6 Estimating Asymptotic Covariance Matrices

The asymptotic normality of the optimization estimators established in chapter 5 provides the fundamental basis for constructing asymptotic confidence intervals and statistics appropriate for testing hypotheses. Such statistics as the Wald and Lagrange multiplier statistic require for their computation either a knowledge of or a consistent estimate of the asymptotic covariance matrix of the estimator. As evident from theorem 5.7, this covariance matrix is of the form $C_n^* \equiv A_n^{*-1} B_n^* A_n^{*-1}$. Because A_n^* and B_n^* are generally unknown, we need consistent estimators for them, say $\hat{A}_n$ and $\hat{B}_n$; a consistent estimator of the asymptotic covariance matrix can then be constructed as $\hat{C}_n = \hat{A}_n^{-1} \hat{B}_n \hat{A}_n^{-1}$. The results of chapter 7 also require consistent estimators for A_n^o and B_n^o. The purpose of this chapter, then, is to provide general conditions compatible with those of the preceding chapters which ensure (where possible) the consistency of useful estimators for A_n^*, A_n^o, B_n^* and B_n^o.

Sufficient structure is already available to give consistent estimators for A_n^o and A_n^*. We have the following result.

Theorem 6.1

Given assumptions DG, OP', MX, SM, DM', NE, and ID

$$\nabla_\theta^2 \hat{Q}_n - A_n^* \to 0 \quad \text{a.s.}$$

$$\nabla_\theta^2 \tilde{Q}_n - A_n^o \to 0 \quad \text{a.s.}$$

where $\nabla_\theta^2 \hat{Q}_n \equiv \nabla_\theta^2 Q_n(\hat{\theta}_n)$ and $\nabla_\theta^2 \tilde{Q}_n \equiv \nabla_\theta^2 Q_n(\tilde{\theta}_n)$. □

Our task here is complete once we have available consistent estimators for B_n^* and B_n^o.

Finding such estimators is an interesting challenge. Recall that

$$B_n^o \equiv \operatorname{var}\left(n^{-1/2} \sum_{t=1}^{n} M_{nt}^o\right)$$

$$= n^{-1} \sum_{t=1}^{n} E(M_{nt}^o M_{nt}^{o\prime}) + n^{-1} \sum_{\tau=1}^{n-1} \sum_{t=\tau+1}^{n} [E(M_{nt}^o M_{n,t-\tau}^{o\prime})$$

$$+ E(M_{n,t-\tau}^o M_{nt}^{o\prime})].$$

Thus B_n^o is the sum of $n(n+1)/2$ terms, while we have available only n observations. Without further information, it will generally not be possible to estimate B_n^o consistently.

A variety of different circumstances arise which do provide us with information which will allow consistent estimation of an important component of B_n^o. The simplest such circumstance arises when $\{M_{nt}^o, F^t\}$ is a martingale difference sequence. In this case, the elements of M_{nt}^o are measurable-F_t/B and $E(M_{nt}^o \mid F^{t-1}) = 0$ a.s. Using the law of iterated expectations, we have for all $\tau > 0$

$$E(M_{nt}^o M_{n,t-\tau}^{o\prime}) = E(E(M_{nt}^o M_{n,t-\tau}^{o\prime} \mid F^{t-1}))$$

$$= E(E(M_{nt}^o \mid F^{t-1}) M_{n,t-\tau}^{o\prime})$$

$$= 0.$$

Thus in this case, we simply have

$$B_n^o = n^{-1} \sum_{t=1}^{n} E(M_{nt}^o M_{nt}^{o\prime}).$$

Such situations arise when one has independent observations, or in certain cases in which the model under consideration is not subject to dynamic misspecification.

To consider consistent estimation of B_n^o in this case we decompose M_{nt}^o as

$$M_{nt}^o = S_{nt}^o - E(S_{nt}^o)$$

where S_{nt}^o is the generalized score,

$$S_{nt}^o = \nabla_\theta \bar{\nu}_n^{o\prime} \nabla_\psi^2 g_n(\bar{\psi}_n^o) q_{nt}^o + \nabla_\theta q_{nt}^{o\prime} \nabla_\psi g_n(\bar{\psi}_n^o)'.$$

In general, $E(S_{nt}^o)$ need not equal zero, although in special cases in which the model is correctly specified or in which the observations are generated by a stationary process we may have $E(S_{nt}^o) = 0$. Otherwise, $E(S_{nt}^o)$ will be unknown, and as pointed out by Chow (1981) it will not be possible to estimate that component of B_n^o which we now write as

$$U_n^o = n^{-1} \sum_{t=1}^{n} E(S_{nt}^o)E(S_{nt}^{o\prime}).$$

However, it will generally be possible to estimate

$$B_n^o + U_n^o = n^{-1} \sum_{t=1}^{n} E(S_{nt}^o S_{nt}^{o\prime}).$$

As in Eicker (1967), the basic intuition is that if S_{nt}^o were observable, then $B_n^o + U_n^o$ could be consistently estimated by

$$B_n \equiv n^{-1} \sum_{t=1}^{n} S_{nt}^o S_{nt}^{o\prime}.$$

Unfortunately S_{nt}^o is not observable because θ_n^o is unknown; however, it can be estimated by

$$\tilde{S}_{nt} \equiv \nabla_\theta \tilde{\psi}_n' \nabla_\psi^2 g_n(\tilde{\psi}_n)\tilde{q}_{nt} + \nabla_\theta \tilde{q}_{nt}' \nabla_\psi g_n(\tilde{\psi}_n)'$$

where $\tilde{q}_{nt} \equiv q_t(\tilde{\theta}_n)$, $\nabla_\theta \tilde{q}_{nt} \equiv \nabla_\theta q_t(\tilde{\theta}_n)$, $\tilde{\psi}_n \equiv \psi_n(\tilde{\theta}_n)$, and $\nabla_\theta \tilde{\psi}_n \equiv \nabla_\theta \psi_n(\tilde{\theta}_n)$. This suggests consideration of the estimator

$$\tilde{B}_n \equiv n^{-1} \sum_{t=1}^{n} \tilde{S}_{nt} \tilde{S}_{nt}'.$$

In fact, when $\{M_{nt}^o, F^t\}$ is a martingale difference sequence, this estimator is consistent for $B_n^o + U_n^o$ under conditions given below, and similarly

$$\hat{B}_n \equiv n^{-1} \sum_{t=1}^{n} \hat{S}_{nt} \hat{S}_{nt}'$$

is consistent for $B_n^* + U_n^*$, where

$$\hat{S}_{nt} \equiv \nabla_\theta \hat{\psi}_n' \nabla_\psi^2 g_n(\hat{\psi}_n)\hat{q}_{nt} + \nabla_\theta \hat{q}_{nt}' \nabla_\psi g_n(\hat{\psi}_n)'$$

with $\hat{q}_{nt} \equiv q_t(\hat{\theta}_n)$, $\nabla_\theta \hat{q}_{nt} \equiv \nabla_\theta q_t(\hat{\theta}_n)$, $\hat{\psi}_n \equiv \psi_n(\hat{\theta}_n)$, and $\nabla_\theta \hat{\psi}_n \equiv \nabla_\theta \psi_n(\hat{\theta}_n)$; and

$$U_n^* = n^{-1} \sum_{t=1}^{n} E(S_{nt}^*)E(S_{nt}^{*\prime}),$$

with

$$S_{nt}^* \equiv \nabla_\theta \psi_n^{*\prime} \nabla_\psi^2 g_n(\psi_n^*)q_{nt}^* + \nabla_\theta q_{nt}^{*\prime} \nabla_\psi g_n(\psi_n^*)'.$$

In order to prove the consistency of $\tilde{B}_t$ and $\hat{B}_n$ further structure is required. The nature of this structure can be appreciated by considering the computationally infeasible estimator B_n suggested above,

$$B_n \equiv n^{-1} \sum_{t=1}^{n} S_{nt}^o S_{nt}^{o\prime}.$$

It is reasonable to expect that $B_n - E(B_n) \xrightarrow{p} 0$ (note $E(B_n) = B_n^o + U_n^o$) by some suitable law of large numbers, and indeed this convergence underlies our proof of the consistency of $\tilde{B}_n$. In fact, no such law of large numbers is yet available here. With the structure presently available, we have that $\{S_{nt}^o\}$ is mixingale of size -1; however, we need a law of large numbers for the elements of $\{S_{nt}^o S_{nt}^{o\prime}\}$ and these are not necessarily mixingales. Even if we succeed in ensuring that the elements of $\{S_{nt}^o S_{nt}^{o\prime}\}$ are mixingales, we still face a further problem: the summands are doubly indexed, while the strong law of large numbers for mixingales provided by theorem 3.15 applies to singly indexed summands. A strong law for doubly indexed sequences is not easily available; however, a weak law of large numbers can be proven quite easily under weak conditions. Because convergence in probability suffices for subsequent results, we use this weak law.

In order to ensure that the elements of $\{S_{nt}^o S_{nt}^{o\prime}\}$ are indeed mixingales of the proper size, we rely on corollary 4.3(b), which gives sufficient conditions for products of near epoch dependent functions also to be near epoch dependent functions. We apply that result by letting $Y_{nt}Z_{nt}$ be products and cross-products of the elements of q_{nt}^o and ∇q_{nt}^o. In order to satisfy the integrability conditions of corollary 4.3(b), we strengthen assumption DM′.

Assumption DM″

(i) The elements of $\{q_t(\theta)\}$ are $2r$-dominated on Θ uniformly in $t = 1, 2, \ldots, r > 2$.

(ii) The elements of $\{\nabla_\theta q_t(\theta)\}$ are $2r$-dominated on Θ uniformly in $t = 1, 2, \ldots, r > 2$.

(iii) The elements of $\{\nabla_\theta^2 q_t(\theta)\}$ are $2r$-dominated on Θ uniformly in $t = 1, 2, \ldots, r > 2$. □

Assumption DM″(iii) is not needed immediately; however, it plays a crucial although somewhat different role in establishing the consistency of $\tilde{B}_n$, and we impose it here for convenience.

A weak law of large numbers for double arrays of mixingales is the following.

Theorem 6.2

Suppose $\{Z_{nt}\}$ is a double array of random scalars such that $\|Z_{nt}\|_r \leqslant \Delta < \infty$ for some $r \geqslant 2$, and $E(Z_{nt}) = 0$, $n, t = 1, 2, \ldots$. If $\{Z_{nt}\}$ is near epoch dependent on $\{V_t\}$ of size $-1/2$, where $\{V_t\}$ is a mixing process with ϕ_m of size $-r/(2r-2)$, $r \geqslant 2$ or α_m of size $-r/(r-2)$, $r > 2$, then $n^{-1}\sum_{t=1}^{n} Z_{nt} \xrightarrow{p} 0$. □

Recently, Andrews (1987) has given a more general weak law of large numbers for double arrays of L_p-mixingales $1 \leqslant p < \infty$. Although we do not pursue the implications of Andrews's (1987) results here, it appears that these results may allow proof of versions of the results which follow under weaker moment and memory conditions.

