Econometric Theory, **18**, 2002, 1449–1459. Printed in the United States of America.
DOI: 10.1017/S0266466602186087

CONSISTENT COVARIANCE MATRIX ESTIMATION FOR LINEAR PROCESSES

Michael Jansson
University of California, Berkeley

Consistency of kernel estimators of the long-run covariance matrix of a linear process is established under weak moment and memory conditions. In addition, it is pointed out that some existing consistency proofs are in error as they stand.

1. INTRODUCTION

Suppose $\{V_t : t \geq 1\}$ is a sequence of random n-vectors generated by the linear process

$$V_t = \sum_{l=0}^{\infty} C_l e_{t-l}.$$

This paper considers estimation of $\Omega = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} E(V_t V_s')$, the long-run covariance matrix of V_t. Consistency of kernel estimators of Ω is established under weak conditions on $\{C_l : l \geq 0\}$ and $\{e_t : t \in \mathbb{Z}\}$. In addition, it is pointed out that some existing consistency proofs are in error as they stand.

2. RESULTS

Let $\|\cdot\|$ denote the Euclidean norm, let $\overset{a.s.}{=}$ signify almost sure equality, and let $E_{t-1}(\cdot)$ denote conditional expectation with respect to the σ-algebra generated by $\{e_s : s \leq t-1\}$. The development of formal results proceeds under the following assumptions.

A1. $\sum_{l=0}^{\infty} \|C_l\| < \infty$.

A2. $E_{t-1}(e_t) \overset{a.s.}{=} 0$, $E_{t-1}(e_t e_t') \overset{a.s.}{=} I_n$, and $\{\|e_t\|^2\}$ is uniformly integrable.

The assumption $E_{t-1}(e_t e_t') \overset{a.s.}{=} I_n$ implies the conditional homoskedasticity restriction $E_{t-1}(V_t V_t' - E_{t-1}(V_t V_t')) \overset{a.s.}{=} C_0 C_0'$ but does not impose restrictions on the form of the conditional covariance matrix because $C_0 = I_n$ is not assumed.

I am grateful to Don Andrews (the co-editor) and two anonymous referees for helpful comments. Address correspondence to: Michael Jansson, Department of Economics, University of California, Berkeley, 549 Evans Hall #3880, Berkeley, CA 94720-3880, USA; e-mail: mjansson@econ.berkeley.edu.

 0266-4666/02 $9.50

The uniform integrability condition is satisfied whenever $e_t \sim i.i.d.(0, I_n)$ or $\sup_{t\in\mathbb{Z}} E\|e_t\|^r < \infty$ for some $r > 2$. The best currently available consistency results for linear processes would appear to be those of Robinson (1991) and de Jong and Davidson (2000). The moment and memory assumptions of these papers are stronger than A1 and A2 in the leading special case where $\{e_t\}$ is independent and identically distributed (i.i.d.). When the bandwidth expansion rates recommended by Andrews (1991) are employed, Robinson (1991) requires A1 and $\sup_{t\in\mathbb{Z}} E\|e_t\|^r < \infty$ for some $r > \frac{5}{2}$. On the other hand, de Jong and Davidson (2000) only require two finite moments but do require L_2-near epoch dependence of size $-\frac{1}{2}$. In the linear process case, this condition implies $\sum_{l=1}^{\infty} l^{\varphi}\|C_l\|^2 < \infty$ for some $\varphi > 1$, a stronger requirement than A1.

Defining $\Gamma = \lim_{T\to\infty} T^{-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1} E(V_t V_s')$ and $\Sigma = \lim_{T\to\infty} T^{-1} \times \sum_{t=1}^{T} E(V_t V_t')$, the matrix Ω can be decomposed as $\Omega = \Gamma + \Gamma' + \Sigma$. In some applications in nonstationary time series analysis, the matrix Γ is of interest in its own right. Obvious examples include the cointegration procedures of Phillips and Hansen (1990) and Park (1992). In recognition of this fact, the present paper focuses explicitly on consistent estimation of Γ. It is assumed that Γ is estimated by a kernel estimator of the form

$$\hat{\Gamma}_T = T^{-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1} k\left(\frac{|t-s|}{b_T}\right) V_t V_s',$$

where $k(\cdot)$ is a (measurable) kernel function and $\{b_T : T \geq 1\}$ is a sequence of bandwidth parameters. The corresponding estimator of Ω is

$$\hat{\Omega}_T = T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k\left(\frac{|t-s|}{b_T}\right) V_t V_s',$$

which can be written as $\hat{\Gamma}_T + \hat{\Gamma}_T' + \hat{\Sigma}_T$, where $\hat{\Sigma}_T = T^{-1}\sum_{t=1}^{T} V_t V_t'$. Because $\hat{\Sigma}_T \to_p \Sigma$ under A1 and A2, $\hat{\Omega}_T$ is a consistent estimator of Ω whenever $\hat{\Gamma}_T$ is a consistent estimator of Γ. Consider the following assumptions on $k(\cdot)$ and $\{b_T\}$.

