5 Asymptotic Normality

In this chapter we present results which allow us to state conditions ensuring the asymptotic normality of the unconstrained optimization estimators considered in chapters 2 and 3. These results provide the basis for constructing confidence intervals and statistical hypothesis tests of specified size asymptotically. We also present related results which allow us to examine the power properties of these tests.

The first result we present is a version of the asymptotic normality result of Domowitz and White (1982, theorem 2.4).

Theorem 5.1

Given (Ω, F, P) and a compact set $\Theta \subset \mathbb{R}^k$ let $Q_n:\Omega\times\Theta\to\mathbb{R}$ be a random function continuously differentiable of order 2 on Θ a.s., $n=1,2,\ldots$. Let $\hat{\theta}_n:\Omega\to\Theta$ be a function measurable-$F/B(\mathbb{R}^k)$, $n=1,2,\ldots$ which solves $\min_\Theta Q_n(\theta)$ a.s., and suppose $\hat{\theta}_n-\theta_n^*\to 0$ a.s., where $\{\theta_n^*\}$ is interior to Θ uniformly in n. Suppose there exists a nonstochastic sequence of $k\times k$ matrices $\{B_n^*\}O(1)$ and uniformly positive definite such that

$$B_n^{*-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime} \overset{A}{\sim} N(0,I_k),$$

where $\nabla_\theta Q_n^* \equiv \nabla_\theta Q_n(\theta_n^*)$. If there exists a nonstochastic sequence $\{A_n:\Theta\to\mathbb{R}^{k\times k}\}$ such that $\{A_n(\theta)\}$ is continuous on Θ uniformly in n, $\nabla_\theta^2 Q_n(\theta)-A_n(\theta)\to 0$ a.s. uniformly on Θ, and $\{A_n^*\equiv A_n(\theta_n^*)\}$ is $O(1)$ and uniformly positive definite, then

$$B_n^{*-1/2}A_n^*\sqrt{(n)}(\hat{\theta}_n-\theta_n^*) \overset{A}{\sim} N(0,I_k). \quad \square$$

The asymptotic normality result which we present for our optimization estimator is obtained by verifying the conditions of this result under assumptions compatible with those of chapters 2 and 3.

The condition that $\{\theta_n^*\}$ is interior to Θ uniformly in n means that there exists $\varepsilon>0$ (not depending on n) sufficiently small such that for all

n sufficiently large $\{\theta\in\mathbb{R}^k:|\theta-\theta_n^*|<\varepsilon\}=\{\theta\in\Theta:|\theta-\theta_n^*|<\varepsilon\}$, so that θ_n^* is prevented from getting too close to the boundary of Θ. The condition that $\{B_n^*\}$ is uniformly positive definite means that for each n, B_n^* is positive definite and there exists $\varepsilon>0$ sufficiently small such that for all n sufficiently large $\det B_n^*>\varepsilon$. When $\{B_n^*\}$ is $O(1)$, uniform positive definiteness implies that $\{B_n^{*-1}\}$ is $O(1)$ and uniformly positive definite.

The condition that

$$B_n^{*-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime} \overset{A}{\sim} N(0,I_k) \tag{5.1}$$

together with a mean value expansion of $\nabla Q_n(\hat{\theta}_n)$ around θ_n^* plays the fundamental role in establishing asymptotic normality of $\hat{\theta}_n$. This condition also plays a crucial role in establishing the foundation for our later results on the power of various test statistics, as does an analogous condition for $\tilde{\theta}_n$. This analogous condition is that there exists a nonstochastic sequence $\{B_n^o\}$ $O(1)$ and uniformly positive definite such that

$$B_n^{o\,-1/2}\sqrt{(n)}(\nabla_\theta Q_n^o-\nabla_\theta\bar{Q}_n^o)' \overset{A}{\sim} N(0,I_k), \tag{5.2}$$

