'v_i\mathbf{b}'[\text{vech}(w_j - E(w_j)) \quad v_k]\} \\
&\quad + 3\lim_{T\to\infty}\frac{1}{T}\sum_{i,j,k,l=-T}^{T}E\{\mathbf{a}'v_0\mathbf{a}'v_i\mathbf{c}'[\text{vech}(v_jv_{j-k}' - \Gamma_k) \quad v_l\} \\
&\quad + 3\lim_{T\to\infty}\frac{1}{T^2}\sum_{i,j,k=-T}^{T}E\{\mathbf{a}'v_0\mathbf{a}'v_i\mathbf{c}'\text{vech}[\nabla\bar{S}(E(w_0)'VE(w_0))^{-1}E(w_0)'Vv_j] \quad v_k\},
\end{aligned}$$

$$\begin{aligned}
&\zeta_\infty \\
&= \frac{4}{\ell T}\sum_{i,j,k,m=-T}^{T}\sum_{l=-\ell}^{\ell}E\{\mathbf{a}'v_0\mathbf{a}'v_i\mathbf{a}'v_j\mathbf{c}'[\text{vech}(v_kv_{k-l}' - \Gamma_l) \quad v_m]\} \\
&\quad + \lim\frac{4}{\ell T^2}\sum_{i,j,k,m,o=-T}^{T}\sum_{l,n=-\ell}^{\ell}E\{\mathbf{a}'v_0\mathbf{a}'v_i\mathbf{a}'v_j\mathbf{e}'[\text{vech}(v_kv_{k-l}' - \Gamma_l) \quad \text{vech}(v_mv_{m-n}' - \Gamma_n) \quad v_o]\} \\
&\quad + \lim\frac{6}{\ell T^2}\sum_{i,j,l,m,o=-T}^{T}\sum_{k,n=-\ell}^{\ell}E\{\mathbf{a}'v_0\mathbf{a}'v_i\mathbf{c}'[\text{vech}(v_jv_{j-k}' - \Gamma_k) \quad v_l]\mathbf{c}'[\text{vech}(v_mv_{m-n}' - \Gamma_n) \quad v_o]\}
\end{aligned}$$

$$-12 \lim \frac{1}{\ell} \sum_{j,l=-T}^{T} \sum_{k=-\ell}^{\ell} E\{\mathbf{a}' v_0 \mathbf{c}' [\mathrm{vech}(v_j v_{j-k}' - \Gamma_k) \quad v_l]\}$$

$$-12 \lim \frac{1}{\ell T} \sum_{j,l,n=-T}^{T} \sum_{k,m=-\ell}^{\ell} E\{\mathbf{a}' v_0 \mathbf{e}' [\mathrm{vech}(v_j v_{j-k}' - \Gamma_k) \quad (v_l v_{l-m}' - \Gamma_m) \quad v_n]\}$$

$$-6 \lim \frac{1}{\ell T^2} \sum_{j,k,m=-T}^{T} \sum_{i,l=-\ell}^{\ell} E\{\mathbf{c}' [(v_0 v_{-i}' - \Gamma_i) \quad v_j] \mathbf{c}' [(v_k v_{k-l}' - \Gamma_l) \quad v_m]\}.$$

Proof of Lemma A.3: First, we will prove (A.17). By Hölder's inequality and Lemma A.2, it suffices to show that

$$T^{1/2} E[(G_T - G_0) \quad m_T] = \sum_{i=-\infty}^{\infty} E[w_0 \quad v_i] + O(T^{-1}), \tag{A.21}$$

$$T^{1/2} E[\mathrm{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T] = \sum_{i,j=-\infty}^{\infty} E[\mathrm{vech}(v_0 v_i') \quad v_j] + O(\ell^{-q}) + O(\ell T^{-1}), \tag{A.22}$$

$$T^{1/2} E[\mathrm{vech}(\hat{S}_T - \tilde{S}_T) \quad m_T] = \sum_{i=-\infty}^{\infty} E\{\mathrm{vech}[\nabla \bar{S} (E(w_0)' V E(w_0))^{-1} E(w_0)' V v_0] \quad v_i\}$$
$$+ O(\ell^{1/2} T^{-1/2}), \tag{A.23}$$

$$(T/\ell) E[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad \mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T] = o(1). \tag{A.24}$$

First, (A.21) follows from several applications of the mixing inequality. Second, we will show (A.22). We have

$$
\begin{aligned}
& T^{\frac{1}{2}} E[\sum_{j=0}^{\ell} \omega_j \mathrm{vech}(\tilde{\Gamma}_j - \Gamma_j) \quad m_T] \\
= & \sum_{i=0}^{\ell} \omega_i \sum_{j=-\ell-T+1}^{T-1} \frac{T - i 1(j > i) - |j| 1(j > 0 \text{ or } j \le -i)}{T} E[\mathrm{vech}(v_0 v_{-i}') \quad v_j] \\
= & \sum_{i=0}^{\ell} \omega_i \sum_{j=-\ell-T+1}^{T-1} E[\mathrm{vech}(v_0 v_{-i}') \quad v_j] + O(\ell T^{-1}) \\
= & \sum_{i=0}^{\ell} \sum_{j=-\ell-T+1}^{T-1} E[\mathrm{vech}(v_0 v_{-i}') \quad v_j] + O(\ell^{-q}) + O(\ell T^{-1}) \\
= & \sum_{i=0}^{\infty} \sum_{j=-\infty}^{\infty} E[\mathrm{vech}(v_0 v_{-i}') \quad v_j] + O(\ell^{-q}) + O(\ell T^{-1}).
\end{aligned} \tag{A.25}
$$

The first equality follows from strict stationarity. Repeated applications of the moment inequality of Yokoyama (1980) produce

$$
\begin{aligned}
& \sum_{i=0}^{\ell} \omega_i \sum_{j=-\ell-T+1}^{T-j} \frac{T - i 1(j > i) - |j| 1(j > 0 \text{ or } j \le -i)}{T} E[\mathrm{vech}(v_0 v_{-i}') \quad v_j] \\
= & O\left(T^{-1} \sum_{i=0}^{\ell} \omega_i \left[\sum_{j=-\ell-T}^{-2j-1} |j| \alpha_{-i-j}^{r'} + \sum_{j=-2j}^{-j} |j| \alpha_i^{r'} + \sum_{j=-i}^{-(1/2)i} i \alpha_{-j} \right.\right. \\
& \left.\left. + \sum_{j=-(1/2)i+1}^{-1} i \alpha_{i+j}^{r'} + \sum_{j=0}^{i} (i+j) \alpha_i^{r'} + \sum_{j=i+1}^{T-1} (i+j) \alpha_j^{r'} \right] \right) \\
= & O(\ell T^{-1}).
\end{aligned} \tag{A.26}
$$

for some $r' \in (0,1)$, from which the second equality follows. Arguments analogous to the proof of Theorem 10 of Hannan (1970, pp.283-284) yield the last two equalities. By symmetric arguments, it follows that

$$\begin{aligned} & T^{\frac{1}{2}} E[\sum_{j=-\ell}^{-1} \omega_j \text{vech}(\tilde{\Gamma}_j - \Gamma_j) \quad m_T] \\ = & \sum_{i=-\infty}^{-1} \sum_{j=-\infty}^{\infty} E[\text{vech}(v_0 v'_{-i}) \quad v_j] + O(\ell^{-q}) + O(\ell T^{-1}). \qquad \text{(A.27)} \end{aligned}$$

Hence, (A.23) follows from (A.25) and (A.27). Third, we will show (A.23). It follows from (A.14), Assumption 1(i) and Lemma A.2 that

$$\begin{aligned} & T^{\frac{1}{2}} E[\text{vech}(\hat{S}_T - \tilde{S}_T) \quad m_T] \\ = & T^{\frac{1}{2}} E[\text{vech}(\nabla \tilde{S}_T(\tilde{\beta}_T - \beta_0) + \nabla^2 \tilde{S}_T(\tilde{\beta}_T - \beta_0)^2) \quad m_T] \\ = & T^{\frac{1}{2}} E[\text{vech}((\nabla \tilde{S}_T - \nabla \bar{S}_T)(\tilde{\beta}_T - \beta_0) \quad m_T)] + T^{\frac{1}{2}} E[\text{vech}(\nabla \bar{S}_T(\tilde{\beta}_T - \beta_0) \quad m_T)] \\ & + T^{\frac{1}{2}} E[\text{vech}((\nabla^2 \tilde{S}_T - \nabla^2 \bar{S}_T)(\tilde{\beta}_T - \beta_0)^2) \quad m_T] \\ & + T^{\frac{1}{2}} E[\text{vech}((\nabla^2 \nabla^2 \bar{S}_T(\tilde{\beta}_T - \beta_0)^2) \quad m_T)] \\ = & \sum_{i=-\infty}^{\infty} E\{\text{vech}[\nabla \bar{S}(E(w_0)' V E(w_0))^{-1} E(w_0)' V v_0 \quad v_i]\} + O(\ell^{1/2} T^{-1/2}), \qquad \text{(A.28)} \end{aligned}$$

which completes the proof of (A.23). Lastly, we will show (A.24).

$$\begin{aligned} & (T/\ell) E[\text{vech}(\hat{S}_T - \bar{S}_T) \quad \text{vech}(\hat{S}_T - \bar{S}_T) \quad m_T] \\ = & (T/\ell) E[\text{vech}(\tilde{S}_T - \bar{S}_T) \quad \text{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T] + o(1) \\ = & \ell^{-1} T^{-3/2} \sum_{i,j=-\ell}^{\ell} \sum_{t,s,u=1}^{T} E[\text{vech}(v_{t+i} v'_t - \Gamma_i) \quad \text{vech}(v_{s+j} v_s - \Gamma_j) \quad v_u] + o(1) \\ = & O(\ell^2 T^{-1/2}) \; = \; o(1). \qquad \text{(A.29)} \end{aligned}$$

Therefore, (A.17) follows from (A.21)–(A.24).

Next, we will prove (A.18). It follows from (A.17), Hölder's inequality and Lemma A.2 that

$$\begin{aligned} \kappa_2(g_T) - 1 &= E(g_T^2) - [E(g_T)]^2 - 1 \\ &= 2E\{\mathbf{a}' m_T \mathbf{b}'[(G_T - G_0) \quad m_T]\} + 2E\{\mathbf{a}' m_T \mathbf{c}'[\text{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T]\} \\ & \quad + 2E\{\mathbf{a}' m_T \mathbf{e}'[\text{vech}(\tilde{S}_T - \bar{S}_T) \quad \text{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T]\} \\ & \quad + E\{\mathbf{c}'[\text{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\}^2 + O(\ell^{1/2} T^{-1}). \qquad \text{(A.30)} \end{aligned}$$

Thus, we only need to analyze the first four terms on the RHS of (A.30). First, by repeated applications of the mixing inequality as in the proof of moment inequalities (e.g, the proof of Lemma 4 of Billingsley, 1968, pp.172–174), one can show that

$$T E\{\mathbf{a}' m_T \mathbf{b}'[(G_T - G_0) \quad m_T]\} \; = \; O(1). \qquad \text{(A.31)}$$