A3.

(i) $k(0) = 1, k(\cdot)$ is continuous at zero and $\sup_{x\geq 0}|k(x)| < \infty$.
(ii) $\int_{[0,\infty)} \bar{k}(x)\,dx < \infty$, where $\bar{k}(x) = \sup_{y\geq x}|k(y)|$.

A4. $\{b_T\} \subseteq (0,\infty)$ and $\lim_{T\to\infty}(b_T^{-1} + T^{-1/2} b_T) = 0$.

Assumption A3 generalizes Robinson's (1991) assumption A2(0) and would appear to be satisfied by any kernel in actual use.[1] For instance, it holds for the 15 kernels studied by Ng and Perron (1996). Moreover, it holds for all kernels in the class $\mathcal{K}_3$ of Andrews (1991) and Andrews and Monahan (1992) and for all kernels satisfying Assumptions 1 and 3 of Newey and West (1994). Assumption A4 is standard and holds whenever the bandwidth expansion rate coincides with the optimal (under a mean squared error criterion) rate reported in Andrews (1991, p. 830).

An important implication of A3(ii) is the following lemma.

LEMMA 1. *Suppose $k(\cdot)$ satisfies A3(ii) and suppose $\{b_T\} \subseteq (0,\infty)$. Then*

$$\overline{\lim}_{T\to\infty} \sup_{0<\alpha\le\alpha_u} b_T^{-1} \sum_{i=1}^{T-1} \left| k\left(\frac{i}{\alpha b_T}\right)\right| < \infty$$

for any $0 < \alpha_u < \infty$.

Results similar to Lemma 1 have been stated (without proof) by Andrews (1991, p. 852), Hansen (1992, p. 970), and Hall (2000, Lemma 2). As demonstrated by the examples that follow, the assumptions made in the cited papers do not imply $\overline{\lim}_{T\to\infty} b_T^{-1} \sum_{i=1}^{T-1} |k(i/b_T)| < \infty$. Therefore, the proofs of Theorem 1 of Andrews (1991) and Theorems 1–3 of Hansen (1992) are in error as they stand, as are the proofs of Theorems 2 and 3 of Hall (2000).

The sequence $\{b_T^{-1} \sum_{i=1}^{T-1} |k(i/b_T)|\}$ depends on $k(\cdot)$ through $\{k(x) : x \in \mathcal{D}\}$, where $\mathcal{D} = \cup_{T\ge1} \cup_{1\le i\le T-1} \{i/b_T\}$. The set $\mathcal{D}$ is countable, so it is possible to have $k(x) = 1\ \forall x \in \mathcal{D}$ under Hansen's (1992) Condition (K) and Hall's (2000) Assumption 5. In particular, if $\lim_{T\to\infty} T^{-1} b_T = 0$ it is possible to have

$$\underline{\lim}_{T\to\infty} b_T^{-1} \sum_{i=1}^{T-1} \left| k\left(\frac{i}{b_T}\right)\right| = \underline{\lim}_{T\to\infty} \frac{T-1}{b_T} = \infty.$$

Moreover, a kernel can have $k(x) = 1\ \forall x \in \mathbb{N}$ and belong to the class $\mathcal{K}_1$ of Andrews (1991) and Andrews and Monahan (1992). For any such kernel and any $\{b_T\} \subseteq \mathbb{N}$ with $\lim_{T\to\infty} T^{-1/2} b_T = 0$,

$$\begin{aligned}\underline{\lim}_{T\to\infty} b_T^{-1} \sum_{i=1}^{T-1} \left| k\left(\frac{i}{b_T}\right)\right| &\ge \underline{\lim}_{T\to\infty} b_T^{-1} \sum_{x=1}^{\lfloor (T-1)/b_T \rfloor} |k(x)| \\ &= \underline{\lim}_{T\to\infty} \frac{\lfloor (T-1)/b_T \rfloor}{b_T} = \infty,\end{aligned}$$

where $\lfloor \cdot \rfloor$ denotes the integer part of the argument.[2] Assumption A3(ii) rules out pathological cases such as these.