Theorem 6.2 will be applied to products and cross-products of the elements of q_{nt}^o and $\nabla_\theta q_{nt}^o$. In order to ensure that these are near epoch dependent on $\{V_t\}$ of the appropriate size $(-1/2)$, we strengthen assumption NE′.

Assumption NE″

(i) The elements of $\{q_t(\theta)\}$ are near epoch dependent on $\{V_t\}$ of size $-(r-1)/(r-2)$ uniformly on (Θ, ρ).

(ii) The elements of $\{\nabla_\theta q_t(\theta)\}$ are near epoch dependent on $\{V_t\}$ of size $-(r-1)/(r-2)$ uniformly on (Θ, ρ). □

Note that as $r \to \infty$, we approach the size requirement (-1) imposed in assumption NE′.

Because the proof of consistency of $\tilde{B}_n$ relies on being able to take mean value expansions around θ_n^o, we complete assumption ID′ as follows.

Assumption ID′

(ii) The sequence $\{\bar{Q}_n(\theta)\}$ has identifiably unique minimizers $\{\theta_n^o\}$ on $\{\Theta_n\}$, where $\{\theta_n^o\}$ is interior to Θ uniformly in n. □

The need for assumption DM″(iii) arises from the appearance of second derivatives of $q_t(\theta)$ after taking the mean value expansion around θ_n^o.

We now have sufficient conditions available to state our first consistency result for $\tilde{B}_n$ and $\hat{B}_n$.

Theorem 6.3

Given assumptions DG, OP′, MX, SM(i) and (ii), DM″, NE″, and ID′:

(a) If $\{M_{nt}^o, F^t\}$ is a martingale difference sequence for all $n = 1, 2, \ldots$, then $\{U_n^o\}$ is $O(1)$, positive semidefinite, and

$$\tilde{B}_n - (B_n^o + U_n^o) \xrightarrow{p} 0,$$

where $\tilde{B}_n = n^{-1}\sum_{t=1}^{n} \tilde{S}_{nt}\tilde{S}_{nt}'$ and $U_n^o = n^{-1}\sum_{t=1}^{n} E(S_{nt}^o)E(S_{nt}^{o\prime})$.

(b) If $\{M_{nt}^*, F^t\}$ is a martingale difference sequence for all $n = 1, 2, \ldots$, then $\{U_n^*\}$ is $O(1)$, positive semidefinite, and

$$\hat{B}_n - (B_n^* + U_n^*) \xrightarrow{p} 0$$

where $\hat{B}_n = n^{-1}\sum_{t=1}^{n} \hat{S}_{nt}\hat{S}_{nt}'$ and $U_n^* = n^{-1}\sum_{t=1}^{n} E(S_{nt}^*)E(S_{nt}^{*\prime})$. □

Because U_n^o and U_n^* are positive semidefinite and do not vanish in the presence of heterogeneous observations and misspecification, this result provides at worst a basis for constructing conservative hypothesis tests. The reason for this is that $\hat{A}_n^{-1}\hat{B}_n\hat{A}_n^{-1}$ will be consistent for $A_t^{*-1}B_n^*A_n^{*-1} + A_n^{*-1}U_n^*A_n^{*-1}$, which will always overestimate $C_n^* \equiv A_t^{*-1}B_n^*A_n^{*-1}$ by the positive semidefinite matrix $D_n^* = A_n^{*-1}U_n^*A_n^{*-1}$. (The use of the symbol D_n^* can be thought of as a mnemonic for "discrepancy.") This point is incorrectly treated in White (1983).

The present result generalizes results of Eicker (1967) and White (1980) for the linear, independent, correctly specified case and results of Nicholls and Pagan (1983) for the linear, martingale difference, correctly specified case to the nonlinear, martingale difference, possibly misspecified case. The underlying method of proof is essentially the same, however.

Similar results are available for situations in which $E(M_{nt}^o M_{n,t-\tau}^{o\prime}) = 0$ for all $\tau > m$, where m is a *known* finite integer, $m \geqslant 1$. For this case we choose

$$\tilde{B}_n = n^{-1}\sum_{t=1}^{n} \tilde{S}_{nt}\tilde{S}_{nt}' + n^{-1}\sum_{\tau=1}^{m}\sum_{t=\tau+1}^{n} [\tilde{S}_{nt}\tilde{S}_{n,t-\tau}' + \tilde{S}_{n,t-\tau}\tilde{S}_{nt}'] \quad (6.1)$$

$$\hat{B}_n = n^{-1}\sum_{t=1}^{n} \hat{S}_{nt}\hat{S}_{nt}' + n^{-1}\sum_{\tau=1}^{m}\sum_{t=\tau+1}^{n} [\hat{S}_{nt}\hat{S}_{n,t-\tau}' + \hat{S}_{n,t-\tau}\hat{S}_{nt}']. \quad (6.2)$$

The result for this case is the following.

Theorem 6.4

Given assumptions DG, OP′, MX, SM(i) and (ii), DM″, NE″, and ID′:

(a) If $E(M^o_{nt}M^{o'}_{n,t-\tau})=0$ for all $\tau > m, n=1,2,\ldots$, then $\{U^o_n\}$ is $O(1)$ and

$$\tilde{B}_n-(B^o_n+U^o_n)\xrightarrow{P}0,$$

where $\tilde{B}_n$ is given by (6.1) and

$$U^o_n=n^{-1}\sum_{t=1}^{n}E(S^o_{nt}S^{o'}_{nt})+n^{-1}\sum_{\tau=1}^{m}\sum_{t=\tau+1}^{n}[E(S^o_{nt})E(S^{o'}_{n,t-\tau})$$
$$+E(S^o_{nt-\tau})E(S^{o'}_{nt})].$$

(b) If $E(M^*_{nt}M^{*'}_{n,t-\tau})=0$ for all $\tau > m$, $n=1,2,\ldots$, then $\{U^*_n\}$ is $O(1)$ and

$$\hat{B}_n-(B^*_n+U^*_n)\xrightarrow{P}0,$$

where $\hat{B}_n$ is given by (6.2) and

$$U^*_n=n^{-1}\sum_{t=1}^{n}E(S^*_{nt})E(S^{*'}_{nt})$$
$$+n^{-1}\sum_{\tau=1}^{m}\sum_{t=\tau+1}^{n}[E(S^*_{nt})E(S^{*'}_{n,t-\tau})+E(S^*_{n,t-\tau})E(S^{*'}_{nt})].\quad\square$$

Estimators of the form (6.1) and (6.2) have been proposed by Hansen (1982) for the nonlinear, stationary ergodic, correctly specified case. The present result applies in the nonlinear, dependent heterogeneous, possibly misspecified case.

This case presents practical difficulties not encountered in theorem 6.3. Specifically, U^o_n and U^*_n are not guaranteed to be positive semidefinite for any n, so that conservative inferences based on $\hat{A}_n^{-1}\hat{B}_n\hat{A}_n^{-1}$ are no longer necessarily possible. Further, nothing ensures that $\tilde{B}_n$ or $\hat{B}_n$ are positive semidefinite, although when $U^*_n=0$ ($U^o_n=0$) then $\hat{B}_n$ ($\tilde{B}_n$) will be positive definite for all n sufficiently large, almost surely. Nevertheless, in finite samples $\hat{B}_n$ and $\tilde{B}_n$ can be quite badly behaved. Another serious practical difficulty is that m is required to be known. Such knowledge may be available in special cases, such as when the investigator exploits the m-period ahead prediction errors of a correctly specified forecasting equation to estimate parameters of interest, but in general m will not be known, nor will it necessarily be known that such an m exists.

In the general case in which $E(M^o_{nt}M^{o'}_{n,t-\tau})$ does not equal zero after a finite number of lags, it is nevertheless possible to obtain useful estimators by using the fact that the near epoch dependence and mixing conditions imposed earlier imply that $E(M^o_{nt}M^{o'}_{n,t-\tau})$ converges to zero as τ becomes arbitrarily large. This suggests that it may be possible to obtain a useful estimator by neglecting large values of τ in forming $\tilde{B}_n$ and $\hat{B}_n$, as in the estimator suggested by Domowitz and White (1982):

$$\tilde{B}_n=n^{-1}\sum_{t=1}^{n}\tilde{S}_{nt}\tilde{S}'_{nt}+n^{-1}\sum_{\tau=1}^{m_n}\sum_{t=\tau+1}^{n}[\tilde{S}_{nt}\tilde{S}'_{n,t-\tau}+\tilde{S}_{n,t-\tau}\tilde{S}'_{nt}]$$

where $m_n\to\infty$ as $n\to\infty$. By requiring that m_n grows at the proper rate, it is possible to ensure that the neglected terms never become too important. Just as in the immediately preceding case, $\tilde{B}_n$ is not guaranteed to be positive semidefinite. However, Newey and West (1987) have shown that it is possible to guarantee positive semidefiniteness by introducing appropriate weights $\{w_{n\tau}\}$ and forming estimators

$$\tilde{B}_n=w_{no}n^{-1}\sum_{t=1}^{n}\tilde{S}_{nt}\tilde{S}'_{nt}+n^{-1}\sum_{\tau=1}^{m_n}w_{n\tau}\sum_{t=\tau+1}^{n}[\tilde{S}_{nt}\tilde{S}'_{n,t-\tau}+\tilde{S}_{n,t-\tau}\tilde{S}'_{nt}]\tag{6.3}$$

$$\hat{B}_n=w_{no}n^{-1}\sum_{t=1}^{n}\hat{S}_{nt}\hat{S}'_{nt}+n^{-1}\sum_{\tau=1}^{m_n}w_{n\tau}\sum_{t=\tau+1}^{n}[\hat{S}_{nt}\hat{S}'_{n,t-\tau}+\hat{S}_{n,t-\tau}\hat{S}'_{nt}].\tag{6.4}$$

To be useful, the weights must have two properties: they must ensure the nonnegativity of $\lambda'\tilde{B}_n\lambda$ and $\lambda'\hat{B}_n\lambda$ for any $\lambda\in\mathbb{R}^k$, and they must not interfere with the convergence of $\tilde{B}_n$ and $\hat{B}_n$ to the appropriate limit. The following lemma allows the construction of weights which will ensure the nonnegativity requirement.

Lemma 6.5

Let $\{Z_{nt}\}$ be an arbitrary double array, and let $\{a_{ni}\}$, $n=1,2,\ldots$, $i=1,\ldots,m_n+1$ be a triangular array of real numbers. Then for any triangular array of weights

$$w_{n\tau}=\sum_{i=\tau+1}^{m_n+1}a_{ni}a_{n,i-\tau},\quad n=1,2,\ldots,\quad \tau=1,\ldots,m_n$$

we have

$$\Psi_n \equiv w_{no} \sum_{t=1}^{n} Z_{nt}^2 + 2 \sum_{\tau=1}^{m_n} w_{n\tau} \sum_{t=\tau+1}^{n} Z_{nt} Z_{n,t-\tau} \geqslant 0. \qquad \square$$

We apply this result by setting $Z_{nt} = \lambda' \tilde{S}_{nt}$ for arbitrary $\lambda \in \mathbb{R}^k$. For this case $\Psi_n = \lambda' \tilde{B}_n \lambda \geqslant 0$ so that $\tilde{B}_n$ is positive definite as required, and similarly for $\hat{B}_n$.