Second, it follows from arguments similar to the one used in the proof of (A.17) that

$$\begin{aligned} & (T/\ell) E\{\mathbf{a}' m_T \mathbf{c}'[\text{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T]\} \\ = & (\ell T)^{-1} \sum_{j=-\ell}^{\ell} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{u=1}^{T} \omega_j E\{\mathbf{a}' v_t \mathbf{c}'[\text{vech}(v_s v'_{s-j} - \Gamma_j) \quad v_u]\} \\ = & \ell^{-1} \sum_{j=-\ell}^{\ell} \sum_{i,k=-T+1}^{T-1} \omega_j (1 - \tau_{i,k}) E\{\mathbf{a}' v_0 \mathbf{c}'[\text{vech}(v_i v'_{i-j} - \Gamma_j) \quad v_k]\} \end{aligned}$$

$$
\begin{aligned}
&= \ell^{-1} \sum_{j=-\ell}^{\ell} \sum_{i,k=-T+1}^{T-1} \omega_j E\{\mathbf{a}' v_0 \mathbf{c}'[\mathrm{vech}(v_i v_{i-j}' - \Gamma_j) \quad v_k]\} + O(\ell T^{-1}) \\
&= \ell^{-1} \sum_{j=-\ell}^{\ell} \sum_{i,k=-T+1}^{T-1} E\{\mathbf{a}' v_0 \mathbf{c}'[\mathrm{vech}(v_i v_{i-j}' - \Gamma_j) \quad v_k]\} + O(\ell^{-q_w}) + O(\ell T^{-1}) \\
&= \lim_{T\to\infty} \ell^{-1} \sum_{j=-\ell}^{\ell} \sum_{t=-T+1}^{T-1} \sum_{s=-T+1}^{T-1} E\{\mathbf{a}' v_0 \mathbf{c}'[\mathrm{vech}(v_t v_{t-j}' - \Gamma_j) \quad v_s]\} + O(\ell^{-1}), \quad \text{(A.32)}
\end{aligned}
$$

$$
\begin{aligned}
& (T/\ell) E\{\mathbf{a}' m_T \mathbf{e}'[\mathrm{vech}(\tilde{S}_T - \bar{S}_T) \quad \mathrm{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T]\} \\
=\ & \frac{1}{\ell T^2} \sum_{i,j=-\ell}^{\ell} \sum_{r,s,t,u=1}^{T} \omega_i \omega_j E\{\mathbf{a}' v_r \mathbf{e}'[\mathrm{vech}(v_s v_{s-i}' - \Gamma_i) \quad \mathrm{vech}(v_t v_{t-j}' - \Gamma_j) \quad v_u]\} \\
=\ & \frac{1}{\ell T} \sum_{i,j=-\ell}^{\ell} \sum_{s,t,u=-T}^{T} \omega_i \omega_j (1 - \tau_{s,t,u}) E\{\mathbf{a}' v_0 \mathbf{e}'[\mathrm{vech}(v_s v_{s-i}' - \Gamma_i) \quad \mathrm{vech}(v_t v_{t-j}' - \Gamma_j) \quad v_u]\} \\
=\ & \frac{1}{\ell T} \sum_{i,j=-\ell}^{\ell} \sum_{s,t,u=-T}^{T} \omega_i \omega_j E\{\mathbf{a}' v_0 \mathbf{e}'[\mathrm{vech}(v_s v_{s-i}' - \Gamma_i) \quad \mathrm{vech}(v_t v_{t-j}' - \Gamma_j) \quad v_u]\} \\
& + O(\ell^2 T^{-1}) \\
=\ & \frac{1}{\ell T} \sum_{i,j=-\ell}^{\ell} \sum_{s,t,u=-T}^{T} E\{\mathbf{a}' v_0 \mathbf{e}'[\mathrm{vech}(v_s v_{s-i}' - \Gamma_i) \quad \mathrm{vech}(v_t v_{t-j}' - \Gamma_j) \quad v_u]\} \\
& + O(\ell^{-q}) + O(\ell^2 T^{-1}) \\
=\ & \lim_{T\to\infty} \frac{1}{\ell T} \sum_{i,j=-\ell}^{\ell} \sum_{s,t,u=-T}^{T} E\{\mathbf{a}' v_0 \mathbf{e}'[\mathrm{vech}(v_s v_{s-i}' - \Gamma_i) \quad \mathrm{vech}(v_t v_{t-j}' - \Gamma_j) \quad v_u]\} \\
& + O(\ell^{-1}), \qquad \text{(A.33)}
\end{aligned}
$$

and

$$
\begin{aligned}
& (T/\ell) E\{\mathbf{c}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\}^2 \\
=\ & \ell^{-1} T^{-2} \sum_{t,s,u,v=1}^{T} \sum_{i,j=-\ell}^{\ell} \omega_i \omega_j E\{\mathbf{c}'[\mathrm{vech}(v_s v_{s-i}' - \Gamma_i) \quad v_t] \mathbf{c}'[\mathrm{vech}(v_u v_{u-j}' - \Gamma_j) \quad v_v]\} \\
=\ & (\ell T)^{-1} \sum_{j,k,m=-T}^{T} \sum_{i,l=-\ell}^{\ell} \omega_i \omega_j (1 - \tau_{j,k,m}) E\{\mathbf{c}'[\mathrm{vech}(v_0 v_{-i}' - \Gamma_i) \quad v_j] \\
& \times \mathbf{c}'[\mathrm{vech}(v_k v_{k-l}' - \Gamma_k) \quad v_m]\} \\
=\ & (\ell T)^{-1} \sum_{j,k,m=-T}^{T} \sum_{i,l=-\ell}^{\ell} \omega_i \omega_j E\{\mathbf{c}'[\mathrm{vech}(v_0 v_{-i}' - \Gamma_i) \quad v_j] \mathbf{c}'[\mathrm{vech}(v_k v_{k-l}' - \Gamma_k) \quad v_m]\} \\
& + O(\ell T^{-1}) \\
=\ & (\ell T)^{-1} \sum_{j,k,m=-T}^{T} \sum_{i,l=-\ell}^{\ell} E\{\mathbf{c}'[\mathrm{vech}(v_0 v_{-i}' - \Gamma_i) \quad v_j] \mathbf{c}'[\mathrm{vech}(v_k v_{k-l}' - \Gamma_k) \quad v_m]\} \\
& + O(\ell^{-q}) + O(\ell T^{-1}) \\
=\ & \lim_{T\to\infty} \ell^{-1} T^{-1} \sum_{j,k,m=-T}^{T} \sum_{i,l=-\ell}^{\ell} E\{\mathbf{c}'[\mathrm{vech}(v_0 v_{-i}' - \Gamma_i) \quad v_j] \mathbf{c}'[\mathrm{vech}(v_k v_{k-l}' - \Gamma_k) \quad v_m]\} \\
& + O(\ell^{-1}), \qquad \text{(A.34)}
\end{aligned}
$$

where $\tau_{i,k} = (1/T)\min(\max(|i|, |k|, |i-k|), T)$ and $\tau_{s,t,u} = (1/T)\min(\max(|s|, |t|, |u|, |s-t|, |t-u|, |u-s|), T)$. The proofs of (A.32), (A.33) and (A.34) are similar to that of (A.17) and thus details are omitted. Therefore, (A.18) follows from (A.30)–(A.33).

Third, we will prove (A.19). By (A.17), (A.18) and

$$\kappa_3(g_T) \;=\; E(g_T^3) - 3E(g_T^2)E(g_T) + 2(E(g_T))^3, \tag{A.35}$$

it suffices to show that

$$T^{1/2}E(g_T^3) \;=\; \kappa_\infty + O(\ell^{-q}) + o(\ell T^{-1/2}). \tag{A.36}$$

It follows from Assumption 1(i), Hölder's inequality and Lemma A.2 that

$$\begin{aligned} E(g_T^3) \;&=\; E[(\mathbf{a}'m_T)^3] + 3E\{(\mathbf{a}'m_T)^2\mathbf{b}'[(G_T - G_0)' \quad m_T]\} \\ &\quad +3E\{(\mathbf{a}'m_T)^2\mathbf{c}'[\mathrm{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T]\} \\ &\quad +3E\{(\mathbf{a}'m_T)^2\mathbf{c}'[\mathrm{vech}(\hat{S}_T - \tilde{S}_T) \quad m_T]\} + o(\ell T^{-1}). \end{aligned} \tag{A.37}$$

The rest of the proof is similar to that of (A.17), and thus we will only show that

$$\begin{aligned} & T^{\frac{1}{2}}E\{(\mathbf{a}'m_T)^2\mathbf{c}'[\sum_{j=-\ell}^{\ell}\mathrm{vech}(\tilde{\Gamma}_j - \Gamma_j) \quad m_T]\} \\ =\; & \lim_{T\to\infty}(1/T)\sum_{\tau,t,s,k=-T+1}^{T-1} E\{\mathbf{a}'v_0\mathbf{a}'v_\tau\mathbf{c}'[\mathrm{vech}(v_t v_{t-k}' - \Gamma_k) \quad v_s]\}. \end{aligned} \tag{A.38}$$

It follows from arguments similar to the proof of (A.21) that

$$\begin{aligned} & T^{\frac{1}{2}}E\{(\mathbf{a}'m_T)^2\mathbf{c}'[\mathrm{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T]\} \\ =\; & (1/T)\sum_{s,t,u=-T+1}^{T-1}\sum_{j=-\ell}^{\ell}\omega_j(1-\tau_{s,t,u})E\{\mathbf{a}'v_0\mathbf{a}'v_s\mathbf{c}'[\mathrm{vech}(v_t v_{t-j} - \Gamma_j) \quad v_u]\} \\ =\; & (1/T)\sum_{s,t,u=-T+1}^{T-1}\sum_{j=-\ell}^{\ell}\omega_j E\{\mathbf{a}'v_0\mathbf{a}'v_s\mathbf{c}'[\mathrm{vech}(v_t v_{t-j} - \Gamma_j) \quad v_u]\} + O(T^{-1}) \\ =\; & (1/T)\sum_{s,t,u=-T+1}^{T-1}\sum_{j=-\ell}^{\ell} E\{\mathbf{a}'v_0\mathbf{a}'v_s\mathbf{c}'[\mathrm{vech}(v_t v_{t-j} - \Gamma_j) \quad v_u]\} + O(\ell^{-q}) \\ =\; & \lim_{T\to\infty}T^{-1}\sum_{\tau,t,s=-T+1}^{T-1}\sum_{j=-\ell}^{\ell} E\{\mathbf{a}'v_0\mathbf{a}'v_\tau\mathbf{c}'[\mathrm{vech}(v_t v_{t-j}' - \Gamma_j)' \quad v_s]\} + O(\ell^{-q}). \end{aligned} \tag{A.39}$$

By arguments similar to the proof of Lemma 1 of Andrews (1991, pp.850–851), one can show that the RHS of (A.39) equals the infinite sum of the product of two expectations plus some finite number. By the mixing inequality, it follows that the infinite sum of the product of two expectations is finite. Therefore, the RHS of (A.39) is well defined.

Lastly, we will show (A.20).

$$\begin{aligned} \kappa_4(g_T) - 3 \;&=\; 4E\{(\mathbf{a}'m_T)^3\mathbf{c}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\} \\ &\quad +4E\{(\mathbf{a}'m_T)^3\mathbf{e}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad \mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\} \\ &\quad +6E\left((\mathbf{a}'m_T)^2\{\mathbf{c}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\}^2\right) \\ &\quad -12E\{\mathbf{a}'m_T\mathbf{c}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\} \\ &\quad -12E\{\mathbf{a}'m_T\mathbf{e}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad \mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\} \\ &\quad -6E\{\mathbf{c}'[\mathrm{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\}^2 + O(\ell^{1/2}T^{-1}), \end{aligned} \tag{A.40}$$

from which the desired result follows by similar arguments. *Q.E.D.*

Lemma A.4:

$$\begin{aligned} & \psi_{g,T}(x) \\ = \; & \exp\left[-\frac{1}{2}\theta^2 + T^{-\frac{1}{2}}(\alpha_\infty(i\theta) - \frac{i\theta^3}{6}(\kappa_\infty - 3\alpha_\infty)) - \frac{\ell}{T}(\frac{\theta^2}{2}\gamma_\infty + \frac{\theta^4}{24}\zeta_\infty) + o(\frac{\ell}{T})\right] \end{aligned} \quad (A.41)$$

$$P(g_T \le x) \;=\; \Psi(x) + T^{-1/2}p_1(x) + (\ell/T)p_2(x) + o(\ell/T). \quad (A.42)$$

Proof of Lemma A.4: The proof of (A.41) follows from the standard arguments. (A.42) can be obtained by inverting (A.41). *Q.E.D.*

Lemma A.5: Following Götze and Künsch (1996), define a truncation function by

$$\tau(x) \;=\; T^\gamma x f(T^{-\gamma}\|x\|)/\|x\|$$

where $\gamma \in (2/r, 1/2)$ and $f \in C^\infty(0,\infty)$ satisfies (i) $f(x) = x$ for $x \le 1$; (ii) f is increasing; and (iii) $f(x) = 2$ for $x \ge 2$. Let $f_T^\dagger$ denote f_T with $\bar{R}_t \equiv (v_t', \tilde{v}_t, \mathrm{vec}(w_t)')$ replaced by

$$\bar{R}_t^\dagger = (v_t^{\dagger\prime}, \tilde{v}_t^{\dagger\prime}, \mathrm{vec}(w_t^\dagger)')' \;=\; \tau\left((v_t', \tilde{v}_t', \mathrm{vec}(w_t)')'\right).$$

Let $\Psi_T^\dagger$ and $\Psi_{g,T}^\dagger$ denote the Edgeworth expansions of $f_T^\dagger$ and $g_T^\dagger$, respectively. Let $\psi_{g,T}^\dagger(x)$ and $\tilde{\psi}_{g,T}^\dagger(x)$ denote the characteristic functions of $g_T^\dagger$ and $\Psi_{g,T}^\dagger$, respectively. Then

$$\sup_x |P(f_T \le x) - \Psi_T(x)| \le C\int_{|\theta|<T^{1-2/r}} |\psi_{g,T}^\dagger(\theta) - \tilde{\psi}_{g,T}^\dagger(\theta)||\theta|^{-1}d\theta + O(\ell^{-q}) + o(\ell T^{-1}). \quad (A.43)$$

Proof of Lemma A.5: First, we will show that

$$\sup_{-\infty<x<\infty} |P(f_T \le x) - \Psi_T(x)| \;=\; \sup_{-\infty<x<\infty} |P(f_T^\dagger \le x) - \Psi_T^\dagger(x)| + o(\ell T^{-1}). \quad (A.44)$$