The main result of the paper is the following theorem.

THEOREM 2. *Suppose A1–A4 hold. Then* $\hat{\Gamma}_T \to_p \Gamma$ *and* $\hat{\Omega}_T \to_p \Omega$.

In applications, the vectors $\{V_t\}$ are often functions of an unknown parameter vector θ (say), $V_t = V_t(\theta_0)$, where θ_0 denotes the true value of θ. Consider the estimators

$$\hat{\Gamma}_T(\hat{\theta}_T) = T^{-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1} k\left(\frac{|t-s|}{b_T}\right)V_t(\hat{\theta}_T)V_s(\hat{\theta}_T)',$$

$$\hat{\Omega}_T(\hat{\theta}_T) = T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k\left(\frac{|t-s|}{b_T}\right)V_t(\hat{\theta}_T)V_s(\hat{\theta}_T)',$$

where $\hat{\theta}_T$ is an estimator of θ_0 satisfying the following assumption.

A5. Either

(i) $V_t(\theta) = V_t(\theta_0) - (\theta - \theta_0)X_t$, where $T^{1/2}(\hat{\theta}_T - \theta_0)\delta_T^{-1} = O_p(1)$ and $\max_{1\le t\le T} \|\delta_T X_t\| = O_p(1)$ for some sequence $\{\delta_T\}$ of nonsingular matrices or

(ii) $T^{1/2}(\hat{\theta}_T - \theta_0) = O_p(1)$ and $\sup_{t\ge 1} E(\sup_{\theta\in\mathcal{N}}\|(\partial/\partial\theta')V_t(\theta)\|^2) < \infty$ for some neighborhood $\mathcal{N}$ of θ_0.

Assumption A5(i) is Condition (V3) of Hansen (1992) whereas A5(ii) is equivalent to Assumption B of Andrews (1991) under A1 and A2. As in Hansen (1992), the following corollary is an immediate consequence of Theorem 2.

COROLLARY 3. *Suppose A1–A5 hold. Then* $\hat{\Gamma}_T(\hat{\theta}_T) \to_p \Gamma$ *and* $\hat{\Omega}_T(\hat{\theta}_T) \to_p \Omega$.

Sample-dependent bandwidth parameters can also be accommodated. Let

$$\hat{\Gamma}_T(\hat{\theta}_T, \hat{b}_T) = T^{-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1} k\left(\frac{|t-s|}{\hat{b}_T}\right)V_t(\hat{\theta}_T)V_s(\hat{\theta}_T)',$$

$$\hat{\Omega}_T(\hat{\theta}_T, \hat{b}_T) = T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} k\left(\frac{|t-s|}{\hat{b}_T}\right)V_t(\hat{\theta}_T)V_s(\hat{\theta}_T)',$$

where $\{\hat{b}_T : T \ge 1\}$ is a sequence of (possibly) stochastic bandwidth parameters satisfying the following assumption.

A4′. $\hat{b}_T = \hat{\alpha}_T b_T$, where $\hat{\alpha}_T > 0, \hat{\alpha}_T + \hat{\alpha}_T^{-1} = O_p(1)$, and $\{b_T\}$ satisfies A4.

COROLLARY 4. *Suppose A1–A3, A4′, and A5 hold. Then* $\hat{\Gamma}_T(\hat{\theta}_T, \hat{b}_T) \to_p \Gamma$ *and* $\hat{\Omega}_T(\hat{\theta}_T, \hat{b}_T) \to_p \Omega$.

3. PROOFS

Using change of variables, $\hat{\Gamma}_T$ can be written as

$$\hat{\Gamma}_T = \sum_{i=1}^{T-1} k\left(\frac{i}{b_T}\right)\left(T^{-1}\sum_{j=1}^{T-i} V_{j+i}V_j'\right).$$

The proofs of Theorem 2 and its corollaries are based on this representation, Lemma 1, and the following lemmas.

LEMMA 5. *Suppose A1 and A2 hold. Then*

$$E\left\| T^{-1}\sum_{j=1}^{T-i}[V_{j+i}V_j' - E(V_{j+i}V_j')]\right\| \le \beta_i\psi_T + T^{-1/2}\eta_T, \qquad 0 \le i \le T-1,$$

where $\{\beta_i : i \ge 0\}$ *and* $\{\psi_T, \eta_T : T \ge 1\}$ *are nonnegative sequences with* $\sum_{i=1}^{\infty}\beta_i < \infty$, $\lim_{T\to\infty}\psi_T = 0$, *and* $\overline{\lim}_{T\to\infty}\eta_T < \infty$.