To ensure that $\{w_{n\tau}\}$ does not interfere with the consistency of $\tilde{B}_n$, further conditions must be imposed. It is reasonable to anticipate that such conditions will include the requirement that for each $\tau = 1, 2, \ldots$, $w_{n\tau} \to 1$ as $n \to \infty$. One such sequence of weights, related to the Bartlett (1950) sequence, is given by Newey and West (1987), namely

$$w_{n\tau} = 1 - \tau/(m_n + 1),$$

which arises from the choice $a_{ni} = (m_n + 1)^{1/2}$ for all $i = 1, \ldots, m_n + 1$. In fact, this choice of weights does yield a consistent estimator under appropriate conditions.

In this stationary case, the problem of estimating B_n is essentially the problem of estimating the spectrum of a time series at zero frequency. Anderson (1971, chapter 8) discusses a variety of different approaches for stationary time series, each of which essentially involves different choices of weights (windows). Here we seek results for heterogeneous processes.

The strategy of our consistency proof for $\tilde{B}_n$ is straightforward. Let $\{Z_{nt}\}$ be an arbitrary double array of random $k \times 1$ vectors. We define

$$B_n \equiv \text{var}\left(n^{-1/2} \sum_{t=1}^{n} Z_{nt}\right)$$

and impose conditions on $\{Z_{nt}\}$, m_n, and weights $\{w_{n\tau}\}$ to ensure that

$$B_n - \ddot{B}_n \to 0 \quad \text{as } n \to \infty$$

where

$$\ddot{B}_n = w_{no} n^{-1} \sum_{t=1}^{n} E(Z_{nt} Z_{nt}') + n^{-1} \sum_{\tau=1}^{m_n} w_{n\tau} \sum_{t=\tau+1}^{n} E(Z_{nt} Z_{n,t-\tau}')$$
$$+ E(Z_{n,t-\tau} Z_{nt}').$$

Then, setting $Z_{nt} = S_{nt}^o$ we show that $\tilde{B}_n - \ddot{B}_n \xrightarrow{p} 0$ under appropriate conditions. Our first result establishes conditions ensuring that $B_n - \ddot{B}_n \to 0$.

Lemma 6.6

Let $\{Z_{nt}\}$ be a double array of random $k \times 1$ vectors, $k \in \mathbb{N}$, such that $\|Z_{ti} Z_{nt}\|_{r/2} \leqslant \Delta < \infty$ for some $r > 2$, $E(Z_{nt}) = 0$, $n, t = 1, 2, \ldots$ and $\{Z_{nt}\}$ is near epoch dependent on $\{V_t\}$ of size -1, where $\{V_t\}$ is a mixing sequence with ϕ_m of size $-r/(r-1)$ or α_m of size $-2r/(r-2)$. Define

$$B_n \equiv \text{var}\left(n^{-1/2} \sum_{t=1}^{n} Z_{nt}\right),$$

and for any sequence $\{m_n\}$ of integers and any triangular array $\{w_{n\tau}: n = 1, 2, \ldots, \tau = 1, \ldots, m_n\}$ define

$$\ddot{B}_n = w_{no} n^{-1} \sum_{t=1}^{n} E(Z_{nt} Z_{nt}')$$
$$+ n^{-1} \sum_{\tau=1}^{m_n} w_{n\tau} \sum_{t=\tau+1}^{n} [E(Z_{nt} Z_{n,t-\tau}') + E(Z_{n,t-\tau} Z_{nt}')].$$

If $m_n \to \infty$ as $n \to \infty$, if $|w_{n\tau}| \leqslant \Delta$, $n = 1, 2, \ldots$, $\tau = 1, \ldots, m_n$, and if for each τ, $w_{n\tau} \to 1$ as $n \to \infty$, then

$$B_n - \ddot{B}_n \to 0 \quad \text{as } n \to \infty. \qquad \square$$

Note that the moment, near epoch dependence, and mixing conditions imposed here are the same as we impose in stating the central limit theorem. Given these moment and memory conditions, the conditions on m_n and $\{w_{n\tau}\}$ are as previously anticipated: we require $m_n \to \infty$ and for each τ, $w_{n\tau} \to 1$. Also note that we allow dependence of Z_{nt} on past as well as future values of V_t.

Our next result is an intermediate lemma analogous to lemma 6.19 of White (1984). Part (b) of the present result corrects an error in White's lemma pointed out by Newey and West (1987) and Phillips (1985); this error underlies the incorrect rate for m_n given by Domowitz and White (1982) and White (1984, theorem 6.20).

Lemma 6.7

Let $\{Z_{nt}\}$ be a double array of random $k \times 1$ vectors, $k \in \mathbb{N}$, such that

$\|Z'_{nt}Z_{nt}\|_r \leqslant \Delta$ for some $r > 2$, $n, t = 1, 2, \ldots$ and let

$$v_m \equiv \sup_n \sup_t \|Z_{nt} - E_{t-m}^{t+m}(Z_{nt})\|_2$$

where $E_{t-m}^{t+m}(\cdot) \equiv E(\cdot \mid F_{t-m}^{t+m})$, $F_{t-m}^{t+m} = \sigma(V_{t-m}, \ldots, V_{t+m})$ for a given sequence of random vectors $\{V_t\}$. For $i, j = 1, \ldots, k$, define

$$\xi_{nt\tau}^{ij} \equiv Z_{nti}Z_{n,t-\tau,j} - E(Z_{nti}Z_{n,t-\tau,j}).$$

Then

(a) Letting ϕ_m and α_m represent the mixing coefficients associated with $\{V_t\}$, for fixed τ and all $m > 6\tau$

$$|E(\xi_{nt\tau}^{ij}\xi_{n,t-m+\tau,\tau}^{ij})| \leqslant K_o(\phi_{l_{m\tau}}^{1-1/r} + v_{l_{m\tau}}^{(r-2)/2(r-1)}) \quad \text{or}$$
$$\leqslant K_o(\alpha_{l_{m\tau}}^{1/2-1/r} + v_{l_{m\tau}}^{(r-2)/2(r-1)})$$

where $K_o < \infty$ and $l_{m\tau} \equiv [([m/2] - 3\tau)/2]$.

Suppose further that $\{Z_{nt}\}$ is near epoch dependent on $\{V_t\}$ of size $-2(r-1)/(r-2)$ and that $\{V_t\}$ is mixing with ϕ_m of size $-r/(r-1)$ or α_m of size $-2r/(r-2)$. Then

(b) For all $\tau \geqslant 0$, $i, j = 1, \ldots, k$, $n = 1, 2, \ldots$

$$E\left(\left[\sum_{t=\tau+1}^{n} \xi_{nt\tau}^{ij}\right]^2\right) \leqslant (\tau+2)n\Delta_o$$

where $\Delta_o < \infty$.

(c) For all $m = 1, 2, \ldots$, $i, j = 1, \ldots, k$, given any $\varepsilon > 0$

$$P\left[n^{-1}\sum_{\tau=1}^{m}\left|\sum_{t=\tau+1}^{n} \xi_{nt\tau}^{ij}\right| \geqslant \varepsilon\right] \leqslant \Delta m^4/(n\varepsilon^2) + \Delta m^3/(n\varepsilon^2).$$

Finally, if $m_n = o(n^{1/4})$ and $|w_{n\tau}| \leqslant \Delta$, $n = 1, 2, \ldots$, $\tau = 1, 2, \ldots, m_n$, then

(d) For all $i, j = 1, \ldots, k$

$$n^{-1}\sum_{\tau=1}^{m_n} w_{n\tau} \sum_{t=\tau+1}^{n} \xi_{nt\tau}^{ij} \xrightarrow{P} 0. \qquad \square$$

We now have available results which will allow us to prove consistency of $\tilde{B}_n$ and the analogous estimator $\hat{B}_n$ for $B_n^o + U_n^o$ and $B_n^* + U_n^*$ respectively. These results require us to strengthen our near epoch dependence conditions in the following way.

Assumption NE‴

(i) The elements of $\{q_t(\theta)\}$ are near epoch dependent on $\{V_t\}$ of size $-2(r-1)/(r-2)$ uniformly on (Θ, ρ).
(ii) The elements of $\{\nabla_\theta q_t(\theta)\}$ are near epoch dependent on $\{V_t\}$ of size $-2(r-1)/(r-2)$ uniformly on (Θ, ρ). $\square$

Our conditions on the truncation lag m_n and on the weights $\{w_{n\tau}\}$ are formally expressed in the following way.

Assumption TL (truncation lag)

$\{m_n\}$ is a sequence of integers such that $m_n \to \infty$ as $n \to \infty$ and $m_n = O(n^{1/4})$. $\square$

This assumption replaces the incorrect rate $m_n = o(n^{1/3})$ of Domowitz and White (1982) and White (1984) with the appropriate rate $m_n = o(n^{1/4})$ given by Newey and West (1987) and Phillips (1985).

Assumption WT (weights)

For a given sequence $\{m_n\}$ define

$$w_{n\tau} = \sum_{\lambda=\tau+1}^{m_n+1} a_{n\lambda}a_{n\lambda-\tau},$$

where $\{a_{n\lambda}\}$, $n = 1, 2, \ldots$, $\lambda = 1, \ldots, m_n + 1$ is any triangular array such that $|w_{n\tau}| \leqslant \Delta < \infty$, $n = 1, 2, \ldots$, $\tau = 1, \ldots, m_n$, and for each τ, $w_{n\tau} \to 1$ as $n \to \infty$. $\square$

The desired consistency result can now be stated.