Since

$$P\left(\max_{1\le t\le T}\|\bar{R}_t\| > T^\gamma\right) \;\le\; \sum_{t=1}^T P(\|\bar{R}_t\| > T^\gamma) \;=\; O(T^{1-\gamma r}), \quad (A.45)$$

it follows that

$$\sup_{-\infty<x<\infty}\left|P(f_T \le x) - P(f_T^\dagger \le x)\right| \;=\; O(T^{1-\gamma r}) \;=\; O(T^{-1}). \quad (A.46)$$

Then it follows from Lemma A.2 and (A.45) that

$$\begin{aligned} E\|m_T^\dagger - m_T\|^j &\le 2^j E[\|m_T\|^j I(\max_{1\le t\le T}\|R_t\| > T^\gamma)] \\ &\le 2^j (E\|m_T\|^{2j})^{1/2} P(\max_{1\le t\le T}\|R_t\| > T^\gamma)^{1/2} \\ &= o(T^{-1/2}) \end{aligned} \quad (A.47)$$

for $j \le r/2$. Similarly, we obtain that

$$E\|T^{1/2}[(G_T^\dagger - G_0^\dagger) - (G_T - G_0)]\|^j \;=\; o(T^{-1/2}), \quad (A.48)$$

$$E\|(T/\ell)^{1/2}[\mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) - \mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger)]\|^j \;=\; o(T^{-1/2}), \quad (A.49)$$

$$E\|(T/\ell)^{1/2}[\mathrm{vech}(\nabla\tilde{S}_T - \nabla\bar{S}_T) - \mathrm{vech}(\nabla\tilde{S}_T - \nabla\bar{S}_T)]\|^j \;=\; o(T^{-1/2}), \quad (A.50)$$

$$E\|T^{1/2}[\mathrm{vech}(\hat{S}_T^\dagger - \tilde{S}_T^\dagger) - \mathrm{vech}(\hat{S}_T^\dagger - \tilde{S}_T^\dagger)]\|^j \;=\; o(T^{-1/2}), \quad (A.51)$$

for $j \leq r/2$. Thus it follows from Lemma A.2, (A.45), (A.47)-(A.51) that

$$\sup_{-\infty<x<\infty} \left| \Psi_T(x) - \Psi_T^\dagger(x) \right| \;=\; o(\ell T^{-1}). \tag{A.52}$$

Therefore (A.44) follows from (A.46) and (A.52).

Next, we will show that

$$\sup_x |P(f_T^\dagger \leq x) - \Psi_T^\dagger(x)| = \sup_x |P(g_T^\dagger \leq x) - \Psi_{g,T}^\dagger(x)| + O(\ell^{-q}) + O(\ell^{3/2}T^{-3/2}). \tag{A.53}$$

Let

$$\begin{aligned} h_T^\dagger \;=\;& g_T^\dagger + \mathbf{c}'[\mathrm{vech}(\bar{S}_T^\dagger - S_0^\dagger) \quad m_T^\dagger] + \mathbf{d}'[(G_T - G_0) \quad \mathrm{vech}(\bar{S}_T^\dagger - S_0^\dagger) \quad m_T^\dagger] \\ &+\mathbf{e}'[\mathrm{vech}(\hat{S}_T^\dagger - \bar{S}_T^\dagger) \quad \mathrm{vech}(\bar{S}_T^\dagger - S_0^\dagger) \quad m_T^\dagger] \\ &+\mathbf{e}'[\mathrm{vech}(\bar{S}_T^\dagger - S_0^\dagger) \quad \mathrm{vech}(\hat{S}_T^\dagger - \bar{S}_T^\dagger) \quad m_T^\dagger] \\ &+\mathbf{e}'[\mathrm{vech}(\bar{S}_T^\dagger - S_0^\dagger) \quad \mathrm{vech}(\bar{S}_T^\dagger - S_0^\dagger) \quad m_T^\dagger], \end{aligned}$$

and let $\Psi_{h,T}(x)$ denote its Edgeworth expansion. Using the definition of Taylor series expansions, Lemma A.2 and Markov's inequality, $P(|f_T^\dagger - h_T^\dagger| > \ell^{3/2}T^{-3/2})$ can be made arbitrarily small. Thus we have

$$\sup_x |P(f_T^\dagger \leq x) - \Psi_T^\dagger(x)| \;=\; \sup_x |P(h_T^\dagger \leq x) - \Psi_{h,T}^\dagger(x)| + O(\ell^{3/2}T^{-3/2}). \tag{A.54}$$

Since the difference between the Edgeworth expansions of $g_T^\dagger$ and of $h_T^\dagger$ is $O(\bar{S}_T^\dagger - S_T^\dagger)$, it follows that

$$\sup_x |P(h_T^\dagger \leq x) - \Psi_{h,T}^\dagger(x)| \;=\; \sup_x |P(g_T^\dagger \leq x) - \Psi_{g,T}^\dagger(x)| + O(\ell^{-q}). \tag{A.55}$$

Therefore, (A.53) follows from (A.54) and (A.55).

Lastly, it follows from the so-called smoothing lemma (e.g., Proposition C1 of Fan and Linton, 1997) that

$$\sup_x |P(g_T^\dagger \leq x) - \Psi_{g,T}^\dagger(x)| \leq C \int_{|\theta|<T^{1-2/r}} |\psi_{g,T}^\dagger(\theta) - \tilde{\psi}_{g,T}^\dagger(\theta)||\theta|^{-1} d\theta + O(T^{-1+2/r}). \tag{A.56}$$

Therefore, Lemma A.5 follows from (A.44), (A.53) and (A.56) as $r > 12$. *Q.E.D.*

Lemma A.6: For $0 < \varepsilon < 1/6$,

$$\int_{|\theta|\leq T^\varepsilon} |\psi_{g,T}^\dagger(\theta) - \tilde{\psi}_{g,T}^\dagger(\theta)||\theta|^{-1} d\theta \;=\; o(\ell T^{-1}). \tag{A.57}$$

Proof of Lemma A.6: Write $g_T^\dagger$ as

$$\begin{aligned} g_T^\dagger \;=\;& \mathbf{a}'m_T^\dagger + \mathbf{b}'[(G_T^\dagger - G_0^\dagger) \quad m_T^\dagger] + \mathbf{c}'[\mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T) \quad m_T^\dagger] \\ &+\mathbf{c}'[\mathrm{vech}(\hat{S}_T^\dagger - \tilde{S}_T) \quad m_T^\dagger] + \mathbf{d}'[(G_T^\dagger - G_0^\dagger) \quad \mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) \quad m_T^\dagger] \\ &+\mathbf{d}'[(G_T^\dagger - G_0^\dagger) \quad \mathrm{vech}(\hat{S}_T^\dagger - \tilde{S}_T^\dagger) \quad m_T^\dagger] \\ &+\mathbf{e}'[\mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) \quad \mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) \quad m_T^\dagger] \\ &+\mathbf{e}'[\mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) \quad \mathrm{vech}(\hat{S}_T^\dagger - \tilde{S}_T^\dagger) \quad m_T^\dagger] \\ &+\mathbf{e}'[\mathrm{vech}(\hat{S}_T^\dagger - \tilde{S}_T^\dagger) \quad \mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) \quad m_T^\dagger] \\ &+\mathbf{e}'[\mathrm{vech}(\hat{S}_T^\dagger - \tilde{S}_T^\dagger) \quad \mathrm{vech}(\hat{S}_T^\dagger - \tilde{S}_T^\dagger) \quad m_T^\dagger] \\ \equiv\;& g_{T,1}^\dagger + g_{T,2}^\dagger + \dots + g_{T,10}^\dagger. \end{aligned}$$

Then a Taylor series expansion of $E(\exp(i\theta g_T^\dagger))$ around $g_{T,2}^\dagger + g_{T,3}^\dagger + ... + g_{T,10}^\dagger = 0$ yields

$$\begin{aligned} E(\exp(i\theta g_T^\dagger)) &= E(\exp(i\theta g_{T,1}^\dagger) + i\theta E[\exp(i\theta g_{T,1}^\dagger)(g_{T,2}^\dagger + g_{T,3}^\dagger + g_{T,4}^\dagger)] \\ &\quad + \frac{(i\theta)^2}{2} E[\exp(i\theta g_{T,1}^\dagger)(2g_{T,1}^\dagger g_{T,3}^\dagger + 2g_{T,1}^\dagger g_{T,7}^\dagger + g_{T,3}^{\dagger 2})] \\ &\quad + \frac{(i\theta)^3}{6} E[\exp(i\theta g_{T,1}^\dagger)(3g_{T,1}^{\dagger 2} g_{T,2}^\dagger + 3g_{T,1}^{\dagger 2} g_{T,3}^\dagger + 3g_{T,1}^{\dagger 2} g_{T,4}^\dagger)] \\ &\quad + \frac{(i\theta)^4}{24} E[\exp(i\theta g_{T,1}^\dagger)(4g_{T,1}^{\dagger 3} g_{T,3}^\dagger + 4g_{T,1}^{\dagger 3} g_{T,7}^\dagger + 6g_{T,1}^{\dagger 2} g_{T,3}^{\dagger 2})] \\ &\quad + O(\theta^4[E(g_{T,2}^{\dagger 4}) + E(g_{T,3}^{\dagger 4}) + ... + E(g_{T,10}^{\dagger 4})]). \end{aligned} \tag{A.58}$$

We will analyze each term on the RHS of (A.58) in turn. First, it follows from Lemma 3.33 of Götze and Hipp (1983) that

$$\begin{aligned} & E\left\{\exp(i\theta g_{T,1}^\dagger) - \left[1 + \frac{(i\theta)^3}{6} E(\mathbf{a}' m_T^\dagger)^3 + \frac{\theta^4}{24}(E(\mathbf{a}' m_T^\dagger)^4 - 3) - \frac{\theta^6}{72}(E(\mathbf{a}' m_T^\dagger)^3)^2\right] \exp(-\frac{\theta^2}{2})\right\} \\ = \; & O((1+|\theta|^9)\exp(-\theta^2)T^{-1-\varepsilon}). \end{aligned} \tag{A.59}$$

Second, let $\tilde{\psi}_X$ denote the multivariate expansion of $E(\exp(ic' T^{-1/2} \sum_{t=1}^T X_t^\dagger))$ where $X_t^\dagger = (\mathbf{a}' v_t^\dagger, v_t^{\dagger\prime}, (w_t^\dagger - G_0^\dagger)')'$. Then an application of Lemma 3.33 of Götze and Hipp (1983) with $\vartheta = (\theta, 0, ..., 0)'$ yields

$$\begin{aligned} & |E\{\exp(i\theta g_{T,1}^\dagger)[i\theta g_{T,2}^\dagger + \frac{(i\theta)^3}{2} g_{T,1}^{\dagger 2} g_{T,2}^\dagger]\} \\ & - \left((i\theta - \frac{(i\theta)^3}{2}) E\{\mathbf{b}'[(G_T^\dagger - G_0^\dagger) \quad m_T^\dagger]\} + \frac{(i\theta)^3}{2} E\{(\mathbf{a}' m_T^\dagger)^2 \mathbf{b}'[(G_T^\dagger - G_0^\dagger) \quad m_T^\dagger]\}\right) \exp(-\frac{\theta^2}{2})| \\ \leq \; & T^{-1/2} \sum_\alpha |c_\alpha| |D^\alpha [E(\exp(i\vartheta' T^{-1/2} \sum_{t=1}^T X_{2t}^\dagger)) - \tilde{\psi}_X]| \\ = \; & O((1+|\theta|^8 + |\theta|^{10})\exp(-\theta^2)T^{-1-\varepsilon}), \end{aligned} \tag{A.60}$$

where c_α are the corresponding elements of a, b and G_0.