LEMMA 6. *Suppose A1 and A2 hold. Moreover, suppose* $k(\cdot)$ *satisfies A3(i) and suppose* $\{b_T\} \subseteq (0,\infty)$ *with* $\lim_{T\to\infty} b_T^{-1} = 0$. *Then*

$\lim_{T\to\infty}\sup_{\alpha\ge\alpha_l}\|E[\hat{\Gamma}_T(\theta_0, \alpha b_T)] - \Gamma\| = 0$

for any $0 < \alpha_l < \infty$.

Proof of Lemma 1. Let $\bar{k}$ be defined as in A3(ii). By monotonicity of $\bar{k}$,

$$\left|k\left(\frac{i}{\alpha b_T}\right)\right| \le \bar{k}\left(\frac{i}{\alpha b_T}\right) \le \bar{k}\left(\frac{i}{\alpha_u b_T}\right) \le \alpha_u b_T\int_{[(i-1)/\alpha_u b_T,\, i/\alpha_u b_T)}\bar{k}(x)\,dx$$

for any $1 \le i \le T-1$ and any $0 < \alpha \le \alpha_u$. As a consequence,

$$\sup_{0<\alpha\le\alpha_u} b_T^{-1}\sum_{i=1}^{T-1}\left|k\left(\frac{i}{\alpha b_T}\right)\right| \le \alpha_u\int_{[0,(T-1)/\alpha_u b_T)}\bar{k}(x)\,dx$$

$$\le \alpha_u\int_{[0,\infty)}\bar{k}(x)\,dx < \infty. \qquad \blacksquare$$

Proof of Lemma 5. Because

$$V_{j+i}V_j' = \left(\sum_{l=0}^{\infty}C_l e_{j+i-l}\right)\left(\sum_{m=0}^{\infty}C_m e_{j-m}\right)'$$

$$= \sum_{m=0}^{\infty}C_{m+i}e_{j-m}e_{j-m}'C_m' + \sum_{l=0}^{\infty}\sum_{m=0}^{\infty}1\{l \ne m+i\}C_l e_{j+i-l}e_{j-m}'C_m',$$

where $1\{\cdot\}$ is the indicator function, it follows that

$$T^{-1}\sum_{j=1}^{T-i}(V_{j+i}V_j' - E(V_{j+i}V_j'))$$

$$= \sum_{m=0}^{\infty}C_{m+i}\left(T^{-1}\sum_{j=1}^{T-i}(e_{j-m}e_{j-m}' - I_n)\right)C_m'$$

$$+ \sum_{l=0}^{\infty}\sum_{m=0}^{\infty}C_l\left(1\{l \ne m+i\}T^{-1}\sum_{j=1}^{T-i}e_{j+i-l}e_{j-m}'\right)C_m'$$

under A1 and A2. Using subadditivity of $\|\cdot\|$ and the fact that $\|AB\| \le \|A\|\cdot\|B\|$ for conformable A and B, this expression can be bounded as follows:

$$\begin{aligned}
&\left\| T^{-1} \sum_{j=1}^{T-i} (V_{j+i}V_j' - E(V_{j+i}V_j')) \right\| \\
&\quad \le \sum_{m=0}^{\infty} \left\| T^{-1} \sum_{j=1}^{T-i} (e_{j-m}e_{j-m}' - I_n) \right\| \|C_m\|\|C_{m+i}\| \\
&\qquad + \sum_{l=0}^{\infty}\sum_{m=0}^{\infty} \left\| 1\{l \neq m+i\} T^{-1} \sum_{j=1}^{T-i} e_{j+i-l}e_{j-m}' \right\| \|C_l\|\|C_m\|.
\end{aligned}$$

Therefore, $E\|T^{-1}\sum_{j=1}^{T-i}[V_{j+i}V_j' - E(V_{j+i}V_j')]\| \le \beta_i \psi_T + T^{-1/2}\eta_T$, where

$$\beta_i = \sum_{m=0}^{\infty} \|C_m\|\|C_{m+i}\|,$$

$$\psi_T = \sup_{m\ge 0} \max_{0\le i\le T-1} E\left\| T^{-1} \sum_{j=1}^{T-i} (e_{j-m}e_{j-m}' - I_n) \right\|,$$

$$\eta_T = \left(\max_{0\le i\le T-1} \sup_{l,m\ge 0} E\left\| 1\{l\neq m+i\} T^{-1/2} \sum_{j=1}^{T-i} e_{j+i-l}e_{j-m}' \right\| \right) \left(\sum_{l=0}^{\infty} \|C_l\| \right)^2.$$