where $\nabla_\theta\bar{Q}_n^o\equiv\nabla_\theta\bar{Q}_n(\theta_n^o)$. The presence of this latter term arises because $\sqrt{(n)}\nabla_\theta Q_n(\theta_n^o)$ generally will not have zero mean without appropriate centering. In the case of $\hat{\theta}_n$ we have $\sqrt{(n)}\nabla_\theta\bar{Q}_n^*=0$ (where $\nabla_\theta\bar{Q}_n^*\equiv\nabla_\theta\bar{Q}_n(\theta_n^*)$) for all n sufficiently large, because for such n, θ_n^* will minimize $\bar{Q}_n$ interior to Θ.

Our first goal, therefore, is to obtain conditions ensuring the validity of (5.1) and (5.2). We establish (5.2); (5.1) will then follow by putting $\Theta_n\equiv\Theta$, so that $\theta_n^o\equiv\theta_n^*$.

We begin by placing conditions on g_n and ψ_n ensuring that $Q_n(\theta)\equiv g_n(\psi_n(\theta))$ satisfies appropriate measurability and continuity requirements, specifically those of theorem 5.1. It suffices to modify the optimand assumption in the following way.

Assumption OP′

Let Θ be a compact subset of $\mathbb{R}^k$. For $n=1,2,\ldots$ define the optimand $Q_n:\Omega\times\Theta\to\mathbb{R}$ as

$$Q_n(\omega,\theta)\equiv g_n(\psi_n(\omega,\theta)),$$

where $\psi_n(\omega,\theta)\equiv n^{-1}\sum_{t=1}^n q_t(\omega,\theta)$, and

(i) $\{g_n : \mathbb{R}^l \to \mathbb{R}\}$ is continuously differentiable of order 2 on compact subsets of $\mathbb{R}^l$ uniformly in n;
(ii) $q_t : \Omega \times \Theta \to \mathbb{R}^l$ is a random function continuously differentiable of order 2 on Θ a.s., $t = 1, 2, \ldots$. □

Lemma 5.2

Given assumptions DG and OP′, $Q_n \equiv g_n \circ \psi_n : \Omega \times \Theta \to \mathbb{R}$ is a random function continuously differentiable of order 2 on Θ a.s., $n = 1, 2, \ldots$ with

$$\underset{1\times k}{\nabla_\theta Q_n} = (\underset{1\times l}{\nabla_\psi g_n} \circ \psi_n)\, \underset{l\times k}{\nabla_\theta \psi_n}$$

and

$$\underset{k\times k}{\nabla_\theta^2 Q_n} = \underset{k\times l}{\nabla_\theta \psi_n'}\, (\underset{l\times k}{\nabla_\psi^2 g_n} \circ \psi_n)\, \underset{l\times k}{\nabla_\theta \psi_n} + (\underset{k\times lk}{\nabla_\psi g_n \circ \psi_t} \otimes I_k)\, \underset{lk\times k}{\nabla_\theta^2 \psi_n},$$

where $\nabla_\theta^2 \psi_n \equiv [\nabla_\theta^2 \psi_{1n}', \ldots, \nabla_\theta^2 \psi_{ln}']'$. □

Thus, $\{\nabla_\theta Q_n(\theta_n^o)\}$ is a sequence of measurable functions. As we establish in the proof of theorem 5.4 below, $\bar{Q}_n(\theta)$ is also differentiable on Θ, and

$$\nabla_\theta \bar{Q}_n = (\nabla_\psi g_n \circ \bar{\psi}_n) \nabla_\theta \bar{\psi}_n,$$

$$\nabla_\theta \bar{\psi}_n(\theta) \equiv n^{-1} \sum_{t=1}^{n} \nabla_\theta E(q_t(\theta)) = n^{-1} \sum_{t=1}^{n} E(\nabla_\theta q_t(\theta))$$

under appropriate domination conditions on $\nabla_\theta q_t(\theta)$. This ensures that consideration of the asymptotic distribution of $\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \bar{Q}_n^o)$ is meaningful.