Third, we will show that

$$\begin{aligned} & |i\theta E[\exp(i\theta g_{T,1}^\dagger)[i\theta g_{T,3}^\dagger + (i\theta)^2 g_{T,1}^\dagger g_{T,3}^\dagger + \frac{(i\theta)^3}{2} g_{T,1}^{\dagger 2} g_{T,3}^\dagger + \frac{(i\theta)^4}{6} g_{T,1}^{\dagger 3} g_{T,3}^\dagger]] \\ - \; & \Big((i\theta - \frac{1}{2}(i\theta)^3) E\{\mathbf{c}'[\mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) \quad m_T^\dagger]\} + (i\theta)^2 E\{\mathbf{a}' m_T^\dagger \mathbf{c}'[\mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) \quad m_T^\dagger]\} \\ & + \frac{(i\theta)^3}{2} E\{(\mathbf{a}' m_T^\dagger)^2 \mathbf{c}'[\mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) \quad m_T^\dagger]\} \\ & + \frac{(i\theta)^4}{6} E\{(\mathbf{a}' m_T^\dagger)^3 \mathbf{c}'[\mathrm{vech}(\tilde{S}_T^\dagger - \bar{S}_T^\dagger) \quad m_T^\dagger]\}\Big) \exp(-\frac{1}{2}\theta^2)| \\ = \; & O((1+|\theta|^6)\exp(-\theta^2)\ell T^{-1-\varepsilon}). \end{aligned} \tag{A.61}$$

Note that the first term of (A.61) can be written as a weighted sum of

$$E\{\exp(i\theta g_{T,1}^\dagger)[i\theta + (i\theta)^2 \mathbf{a}' m_T^\dagger + \frac{(i\theta)^3}{2}(\mathbf{a}' m_T^\dagger)^2 + \frac{(i\theta)^4}{6}(\mathbf{a}' m_T^\dagger)^3]\mathbf{c}'[\mathrm{vech}(\tilde{\Gamma}_j^\dagger - \Gamma_j^\dagger) \quad m_T^\dagger]\} \tag{A.62}$$

and that the rest of the terms can be written as a weighted sum of

$$E\{[i\theta - \frac{1}{2}(i\theta)^3 + (i\theta)^2 \mathbf{a}' m_T^\dagger + \frac{(i\theta)^3}{2}(\mathbf{a}' m_T^\dagger)^2 + \frac{(i\theta)^4}{6}(\mathbf{a}' m_T^\dagger)^3]\mathbf{c}'[\mathrm{vech}(\tilde{\Gamma}_j^\dagger - \Gamma_j^\dagger) \quad m_T^\dagger]\} \exp(-\frac{\theta^2}{2}) \tag{A.63}$$

We will apply Lemma 3.33 of Götze and Hipp (1983) to (A.62) and (A.63). Let $\tilde{\psi}_Y$ denote the multivariate expansion of $E(\exp(i\vartheta' T^{-1/2}\sum_{t=1}^{T} Y_t^\dagger))$ where $\vartheta = (\theta, 0, ..., 0)$ and

$$Y_t^\dagger = (\mathbf{a}' m_T^\dagger, m_T^{\dagger\prime}, T^{-1/2}\sum_{t=1}^{T}\text{vech}[v_t v_{t-j}' - E(v_t v_{t-j}')]')'.$$

Then the difference between (A.61) and (A.62) are bounded by

$$T^{-1/2}\sum_{\alpha}|c_\alpha| D^\alpha \left| E(\exp(i\vartheta' T^{-1/2}\sum_{t=1}^{T} Y_t^\dagger)) - \tilde{\psi}_Y \right| = O((1+|\theta|^6)\exp(-\theta^2)T^{-1-\varepsilon}), \quad \text{(A.64)}$$

where c_α are the corresponding linear combinations of a and c. Thus (A.61) follows.

Fourth, by arguments analogous to the proof of (A.61), one can show that

$$\begin{aligned}
& |E[\exp(i\theta g_{T,1}^\dagger)(i\theta g_{T,4}^\dagger + \frac{(i\theta)^3}{2} g_{T,1}^{\dagger 2} g_{T,4}^\dagger)] \\
- \ & ((i\theta - \frac{1}{2}(i\theta)^3) E\{\mathbf{c}'[\text{vech}(\hat{S}_T^\dagger - \tilde{S}_T^\dagger) \quad m_T^\dagger]\} + (i\theta)^3 E\{(\mathbf{a}' m_T^\dagger)^2 \mathbf{c}'[\text{vech}(\hat{S}_T^\dagger - \tilde{S}_T^\dagger) \quad m_T^\dagger]\}) \\
& \times \exp(-\frac{\theta^2}{2})| \\
= \ & O((1+|\theta|^6)\exp(-\theta^2)\ell T^{-1-\varepsilon}), \qquad \text{(A.65)}
\end{aligned}$$

and

$$\begin{aligned}
& |E[\exp(i\theta g_{T,1}^\dagger)[\frac{(i\theta)^2}{2}(2 g_{T,1}^\dagger g_{T,7}^\dagger + g_{T,3}^2) + \frac{(i\theta)^4}{6} g_{T,1}^{\dagger 3} g_{T,7}^\dagger + \frac{(i\theta)^4}{4} g_{T,1}^{\dagger 2} g_{T,3}^{\dagger 2}] \\
- \ & (\frac{(i\theta)^2}{2}(2E\{\mathbf{a}' m_T \mathbf{c}'[\text{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T]\} + E\{\mathbf{c}'[\text{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\}^2) \\
& + \frac{(i\theta)^4}{6} E\{\mathbf{a}' m_T^\dagger \mathbf{e}'[\text{vech}(\hat{S}_T^\dagger - \bar{S}_T^\dagger) \quad \text{vech}(\hat{S}_T^\dagger - \bar{S}_T^\dagger) \quad m_T^\dagger]\} \\
& + \frac{(i\theta)^4}{6} E\{\mathbf{c}'[\text{vech}(\hat{S}_T - \bar{S}_T) \quad m_T]\}^2)\exp(-\frac{\theta^2}{2})| \\
= \ & O((1+|\theta|^6)\exp(-\theta^2)\ell^2 T^{-3/2-\varepsilon}). \qquad \text{(A.66)}
\end{aligned}$$

Lastly, it follows from Lemma A.2 that

$$\theta^4[E(g_{T,2}^{\dagger 4}) + E(g_{T,3}^{\dagger 4}) + \ldots + E(g_{T,4}^{\dagger 4})] \ = \ O(\theta^4 \ell^2 T^{-2}). \qquad \text{(A.67)}$$

Combining and integrating (A.59), (A.60), (A.61), (A.65), (A.66) and (A.67) produces the desired result. *Q.E.D.*

Lemma A.7:

$$\int_{T^\varepsilon < |\theta| < T^{1-2/r}} |\psi_{g,T}^\dagger(\theta) - \tilde{\psi}_{g,T}^\dagger(\theta)||\theta|^{-1} d\theta = o(\ell T^{-1}). \qquad \text{(A.68)}$$

Proof of Lemma A.7: We closely follow the proof of Götze and Hipp (1996, pp.1927–1930). To simplify the notation, we will omit the superscript †. Let $m = M\log T$ for some $M > 0$. Let $N = [(T/\theta^2 + 1)m^2]$ for $T^\varepsilon < |\theta| < T^{1-2/r}$. Then $m \le N \le T$ for sufficiently large T. Define

$$m_N = T^{-1/2}\sum_{t=1}^{N} v_t, \qquad m_{T-N} = T^{-1/2}\sum_{t=N+1}^{T} v_t,$$

$$G_N - E(G_N) = (1/T)\sum_{t=1}^{N}(w_t - E(w_t)), \qquad G_{T-N} - E(G_{T-N}) = (1/T)\sum_{t=N+1}^{T}(w_t - E(w_t)),$$

$$\tilde{S}_N - \bar{S}_N = \sum_{j=-\ell}^{\ell} \omega_j(\tilde{\Gamma}_{j,N} - \Gamma_j), \qquad \tilde{S}_{T-N} - \bar{S}_{T-N} = \sum_{j=-\ell}^{\ell} \omega_j(\tilde{\Gamma}_{j,T-N} - \Gamma_j),$$

$$\hat{S}_N - \tilde{S}_N = \sum_{j=-\ell}^{\ell} \omega_j(\hat{\Gamma}_{j,N} - \tilde{\Gamma}_{j,N}), \qquad \hat{S}_{T-N} - \tilde{S}_{T-N} = \sum_{j=-\ell}^{\ell} \omega_j(\hat{\Gamma}_{j,T-N} - \tilde{\Gamma}_{j,T-N})$$

so that

$$\begin{aligned} m_T &= m_N + m_{T-N}, \\ G_T - G_0 &= G_N - E(G_N) + G_{T-N} - E(G_{T-N}), \\ \tilde{S}_T - \bar{S}_T &= \tilde{S}_N - \bar{S}_N + \tilde{S}_{T-N} - \bar{S}_{T-N}, \\ \hat{S}_T - \tilde{S}_T &= \hat{S}_N - \tilde{S}_N + \hat{S}_{T-N} - \tilde{S}_{T-N}. \end{aligned}$$

Write

$$g_T = \mathbf{a}' m_T + Q(m_T, G_T, \hat{S}_T, \tilde{S}_T, \bar{S}_T).$$

Then a Taylor series expansion of Q around $v_t = 0$ and $w_t = 0$ for $t = 1, 2, ..., N$ yields

$$\begin{aligned} & E\exp(i\theta g_T) \\ = \; & E[\exp(i\theta \mathbf{a}' m_T + i\theta Q(m_{T-N}, G_{T-N}, \hat{S}_{T-N}, \tilde{S}_{T-N}, \bar{S}_{T-N}) \\ & \times \sum_{\alpha,\beta} v^\mu w^\nu Q_{\mu\nu}(m_{T-N}, G_{T-N}, \hat{S}_{T-N}, \tilde{S}_{T-N}, \bar{S}_{T-N})] \\ & + O(|\theta|^r E|Q(m_T, G_T, \hat{S}_T, \tilde{S}_T, \bar{S}_T) - Q(m_{T-N}, G_{T-N}, \hat{S}_{T-N}, \tilde{S}_{T-N}, \bar{S}_{T-N})|^r) \end{aligned} \quad \text{(A.69)}$$

where the power is element-by-element and the indices satisfy

$$\mu \;=\; (\mu_1, ..., \mu_{N+\ell-1}, 0, ..., 0), \;\; \nu \;=\; (\nu_1, ..., \nu_N, 0, ..., 0), \;\; |\mu| + |\nu| \;\leq\; 5(r-1).$$

First, we will consider the expansion terms in (A.69). Let

$$\begin{aligned} \{j_1^0, ..., j_{5(r-1)}^0\} &= \{j : \mu_j \text{ or } \nu_j > 0\}, \\ I &= \{j \in \{1, ..., N-m\} : |j - j_k^0| \geq 3m, k = 1, ..., 5(r-1)\}, \\ j_{k+1} &= \inf\{j \in I : j \geq j_k + 7m\} \end{aligned}$$

and $j_1 = \inf I$. Let s denote the smallest integer for which the inf is undefined. Let

$$\begin{aligned} A_k &= \prod\{\exp(i\theta T^{-1/2}\mathbf{a}' v_t : j \in I, |j - j_k| \leq m\}, \; k = 1, ..., s, \\ B_k &= \prod\{\exp(i\theta T^{-1/2}\mathbf{a}' v_t : j \in I, j_k + m + 1 \leq j \leq j_{k+1} - m - 1\}, k = 1, ..., s-1, \\ B_s &= \prod\{\exp(i\theta T^{-1/2}\mathbf{a}' v_t : j \in I, j \geq j_s - m - 1\}, \\ R &= \prod_{j \notin I} \exp(i\theta T^{-1/2}\mathbf{a}' v_t) \exp(i\theta Q(m_{T-N}, G_{T-N}, \hat{S}_{T-N}, \tilde{S}_{T-N}, \bar{S}_{T-N})) v^\mu w^\nu Q_{\mu\nu}. \end{aligned}$$

Then we can write

$$\begin{aligned} & E[\exp(i\theta \mathbf{a}' m_T + i\theta Q(m_{T-N}, G_{T-N}, \hat{S}_{T-N}, \tilde{S}_{T-N}, \bar{S}_{T-N}) \\ & \quad \times \sum_{\alpha,\beta} v^\mu w^\nu Q_{\mu\nu}(m_{T-N}, G_{T-N}, \hat{S}_{T-N}, \tilde{S}_{T-N}, \bar{S}_{T-N})] \;=\; \prod_{k=1}^{s} A_k B_k R. \end{aligned} \quad \text{(A.70)}$$

Note that $|A_k| \leq 1$, $|B_k| \leq 1$, $|R| \leq T^{\gamma(s-1)r}$, and that A_k, B_k and R are measurable with respect to $\mathcal{F}_{j_k-2m}^{j_k+2m}$, $\mathcal{F}_{j_k-1}^{j_k+1}$, $\{\mathcal{F}_l : \exists j \notin I, \; |l - j| \leq m\}$, respectively. By Assumption 1(d), it follows that

$$
\begin{aligned}
& |E[\prod_{k=1}^{s} A_k B_k R] - E[\prod_{k=1}^{s} E(A_k|\mathcal{F}_j : |j-j_k| \le 3m) B_k R]| \\
\le & \sum_{j=1}^{s} |E[\prod_{k=1}^{j-1} A_k B_k (A_j - E(A_j|\mathcal{F}_j : |j-j_k| \le 3m)) \prod_{l=j+1}^{s} E(A_l|\mathcal{F}_j, |j-j_l| \le 3m) B_l| \\
= & \sum_{j=1}^{s} |E[\prod_{k=1}^{j-1} A_k B_k (E(A_j|\mathcal{F}_{-\infty}^{j_k-1} \cup \mathcal{F}_{j_k+1}^{\infty}) - E(A_j|\mathcal{F}_j : |j-j_k| \le 3m)) \\
& \times \prod_{l=j+1}^{s} E(A_l|\mathcal{F}_j : |j-j_l| \le 3m) B_l| \\
= & O(T^{c_1}\exp(-dm)) \;=\; o(T^{-c_2}) \qquad \text{(A.71)}
\end{aligned}
$$

for any arbitrary $c_2 > 0$ by choosing sufficiently large M. By the mixing inequality of Hall and Heyde (1980), we obtain

$$
\begin{aligned}
& |E[R \prod_{k=1}^{s} E(A_k|\mathcal{F}_j : |j-j_k| \le 3m) B_k]| \\
\le & T^{c_3} E \prod_{j=1}^{s} |E(A_k|\mathcal{F}_k : 0 < |j-j_k| \le 3m)| \\
& + T^{c_3} \prod_{j=1}^{s} E|E(A_k|\mathcal{F}_j : 0 < |j-j_k| \le 3m)| + 4T^{c_3}(q/d)\exp(-dm) \qquad \text{(A.72)}
\end{aligned}
$$

for some $c_3 > 0$. For $|\theta| \ge d$, we have $E|E(A_k|\mathcal{F}_j, j \ne j_k)| \le \exp(-d)$. Thus by Lemma 3.2 of Götze and Hipp (1983) and Assumption 1(d), it follows that

$$
\begin{aligned}
E|E(A_k|\mathcal{F}_j, |j-j_k| \le 3m)| & \le E|E(A_k|\mathcal{F}_j : |j-j_k| \ne 0)| + O(T^{c}\exp(-dm)) \\
& \le \max(\exp(-d\theta^2/T), \exp(-d)) + O(T^{c_3}\exp(-dm)) \qquad \text{(A.73)}
\end{aligned}
$$

$$
E[\prod_{k=1}^{s} A_k B_k R] = O(T^{-c}) \qquad \text{(A.74)}
$$

for arbitrary $c > 0$ by choosing sufficiently large M.