By A1,

$$\sum_{i=1}^{\infty} \beta_i = \sum_{i=1}^{\infty}\sum_{m=0}^{\infty} \|C_m\|\|C_{m+i}\| \le \left(\sum_{m=0}^{\infty} \|C_m\| \right)^2 < \infty.$$

Each element of $\{e_{j-m}e_{j-m}' - I_n : j \ge 1\}$ is a uniformly integrable martingale difference sequence under A2. As a consequence, for any $\varepsilon > 0$ there is a finite constant λ_ε (independent of i and m) such that

$$\begin{aligned}
E\left\| T^{-1} \sum_{j=1}^{T-i} (e_{j-m}e_{j-m}' - I_n) \right\| &\le (T-i)^{1/2}T^{-1}\lambda_\varepsilon + (T-i)T^{-1}\varepsilon \\
&\le T^{-1/2}\lambda_\varepsilon + \varepsilon,
\end{aligned}$$

where the first inequality is obtained by proceeding as in the proof of Hall and Heyde (1980, Theorem 2.22). Therefore, $\overline{\lim}_{T\to\infty} \psi_T \le \varepsilon$ for any $\varepsilon > 0$, so $\psi_T \to 0$. Finally,

$$\begin{aligned}
&\left\| 1\{l\neq m+i\} T^{-1/2} \sum_{j=1}^{T-i} e_{j+i-l}e_{j-m}' \right\|^2 \\
&\quad = 1\{l\neq m+i\} \mathrm{tr}\left[\left(T^{-1/2} \sum_{j_1=1}^{T-i} e_{j_1+i-l}e_{j_1-m}' \right)' \left(T^{-1/2} \sum_{j_2=1}^{T-i} e_{j_2+i-l}e_{j_2-m}' \right) \right] \\
&\quad = T^{-1} \sum_{j_1=1}^{T-i}\sum_{j_2=1}^{T-i} 1\{l\neq m+i\} e_{j_1-m}' e_{j_2-m} e_{j_1+i-l}' e_{j_2+i-l}.
\end{aligned}$$

By A2, $E(1\{l \neq m+i\}e'_{j_2-m}e_{j_1-m}e'_{j_1+i-l}e_{j_2+i-l}) = n^2 1\{j_1 = j_2\}1\{l \neq m+i\}$, because, e.g.,

$$E(e'_{j_2-m}e_{j_1-m}e'_{j_1+i-l}e_{j_2+i-l}) = E[E_{j_2-m}(e'_{j_1-m})e_{j_2-m}e'_{j_1+i-l}e_{j_2+i-l}] = 0$$

when $j_1 > j_2$ and $l > m+i$, whereas

$$E(e'_{j_2-m}e_{j_1-m}e'_{j_1+i-l}e_{j_2+i-l}) = E[E_{j+i-l}(e'_{j-m}e_{j-m})e'_{j+i-l}e_{j+i-l}] = n^2$$

when $j_1 = j_2 = j$ and $l > m+i$. Therefore,

$$\begin{aligned}
E&\left\|1\{l \neq m+i\}T^{-1/2}\sum_{j=1}^{T-i} e_{j+i-l}e'_{j-m}\right\| \\
&\leq E\left(\left\|1\{l \neq m+i\}T^{-1/2}\sum_{j=1}^{T-i} e_{j+i-l}e'_{j-m}\right\|^2\right)^{1/2} \\
&= \left(T^{-1}\sum_{j_1=1}^{T-i}\sum_{j_2=1}^{T-i} n^2 1\{j_1 = j_2\}1\{l \neq m+i\}\right)^{1/2} \\
&= [1\{l \neq m+i\}n^2 T^{-1}(T-i)]^{1/2} \\
&\leq n,
\end{aligned}$$

where the first inequality uses the Cauchy–Schwarz inequality. In particular,

$$\overline{\lim}_{T\to\infty}\eta_T \leq n \cdot \left(\sum_{l=0}^{\infty}\|C_l\|\right)^2 < \infty,$$

as was to be shown. ■

Proof of Lemma 6. Under A1 and A2,

$$E[\hat{\Gamma}_T(\theta_0, \alpha b_T)] = \sum_{i=1}^{\infty} 1\{i \leq T-1\}k\left(\frac{i}{\alpha b_T}\right)\frac{T-i}{T}E(V_{1+i}V'_1)$$