To obtain the asymptotic distribution of $\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \bar{Q}_n^o)$ we follow Bates (1984) and define the random function

$$\tilde{Q}_n(\theta) \equiv (\bar{Q}_n(\theta) + (\psi_n^o - \bar{\psi}_n^o)' \nabla_\psi g_n(\bar{\psi}_n(\theta))' + \nabla_\psi g_n(\bar{\psi}_n^o)(\psi_n(\theta) - \bar{\psi}_n(\theta)),$$

where $\psi_n^o \equiv \psi_n(\theta_n^o)$, $\bar{\psi}_n^o \equiv \bar{\psi}_n(\theta_n^o)$, and we show tha:

$$B_n^{o\,-1/2} \sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \bar{Q}_n^o) - B_n^{o\,-1/2} \sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o) \xrightarrow{P} 0,$$

where $\nabla \tilde{Q}_n^o \equiv \nabla \tilde{Q}_n(\theta_n^o)$, and that

$$B_n^{o\,-1/2} \sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o)' \overset{A}{\sim} N(0, I_k).$$

These facts imply that $B_n^{o\,-1/2} \sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \bar{Q}_n^o)' \overset{A}{\sim} N(0, I_k)$, as desired. A little algebra yields

$$\sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o)' = n^{-1/2} \sum_{t=1}^{n} M_{nt}^o$$

where

$$\underset{k\times 1}{M_{nt}^o} \equiv \underset{k\times l}{\nabla_\theta \bar{\psi}_n^{o\prime}}\, \underset{l\times l}{\nabla_\psi^2 g_n}(\bar{\psi}_n^o)[\underset{l\times 1}{q_{nt}^o - E(q_{nt}^o)}]$$
$$- [\underset{k\times l}{\nabla_\theta q_{nt}^o - E(\nabla_\theta q_{nt}^o)}]'\, \underset{l\times 1}{\nabla_\psi g_n(\bar{\psi}_n^o)'},$$

with $q_{nt}^o \equiv q_t(\theta_n^o)$. To obtain asymptotic normality, we define

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

and seek conditions which will ensure the asymptotic unit normality of

$$n^{-1/2} \sum_{t=1}^{n} Z_{nt} \equiv n^{-1/2} \sum_{t=1}^{n} \lambda' B_n^{o\,-1/2} M_{nt}^o$$

for arbitrary $\lambda \in \mathbb{R}^k$ such that $\lambda'\lambda = 1$. This will establish that $B_n^{o\,-1/2} \sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o) \overset{A}{\sim} N(0, I_k)$ by the Cramér-Wold device.

Before stating a central limit theorem which will allow such conditions to be imposed in a straightforward manner, it is helpful to examine the form of M_{nt}^o a little more closely. This form simplifies in some important special cases. For example, suppose that $l = 1$ and $g_n(\psi) = -\psi$, as in the case of quasi-maximum likelihood estimation. Then $\nabla_\psi g_n(\psi) = -1$ for all $\psi \in \mathbb{R}$ while $\nabla_\psi^2 g_n(\psi) = 0$ for all $\psi \in \mathbb{R}$. In this case we have

$$M_{nt}^o = -\nabla_\theta q_{nt}^{o\prime} + E(\nabla_\theta q_{nt}^{o\prime}).$$

Alternatively, suppose $g_n(\psi) = \psi' P_n \psi$ as in the case of method of moments estimation. When the moment conditions are correctly specified (i.e. $\bar{\psi}_n(\theta_o) = 0$ for some θ_o in Θ) and the constraints imposed are correct (so that θ_o is in Θ_n), then $\theta_n^o = \theta_o$ and $\bar{\psi}_n^o = 0$. Now $\nabla_\psi g_n(\psi) = 2P_n\psi$, so that in this case $\nabla_\psi g_n(\bar{\psi}_n^o) = 0$, while $\nabla_\psi^2 g_n(\psi) = 2P_n$ for all $\psi \in \mathbb{R}^l$, implying

$$M_{nt}^o = 2\nabla_\theta \bar{\psi}_n^{o\prime} P_n q_{nt}^o.$$

On the other hand, when the moment conditions are incorrect or the constraints are invalid, both terms of M_{nt}^o play an important role in determining the asymptotic distribution of $\sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o)$, and thus of $\sqrt{(n)}(\hat{\theta}_n - \theta_n^*)$.