Next, consider the remainder term in (A.69). It follows from Lemma A.2 that

$$
\begin{aligned}
E|m_N|^r &= O((N/T)^r), && \text{(A.75)} \\
E|T^{1/2}(G_N - G_0)|^r &= O((N/T)^r), && \text{(A.76)} \\
E|(T/\ell)^{1/2}\mathrm{vech}(\tilde{S}_N - \bar{S}_N)|^r &= O((N/T)^{r/2}), && \text{(A.77)} \\
E|(T/\ell)^{1/2}\mathrm{vech}(\nabla\tilde{S}_N - \nabla\bar{S}_N)|^r &= O((N/T)^{r/2}), && \text{(A.78)} \\
E|T^{1/2}\mathrm{vech}(\hat{S}_N - \tilde{S}_N)|^r &= O((N/T)^{r/2}). && \text{(A.79)}
\end{aligned}
$$

Using the definition of N and $\varepsilon r > 2$, we obtain that

$$
\begin{aligned}
& |\theta|^r E|Q(m_T, G_T, \hat{S}_T, \tilde{S}_T, \bar{S}_T) - Q(m_{T-N}, G_{T-N}, \hat{S}_{T-N}, \tilde{S}_{T-N}, \bar{S}_{T-N})|^r \\
= & O(\ell^{r/2}|\theta|^r N^{r/2} T^{-r}) \\
= & \begin{cases} O(\ell^{r/2} m^r T^{-r/2}) & \text{for } |\theta| \le T^{1/2} \\ O(|\theta|^r \ell^{r/2} m^r T^{-r}) & \text{for } T^{1/2} < |\theta| \le \ell^{-1/2} T^{1-\varepsilon} \end{cases} \\
= & o(\ell T^{-1}). \qquad \text{(A.80)}
\end{aligned}
$$

Lastly, it follows from (A.69), (A.71)-(A.73) and (A.80) that

$$\begin{aligned} E\exp(i\theta g_T) &= T^c \max(\exp(-d\theta^2/T), \exp(-d))^{N/M} + O(T^c \exp(-dm)) + o(\ell T^{-1}) \\ &= o(\ell T^{-1}) \end{aligned} \quad \text{(A.81)}$$

for $s \geq N/M$ and sufficiently large M, which completes the proof. *Q.E.D.*

Lemma B.1: For $1 \leq s \leq r/2$,

$$E^*[\|\mathrm{vec}(F_{N_j})\|^s] - E\left\{E^*[\|\mathrm{vec}(F_{N_j})\|^s]\right\} = O_p(b^{-1/2}), \quad \text{(A.82)}$$

$$E^*[\|B_{N_j}\|^s] - E\left\{E^*[\|B_{N_j}\|^s]\right\} = O_p(b^{-1/2}). \quad \text{(A.83)}$$

Proof of Lemma B.1: First we will prove (A.82). We can write the LHS of (A.82) as

$$(1/(T-\ell+1)) \sum_{t=0}^{T-\ell} \|\mathrm{vec}(F_j)\|^s - E[\|\mathrm{vec}(F_j)\|^s] = (1/(T-\ell+1))(1/\ell) \sum_{\nu=1}^{\ell} f_{s,\nu}, \quad \text{(A.84)}$$

where

$$f_{s,\nu} = (1/b) \sum_{\mu=0}^{b-1} \left(\|\mathrm{vec}(F_{\mu\ell+\nu})\|^s - E(\|\mathrm{vec}(F_{\mu\ell+\nu})\|^s)\right).$$

Note that $\{\mathrm{vec}(F_{\mu\ell+\nu})\}_{\mu=0}^{b-1}$ is a triangular array of strong mixing sequence with mixing coefficients given by $\{\alpha_{\mu\ell}\}$ where α_m is the mixing coefficient of the original variables. So is $\|\mathrm{vec}(F_{\mu\ell+\nu})\|^s$. Thus it follows that

$$f_{s,\nu} = O_p(b^{-1/2}). \quad \text{(A.85)}$$

Since the decay rate of the mixing coefficients is uniform in ν, (A.85) also holds uniformly in ν. Hence (A.82) follows from (A.84) and (A.85).

Next we will prove (A.83). Note that the LHS of (A.83) is bounded by

$$O\left((1/(T-\ell+1)) \sum_{t=0}^{T-\ell} \|\tilde{B}_t\|^s - E\|\tilde{B}_t\|^s\right) \quad \text{(A.86)}$$

$$+O\left((1/(T-\ell+1)) \sum_{t=0}^{T-\ell} \|\hat{B}_t\|^s - \|\tilde{B}_t\|^s - E(\|\hat{B}_t\|^s - \|\tilde{B}_t\|^s)\right) \quad \text{(A.87)}$$

$$+O\left(\|\mu_T^*\|^s - E(\|\mu_T^*\|^s)\right), \quad \text{(A.88)}$$

where $\hat{B}_t = \ell^{-1/2} \sum_{j=1}^{\ell} \tilde{v}_{t+j}$ and $\tilde{B}_t = \ell^{-1/2} \sum_{j=1}^{\ell} v_{t+j}$. First, the proof that (A.86) is $O_p(b^{-1/2})$ is analogous to the proof of (A.82) and thus is omitted. Second, we will prove that (A.87) is $O_p(b^{-1/2})$. A Taylor series expansion yields

$$\|\hat{B}_t\| - \|\tilde{B}_t\| = s\|\tilde{B}_t\|^{s-2} F_t \ell^{\frac{1}{2}} (\tilde{\beta}_T - \beta_0). \quad \text{(A.89)}$$

Thus we have

$$(1/(T-\ell+1)) \sum_{t=0}^{T-\ell} \|\hat{B}_t\|^s - \|\tilde{B}_t\|^s = (1/(T-\ell+1)) \sum_{t=0}^{T-\ell} s\|\tilde{B}_t\|^{s-2} F_t \ell^{\frac{1}{2}} (\tilde{\beta}_T - \beta_0). \quad \text{(A.90)}$$

By using arguments analogous to the one used in the proof of (A.82), it follows from the ergodic theorem that

$$(1/(T-\ell+1)) \sum_{t=0}^{T-\ell} s\|\tilde{B}_t\|^{s-2} F_t = O_{as}(1). \quad \text{(A.91)}$$

Thus it follows from Assumption 1(i) that

$$(1/(T-\ell+1))\sum_{t=0}^{T-\ell}(\|\hat{B}_t\|^s - \|\tilde{B}_t\|^s) = O_p(b^{-1/2}). \tag{A.92}$$

Similarly we obtain

$$(1/(T-\ell+1))\sum_{t=0}^{T-\ell}E(\|\hat{B}_t\|^s - \|\tilde{B}_t\|^s) = O(b^{-1/2}). \tag{A.93}$$

Hence it follows from (A.92) and (A.93) that (A.87) is $O_p(b^{-1/2})$. Third, we will prove that (A.88) is $O_p(b^{-1/2})$. We can write μ_T^* as

$$\begin{aligned}
\mu_T^* &= (1/(T-\ell+1))\sum_{t=0}^{T-\ell}\hat{B}_t \\
&= (1/(T-\ell+1))\sum_{t=0}^{T-\ell}\tilde{B}_t + (1/(T-\ell+1))\sum_{t=0}^{T-\ell}F_t\ell^{1/2}(\tilde{\beta}_T-\beta_0) \\
&\quad +(1/(T-\ell+1))\sum_{t=0}^{T-\ell}H_t\ell^{1/2}(\tilde{\beta}_T-\beta_0)^2.
\end{aligned} \tag{A.94}$$

Thus we obtain

$$\begin{aligned}
&\|\mu_T^*\|^s - E\|\mu_T^*\|^s \\
= \ &O((1/(T-\ell+1))\sum_{t=0}^{T-\ell}\|\tilde{B}_t\|^s - E\|\tilde{B}_t\|^s) \\
&+O((1/(T-\ell+1))\sum_{t=0}^{T-\ell}\|F_t\ell^{1/2}(\tilde{\beta}_T-\beta_0)\|^s - E\|F_t\ell^{1/2}(\tilde{\beta}_T-\beta_0)\|^s).
\end{aligned} \tag{A.95}$$

The rest of the proof is analogous to the proofs of (A.86) and (A.87). Therefore (A.83) follows from (A.86), (A.87) and (A.88). *Q.E.D.*

Lemma B.2: Let $G_0^* = E^*(G_T^*)$ and B_T^* and C_T^* denote the bootstrap version of B and C in Lemma A.1 with S_0 replaced by S_T^*, respectively. Then

$$G_0^* = G_0 + O_p(T^{-1/2}), \tag{A.96}$$
$$S_T^* = S + O(\ell^{-1}) + O_p(b^{-1/2}). \tag{A.97}$$

Proof of Lemma B.2: First, we will prove (A.96).

$$\begin{aligned}
G_0^* &= E^*[G_T^*] = E^*[\frac{1}{b}\sum_{k=1}^{b}F_{N_k}] = E^*[F_{N_1}] \\
&= \frac{1}{T-\ell+1}\sum_{t=0}^{T-\ell}F_t = \frac{1}{T-\ell+1}\sum_{t=0}^{T-\ell}\frac{1}{\ell}\sum_{i=1}^{\ell}w_{t+i} \\
&= \frac{1}{T}\sum_{t=1}^{T}w_t + O_p(\ell T^{-1}) = G_T + O_p(\ell T^{-1}).
\end{aligned}$$

Therefore, (A.96 follows from $G_T - G_0 = O_p(T^{-1/2})$.