for any $\alpha > 0$, whereas $\Gamma = \sum_{i=1}^{\infty} E(V_{1+i}V'_1)$. Therefore, by subadditivity of $\|\cdot\|$,

$$\begin{aligned}
\|E&[\hat{\Gamma}_T(\theta_0, \alpha b_T)] - \Gamma\| \\
&\leq \sum_{i=1}^{I}\left|1\{i \leq T-1\}k\left(\frac{i}{\alpha b_T}\right)\frac{T-i}{T} - 1\right| \cdot \|E(V_{1+i}V'_1)\| \\
&\quad + \sum_{i=I+1}^{\infty}\left|1\{i \leq T-1\}k\left(\frac{i}{\alpha b_T}\right)\frac{T-i}{T} - 1\right| \cdot \|E(V_{1+i}V'_1)\| \\
&\leq \max_{1\leq i\leq I}\left|1\{i \leq T-1\}k\left(\frac{i}{\alpha b_T}\right)\frac{T-i}{T} - 1\right| \cdot \sum_{i=1}^{I}\|E(V_{1+i}V'_1)\| \\
&\quad + \left(\sup_{x\geq 0}|k(x)| + 1\right) \cdot \sum_{i=I+1}^{\infty}\|E(V_{1+i}V'_1)\|
\end{aligned}$$

for any $I \geq 1$. Because $\sum_{i=0}^{\infty} \| E(V_{1+i} V_1') \| < \infty$,

$$\left(\sup_{x \geq 0} |k(x)| + 1 \right) \cdot \sum_{i=I+1}^{\infty} \| E(V_{1+i} V_1') \|$$

can be made arbitrarily small (under A3(i)) by taking I large enough. For any given I,

$$\begin{aligned} &\sup_{\alpha \geq \alpha_l} \max_{1 \leq i \leq I} \left| 1\{i \leq T-1\} k\left(\frac{i}{\alpha b_T} \right) \frac{T-i}{T} - 1 \right| \\ &\quad = \sup_{\alpha \geq \alpha_l} \max_{1 \leq i \leq I} \left| k\left(\frac{i}{\alpha b_T} \right) \frac{T-i}{T} - 1 \right| \\ &\quad \leq \sup_{\alpha \geq \alpha_l} \max_{1 \leq i \leq I} \left(\left| k\left(\frac{i}{\alpha b_T} \right) - 1 \right| + T^{-1} i \left| k\left(\frac{i}{\alpha b_T} \right) \right| \right) \\ &\quad \leq \sup_{0 \leq x \leq I/(\alpha_l b_T)} |k(x) - 1| + T^{-1} I \sup_{x \geq 0} |k(x)| \end{aligned}$$

whenever $0 < \alpha_l < \infty$ and $T \geq I + 1$. Lemma 6 now follows because the expression on the last line tends to zero whenever A3(i) holds and $\lim_{T \to \infty} b_T^{-1} = 0$. ∎

Proof of Theorem 2. By subadditivity of $\| \cdot \|$,

$$\| \hat{\Gamma}_T - \Gamma \| \leq \| \hat{\Gamma}_T - E(\hat{\Gamma}_T) \| + \| E(\hat{\Gamma}_T) - \Gamma \|,$$

$$\| \hat{\Omega}_T - \Omega \| \leq \| \hat{\Omega}_T - E(\hat{\Omega}_T) \| + \| E(\hat{\Omega}_T) - \Omega \|.$$

Now, $E(\hat{\Sigma}_T) = \Sigma$, so $\| E(\hat{\Omega}_T) - \Omega \| \leq 2 \cdot \| E(\hat{\Gamma}_T) - \Gamma \| \to 0$ by Lemma 6. Moreover,

$$\| \hat{\Omega}_T - E(\hat{\Omega}_T) \| \leq 2 \cdot \| \hat{\Gamma}_T - E(\hat{\Gamma}_T) \| + \| \hat{\Sigma}_T - E(\hat{\Sigma}_T) \|.$$

By Lemma 5,

$$E \| \hat{\Sigma}_T - E(\hat{\Sigma}_T) \| = E \left\| T^{-1} \sum_{j=1}^{T} [V_j V_j' - E(V_j V_j')] \right\| \to 0.$$