Theorem 6.8

Given assumptions DG, OP′, MX′, SM, DM″, NE‴, ID′, TL, and WT:

(a) For all $n = 1, 2, \ldots$, the matrix $\tilde{B}_n$ given by (6.3) is positive semidefinite, U_n^o is positive semidefinite, and provided that $\sqrt{(n)}(\tilde{\theta}_n - \theta_n^o) = O_p(1)$,

$$\tilde{B}_n - (B_n^o + U_n^o) \xrightarrow{P} 0,$$

where

$$U_n^o = w_{no} n^{-1} \sum_{t=1}^{n} E(S_{nt}^o)E(S_{nt}^{o\prime})$$

$$+ n^{-1} \sum_{\tau=1}^{m_n} w_{n\tau} \sum_{t=\tau+1}^{n} [E(S_{nt}^o)E(S_{n,t-\tau}^{o\prime}) + E(S_{n,t-\tau}^o)E(S_{nt}^{o\prime})].$$

(b) For all $n = 1, 2, \ldots$, the matrix $\hat{B}_n$ given by (6.4) is positive semidefinite, U_n^* is positive semidefinite, and

$$\hat{B}_n - (B_n^* + U_n^*) \xrightarrow{P} 0,$$

where

$$U_n^* = w_{no} n^{-1} \sum_{t=1}^{n} E(S_{nt}^*)E(S_{nt}^{*\prime})$$

$$+ n^{-1} \sum_{\tau=1}^{m_n} w_{n\tau} \sum_{t=\tau+1}^{n} [E(S_{nt}^*)E(S_{n,t-\tau}^{*\prime})$$

$$+ E(S_{n,t-\tau}^*)E(S_{nt}^{*\prime})]. \quad \square$$

Note that conclusion (a) requires the additional condition that $\sqrt{(n)}(\tilde{\theta}_n - \theta_n^o)$ be bounded in probability. This condition is automatically satisfied for $\sqrt{(n)}(\hat{\theta}_n - \theta_n^*)$ as a consequence of asymptotic normality; however, asymptotic normality for $\sqrt{(n)}(\tilde{\theta}_n - \theta_n^o)$ has not been established. Nevertheless, under general conditions $\sqrt{(n)}(\tilde{\theta}_n - \theta_n^o)$ will be bounded in probability as required. These are given in the next chapter.

Note also that although U_n^o and U_n^* are guaranteed to be positive semidefinite for all n, there is nothing to ensure that $\{U_n^o\}$ or $\{U_n^*\}$ are $O(1)$. Thus, although conservative inferences will be feasible, the actual (and unknown) size of a given test will decrease with n. Sufficient conditions for $\{U_n^o\}$ or $\{U_n^*\}$ to be $O(1)$ are that $\{X_t\}$ be a stationary sequence and $q_t(\theta)$ depends on t only through a measure preserving shift transformation T, i.e. $q_t(\omega, \theta) = Q(T^t\omega, \theta)$, or that the model is correctly specified. Either of these conditions ensures that $E(S_{nt}^*) = 0$ or $E(S_{nt}^o) = 0$ so that U_n^o and U_n^* vanish.

While the present results establish some minimal conditions on m_n which ensure consistency of $\tilde{B}_n$ and $\hat{B}_n$, they provide almost no practical guidance as to how m_n might usefully be chosen in applications. It is possible that some sort of cross-validation technique (Stone 1974) might prove helpful. However, because of the very large samples which might be required even for cross-validation to be helpful, it may be more advisable to attempt to improve the dynamic specification of the model so that M_{nt}^o is more nearly a martingale difference sequence, rather than attempting to adjust for a poor dynamic specification by using theorem 6.8 to estimate B_n^o or B_n^* consistently.

MATHEMATICAL APPENDIX

Proof of theorem 6.1

Given assumptions DG, OP′, MX, SM, DM′, and NE, it follows from theorem 5.6(c) that $\nabla_\theta^2 Q_n(\theta) - A_n(\theta) \to 0$ a.s. uniformly on Θ, and A_n is continuous on Θ uniformly in n. Given assumption ID also, it follows that $\tilde{\theta}_n - \theta_n^o \to 0$ a.s. and $\hat{\theta}_n - \theta_n^* \to 0$ a.s. It then follows from theorem 2.3 of Domowitz and White (1982) that $\nabla_\theta^2 \tilde{Q}_n - A_n^o \to 0$ a.s. and $\nabla_\theta^2 \hat{Q}_n - A_n^* \to 0$ a.s. $\square$

Proof of theorem 6.2

By Chebyshev's inequality, for any $\varepsilon > 0$

$$P\left[\left|n^{-1}\sum_{t=1}^{n} Z_{nt}\right| \geq \varepsilon\right] \leq E\left(\left[\sum_{t=1}^{n} Z_{nt}\right]^2\right)\Big/ n^2\varepsilon^2$$

$$\leq E\left(\max_{1\leq j\leq n}\left[\sum_{t=1}^{j} Z_{nt}\right]^2\right)\Big/ n^2\varepsilon^2.$$

Given the conditions of the theorem, it follows from lemma 3.14 that $\{Z_{nt}\}$ is a mixingale of size $-1/2$ with $c_{nt} \leq \Delta < \infty$ for all n, t. By McLeish's inequality

$$E\left(\max_{1\leq j\leq n}\left[\sum_{t=1}^{j} Z_{nt}\right]^2\right) \leq Kn\Delta^2$$

so that for arbitrary $\varepsilon > 0$

$$P\left[\left|n^{-1}\sum_{t=1}^{n} Z_{nt}\right| \geq \varepsilon\right] \leq K\Delta^2/n\varepsilon^2 \to 0$$

as $n \to \infty$, that is, $n^{-1}\sum_{t=1}^{n} Z_{nt} \xrightarrow{P} 0$. $\square$

Proof of theorem 6.3(*a*)

Because $\{M^o_{nt}, F_t\}$ is a martingale difference sequence

$$B^o_n = \text{var}\left(n^{-1/2}\sum_{t=1}^{n} M^o_{nt}\right) = n^{-1}\sum_{t=1}^{n} E(M^o_{nt}M^{o\prime}_{nt})$$

$$= n^{-1}\sum_{t=1}^{n} E(S^o_{nt}S^{o\prime}_{nt}) - n^{-1}\sum_{t=1}^{n} E(S^o_{nt})E(S^{o\prime}_{nt})$$

$$= n^{-1}\sum_{t=1}^{n} E(S^o_{nt}S^{o\prime}_{nt}) - U^o_n,$$

where $U^o_n \equiv n^{-1}\Sigma^n_{t=1} E(S^o_{nt})E(S^{o\prime}_{nt})$. That $\{U^o_n\}$ is $O(1)$ follows immediately from the fact that $|E(S^o_{nt})| < \Delta$ for all $n, t = 1, 2, \ldots$ given assumptions OP′ and DM″. That U^o_n is positive semidefinite follows because U^o_n is the average of positive semidefinite matrices $E(S^o_{nt})E(S^{o\prime}_{nt})$. It follows from the last equation above that

$$\tilde{B}_n - (B^o_n + U^o_n) = n^{-1}\sum_{t=1}^{n} \tilde{S}_{nt}\tilde{S}'_{nt} - n^{-1}\sum_{t=1}^{n} E(S^o_{nt}S^{o\prime}_{nt}).$$

Now

$$\tilde{S}_{nt} \equiv \nabla_\theta\tilde{\psi}'_n\nabla^2_\psi\tilde{g}_n\tilde{q}_{nt} + \nabla_\theta\tilde{q}'_{nt}\nabla_\psi\tilde{g}'_n$$
$$= \nabla_\theta\tilde{\psi}'_n\nabla^2_\psi\tilde{g}_n\tilde{q}_{nt} + (\nabla_\psi\tilde{g}_n \otimes I_k)\tilde{r}_{nt},$$

by applying the equality $\text{vec}(ABC) = (C' \otimes A)\,\text{vec}\,B$ to the second term of $\tilde{S}_{nt}$, and writing $\nabla_\psi\tilde{g}_n \equiv \nabla_\psi g_n(\tilde{\psi}_n)$, $\nabla^2_\psi\tilde{g}_n \equiv \nabla^2_\psi g_n(\tilde{\psi}_n)$, and $\tilde{r}_{nt} = \text{vec}\,\nabla_\theta\tilde{q}'_{nt}$. This can be written more compactly as

$$\tilde{S}_{nt} = \tilde{G}_n\tilde{s}_{nt}$$

where $\tilde{G}_n \equiv [\nabla_\theta\tilde{\psi}'_n\nabla^2_\psi\tilde{g}_n, \nabla_\psi\tilde{g}_n \otimes I_k]$ and $\tilde{s}_{nt} = [\tilde{q}'_{nt}, \tilde{r}'_{nt}]$.

This allows us to write

$$\text{vec}\, n^{-1}\sum_{t=1}^{n} \tilde{S}_{nt}\tilde{S}'_{nt} = \text{vec}\, n^{-1}\sum_{t=1}^{n} \tilde{G}_n\tilde{s}_{nt}\tilde{s}'_{nt}\tilde{G}'_n$$

$$= n^{-1}\sum_{t=1}^{n} (\tilde{G}'_n \otimes \tilde{G}_n)\,\text{vec}\,\tilde{s}_{nt}\tilde{s}'_{nt}$$

$$= (\tilde{G}'_n \otimes \tilde{G}_n)n^{-1}\sum_{t=1}^{n} \text{vec}\,\tilde{s}_{nt}\tilde{s}'_{nt}.$$

Similarly, we can write

$$\text{vec}\, n^{-1}\sum_{t=1}^{n} E(S^o_{nt}S^{o\prime}_{nt}) = (G^o_n \otimes G^o_n)n^{-1}\sum_{t=1}^{n} \text{vec}\, E(s^o_{nt}s^{o\prime}_{nt})$$

where

$$G'_t \equiv [\nabla_\theta\bar{\psi}^{o\prime}_n\nabla^2_\psi g_n(\bar{\psi}^o_n), \nabla_\psi g_n(\bar{\psi}^o_n) \otimes I_k]$$

and, setting $r^o_{nt} = \text{vec}\,\nabla_\theta q^{o\prime}_{nt}$, we write $s^{o\prime}_{nt} = [q^{o\prime}_{nt}, r^{o\prime}_{nt}]$. Thus

$$\text{vec}\,[\tilde{B}_n - (B^o_n + U_n)]$$
$$= (\tilde{G}'_n \otimes \tilde{G}_n)n^{-1}\sum_{t=1}^{n} \text{vec}\,\tilde{s}_{nt}\tilde{s}'_{nt} - (G^{o\prime}_n \otimes G^o_n)n^{-1}\sum_{t=1}^{n} \text{vec}\, E(s^o_{nt}s^{o\prime}_{nt}).$$

The desired result follows from proposition 2.30 of White (1984) provided that $\tilde{G}_n - G^o_n \xrightarrow{P} 0$ and

$$n^{-1}\sum_{t=1}^{n} \text{vec}\,\tilde{s}_{nt}\tilde{s}'_{nt} - n^{-1}\sum_{t=1}^{n} \text{vec}\, E(s^o_{nt}s^{o\prime}_{nt}) \xrightarrow{P} 0$$

for $O(1)$ sequences $\{G^o_n\}$ and $\{n^{-1}\Sigma^n_{t=1}\,\text{vec}\, E(s^o_{nt}s^{o\prime}_{nt})\}$. Assumptions OP′ and DM″ ensure that these sequences are $O(1)$ as required. Further, $\tilde{G}_n - G^o_n \to 0$ a.s. by lemma 3.4 of this work and theorem 2.3 of Domowitz and White (1982) given assumptions DG, OP′, MX, SM(i) and (ii), DM″, NE″, and ID.