Next, we will prove (A.97). By definition, it follows that

$$
\begin{aligned}
S_T^* &\equiv \mathrm{Var}^*(m_T^*) = \mathrm{Var}^*\Big(\frac{1}{\sqrt{b}}\sum_{k=1}^{b} B_{N_k}\Big) && \text{(A.98)}\\
&= E^*\left(\frac{1}{\sqrt{b}}\sum_{k=1}^{b} B_{N_k} - \sqrt{b}E^*(B_{N_1})\right)\left(\frac{1}{\sqrt{b}}\sum_{k=1}^{b} B_{N_k} - \sqrt{b}E^*(B_{N_1})\right)' && \text{(A.99)}\\
&= E^*\left(\frac{1}{\sqrt{b}}\sum_{k=1}^{b} B_{N_k}\right)\left(\frac{1}{\sqrt{b}}\sum_{k=1}^{b} B_{N_k}\right)' && \text{(A.100)}\\
&= \frac{1}{b}\sum_{k=1}^{b} E^*(B_{N_k}B_{N_k}^{'}) = E^*(B_{N_1}B_{N_1}^{'}) = \frac{1}{T-\ell+1}\sum_{t=0}^{T-\ell} B_t B_t^{'}. && \text{(A.101)}
\end{aligned}
$$

It follows from Lemma B.1 that

$$
S_T^* - E[S_T^*] = O_p(b^{-1/2}). \tag{A.102}
$$

Since $\mu_T^* = O_p(T^{-1/2})$, we have

$$
E[S_T^*] = E[B_t B_t^{'}] = \sum_{j=-\ell}^{\ell}(1-|j|/\ell)E[v_0 v_{-j}'] = S + O(\ell^{-1}). \tag{A.103}
$$

Thus (A.97) follows from (A.101),(A.102) and (A.103). *Q.E.D.*

Lemma B.3: Let

$$
\begin{aligned}
\alpha_T^* &= T^{1/2}\kappa_1^*(g_T^*),\\
\gamma_T^* &= (T/\ell)(\kappa_2^*(g_T^*)-1) = (T/\ell)(E^*(g_T^{*2}) - [E^*(g_T^*)]^2 - 1),\\
\kappa_T^* &= T^{1/2}E^*(g_T^{*3}) = T^{1/2}\{\kappa_3^*(g_T^*) + 3E^*(g_T^{*2})E^*(g_T^*) - 2[E^*(g_T^*)]^3\},\\
\zeta_T^* &= (T/\ell)(\kappa_4^*(g_T^*)-3).
\end{aligned}
$$

Then

$$
\begin{aligned}
\alpha_T^* &= \alpha_\infty + T^{1/2}\mathbf{b}^{*\prime}E^*[(G_T^* - G_0^*)\quad m_T^*] + T^{1/2}\mathbf{c}^{*\prime}E^*[\mathrm{vech}(\tilde{S}_T^* - S_T^*)\quad m_T^*]\\
&\quad + T^{1/2}\mathbf{c}^{*'}E^*[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*)\quad m_T^*] + o_p^*(\ell T^{-1/2})\\
&= \alpha_\infty + O_p(\ell^{-1}) + O_p(b^{-1/2}) + o_p^*(\ell T^{-1/2}), && \text{(A.104)}\\
\gamma_T^* &= \gamma_\infty + 2(T/\ell)E^*\{\mathbf{a}^{*\prime}m_T^*\mathbf{b}^{*\prime}[(G_T^* - G_0^*)\quad m_T^*]\}\\
&\quad + 2(T/\ell)E^*\{\mathbf{a}^{*\prime}m_T^*\mathbf{c}^{*\prime}[\mathrm{vech}(\tilde{S}_T^* - S_T^*)\quad m_T^*]\}\\
&\quad + 2(T/\ell)E^*\{\mathbf{a}^{*\prime}m_T^*\mathbf{e}^{*\prime}[\mathrm{vech}(\tilde{S}_T^* - S_T^*)\quad \mathrm{vech}(\tilde{S}_T^* - S_T^*)\quad m_T^*]\}\\
&\quad + (T/\ell)E^*\{\mathbf{c}^{*\prime}[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*)\quad m_T^*]\}^2 + o_p^*(1),\\
&= \gamma_\infty + o_p(1) + o_p^*(1), && \text{(A.105)}\\
\kappa_T^* &= \kappa_\infty + T^{1/2}E^*[(\mathbf{a}^{*\prime}m_T^*)^3] + 3T^{1/2}E^*\{(\mathbf{a}^{*\prime}m_T^*)^2\mathbf{b}^{*\prime}[(G_T - G_0)'\quad m_T]\}\\
&\quad + 3T^{1/2}E^*\{(\mathbf{a}^{*\prime}m_T^*)^2\mathbf{c}^{*\prime}[\mathrm{vech}(\tilde{S}_T^* - S_T^*)\quad m_T^*]\}\\
&\quad + 3T^{1/2}E^*\{(\mathbf{a}^{*\prime}m_T^*)^2\mathbf{c}^{*\prime}[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*)\quad m_T^*]\} + o_p^*(\ell T^{-1/2})\\
&= \kappa_\infty + O_p(\ell^{-1/2}) + O_p(b^{-1/2}) + o_p^*(\ell T^{-1/2}), && \text{(A.106)}\\
\zeta_T^* &= \zeta_\infty + 4(T/\ell)E^*\{(\mathbf{a}^{*\prime}m_T^*)^3\mathbf{c}^{*\prime}[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*)\quad m_T^*]\}\\
&\quad + 4(T/\ell)E^*\{(\mathbf{a}^{*\prime}m_T^*)^3 e'[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*)\quad \mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*)\quad m_T^*]\}\\
&\quad + 6(T/\ell)E^*\left((\mathbf{a}^{*\prime}m_T^*)^2\{\mathbf{c}^{*\prime}[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*)\quad m_T^*]\}^2\right)
\end{aligned}
$$

$$
\begin{aligned}
& -12(T/\ell)E^*\{\mathbf{a}^{*\prime} m_T^* \mathbf{c}^{*\prime}[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*) \quad m_T^*]\} \\
& -12(T/\ell)E^*\{\mathbf{a}^{*\prime} m_T^* \mathbf{e}^{*\prime}[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*) \quad \mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*) \quad m_T^*]\} \\
& -6(T/\ell)E^*\{\mathbf{c}^{*\prime}[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*) \quad m_T^*]\}^2 + o_p^*(1) \\
= \; & \zeta_\infty + o_p(1) + o_p^*(1), \qquad \text{(A.107)}
\end{aligned}
$$

where α_∞, γ_∞, κ_∞, and ζ_∞ are defined in Lemma A.4.

Proof of Lemma B.3: The first equalities in (A.104)–(A.107) follow from Lemmas B.1 and B.2. Thus we will show that the second equalities hold in the rest of the proof.

Part (a): Proof of (A.104). First, we introduce some notation for the proof. Let

$$
\begin{aligned}
\alpha_{1T}^* &= T^{1/2}\mathbf{b}^{*\prime} E^*[(G_T^* - G_0^*) \quad m_T^*], \\
\alpha_{2T}^* &= T^{1/2}\mathbf{c}^{*\prime} E^*[\mathrm{vech}(\tilde{S}_T^* - S_T^*) \quad m_T^*], \\
\alpha_{3T}^* &= T^{1/2}\mathbf{c}^{*'} E^*[\mathrm{vech}(\hat{S}_T^* - \tilde{S}_T^*) \quad m_T^*], \\
\alpha_{1T} &= T^{1/2}\mathbf{b}' E[(G_T - G_0) \quad m_T], \\
\alpha_{2T} &= T^{1/2}\mathbf{c}' E[\mathrm{vech}(\tilde{S}_T - \bar{S}_T) \quad m_T], \\
\alpha_{3T} &= T^{1/2}\mathbf{c}' E[\mathrm{vech}(\hat{S}_T - \tilde{S}_T) \quad m_T], \\
\alpha_{1\infty} &= \mathbf{b}' \sum_{i=-\infty}^{\infty} E[w_0 \quad v_i], \\
\alpha_{2\infty} &= \mathbf{c}' \sum_{i,j=-\infty}^{\infty} E[\mathrm{vech}(v_0 v_i') \quad v_j], \\
\alpha_{3\infty} &= \mathbf{c}' \sum_{i=-\infty}^{\infty} E\{\mathrm{vech}[\nabla \bar{S}(E(w_0)' V E(w_0))^{-1} E(w_0)' V v_0] \quad v_i\}.
\end{aligned}
$$

Next, we will prove that

$$
\begin{aligned}
\alpha_{1T}^* - \alpha_{1\infty} &= O_p(\ell^{-1}) + O_p(\ell^{1/2} b^{-1/2}), && \text{(A.108)} \\
\alpha_{2T}^* - \alpha_{2\infty} &= O_p(\ell^{-1}) + O_p(\ell^{1/2} b^{-1/2}), && \text{(A.109)} \\
\alpha_{3T}^* - \alpha_{3\infty} &= O_p(\ell^{-1}) + O_p(\ell^{1/2} b^{-1/2}). && \text{(A.110)}
\end{aligned}
$$

Since $\alpha_\infty = \alpha_{1\infty} + \alpha_{2\infty} + \alpha_{3\infty}$ and $\alpha_T^* = \alpha_{1T}^* + \alpha_{2T}^* + \alpha_{3T}^*$, (A.104) follows from (A.108), (A.109) and (A.110).

First, we will prove (A.108). From Lemma B.2, we have $\mathbf{b}^* = \mathbf{b} + O(\ell^{-1}) + O_p(b^{-1/2})$ and thus

$$
\begin{aligned}
\alpha_{1T}^* &= \sqrt{\ell}\mathbf{b}^{*\prime} E^*\{[F_{N_1} - E^*(F_{N_1})] \quad B_{N_1}\} \\
&= \mathbf{b}^{*\prime} E^*[\widetilde{F}_{N_1} \quad B_{N_1}] \\
&= \mathbf{b}' E^*[\widetilde{F}_{N_1} \quad B_{N_1}] + O_p(\ell^{-1}) + O_p(b^{-1/2}) \\
&= \alpha_{1\infty}^* + O_p(\ell^{-1}) + O_p(b^{-1/2}), \qquad \text{say.} \qquad \text{(A.111)}
\end{aligned}
$$

By combining (A.111) with

$$
E\left[\alpha_{1\infty}^*\right] = \mathbf{b}' E\left[\widetilde{F}_t \quad B_t\right] = \sum_{j=-\ell}^{\ell} (1 - |j|/\ell)\mathbf{b}' E[w_0 \quad v_{-j}] = \alpha_{1\infty} + O(\ell^{-1}). \qquad \text{(A.112)}
$$

and $\alpha_{1\infty}^* - E\left[\alpha_{1\infty}^*\right] = O_p(b^{-1/2})$ from Lemma B.1, we obtain (A.108).

Second, we will prove (A.109). Similarly, we have $\mathbf{c}^* = \mathbf{c} + O(\ell^{-1}) + O_p(b^{-1/2})$ from Lemma B.2 and thus

$$
\alpha_{2T}^* = \sqrt{\ell}\mathbf{c}^{*\prime} E^*[\mathrm{vech}(B_{N_1} B_{N_1}' - E^*(B_{N_1} B_{N_1}')) \quad B_{N_1}]
$$

$$
\begin{aligned}
&= \sqrt{\ell}\mathbf{c}'E^*[\mathrm{vech}(B_{N_1}B_{N_1}' - E^*(B_{N_1}B_{N_1}')) \quad B_{N_1}] + O_p(\ell^{-1}) + O_p(b^{-1/2}) \\
&= \sqrt{\ell}\mathbf{c}'E^*[\mathrm{vech}(B_{N_1}B_{N_1}') \quad B_{N_1}] + O_p(\ell^{-1}) + O_p(b^{-1/2}) \\
&= \alpha_{2\infty}^* + O_p(\ell^{-1}) + O_p(b^{-1/2}), \qquad \text{say.}
\end{aligned} \tag{A.113}
$$

By combining (A.113) with

$$
\begin{aligned}
E\left[\alpha_{2\infty}^*\right] &= \sqrt{\ell}\mathbf{c}'E[\mathrm{vech}\left(B_tB_t'\right) \quad B_t] \\
&= \mathbf{c}'\sum_{i,j=-\ell}^{\ell}\left(1 - \frac{\min\left((\max|i|,|j|)(i\cdot j>0) + (|i|+|j|)(i\cdot j\leq 0),\ell\right)}{\ell}\right) \\
&\quad \times E[\mathrm{vech}\left(v_0v_{-i}'\right) \quad v_{-j}] \\
&= \alpha_{2\infty} + O(\ell^{-1}).
\end{aligned} \tag{A.114}
$$

and $\alpha_{2\infty}^* - E\left[\alpha_{2\infty}^*\right] = O_p(b^{-1/2})$ from Lemma B.1, we obtain (A.109).

Lastly, we will prove (A.110). Note that

$$
\begin{aligned}
\hat{S}_T^* - \widetilde{S}_T^* &= \frac{1}{b}\sum_{k=1}^{b}\left(\widehat{B}_{N_k}\widehat{B}_{N_k}' - B_{N_k}B_{N_k}'\right) \\
&= \bigtriangledown\widetilde{S}_T^*(\widetilde{\beta}^* - \hat{\beta}) + \bigtriangledown^2\widetilde{S}_T^*(\widetilde{\beta}^* - \hat{\beta})^2
\end{aligned} \tag{A.115}
$$

where

$$
\begin{aligned}
\bigtriangledown\widetilde{S}_T^* &= \frac{\sqrt{\ell}}{b}\sum_{k=1}^{b}\left(F_{N_k}B_{N_k}' + B_{N_k}F_{N_k}'\right), \\
\bigtriangledown^2\widetilde{S}_T^* &= \frac{\ell}{b}\sum_{k=1}^{b}\left(F_{N_k}F_{N_k}'\right), \\
\widetilde{\beta}^* - \hat{\beta} &= \left[G_T^{*\prime}V_TG_T^*\right]^{-1}G_T^{*\prime}V_T\frac{1}{\sqrt{T}}m_T^*.
\end{aligned}
$$

First, note that

$$
\begin{aligned}
\bigtriangledown\widetilde{S}_T^* &= E^*\left[\bigtriangledown\widetilde{S}_T^*\right] + O_p^*(b^{-1/2}), \\
\ell^{-1}\bigtriangledown^2\widetilde{S}_T^* &= \ell^{-1}E^*\left[\bigtriangledown^2\widetilde{S}_T^*\right] + O_p^*(b^{-1/2}),
\end{aligned}
$$

where

$$
\begin{aligned}
E^*\left[\bigtriangledown\widetilde{S}_T^*\right] &= \sqrt{\ell}E^*[F_{N_1}B_{N_1}' + B_{N_1}F_{N_1}'] = E^*[\widetilde{F}_{N_1}B_{N_1}' + B_{N_1}\widetilde{F}_{N_1}'], \\
E^*\left[\bigtriangledown^2\widetilde{S}_T^*\right] &= \ell E^*[F_{N_1}F_{N_1}'].
\end{aligned}
$$