In particular, $\hat{\Sigma}_T - E(\hat{\Sigma}_T) = o_p(1)$. The proof of Theorem 2 can be completed by showing that $E \| \hat{\Gamma}_T - E(\hat{\Gamma}_T) \| \to 0$. By subadditivity of $\| \cdot \|$,

$$\| \hat{\Gamma}_T - E(\hat{\Gamma}_T) \| \leq \sum_{i=1}^{T-1} \left| k\left(\frac{i}{b_T} \right) \right| \cdot \left\| T^{-1} \sum_{j=1}^{T-i} [V_{j+i} V_j' - E(V_{j+i} V_i')] \right\|.$$

Using Lemmas 1 and 5 and the notation from Lemma 5,

$$
\begin{aligned}
E\|\hat{\Gamma}_T - E(\hat{\Gamma}_T)\| &\leq \sum_{i=1}^{T-1}\left|k\left(\frac{i}{b_T}\right)\right| \cdot E\left\|T^{-1}\sum_{j=1}^{T-i}[V_{j+i}V_j' - E(V_{j+i}V_i')]\right\| \\
&\leq \bar{k}(0)\left(\sum_{i=1}^{T-1}\beta_i\right)\psi_T + T^{-1/2}\eta_T\sum_{i=1}^{T-1}\left|k\left(\frac{i}{b_T}\right)\right| \\
&\leq \bar{k}(0)\left(\sum_{i=1}^{\infty}\beta_i\right)\psi_T + (T^{-1/2}b_T\eta_T)\left(b_T^{-1}\sum_{i=1}^{T-1}\left|k\left(\frac{i}{b_T}\right)\right|\right) \\
&\to 0. \qquad \blacksquare
\end{aligned}
$$

Proof of Corollaries 3 and 4. Corollary 3 is a special case (with $\hat{\alpha}_T = 1$) of Corollary 4, so it suffices to prove the latter. Under $A4'$, $\inf_{\tau\geq T}\Pr(\alpha_l \leq \hat{\alpha}_\tau \leq \alpha_u)$ can be made arbitrarily close to unity for sufficiently large T and some $0 < \alpha_l \leq \alpha_u < \infty$. Consequently, it suffices to show that for any $0 < \alpha_l \leq \alpha_u < \infty$,

$$
\sup_{\alpha_l\leq\alpha\leq\alpha_u}\|\hat{\Gamma}_T(\hat{\theta}_T, \alpha b_T) - \Gamma\| = o_{p^*}(1),
$$

which is easily shown to imply $\sup_{\alpha_l\leq\alpha\leq\alpha_u}\|\hat{\Omega}_T(\hat{\theta}_T, \alpha b_T) - \Omega\| = o_{p^*}(1)$, where $o_{p^*}(1)$ denotes convergence to zero in outer probability.[3] Now,

$$
\begin{aligned}
\sup_{\alpha_l\leq\alpha\leq\alpha_u}\|\hat{\Gamma}_T(\hat{\theta}_T, \alpha b_T) - \Gamma\| \leq &\sup_{\alpha_l\leq\alpha\leq\alpha_u}\|\hat{\Gamma}_T(\hat{\theta}_T, \alpha b_T) - \hat{\Gamma}_T(\theta_0, \alpha b_T)\| \\
&+ \sup_{\alpha_l\leq\alpha\leq\alpha_u}\|\hat{\Gamma}_T(\theta_0, \alpha b_T) - E[\hat{\Gamma}_T(\theta_0, \alpha b_T)]\| \\
&+ \sup_{\alpha_l\leq\alpha\leq\alpha_u}\|E[\hat{\Gamma}_T(\theta_0, \alpha b_T)] - \Gamma\|,
\end{aligned}
$$

so the proof can be completed by showing that each term on the right-hand side is $o_{p^*}(1)$. As in the proofs of Theorems 2 and 3 in Hansen (1992), Condition A5 implies that

$$
\|\hat{\Gamma}_T(\hat{\theta}_T, \alpha b_T) - \hat{\Gamma}_T(\theta_0, \alpha b_T)\| \leq (T^{-1/2}b_T)\cdot\left(b_T^{-1}\sum_{i=1}^{T-1}\left|k\left(\frac{i}{\alpha b_T}\right)\right|\right)\cdot Q_T,
$$

for any $\alpha > 0$, where Q_T is $O_p(1)$ and does not depend on α or b_T. Now, $T^{-1/2}b_T \to 0$ and $\overline{\lim}_{T\to\infty}\sup_{\alpha_l\leq\alpha\leq\alpha_u} b_T^{-1}\sum_{i=1}^{T-1}|k(i/(\alpha b_T))| < \infty$ for any $0 < \alpha_l \leq \alpha_u < \infty$ (Lemma 1), so

$$
\sup_{\alpha_l\leq\alpha\leq\alpha_u}\|\hat{\Gamma}_T(\hat{\theta}_T, \alpha b_T) - \hat{\Gamma}_T(\theta_0, \alpha b_T)\| = o_{p^*}(1).
$$