It remains to show that

$$n^{-1}\sum_{t=1}^{n} [\text{vec}\,\tilde{s}_{nt}\tilde{s}'_{nt} - \text{vec}\, E(s^o_{nt}s^{o\prime}_{nt})] \xrightarrow{P} 0.$$

To establish this, we apply lemma 3 of Jennrich (1969) and take a mean value expansion of a typical element of $n^{-1}\Sigma^n_{t=1}\tilde{s}_{nt}\tilde{s}'_{nt}$ around θ^o_n. Let $\tilde{s}_{nti}$ be a typical element of $\tilde{s}_{nt}$. Then given assumptions OP′ and ID′

$$n^{-1}\sum_{t=1}^{n} \tilde{s}_{nti}\tilde{s}_{ntj} = n^{-1}\sum_{t=1}^{n} s^o_{nti}s^o_{ntj} + \bar{\pi}_n(\tilde{\theta}_n - \theta^o_n) \quad \text{a.a. } n \quad \text{a.s.}$$

where $\bar{\pi}_n$ is the $l \times k$ gradient

$$\bar{\pi}_n \equiv n^{-1}\sum_{t=1}^{n} \bar{s}_{nti}\nabla_\theta\bar{s}_{ntj} + \bar{s}_{ntj}\nabla_\theta\bar{s}_{nti},$$

where $\bar{s}_{nti}$, $\bar{s}_{ntj}$, $\nabla_\theta\bar{s}_{nti}$, and $\nabla_\theta\bar{s}_{ntj}$ are evaluated at a mean value lying between $\tilde{\theta}_n$ and θ^o_n. (In the mean value expansion, $\tilde{\theta}_n$ is replaced by a tail equivalent sequence, but for convenience we do not change notation.)

Hence

$$n^{-1}\sum_{t=1}^{n}[\tilde{s}_{nti}\tilde{s}_{ntj}-E(s_{nti}s_{ntj})]$$
$$=n^{-1}\sum_{t=1}^{n}[s^o_{nti}s^o_{ntj}-E(s^o_{nti}s^o_{ntj})]+\bar{\pi}_n(\tilde{\theta}_n-\theta^o_n) \quad \text{a.a. } n \quad \text{a.s.}$$

Now $(\tilde{\theta}_n-\theta^o_n)\to 0$ a.s. from theorem 3.19, and because $\tilde{\theta}_n$ is measurable by theorem 2.2, it follows that $(\tilde{\theta}_n-\theta^o_n)$ is $o_p(1)$. Next, $\bar{\pi}_n$ is $O_p(1)$. This follows because a typical element of $\bar{\pi}_n$, say $\bar{\pi}_{nh}=n^{-1}\sum_{t=1}^n \bar{s}_{nti}(\partial/\partial\theta_h)\bar{s}_{ntj}+\bar{s}_{ntj}(\partial/\partial\theta_h)\bar{s}_{nti}$, has

$$\|\bar{\pi}_{nh}\|_r\leqslant n^{-1}\sum_{t=1}^{n}\|\bar{s}_{nti}\|_{2r}\|(\partial/\partial\theta_h)\bar{s}_{ntj}\|_{2r}+\|\bar{s}_{ntj}\|_{2r}\|(\partial/\partial\theta_h)\bar{s}_{nti}\|_{2r}$$
$$\leqslant 2\Delta^2<\infty$$

where the last inequality follows from the domination conditions imposed by assumption DM″. Boundedness in probability then follows from proposition 2.41 of White (1984). Thus, $\bar{\pi}_n(\tilde{\theta}_n-\theta^o_n)$ is $o_p(1)$ by exercise 2.35 of White (1984).

Finally, consider

$$n^{-1}\sum_{t=1}^{n}s^o_{nti}s^o_{ntj}-E(s^o_{nti}s^o_{ntj}).$$

Given assumption DM″,

$$\|s^o_{nti}s^o_{ntj}-E(s^o_{nti}s^o_{ntj})\|_r\leqslant 2\|s^o_{nti}s^o_{ntj}\|_r$$
$$\leqslant 2\|s^o_{nti}\|_{2r}\|s^o_{ntj}\|_{2r}$$
$$\leqslant 2\Delta^2<\infty$$

for all $n,t=1,2,\ldots$. Further, given assumptions DM″ and NE″ it follows from corollary 4.3(b) that $\{s^o_{nti}s^o_{ntj}-E(s^o_{nti}s^o_{ntj})\}$ is near epoch dependent on $\{V_t\}$ of size $-1/2$. By assumption MX, $\{V_t\}$ satisfies the mixing conditions of corollary 4.3(b). It follows from theorem 6.2 that

$$n^{-1}\sum_{t=1}^{n}s^o_{nti}s^o_{ntj}-E(s^o_{nti}s^o_{ntj})\xrightarrow{P}0.$$

Thus

$$n^{-1}\sum_{t=1}^{n}\tilde{s}_{nti}\tilde{s}_{ntj}-E(s^o_{nti}s^o_{ntj})\xrightarrow{P}0$$

and the proof is complete.

Proof of theorem 6.3(b)

The proof is identical to that for theorem 6.3(a), with $\hat{\theta}_n$ replacing $\tilde{\theta}_n$ and θ^*_n replacing θ^o_n. □

Proof of theorem 6.4(a)

Using the same notation as in the proof of theorem 6.3(a), we have that

$$\operatorname{vec}[\tilde{B}_n-(B^o_n+U^o_n)]$$
$$=\Big[(\tilde{G}'_n\otimes\tilde{G}_n)n^{-1}\sum_{t=1}^{n}\operatorname{vec}\tilde{S}_{nt}\tilde{S}'_{nt}-(G^{o\prime}_n\otimes G^o_n)n^{-1}\sum_{t=1}^{n}\operatorname{vec}E(S^o_{nt}S^{o\prime}_{nt})\Big]$$
$$-\sum_{\tau=1}^{m}\Big[(\tilde{G}'_n\otimes\tilde{G}_n)n^{-1}\sum_{t=\tau+1}^{n}\operatorname{vec}\tilde{S}_{nt}\tilde{S}'_{n,t-\tau}-(G^{o\prime}_n\otimes G^o_n)n^{-1}\sum_{t=1}^{n}\operatorname{vec}E(S^o_{nt}S^{o\prime}_{n,t-\tau})\Big]$$
$$+\sum_{\tau=1}^{m}\Big[(\tilde{G}'_n\otimes\tilde{G}_n)n^{-1}\sum_{t=\tau+1}^{n}\operatorname{vec}\tilde{S}_{n,t-\tau}\tilde{S}'_{nt}-(G^{o\prime}_n\otimes G^o_n)n^{-1}\sum_{t=1}^{n}\operatorname{vec}E(S^o_{n,t-\tau}S^{o\prime}_{nt})\Big].$$

The argument is identical to that for the proof of theorem 6.3(a) with $\tilde{S}_{nt}\tilde{S}'_{n,t-\tau}$ replacing $\tilde{S}_{nt}\tilde{S}'_{nt}$ and $S^o_{nt}S^{o\prime}_{n,t-\tau}$ replacing $S^o_{nt}S^{o\prime}_{nt}$, for $\tau=0,1,\ldots,m<\infty$. This establishes that all the terms in square brackets above vanish in probability, so that

$$\tilde{B}_n-(B^o_n+U^o_n)\xrightarrow{P}0.$$

Proof of theorem 6.4(b)

The proof is identical to that for theorem 6.4(a), with $\hat{\theta}_n$ replacing $\tilde{\theta}_n$ and θ^*_n replacing θ^o_n. □

Proof of lemma 6.5

Given a double array $\{Z_{nt}, t = 1,\dots,n, n = 1,2,\dots\}$, define

$$\underset{(n+m_n)\times(m_n+1)}{Z^n} = \begin{bmatrix} Z_{n,1} & 0 & 0 & \dots & 0 \\ Z_{n,2} & Z_{n,1} & 0 & \dots & 0 \\ Z_{n,3} & Z_{n,2} & Z_{n,1} & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ Z_{n,n} & Z_{n,n-1} & Z_{n,n-2} & \dots & Z_{n,n-m_n} \\ 0 & Z_{n,n} & Z_{n,n-1} & \dots & Z_{n,n-m_n+1} \\ 0 & 0 & Z_{n,n} & \dots & Z_{n,n-m_n+2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & Z_{n,n} \end{bmatrix}$$

and define $a^n = (a_{n,1}, a_{n,2}, \dots, a_{n,m_n+1})'$. We show that $\gamma_n = a^{n\prime} Z^{n\prime} Z^n a^n \geq 0$. Now

$$Z^{n\prime} Z^n = [z^n_{ij}], \quad i,j = 1,\dots,m_n+1,$$

$$z^n_{ij} \equiv \sum_{t=|i-j|+1}^{n} Z_{n,t} Z_{n,t-|i-j|} = z^n_{ji}$$

so that

$$\begin{aligned} a^{n\prime} Z^{n\prime} Z^n a^n &= \sum_{i=1}^{m_n+1} \sum_{j=1}^{m_n+1} a_{ni} a_{nj} z^n_{ij} \\ &= \sum_{i=1}^{m_n+1} a_{ni}^2 z^n_{ii} + 2 \sum_{\tau=1}^{m_n} \sum_{i=\tau+1}^{m_n+1} a_{ni} a_{n,i-\tau} z^n_{i,i-\tau} \\ &= \sum_{i=1}^{m_n+1} a_{ni}^2 \left(\sum_{t=1}^{n} Z_{nt}^2 \right) + 2 \sum_{\tau=1}^{m_n} \sum_{i=\tau+1}^{m_n+1} a_{ni} a_{n,i-\tau} \\ &\quad \times \left(\sum_{t=\tau+1}^{n} Z_{nt} Z_{n,t-\tau} \right) \\ &= w_{no} \sum_{t=1}^{n} Z_{nt}^2 + 2 \sum_{\tau=1}^{m_n} w_{n\tau} \sum_{t=\tau+1}^{n} Z_{nt} Z_{n,t-\tau} \\ &= \gamma_n, \end{aligned}$$

with

$$w_{n\tau} \equiv \sum_{i=\tau+1}^{m_n+1} a_{ni} a_{n,i-\tau}.$$

Hence $\gamma_n \geq 0$. □

Proof of lemma 6.6

Given the definitions of B_n and $\ddot{B}_n$, we have

$$\begin{aligned} B_n - \ddot{B}_n &= (1-w_{no}) n^{-1} \sum_{t=1}^{n} E(Z_{nt} Z'_{nt}) \\ &\quad + \sum_{\tau=1}^{m_n} (1-w_{\tau n}) n^{-1} \sum_{t=\tau+1}^{n} [E(Z_{nt} Z'_{n,t-\tau}) + E(Z_{n,t-\tau} Z'_{nt})] \\ &\quad + n^{-1} \sum_{\tau=m_n+1}^{n-1} \sum_{t=\tau+1}^{n} [E(Z_{nt} Z'_{n,t-\tau}) + E(Z_{n,t-\tau} Z'_{nt})]. \end{aligned}$$

Given that $\|Z'_{nt} Z_{nt}\|_{r/2} \leq \Delta$, $r > 2$, it follows that $\{n^{-1} \sum_{t=1}^{n} E(Z_{nt} Z'_{nt})\}$ is $O(1)$, so that $w_{no} \to 1$ implies that the first term above vanishes as $n \to \infty$. The result follows by showing that the second and third terms vanish.