Second, note that

$$
T^{1/2}(\widetilde{\beta}^* - \hat{\beta}) = \left[E^*[G_T^{*\prime}]V_TE^*[G_T^*]\right]^{-1}E^*[G_T^{*\prime}]V_Tm_T^* + O_p^*(b^{-1/2}) \tag{A.116}
$$

since $G_T^* - E^*[G_T^*] = O_p^*(b^{-1/2})$ and

$$
G_T^{*\prime}V_TG_T^* - E^*[G_T^{*\prime}]V_TE^*[G_T^*] = O_p^*(b^{-1/2}).
$$

Thus it follows from (A.115)–(A.116) that

$$
\alpha_{3T}^* = T^{1/2}\mathbf{c}^{*\prime}E^*[\mathrm{vech}(\hat{S}_T^* - \widetilde{S}_T^*) \quad m_T^*],
$$

$$
\begin{aligned}
&= \mathbf{c}^{*'} E^*[\text{vech}\left(\triangledown \widetilde{S}_T^* T^{1/2}(\widetilde{\beta}^* - \hat{\beta}) + \triangledown^2 \widetilde{S}_T^* T^{1/2}(\widetilde{\beta}^* - \hat{\beta})^2\right) \quad m_T^*] \\
&= \mathbf{c}^{*'} E^*[\text{vech}\left(\triangledown \widetilde{S}_T^* T^{1/2}(\widetilde{\beta}^* - \hat{\beta})\right) \quad m_T^*] + O_p^*(\ell^{1/2}T^{-1/2}) \\
&= \mathbf{c}^{*'} E^*\{\text{vech}\left(E^*[\triangledown \widetilde{S}_T^*]\left[E^*[G_T^{*\prime}]V_T E^*[G_T^*]\right]^{-1} E^*[G_T^{*\prime}]V_T m_T^*\right) \quad m_T^*\} \\
&\quad + O_p^*(\ell^{1/2}T^{-1/2}) \\
&= \mathbf{c}' E^*\{\text{vech}\left(E^*[\widetilde{F}_{N_1} B_{N_1}' + B_{N_1}\widetilde{F}_{N_1}']\left[E^*[F_{N_1}']V E^*[F_{N_1}]\right]^{-1}\right. \\
&\quad \left. \times E^*[F_{N_1}']V B_{N_1}\right) \quad B_{N_1}\} + O_p^*(\ell^{1/2}T^{-1/2}) + O_p(\ell^{-1}) + O_p(b^{-1/2}), \\
&= \alpha_{3\infty}^* + O_p^*(\ell^{1/2}T^{-1/2}) + O_p(\ell^{-1}) + O_p(b^{-1/2}), \qquad \text{say.} \qquad (A.117)
\end{aligned}
$$

Since

$$
\begin{aligned}
E\left[E^*[\widetilde{F}_{N_1} B_{N_1}' + B_{N_1}\widetilde{F}_{N_1}']\right] &= \sqrt{\ell} E[F_t B_t' + B_t F_t'] = \sum_{j=-\ell}^{\ell}(1-|j|/\ell)E[w_0 v_{-j}' + v_0 w_{-j}'] \\
&= \sum_{j=-\infty}^{\infty} E[w_0 v_{-j}' + v_0 w_{-j}'] + O(\ell^{-1}) = \triangledown S + O(\ell^{-1}), \qquad (A.118) \\
E\left[E^*[F_{N_1}]\right] &= E[F_t] \;=\; E[w_0], \qquad (A.119)
\end{aligned}
$$

it follows that

$$
\begin{aligned}
E^*[\widetilde{F}_{N_1} B_{N_1}' + B_{N_1}\widetilde{F}_{N_1}'] - E\left[E^*[\widetilde{F}_{N_1} B_{N_1}' + B_{N_1}\widetilde{F}_{N_1}']\right] &= O_p(b^{-1/2}), \qquad (A.120) \\
E^*[F_{N_1}] - E\left[E^*[F_{N_1}]\right] &= O_p(b^{-1/2}). \qquad (A.121)
\end{aligned}
$$

Hence, it follows from the moment inequality, Lemma B.1, (A.120) and (A.121) that

$$
\begin{aligned}
\alpha_{3\infty}^* &= E\left[\alpha_{3\infty}^*\right] + O_p(b^{-1/2}) \\
&= \mathbf{c}' \sum_{i=-\infty}^{\infty} E\{\text{vech}[\nabla \bar{S}(E(w_0)'VE(w_0))^{-1}E(w_0)'Vv_0] \quad v_i\} + O_p(\ell^{-1}) + O_p(b^{-1/2}) \\
&= \alpha_{3\infty} + O_p(\ell^{-1}) + O_p(b^{-1/2}). \qquad (A.122)
\end{aligned}
$$

Therefore, (A.110) follows from (A.117), and (A.122).

Part (b): Proof of (A.105). Let

$$
\begin{aligned}
\gamma_{1T}^* &= (T/\ell)E^*\{\mathbf{a}^{*\prime} m_T^* \mathbf{b}^{*\prime}[(G_T^* - G_0^*) \quad m_T^*]\}, \\
\gamma_{2T}^* &= (T/\ell)E\{\mathbf{a}^{*\prime} m_T^* \mathbf{c}^{*\prime}[\text{vech}(\tilde{S}_T^* - S_T^*) \quad m_T^*]\}, \\
\gamma_{3T}^* &= (T/\ell)E^*\{\mathbf{a}^{*\prime} m_T^* \mathbf{e}^{*\prime}[\text{vech}(\tilde{S}_T^* - S_T^*) \quad \text{vech}(\tilde{S}_T^* - S_T^*) \quad m_T^*]\}, \\
\gamma_{4T}^* &= (T/\ell)E^*\{\mathbf{c}^{*\prime}[\text{vech}(\hat{S}_T^* - \tilde{S}_T^*) \quad m_T^*]\}^2.
\end{aligned}
$$

From Lemma B.2, we have $\mathbf{a}^* = \mathbf{a} + O(\ell^{-1}) + O_p(b^{-1/2})$ and thus

$$
\gamma_{1T}^* = \ell^{-1/2}E^*\{\mathbf{a}^{*\prime} B_{N_1} \mathbf{b}^{*\prime}[F_{N_1} - E^*(F_{N_1})] \quad B_{N_1}\} = o_p^*(1).
$$

Similarly,

$$
\begin{aligned}
\gamma_{2T}^* &= (T/\ell)E^*\{\mathbf{a}^{*\prime} m_T^* \mathbf{c}^{*\prime}[\text{vech}(\tilde{S}_T^* - S_T^*) \quad m_T^*]\} \\
&= E^*\{\mathbf{a}^{*\prime} B_{N_1} \mathbf{c}^{*\prime}[\text{vech}(B_{N_1}B_{N_1}' - E^*(B_{N_1}B_{N_1}')) \quad B_{N_1}]\} \\
&= E^*\{\mathbf{a}' B_{N_1} \mathbf{c}'[\text{vech}(B_{N_1}B_{N_1}' - E^*(B_{N_1}B_{N_1}')) \quad B_{N_1}]\} + O_p(\ell^{-1}) + O_p(b^{-1/2}) \\
&= \gamma_{2\infty}^* + O_p(\ell^{-1}) + O_p(b^{-1/2}), \qquad \text{say.}
\end{aligned}
$$

It follows from the moment inequality and Lemma B.1 that

$$\begin{aligned}
\gamma^*_{2\infty} &= E\left[\gamma^*_{2\infty}\right] + O_p(b^{-1/2}) \\
&= \lim_{T\to\infty} \frac{1}{\ell} \sum_{j=-\ell}^{\ell} \sum_{i,k=-T}^{T} E\{\mathbf{a}' v_0 \mathbf{c}' [\text{vech}(v_i v'_{i-j} - \Gamma_j) \quad v_k]\} + O_p(\ell^{-1}) + O_p(b^{-1/2}) \\
&= \gamma_{2\infty} + O_p(\ell^{-1}) + O_p(b^{-1/2}). \qquad \text{(A.123)}
\end{aligned}$$

The result for γ^*_{3T} and γ^*_{4T} can be proved using similar arguments, and thus the proof is omitted.

Part (c): Proof of (A.106). Let $\kappa^*_{1T} = T^{1/2} E^*[(\mathbf{a}^{*\prime} m^*_T)^3]$ denote the second term on the RHS of (A.106). Because the proof of Part (c) is analogous to the proofs of Parts (a) and (b), we will only show that

$$\kappa^*_{1T} = \sum_{i,j=-\infty}^{\infty} E(\mathbf{a}' v_0 \mathbf{a}' v_i \mathbf{a}'_j v_j) + O_p(\ell^{-1}) + O_p(b^{-1/2}). \qquad \text{(A.124)}$$

By definition, we have

$$\kappa^*_{1T} \;=\; \ell^{1/2} E^*[(\mathbf{a}^{*\prime} B_{N_1})^3] \;=\; (\ell^{1/2}/(T-\ell+1)) \sum_{t=0}^{T-\ell} [\mathbf{a}^{*\prime}(\tilde{B}_t + \hat{B}_t - \tilde{B}_t - \mu^*_T)]^3, \qquad \text{(A.125)}$$

where $\hat{B}_t$ and $\tilde{B}_t$ are defined in the proof of Lemma B.1. Thus it suffices to show that

$$\begin{aligned}
(\ell^{1/2}/(T-\ell+1)) \sum_{t=0}^{T-\ell} (\mathbf{a}^{*\prime}\tilde{B}_t)^3 &= \sum_{i,j=-\infty}^{\infty} E(\mathbf{a}' v_0 \mathbf{a}' v_i \mathbf{a}' v_j) + o_p(\ell T^{-1/2}), \qquad \text{(A.126)} \\
(\ell^{1/2}/(T-\ell+1)) \sum_{t=0}^{T-\ell} [\mathbf{a}^{*\prime}(\hat{B}_t - \tilde{B}_t)]^3 &= O_p(\ell^{-1}) + O_p(b^{-1/2}), \qquad \text{(A.127)} \\
(\ell^{1/2}/(T-\ell+1)) \sum_{t=0}^{T-\ell} (\mathbf{a}^{*\prime}\mu^*_T)^3 &= O_p(\ell^{-1}) + O_p(b^{-1/2}), \qquad \text{(A.128)}
\end{aligned}$$

First, we will show (A.126). Since a HAC covariance matrix estimator converges at rate $O_p(\ell^{1/2}T^{-1/2})$, it follows that

$$\begin{aligned}
&(1/(T-\ell+1)) \sum_{i=0}^{T-\ell} \ell^{1/2} (\mathbf{a}^{*\prime}\tilde{B}_i)^3 - \ell^{1/2} E(\mathbf{a}^{*\prime}\tilde{B}_i)^3 \\
= \;& O_p\Big(\sum_{i,j=0}^{\ell} (1 - \min(\max(i,j,|i-j|),\ell)/\ell)(1/(T-\ell+1)) \\
& \times \sum_{t=0}^{T-\ell} [a^{*\prime} v_t a^{*\prime} v_{t+i} a^{*\prime} v_{t+j} - E(a^{*\prime} v_t a^{*\prime} v_{t+i} a^{*\prime} v_{t+j})]\Big) \\
= \;& O_p(\ell^{1/2} T^{-1/2}). \qquad \text{(A.129)}
\end{aligned}$$

By the moment inequality, it follows that

$$(1/b) \sum_{i=0}^{b-1} \ell^{1/2} E(\mathbf{a}^{*\prime}\tilde{B}_i)^3 = \sum_{i,j=-\infty}^{\infty} E(\mathbf{a}' v_0 \mathbf{a}' v_i \mathbf{a}' v_j) + o(\ell T^{-1/2}). \qquad \text{(A.130)}$$

Thus (A.126) follows from (A.129) and (A.130). Next we will show (A.127) and (A.128). Using arguments similar to the one used in the proof of Lemma B.1, we obtain

$$(\ell^{\frac{1}{2}}/(T-\ell+1)) \sum_{t=0}^{T-\ell} [\mathbf{a}^{*\prime}(\hat{B}_t - \tilde{B}_t)]^3 = (\ell^2(T-\ell+1)) \sum_{t=0}^{T-\ell} \{\mathbf{a}^{*\prime}[F_t(\tilde{\beta}_T - \beta_0)]\}^3 = O_p(\ell^2 T^{-\frac{3}{2}}) \quad \text{(A.131)}$$

and

$$(\ell^{1/2}/(T-\ell+1))\sum_{t=0}^{T-\ell}(\mathbf{a}^{*\prime}\mu_T^*)^3 = \ell^{1/2}(\mathbf{a}^{*\prime}\mu_T^*)^3 = O_p(\ell^{1/2}b^{-3/2}). \quad \text{(A.132)}$$

Thus (A.127) and (A.128) are satisfied. Therefore, (A.106) follows.