Next, by the properties of $\bar{k}(\cdot)$,

$$
\begin{aligned}
&\|\hat{\Gamma}_T(\theta_0,\alpha b_T) - E[\hat{\Gamma}_T(\theta_0,\alpha b_T)]\| \\
&\qquad \le \sum_{i=1}^{T-1}\left|k\left(\frac{i}{\alpha b_T}\right)\right|\left\|T^{-1}\sum_{j=1}^{T-i}[V_{j+i}V_j' - E(V_{j+i}V_i')]\right\| \\
&\qquad \le \sum_{i=1}^{T-1}\bar{k}\left(\frac{i}{\alpha_u b_T}\right)\left\|T^{-1}\sum_{j=1}^{T-i}[V_{j+i}V_j' - E(V_{j+i}V_i')]\right\|
\end{aligned}
$$

for any $0 < \alpha_l \le \alpha \le \alpha_u < \infty$. By Theorem 2 and its proof, the last line is $o_p(1)$, so

$$\sup_{\alpha_l \le \alpha \le \alpha_u}\|\hat{\Gamma}_T(\theta_0,\alpha b_T) - E(\hat{\Gamma}_T(\theta_0,\alpha b_T))\| = o_{p^*}(1).$$

Finally, $\sup_{\alpha_l \le \alpha \le \alpha_u}\|E[\hat{\Gamma}_T(\theta_0,\alpha b_T)] - \Gamma\| \to 0$ by Lemma 6. ∎

NOTES

1. A3 is weaker than Robinson's (1991) A2(0) because it is not assumed that $\int_{\mathbb{R}}|K(\lambda)|\,d\lambda < \infty$, where $K(\lambda) = (2\pi)^{-1}\int_{\mathbb{R}}k(x)\exp(i\lambda x)\,dx$. A well-known example of a kernel satisfying A3 but violating $\int_{\mathbb{R}}|K(\lambda)|\,d\lambda < \infty$ is the uniform kernel $k(x) = 1\{|x| \le 1\}$, where $1\{\cdot\}$ is the indicator function.

2. One kernel $k \in \mathcal{K}_1$ such that $k(x) = 1\ \forall x \in \mathbb{Z}$ is $k(\cdot) = \sum_{l=0}^{\infty}k_l(\cdot)$, where, for each $l \ge 0$ and any $x \in \mathbb{R}$,

$$k_l(x) = \begin{cases} 1-(l+1)^2(|x|-l), & \text{if } l \le |x| \le l+(l+1)^{-2}, \\ 1-(l+1)^2(l+1-|x|), & \text{if } l+1-(l+1)^{-2} \le |x| \le l+1, \\ 0, & \text{otherwise.} \end{cases}$$

It is easily seen that $k \in \mathcal{K}_1$. In particular, k is continuous and

$$\int_{\mathbb{R}}|k(x)|\,dx = 2\sum_{l=0}^{\infty}\int_{[0,\infty)}k_l(x)\,dx = 2\sum_{l=0}^{\infty}(l+1)^{-2} = \frac{\pi^2}{3} < \infty.$$

3. To avoid measurability complications, convergence in outer probability (rather than convergence in probability) is considered.

REFERENCES

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

Andrews, D.W.K. & J.C. Monahan (1992) An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 60, 953–966.

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

Hall, P. & C.C. Heyde (1980) *Martingale Limit Theory and Its Application*. New York: Academic Press.

Hansen, B.E. (1992) Consistent covariance matrix estimation for dependent heterogeneous processes. *Econometrica* 60, 967–972.

de Jong, R.M. & J. Davidson (2000) Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. *Econometrica* 68, 407–423.

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

Ng, S. & P. Perron (1996) The exact error in estimating the spectral density at the origin. *Journal of Time Series Analysis* 17, 379–408.

Park, J.Y. (1992) Canonical cointegrating regressions. *Econometrica* 60, 119–143.

Phillips, P.C.B. & B.E. Hansen (1990) Statistical inference in instrumental variables regression with I(1) variables. *Review of Economic Studies* 57, 99–125.

Robinson, P.M. (1991) Automatic frequency domain inference on semiparametric and nonparametric models. *Econometrica* 59, 1329–1363.