Let $\xi^{ij}_{nt\tau} \equiv Z_{nti} Z_{n,t-\tau,j}$ be a typical element of the matrix $Z_{nt} Z'_{n,t-\tau}$. To show that the second term vanishes as $n \to \infty$, it suffices to show that for $i,j = 1,\dots,k$

$$\sum_{\tau=1}^{m_n} (1-w_{\tau n}) n^{-1} \sum_{t=\tau+1}^{n} E(\xi^{ij}_{nt\tau}) \to 0.$$

Now

$$\left| \sum_{\tau=1}^{m_n} (1-w_{\tau n}) n^{-1} \sum_{t=\tau+1}^{n} E(\xi^{ij}_{nt\tau}) \right| \leq \sum_{\tau=1}^{m_n} |1-w_{\tau n}| n^{-1} \sum_{t=\tau+1}^{n} |E(\xi^{ij}_{nt\tau})|.$$

Letting $Y_{nt\tau j} \equiv E^{t-\tau+[\tau/2]}_{t-\tau-[\tau/2]}(Z_{n,t-\tau,j})$ we have

$$\begin{aligned} |E(\xi^{ij}_{nt\tau})| &= |E(Z_{nti} Z_{n,t-\tau,j})| \\ &= |E(Z_{nti} Y_{nt\tau j} + Z_{nti}(Z_{n,t-\tau,j} - Y_{nt\tau j}))| \\ &\leq |E(Z_{nti} Y_{nt\tau j})| + |E(Z_{nti}(Z_{n,t-\tau,j} - Y_{nt\tau j}))|. \end{aligned}$$

Now the fact that $Y_{nt\tau j}$ is measurable-$F^{t-\tau+[\tau/2]}$ implies

$$\begin{aligned} |E(Z_{nti} Y_{nt\tau j})| &= |E(E(Z_{nti} Y_{nt\tau j} | F^{t-\tau+[\tau/2]}))| \\ &= |E(E(Z_{nti} | F^{t-\tau+[\tau/2]}) Y_{nt\tau j})| \\ &\leq \| E(Z_{nti} | F^{t-\tau+[\tau/2]}) \|_2 \, \| Y_{nt\tau j} \|_2. \end{aligned}$$

By the law of iterated expectation and the conditional Jensen's inequality $\|Y_{n\tau j}\|_2 \leqslant \|Z_{n,t-\tau,j}\|_2$ and by Jensen's inequality $\|Z_{n,t-\tau,j}\|_2 \leqslant \|Z_{n,t-\tau,j}\|_r \leqslant \Delta < \infty$. Next, equations (3.3) or (3.4) imply

$$\begin{aligned}\|E^{t-\tau+[\tau/2]}(Z_{nti})\|_2 &\leqslant 2\phi_{[\tau/4]}^{1-1/r}\|Z_{nti}\|_r + v_{[\tau/4]} \quad \text{or} \\ &\leqslant 5\alpha_{[\tau/4]}^{1/2-1/r}\|Z_{nti}\|_r + v_{[\tau/4]}.\end{aligned}$$

Thus

$$\begin{aligned}|E(Z_{nti}Y_{n\tau})| &\leqslant \Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + v_{[\tau/4]}) \quad \text{or} \\ &\leqslant \Delta(5\Delta\alpha_{[\tau/4]}^{1/2-1/r} + v_{[\tau/4]}).\end{aligned}$$

Next, equation (3.5) implies

$$\begin{aligned}|E(Z_{nti}(Z_{n,t-\tau,j} - Y_{n\tau j}))| &\leqslant \|Z_{nti}\|_2\|Z_{n,t-\tau,j} - Y_{n\tau j}\|_2 \\ &\leqslant \Delta v_{[\tau/2]} \\ &\leqslant \Delta v_{[\tau/4]}.\end{aligned}$$

The last inequality follows because v_m is decreasing in m. Collecting the inequalities above yields

$$\begin{aligned}|E(\xi_{nt\tau}^{ij})| &\leqslant \Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) \quad \text{or} \\ &\leqslant \Delta(5\Delta\alpha_{[\tau/4]}^{1/2-1/r} + 2v_{[\tau/4]}).\end{aligned}$$

This implies

$$\begin{aligned}n^{-1}\sum_{t=\tau+1}^{n}|E(\xi_{nt\tau}^{ij})| &\leqslant n^{-1}\sum_{t=\tau+1}^{n}\Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) \\ &\leqslant \Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) \quad \text{or} \\ &\leqslant \Delta(5\Delta\alpha_{[\tau/4]}^{1/2-1/r} + 2v_{[\tau/4]}).\end{aligned}$$

It follows that

$$\begin{aligned}&\left|\sum_{\tau=1}^{m_n}(1-w_{n\tau})n^{-1}\sum_{t=\tau+1}^{n}E(\xi_{nt\tau}^{ij})\right| \\ &\quad \leqslant \sum_{\tau=1}^{m_n}|1-w_{n\tau}|\Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) \quad \text{or} \\ &\quad \leqslant \sum_{\tau=1}^{m_n}|1-w_{n\tau}|\Delta(5\Delta\alpha_{[\tau/4]}^{1-2/r} + 2v_{[\tau/4]}).\end{aligned}$$

Taking limits on both sides gives

$$\begin{aligned}\lim_{n\to\infty}&\left|\sum_{\tau=1}^{m_n}(1-w_{n\tau})n^{-1}\sum_{t=\tau+1}^{n}E(\xi_{nt\tau}^{ij})\right| \\ &\leqslant \lim_{n\to\infty}\sum_{\tau=1}^{m_n}|1-w_{n\tau}|\Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) \quad \text{or} \\ &\leqslant \lim_{n\to\infty}\sum_{\tau=1}^{m_n}|1-w_{n\tau}|\Delta(5\Delta\alpha_{[\tau/4]}^{1/2-1/r} + 2v_{[\tau/4]}).\end{aligned}$$

Now

$$\sum_{\tau=1}^{m_n}|1-w_{n\tau}|\Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) = \int_0^\infty f_n(\tau)\,d\mu(\tau)$$

where μ is counting measure on the positive integers and

$$f_n(\tau) \equiv 1_{[\tau \leqslant m_n]}|1-w_{n\tau}|\Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}).$$

We apply the dominated convergence theorem as in Newey and West (1987) to show that $\int_0^\infty f_n(\tau)\,d\mu(\tau)$ converges to zero. Now for each $\tau \in \mathbb{N}$, the requirement that $w_{n\tau} \to 1$ ensures that $f_n(\tau) \to 0$ as $n \to \infty$. Further, because $|w_{n\tau}| \leqslant \Delta$, $|f_n(\tau)| \leqslant |\bar{f}(\tau)|$ for all τ and n, where

$$\bar{f}(\tau) = (\Delta+1)\Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]})$$

is integrable, given the size conditions imposed on ϕ_m and v_m which ensure the finiteness of the sums involved in

$$\int_0^\infty \bar{f}(\tau)\,d\mu(\tau) = \sum_{\tau=1}^{\infty}(\Delta+1)\Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}).$$

It now follows from the dominated convergence theorem (e.g. Bartle 1966, theorem 5.6) that as $n \to \infty$

$$\int_0^\infty f_n(\tau)\,d\mu(\tau) \to 0.$$

A similar argument applies with $\alpha_{[\tau/4]}$ replacing $\phi_{[\tau/4]}$, so that

$$\left|\sum_{\tau=\tau+1}^{n}(1-w_{n\tau})n^{-1}\sum_{t=\tau+1}^{n}E(\xi_{nt\tau}^{ij})\right| \to 0 \quad \text{as } n \to \infty,$$

thus ensuring the convergence to zero of the second term in the expression for $B_n - \ddot{B}_n$.

Now consider the third term,

$$n^{-1}\sum_{\tau=m_n+1}^{n-1}\sum_{t=\tau+1}^{n}[E(Z_{nt}Z'_{n,t-\tau}) + E(Z_{n,t-\tau}Z'_{nt})].$$

It suffices to show that

$$\left| n^{-1} \sum_{\tau=m_n+1}^{n-1} \sum_{t=\tau+1}^{n} E(\xi_{nt\tau}^{ij}) \right| \to 0 \quad \text{as } n \to \infty.$$

Now

$$\begin{aligned} \left| n^{-1} \sum_{\tau=m_n+1}^{n-1} \sum_{t=\tau+1}^{n} E(\xi_{nt\tau}^{ij}) \right| &\leqslant \sum_{\tau=m_n+1}^{n-1} n^{-1} \sum_{t=\tau+1}^{n} |E(\xi_{nt\tau}^{ij})| \\ &\leqslant \sum_{\tau=m_n+1}^{n-1} n^{-1} \sum_{t=\tau+1}^{n} \Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) \quad \text{or} \\ &\leqslant \sum_{\tau=m_n+1}^{n-1} n^{-1} \sum_{t=\tau+1}^{n} \Delta(5\Delta\alpha_{[\tau/4]}^{1/2-1/r} + 2v_{[\tau/4]}), \end{aligned}$$

using the same inequalities as above.

Because

$$\begin{aligned} &\sum_{\tau=m_n+1}^{n-1} n^{-1} \sum_{t=\tau+1}^{n} \Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) \\ &\quad = \sum_{\tau=1}^{n-1} n^{-1} \sum_{t=\tau+1}^{n} \Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) \\ &\quad\quad - \sum_{\tau=1}^{m_n} n^{-1} \sum_{t=\tau+1}^{n} \Delta(2\Delta\phi_{[\tau/4]}^{1-1/r} + 2v_{[\tau/4]}) \end{aligned}$$

and because the size requirements on ϕ_m and v_m ensure the convergence of the two sums on the right above to the same limit provided $m_n \to \infty$ as $n \to \infty$, we have

$$\lim_{n\to\infty} \left| n^{-1} \sum_{t=\tau+1}^{n} E(\xi_{nt\tau}^{ij}) \right| = 0.$$

A similar argument applies with $\alpha_{[\tau/4]}$ replacing $\phi_{[\tau/4]}$, so it follows that the third term in the expression for $B_n - \ddot{B}_n$ converges to zero, and the proof is complete. □

Proof of lemma 6.7(*a*)

For notational convenience in what follows, we write $\xi_{t\tau}$ in place of $\xi_{nt\tau}^{ij}$.