Part (d): Proof of (A.107). Part (d) can be proved using similar arguments and thus the proof is omitted. *Q.E.D.*

Proofs of Main Theorems

Lastly, we will prove the main theorems.

Proof of Theorem 1: The result for the studentized statistic (3.2) follows from Lemmas A.5-A.7. Note that the J test statistic can be written as

$$J_T \;=\; J_T^{1/2\prime} J_T^{1/2} \quad \text{(A.133)}$$

where

$$J_T^{1/2} \;=\; \hat{S}_T^{-1/2}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t(y_t - \hat{\beta}_T' x_t)\right].$$

Then one can show that Lemmas A.1–A.7 with f_T replaced by $J_T^{1/2}$ hold except that a, b, c, d and e now take different values. Thus the distribution of $J_T^{1/2}$ can be approximated by its Edgeworth expansion in a suitable sense. A slight modification of Theorem 1 of Chandra and Ghosh (1979) completes the proof of (3.3). *Q.E.D.*

Proof of Theorem 2: For iid observations, a modification of Theorem 1 with $\ell = 1$ yields

$$\sup_{x\in\Re^p}\left|P(\hat{\Sigma}_T^{-1/2}(\hat{\beta}_T - \beta_0) \le x) - \Psi_T(x)\right| \;=\; o(T^{-1}), \quad \text{(A.134)}$$

$$\sup_{x\ge 0}\left|P(J_T \le x) - \Psi_{J,T}(x)\right| \;=\; o(T^{-1}). \quad \text{(A.135)}$$

under Assumptions 1(b)(c)(d)(i), $\ell = 1$ and Assumption 1(e) replaced by the standard Cramer condition. It suffices to show that the conditions on $R_t = (v_t', \text{vec}(w_t)')'$ required for the Edgeworth expansion of Theorem 1 are also satisfied for $Q_{N_j} = (B_{N_j}', \text{vec}(F_{N_j})')'$ for $j = 1, \ldots, b$ conditionally on the sample $\chi_T = \{(x_t', y_t, z_t')\}_{t=1}^T$, uniformly for all χ_T in a set whose probability tends to 1 as $T \to \infty$. Without loss of generality, we check the conditions using B_{N_j}. For Assumption A1(b), we have

$$E^*[B_{N_j}] = E^*[B_{N_1}] = \frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell} E^*\left(z_i u_i - \mu_T^*\right) = 0. \quad \text{(A.136)}$$

For Assumption A1(c), it follows from Lemma A.0 that

$$E\left[E^*\left|B_{N_j}\right|^{r+\eta}\right] = \frac{1}{T-\ell+1}\sum_{t=0}^{T-\ell} E\left|\frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell} v_{t+i}\right|^{r+\eta} = E\left|\frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell} v_{t+i}\right|^{r+\eta} < \infty. \quad \text{(A.137)}$$

From the proof of Theorem 4.2 of Götze and Künsch (1996),

$$E^*\left|B_{N_j}\right|^{r+\eta} - E\left[E^*\left|B_{N_j}\right|^{r+\eta}\right] = O_p(b^{-1/2}). \quad \text{(A.138)}$$

Combining the two results implies that the probability of $E^*\left|B_{N_j}\right|^{r+\eta} < \infty$ tends to unity.

By construction, the moving block bootstrap sample are based on the independent sampling of B_{N_j}. Therefore, Assumption A1(d) is trivially satisfied (with a probability one) using a sigma-field defined by $\sigma(N_j)$ for $j = 1, \ldots, b$, conditionally on the sample χ_T. By the same reason, we can replace Assumption A1(e) by the standard Cramér condition and we only need to show that the condition holds with probability tends to one. Using an argument that appeared in the proof of Theorem 4.2 of Götze and Künsch (1996), we have that

$$P\left\{\sup_{d<|t|<b^{1/2}}|E^*\exp[itB_{N_1}]|\leq 1-\zeta\right\}=1-o(T^{-1}) \tag{A.139}$$

for some $0<\zeta<1/2$. *Q.E.D.*

Proof of Theorem 3: It follows from Lemmas A.3–A.5, Lemmas B.2-B.3 and Theorems 1 and 2 that

$$\sup_{x\in\Re^p}|P(\tau_T\leq x)-P^*(\tau^*_{1,\alpha}\leq x)| = O_p(\ell T^{-1})+O_p(\ell^{-q}), \tag{A.140}$$

$$\sup_{x\in\Re^p}|P(|\tau_T|\leq x)-P^*(|\tau^*_{2,\alpha}|\leq x)| = o_p(\ell T^{-1})+O_p(\ell^{-q}), \tag{A.141}$$

$$\sup_{x\geq 0}|P(J_T\leq x)-P^*(J^*_T\leq x)| = o(\ell T^{-1})+O_p(\ell^{-q}). \tag{A.142}$$

Then the standard Cornish-Fisher expansions arguments complete the proof. *Q.E.D.*

References

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 Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica,* 60, 953–966.

Billingsley, P. (1968), *Convergence of Probability Measures,* New York: Wiley.

Brown, B.W. and W.K. Newey (1995), "Bootstrapping for GMM," manuscript, Rice University and MIT.

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.

Carlstein, E. (1986), "The Use of Subseries Methods for Estimating the Variance of a General Statistic from a Stationary Time Series," *Annals of Statistics,* 14, 1171–1179.

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

Davison, A.C. and P. Hall (1993), "On Studentizing and Blocking Methods for Implementing the Bootstrap with Dependent Data," *Australian Journal of Statistics,* 35, 215–224.

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

Fan, Y. and O. Linton (1997), "Some Higher Order Theory for a Consistent Nonparametric Model Specification Test," forthcoming, *Journal of Statistical Planning and Inference.*

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

Götze, F. and C. Hipp (1983), "Asymptotic Expansions for Sums of Weakly Dependent Random Vectors," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete,* 64, 211–239.

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

Hahn, J. (1996), "A Note on Bootstrapping Generalized Method of Moments Estimators," *Econometric Theory,* 12, 187–197.

Hall, A.R. (1994), "Testing for a Unit Root in Time Series With Pretest Data-Based Model Selection," *Journal of Business & Economic Statistics,* 12, 461–470.

Hall, P. (1992), *The Bootstrap and Edgeworth Expansion,* New York: Springer.

Hall, P. and C.C. Heyde (1980), *Martingale Limit Theory and Its Application,* San Diego: Academic Press.

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

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

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

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

Lahiri, S.N. (1996), "On Edgeworth Expansion and Moving Block Bootstrap for Studentized M-Estimators in Multiple Linear Regression Models," *Journal of Multivariate Analysis,* 56, 42–59.

Magnus, J.R. and H. Neudecker (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics,* Revised Edition, Chichester, UK: Wiley.

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.

Ng, S. and P. Perron (1995), "Unit Root Tests in ARMA Models With Data-Dependent Methods for the Selection of the Truncated Lag," *Journal of the American Statistical Association,* 90, 268–281.

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

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

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

Tauchen, G. (1985), "Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained From Financial Market Data," *Journal of Business & Economic Statistics,* 4, 397–416.

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

White, H. (1984), *Asymptotic Theory for Econometricians,* New York: Academic Press.

Yokoyama, R. (1980), "Moment Bounds for Stationary Mixing Sequences," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete,* 52, 45–57.

Table 1
Empirical Size of Nominal 10% t and J Tests with Asymptotic and Bootstrap Critical Values

(1) Trapezoidal kernel						
		Asymptotic		Bootstrap		PSD
ρ	T_0+1	t Test	J Test	t Test	J Test	
0.5	64	23.8	10.6	15.3	8.9	10.9
	128	20.9	9.6	14.2	9.9	7.1
0.9	64	44.9	11.5	20.5	9.1	23.3
	128	35.9	13.6	12.2	11.8	10.1
0.95	64	47.5	12.6	24.0	8.4	27.6
	128	42.3	13.6	12.3	10.1	16.3
(2) Parzen (b) kernel						
		Asymptotic		Bootstrap		PSD
ρ	T_0+1	t Test	J Test	t Test	J Test	
0.5	64	22.5	9.4	14.6	7.7	8.0
	128	20.8	8.8	13.7	9.0	4.7
0.9	64	44.3	10.3	22.4	8.1	18.2
	128	35.9	13.6	12.2	11.8	7.8
0.95	64	48.2	14.0	23.8	9.7	21.2
	128	41.5	13.6	11.9	10.6	13.6
(3) Truncated kernel						
		Asymptotic		Bootstrap		PSD
ρ	T_0+1	t Test	J Test	t Test	J Test	
0.5	64	23.0	10.5	15.4	8.2	14.0
	128	20.8	10.1	13.0	9.6	12.2
0.9	64	42.3	8.5	19.8	7.0	33.0
	128	32.6	10.0	11.7	8.7	30.2
0.95	64	45.1	10.2	21.6	7.3	36.8
	128	37.9	10.3	11.0	9.5	34.7
(4) Prewhitened HAC						
		Bartlett		QS		PSD
ρ	T_0+1	t Test	J Test	t Test	J Test	
0.5	64	20.2	11.3	20.7	10.8	—
	128	15.8	10.8	16.4	10.5	—
0.9	64	37.1	13.9	38.2	10.0	—
	128	26.3	13.8	26.2	9.6	—
0.95	64	41.9	20.1	41.4	11.8	—
	128	30.5	15.3	31.0	9.4	—

Notes: Numbers are in percent. "PSD" refers to the frequencies of the positive semidefinite correction procedure described in Section 3.

Table 2
GMM Estimates of the Policy Rule Parameters

(a) Pre-Volcker Period: 1960:1-1972:2

Kernel	β	γ	J
None	0.834	0.274	13.075
	(0.067)	(0.087)	(0.126)
Bartlett	0.871	0.392	22.206
	(0.030)	(0.073)	(0.671)
QS	0.871	0.388	22.242
	(0.030)	(0.073)	(0.673)

(b) Volcker-Greenspan Period: 1979:3-1996:3

Kernel	β	γ	J
None	2.153	0.933	21.376
	(0.379)	(0.454)	(0.625)
Bartlett	2.258	0.854	23.314
	(0.148)	(0.224)	(0.726)
QS	2.280	0.803	34.607
	(0.148)	(0.216)	(0.978)

Notes: Asymptotic standard errors for the estimates of β and γ, and asymptotic p values for the J statistics are in parentheses. For the asymptotic confidence interval based on the Bartlett and QS kernels, the data-dependent bandwidth estimator of Andrews (1991) and the prewhitening procedure of Andrews and Monahan (1992) are used. The estimated bandwidths are reported in Table 3. "None" indicates that the inverse of the variance-covariance matrix is used as the weighting matrix.

Table 3
90% Confidence Intervals of the Policy Rule Parameters

(a) Pre-Volcker Period: 1960:1-1972:2

	Kernel	ℓ	β	γ
Asymptotic	None	0	(0.724, 0.945)	(0.131, 0.416)
	Bartlett	0.640	(0.822, 0.921)	(0.272, 0.512)
	QS	0.944	(0.823, 0.920)	(0.268, 0.507)
HH Bootstrap	None	2	(0.656, 1.013)	(-0.027, 0.575)
IS Bootstrap	Trapezoidal	2	(0.738, 1.191)	(0.177, 0.762)
IS Bootstrap	Parzen (b)	2	(0.693, 1.161)	(0.150, 0.627)
IS Bootstrap	Truncated	2	(0.738, 1.191)	(0.177, 0.762)

(b) Volcker-Greenspan Period: 1979:3-1996:3

	kernels	ℓ	β	γ
Asymptotic	None	0	(1.530, 2.776)	(0.187, 1.680)
	Bartlett	1.227	(2.015, 2.502)	(0.485, 1.222)
	QS	1.460	(2.038, 2.523)	(0.449, 1.158)
HH Bootstrap	None	2	(1.070, 3.263)	(-0.927, 2.738)
IS Bootstrap	Trapezoidal	2	(1.517, 3.026)	(-0.151, 0.996)
IS Bootstrap	Parzen (b)	2	(1.446, 3.204)	(-0.321, 1.273)
IS Bootstrap	Truncated	2	(1.517, 3.026)	(-0.151, 0.996)

Notes: "HH Bootstrap" denotes the bootstrap method of Hall and Horowitz (1996) and "IS Bootstrap" denotes the bootstrap method proposed in the present paper. "None" indicates that the inverse of the variance-covariance matrix is used as the weighting matrix. ℓ denotes the bandwidth for the asymptotic confidence interval and the block length for the bootstrap confidence interval. For the asymptotic confidence interval based on the Bartlett and QS kernels, the data-dependent bandwidth estimator of Andrews (1991) and the prewhitening procedure of Andrews and Monahan (1992) are used. For the bootstrap confidence interval, the data-dependent procedure described in Section 3 is used to select the block length.