For fixed τ and all $m > 6\tau$, set

$$\tilde{\xi}_{t-m+\tau,\tau} \equiv E_{t-m+\tau-[m/2]}^{t-m+\tau+[m/2]}(\xi_{t-m+\tau,\tau})$$

so that

$$\begin{aligned} |E(\xi_{t\tau}\xi_{t-m+\tau,\tau})| &= |E(\xi_{t\tau}\tilde{\xi}_{t-m+\tau,\tau} + \xi_{t\tau}(\xi_{t-m+\tau,\tau} - \tilde{\xi}_{t-m+\tau,\tau}))| \\ &\leqslant |E(\xi_{t\tau}\tilde{\xi}_{t-m+\tau,\tau})| + |E(\xi_{t\tau}(\xi_{t-m+\tau,\tau} - \tilde{\xi}_{t-m+\tau,\tau}))|. \end{aligned}$$

Because $\tilde{\xi}_{t-m+\tau,\tau}$ is measurable-$F^{t-n+\tau+[m/2]}$

$$\begin{aligned} |E(\xi_{t\tau}\tilde{\xi}_{t-m+\tau,\tau})| &\leqslant \|E^{t-m+\tau-[m/2]}(\xi_{t\tau})\|_2 \, \|\tilde{\xi}_{t-m+\tau,\tau}\|_2 \\ &\leqslant \Delta\|E^{t-m+\tau+[m/2]}(\xi_{t\tau})\|_2 \end{aligned}$$

using the same logic as in the proof of lemma 6.6 and using the fact that $\|\tilde{\xi}_{t-n+\tau,\tau}\|_2 \leqslant \|\xi_{t-m+\tau,\tau}\|_2 \leqslant \|\xi_{t-m+\tau,\tau}\|_r \leqslant \Delta$ for $r > 2$ given $\|Z'_{nt}Z_{nt}\|_r \leqslant \Delta$. By the same reasoning as in the proof of lemma 3.14

$$\begin{aligned} \|E^{t-m+\tau+[m/2]}(\xi_{t\tau})\|_2 &\leqslant 2\Delta\phi_{l'_{m\tau}}^{1-1/r} + \|\xi_{t\tau} - E_{t-l'_{m\tau}}^{t+l'_{m\tau}}(\xi_{t\tau})\|_2 \quad \text{or} \\ &\leqslant 5\Delta\alpha_{l'_{m\tau}}^{1/2-1/r} + \|\xi_{t\tau} - E_{t-l'_{m\tau}}^{t+l'_{m\tau}}(\xi_{t\tau})\|_2 \end{aligned}$$

where $l'_{m\tau} = [([m/2]-\tau)/2]$. Let $l_{m\tau} \equiv [([m/2]-3\tau)/2] \leqslant l'_{m\tau}$. Applying lemma 4.1 with the same choices for b and B as in the proof of corollary 4.3(b) with Y_{nt} corresponding to $Z_{n,t-\tau,j}$ and Z_{nt} corresponding to Z_{nti} yields

$$\|\xi_{t\tau} - E_{t-l'_{m\tau}}^{t+l'_{m\tau}}(\xi_{t\tau})\|_2 \leqslant \|\xi_{t\tau} - E_{t-l_{m\tau}}^{t+l_{m\tau}}(\xi_{t\tau})\|_2 \leqslant K_1 v_{l_{m\tau}}^{(r-2)/2(r-1)},$$

where $K_1 < \infty$ is a finite constant given the available moment conditions. Because ϕ_m and α_m are nonincreasing in m and because $l'_{m\tau} \leqslant l_{m\tau}$, we have

$$\begin{aligned} \|E^{t-m-\tau+[m/2]}(\xi_{t\tau})\|_2 &\leqslant \Delta 2\phi_{l'_{m\tau}}^{1-1/r} + K_1 v_{l_{m\tau}}^{(r-2)/2(r-1)} \\ &\leqslant \Delta 2\phi_{l_{m\tau}}^{1-1/r} + K_1 v_{l_{m\tau}}^{(r-2)/2(r-1)} \quad \text{or} \\ &\leqslant \Delta 5\alpha_{l'_{m\tau}}^{1/2-1/r} + K_1 v_{l_{m\tau}}^{(r-2)/2(r-1)} \\ &\leqslant \Delta 5\alpha_{l_{m\tau}}^{1/2-1/r} + K_1 v_{l_{m\tau}}^{(r-2)/2(r-1)}. \end{aligned}$$

Next,

$$\begin{aligned} |E(\xi_{t\tau}(\xi_{t-m+\tau,\tau} - \tilde{\xi}_{t-m+\tau,\tau}))| &\leqslant \|\xi_{t\tau}\|_2 \|\xi_{t-m+\tau,\tau} - \tilde{\xi}_{t-m+\tau,\tau}\|_2 \\ &\leqslant \Delta\|\xi_{t-m+\tau,\tau} - \tilde{\xi}_{t-m+\tau,\tau}\|_2. \end{aligned}$$

By reasoning similar to the proof of corollary 4.3(b), we obtain

$$\begin{aligned} \|\xi_{t-m+\tau,\tau} - \tilde{\xi}_{t-m+\tau,\tau}\|_2 &\leqslant K_2 v_{[m/2]-\tau}^{(r-2)/2(r-2)} \\ &\leqslant K_2 v_{l_{m\tau}}^{(r-2)/2(r-2)} \end{aligned}$$

where the second inequality follows because v_m is decreasing in m and

$[m/2]-\tau \geqslant l_{m\tau}$. Combining the inequalities above yields the desired result

$$|E(\xi_{t\tau}\xi_{t-m+\tau,\tau})| \leqslant K_o(\phi_{l_{m\tau}}^{1-1/r}+v_{l_{m\tau}}^{(r-2)/2(r-2)}) \quad \text{or}$$
$$\leqslant K_o(\alpha_{l_{m\tau}}^{1/2-1/r}+v_{l_{m\tau}}^{(r-2)/2(r-2)})$$

for $K_o < \infty$ sufficiently large.

Proof of lemma 6.7(b)

By the triangle inequality, for fixed τ

$$E\left(\left[\sum_{t=\tau+1}^{n}\xi_{t\tau}\right]^2\right) = E\left(\sum_{t=\tau+1}^{n}\xi_{t\tau}^2+2\sum_{m=\tau+1}^{n-1}\sum_{t=m+1}^{n}\xi_{t\tau}\xi_{t-m+\tau,\tau}\right)$$
$$\leqslant \sum_{t=\tau+1}^{n}E(\xi_{t\tau}^2)+2\sum_{m=\tau+1}^{n-1}\sum_{t=m+1}^{n}|E(\xi_{t\tau}\xi_{t-m+\tau,\tau})|.$$

By lemma 6.7(a) for $m > 6\tau$

$$|E(\xi_{t\tau}\xi_{t-m+\tau,\tau})| \leqslant K_o(\phi_{l_{m\tau}}^{1-1/r}+v_{l_{m\tau}}^{(r-2)/2(r-2)}) \quad \text{or}$$
$$\leqslant K_o(\alpha_{l_{m\tau}}^{1/2-1/2}+v_{l_{m\tau}}^{(r-2)/2(r-2)}),$$

where $l_{m\tau} = [([m/2]-3\tau)/2]$, while for $\tau \leqslant m \leqslant 6\tau$

$$|E(\xi_{t\tau}\xi_{t-m+\tau,\tau})| \leqslant \|\xi_{t\tau}\|_2\,\|\xi_{t-m+\tau,\tau}\|_2$$
$$\leqslant \Delta^2 < \infty.$$

Thus, for $\Delta' < \infty$ sufficiently large

$$E\left(\left[\sum_{t=\tau+1}^{n}\xi_{t\tau}\right]^2\right) \leqslant n\Delta'(\tau+1)+n2K_o\sum_{m=6\tau+1}^{n-1}(\phi_{l_{m\tau}}^{1-1/r}+v_{l_{m\tau}}^{(r-2/2(r-2)}) \quad \text{or}$$
$$\leqslant n\Delta'(\tau+1)+n2K_o\sum_{m=6\tau+1}^{n-1}(\alpha_{l_{m\tau}}^{1/2-1/r}+v_{l_{m\tau}}^{(r-2)/2(r-2)}).$$

Now

$$\sum_{m=6\tau+1}^{n-1}(\phi_{l_{m\tau}}^{1-1/r}+v_{l_{m\tau}}^{(r-2)/2(r-2)}) = \sum_{m=1}^{n-6\tau-1}(\phi_{l'_{m\tau}}^{1-1/r}+v_{l'_{m\tau}}^{(r-2)/2(r-2)})$$

where now $l'_{m\tau} = [m/4]$, so that

$$\sum_{m=6\tau+1}^{n-1}(\phi_{l_{m\tau}}^{1-1/r}+v_{l_{m\tau}}^{(r-2)/2(r-2)}) \leqslant \sum_{m=1}^{\infty}(\phi_{[m/4]}^{1-1/r}+v_{[m/4]}^{(r-2)/2(r-2)}).$$

Similarly,

$$\sum_{m=6\tau+1}^{n-1}(\alpha_{l_{m\tau}}^{1/2-1/r}+v_{l_{m\tau}}^{(r-2)/2(r-2)}) \leqslant \sum_{m=1}^{\infty}(\alpha_{[m/4]}^{1/2-1/r}+v_{[m/4]}^{(r-2)/2(r-2)}).$$

Given the size conditions on v_m and ϕ_m or α_m, it follows that one of these sums is finite. It follows that there exists $\Delta_o < \infty$ sufficiently large that for all $\tau \geqslant 0$

$$E\left(\left[\sum_{t=\tau+1}^{n}\xi_{t\tau}\right]^2\right) \leqslant (\tau+2)n\Delta_o.$$

Proof of lemma 6.7(c)

By the implication rule

$$P\left[\sum_{\tau=1}^{m}\left|n^{-1}\sum_{t=\tau+1}^{n}\xi_{t\tau}\right| \geqslant \varepsilon\right] \leqslant \sum_{\tau=1}^{m}P\left[\left|n^{-1}\sum_{t=\tau+1}^{n}\xi_{t\tau}\right| \geqslant \varepsilon/m\right].$$

By Chebyshev's inequality

$$P\left[\left|n^{-1}\sum_{t=\tau+1}^{n}\xi_{t\tau}\right| \geqslant \varepsilon/m\right] \leqslant E\left(\left[\sum_{t=\tau+1}^{n}\xi_{t\tau}\right]^2\right)m^2/\varepsilon^2n^2.$$

From lemma 6.7(b), $E([\Sigma_{t=\tau+1}^{n}\xi_{t\tau}]^2) \leqslant (\tau+2)n\Delta_o \leqslant (m+2)n\Delta_o$ given that $\tau \leqslant m$. Hence

$$P\left[\left|n^{-1}\sum_{t=\tau+1}^{n}\xi_{t\tau}\right| \geqslant \varepsilon/m\right] \leqslant \Delta_o m^3/\varepsilon^2 n+2\Delta_o m^2/\varepsilon^2 n.$$

It follows that

$$P\left[\sum_{\tau=1}^{m}\left|n^{-1}\sum_{t=\tau+1}^{n}\xi_{t\tau}\right| \geqslant \varepsilon\right] \leqslant \Delta_o m^4/\varepsilon^2 n+2\Delta_o m^3/\varepsilon^2 n.$$

Proof of lemma 6.7(d)

By the triangle inequality and because $|w_{n\tau}| \leqslant \Delta$

$$P\left[\left|n^{-1}\sum_{\tau=1}^{m_n}w_{n\tau}\sum_{t=\tau+1}^{n}\xi_{t\tau}\right| \geqslant \varepsilon\right] \leqslant P\left[\sum_{\tau=1}^{m_n}\Delta\left|n^{-1}\sum_{t=\tau+1}^{n}\xi_{t\tau}\right| \geqslant \varepsilon\right]$$
$$= P\left[\sum_{\tau=1}^{m_n}\left|n^{-1}\sum_{t=\tau+1}^{n}\xi_{t\tau}\right| \geqslant \varepsilon/\Delta\right]$$

where we assume $\Delta > 0$ without loss of generality. From lemma 6.7(c) we have

$$P\left[\sum_{\tau=1}^{m_n}\left|n^{-1}\sum_{t=\tau+1}^{n}\xi_{t\tau}\right|\geqslant \varepsilon/\Delta\right]\leqslant \Delta^2(\Delta_o m_n^4+2m_n^3)/\varepsilon^2 n.$$

Because $m_n = o(n^{1/4})$ we have $m_n^4/n \to 0$ and $m_n^3/n \to 0$. Hence

$$n^{-1}\sum_{\tau=1}^{m_n} w_{n\tau}\sum_{t=\tau+1}^{n}\xi_{t\tau}\xrightarrow{P}0. \qquad \square$$

Proof of theorem 6.8(*a*)

That $\tilde{B}_n$ is positive semidefinite for all $n = 1, 2, \ldots$ follows immediately from lemma 6.5 given assumption WT. Define

$$\ddot{B}_n^o = \left(w_{no}n^{-1}\sum_{t=1}^{n}E(S_{nt}^o S_{nt}^{o\prime}) + n^{-1}\sum_{\tau=1}^{m_n} w_{n\tau}\sum_{t=\tau+1}^{n}[E(S_{nt}^o S_{n,t-\tau}^{o\prime})+E(S_{n,t-\tau}^o S_{nt}^{o\prime})]\right)-U_n^o.$$

That U_n^o is positive semidefinite follows immediately from lemma 6.5 given assumption WT. Given assumptions DG, OP′, MX′, DM″, NE″, TL, and WT the conditions of lemma 6.6 are satisfied, and it follows that

$$\ddot{B}_n^o - B_n^o \to 0 \quad \text{as } n \to \infty.$$

The desired result follows by showing that

$$\tilde{B}_n-(\ddot{B}_n^o+U_n^o)\xrightarrow{P}0.$$

Using the same notation as in the proof of theorem 6.3, we have that

$$\begin{aligned}\operatorname{vec}[\tilde{B}_n-(\ddot{B}_n^o+U_n^o)] &= \left[(\tilde{G}_n'\otimes\tilde{G}_n)w_{no}n^{-1}\sum_{t=1}^{n}\operatorname{vec}\tilde{s}_{nt}\tilde{s}_{nt}' - (G_n^{o\prime}\otimes G_n^o)w_{no}n^{-1}\sum_{t=1}^{n}\operatorname{vec}E(s_{nt}^o s_{nt}^{o\prime})\right]\\ &+(\tilde{G}_n'\otimes\tilde{G}_n)\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}\operatorname{vec}\tilde{s}_{nt}\tilde{s}_{n,t-\tau}'\\ &-(G_n^{o\prime}\otimes G_n^o)\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}\operatorname{vec}E(s_{nt}^o s_{n,t-\tau}^{o\prime})\\ &+(\tilde{G}_n'\otimes\tilde{G}_n)\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}\operatorname{vec}\tilde{s}_{n,t-\tau}\tilde{s}_{nt}'\\ &-(G_n^{o\prime}\otimes G_n^o)\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}\operatorname{vec}E(s_{n,t-\tau}^o s_{nt}^{o\prime}).\end{aligned}$$

The first term converges to zero by argument identical to that used in the proof of theorem 6.3. The desired result follows provided that $\tilde{G}_n-G_n^o\xrightarrow{P}0$, which is valid under the conditions given as previously argued, and if

$$\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}[\operatorname{vec}\tilde{s}_{nt}\tilde{s}_{n,t-\tau}'-\operatorname{vec}E(s_{nt}^o s_{n,t-\tau}^{o\prime})]\xrightarrow{P}0.$$

Taking a mean value expansion of a typical element around θ_n^o (interior to Θ by assumption ID′) gives

$$\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}\tilde{s}_{nti}\tilde{s}_{n,t-\tau,j}=\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}s_{nti}^o s_{n,t-\tau,j}^o + \bar{\pi}_n(\tilde{\theta}_n-\theta_n^o) \quad \text{a.a.}n \quad \text{a.s.}$$

where $\bar{\pi}_n$ is the $1\times k$ gradient

$$\bar{\pi}_n\equiv\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{\tau+1}^{n}\bar{s}_{nti}\nabla_\theta\bar{s}_{n,t-\tau,j}+\bar{s}_{ntj}\nabla_\theta\bar{s}_{nti},$$

where $\bar{s}_{nti}$, $\bar{s}_{ntj}$, $\nabla_\theta\bar{s}_{nti}$, $\nabla_\theta\bar{s}_{ntj}$ are evaluated at a mean value lying between $\tilde{\theta}_n$ and θ_n^o. (As before, $\tilde{\theta}_n$ is replaced by a tail equivalent sequence, but the notation is unchanged.) Hence

$$\begin{aligned}&\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}[\tilde{s}_{nti}\tilde{s}_{n,t-\tau,j}-E(s_{nti}^o s_{n,t-\tau,j}^o)]\\ &=\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}[s_{nti}^o s_{n,t-\tau,j}^o-E(s_{nti}^o s_{n,t-\tau,j}^o)]\\ &\quad+n^{-1/2}\bar{\pi}_n\sqrt{(n)}(\tilde{\theta}_n-\theta_n^o).\end{aligned}$$

Given assumptions DG, OP′, MX′, DM″, NE″, TL, and WT, lemma 6.7(d) applies with $\xi_{nt\tau}^{ij}\equiv s_{nti}^o s_{n,t-\tau,j}^o-E(s_{nti}^o s_{n,t-\tau,j}^o)$ to yield

$$\sum_{\tau=1}^{m_n}w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}[s_{nti}^o s_{n,t-\tau,j}^o-E(s_{nti}^o s_{n,t-\tau,j}^o)]\xrightarrow{P}0.$$

The desired result follows if $n^{-1/2}\bar{\pi}_n\sqrt{(n)}(\tilde{\theta}_n-\theta_n^o)$ is $o_p(1)$. Now $\sqrt{(n)}(\tilde{\theta}_n-\theta_n^o)$ is $O_p(1)$ by assumption, so it suffices that $n^{-1/2}\bar{\pi}_n$ is $o_p(1)$. Consider a typical element of $\bar{\pi}_n$, say

$$\bar{\pi}_{nh}=\sum_{\tau=1}^{m_n} w_{n\tau}n^{-1}\sum_{t=\tau+1}^{n}\bar{s}_{nti}(\partial/\partial\theta_h)\bar{s}_{n,t-\tau,j}+\bar{s}_{n,t-\tau,j}(\partial/\partial\theta_h)\bar{s}_{nti}.$$

Now assumption DM″ ensures that for $\Delta_o<\infty$ sufficiently large

$$\begin{aligned}\|m_n^{-1}\bar{\pi}_{nh}\|_r &\leqslant m_n^{-1}\sum_{\tau=1}^{m_n}\Delta_o n^{-1}\sum_{t=\tau+1}^{n}\|\bar{s}_{nti}\|_{2r}\|(\partial/\partial\theta_h)\bar{s}_{n,t-\tau,j}\|_{2r}\\ &\quad+\|\bar{s}_{n,t-\tau,j}\|_{2r}\|(\partial/\partial\theta_h)\bar{s}_{nti}\|_{2r}\\ &\leqslant m_n^{-1}\sum_{\tau=1}^{m_n}\Delta_o n^{-1}\sum_{t=\tau+1}^{n}2\Delta_o^2\\ &\leqslant m_n^{-1}\sum_{\tau=1}^{m_n}2\Delta_o^3\\ &=2\Delta_o^3,\end{aligned}$$

which implies that $m_n^{-1}\bar{\pi}_{nh}$ is $O_p(1)$. It follows that $n^{-1/2}\bar{\pi}_n = n^{-1/2}m_n(m_n^{-1}\bar{\pi}_{nh})$ is $o_p(1)$, given that m_n is $o(n^{1/4})$, as ensured by assumption TL. The proof is now complete.

Proof of theorem 6.8(*b*)

The proof is identical to that of part (a), except that $\tilde{\theta}_n$ is replaced by $\hat{\theta}_n$, θ_n^o is replaced by θ_n^*, and $\sqrt{(n)}(\hat{\theta}_n-\theta_n^*)$ is $O_p(1)$ under the conditions given as a consequence of theorem 5.7. □

REFERENCES

Anderson, T. W. 1971: *The Statistical Analysis of Time Series.* New York: John Wiley and Sons.

Andrews, D. W. K. 1987: Laws of large numbers for dependent non-identically distributed random variables, California Institute of Technology, Division of Humanities and Social Sciences, unpublished paper.

Bartle, R. 1966: *The Elements of Integration.* New York: John Wiley and Sons.

Bartlett, M. S. 1950: Periodogram analysis and continuous spectra, *Biometrika* 37, 1–16.

Chow, G. 1981: Selection of econometric models by the information criterion, in E. G. Charatsis (ed.), Proceedings of the Econometric Society European Meeting 1979, Amsterdam: North Holland, 199–214.

Domowitz, I. and H. White 1982: Misspecified models with dependent observations, *Journal of Econometrics* 20, 35–58.

Eicker, F. 1967: Limit theorems for regressions with unequal and dependent errors, in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, Berkeley: University of California Press, 59–82.

Hansen, L. P. 1982: Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–54.

Jennrich, R. I. 1969: Asymptotic properties of non-linear least squares estimators, *Annals of Mathematical Statistics* 40, 633–43.

Newey, W. and K. West 1987: A simple positive definite heteroskedasticity and correlation consistent covariance matrix *Econometrica* 55, 703–8.

Nicholls, D. F. and A. Pagan 1983: Heteroskedasticity in models with lagged dependent variables, *Econometrica* 51, 1233–42.

Phillips, P. C. B. 1985: Personal correspondence.

Stone, M. 1974: Cross-validitory choice and assessment of statistical prediction, *Journal of the Royal Statistical Association*, series B 36, 111–47.

White, H. 1980: A heteroskedasticity consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–38.

White, H. 1983: Corrigendum, *Econometrica* 51, 513.

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