We now state a central limit theorem for near epoch dependent functions of a mixing process due to Wooldridge (1986) which allows us to find conditions ensuring the asymptotic normality of the relevant sums. Wooldridge's result is related to results of Withers (1981).

Theorem 5.3

Let $\{Z_{nt}\}$ be a double array such that $\|Z_{nt}\|_r \leqslant \Delta < \infty$ 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 process with φ_m of size $-r/(r-1)$ or α_m of size $-2r/(r-2)$. Define

$$v_n^2 \equiv \operatorname{var}\left(\sum_{t=1}^{n} Z_{nt}\right),$$

and suppose that v_n^{-2} is $O(n^{-1})$. Then

$$v_n^{-1} \sum_{t=1}^{n} Z_{nt} \overset{A}{\sim} N(0,1). \qquad \square$$

This result is analogous to the Liapounov central limit theorem, in that bounds are placed on some moment greater than the second. However, the present result allows for much greater dependence than the independence required in the Liapounov result. Note that the present result dispenses with the asymptotic covariance stationarity condition of results such as that of Serfling (1968) and the versions of Serfling's result used by Domowitz and White (1982). This result thus allows for fairly arbitrary heterogeneity. However, relative to the moment conditions, the mixing and near epoch dependence conditions have been strengthened by a factor of two over those earlier required for laws of large numbers.

As indicated above, we obtain the desired result by setting $Z_{nt} = \lambda' B_n^{o\,-1/2} M_{nt}^o$ for arbitrary $\lambda \in \mathbb{R}^k$, $\lambda'\lambda = 1$. We therefore impose assumptions which ensure that the conditions of theorem 5.3 are met. The first condition is the r-integrability uniformly in n, t of Z_{nt}. For this, the following modification of the domination condition plays a key role.

Assumption DM′

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

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

Assumption DM′(i) ensures that $\{\bar{\psi}_n^o\}$ is $O(1)$, so that given assumption OP′(i), $\{\nabla_\psi g_n(\bar{\psi}_n^o)\}$ and $\{\nabla_\psi^2 g_n(\bar{\psi}_n^o)\}$ are $O(1)$. Assumptions DM′(i) and (ii) together ensure that $\bar{\psi}_n(\theta)$ is differentiable on Θ and that $\{\nabla_\theta \bar{\psi}_n^o\}$ is $O(1)$. That M_{nt}^o is r-integrable uniformly in n, t then follows from the Minkowski inequality.

The r-integrability of Z_{nt} follows if $\{B_n^{o\,-1/2}\}$ is $O(1)$. Now $B_n^o \equiv \operatorname{var}(n^{-1/2} \sum_{t=1}^n M_{nt}^o)$. For $\{B_n^{o\,-1/2}\}$ to be $O(1)$ it suffices that $\{B_n^o\}$ is $O(1)$ and uniformly positive definite. To ensure that $\{B_n^o\}$ is $O(1)$, some restriction on the dependence of the random variables M_{nt}^o is required. The conditions which ensure that Z_{nt} has the appropriate dependence properties for asymptotic normality also allow application of McLeish's inequality to guarantee that $\{B_n^o\}$ is $O(1)$. Thus we strengthen assumptions MX and NE in the following way.

Assumption MX′

$\{V_t\}$ is a mixing sequence such that either ϕ_m is of size $-r/(r-1)$, $r \geqslant 2$ or α_m is of size $-2r/(r-2)$, $r > 2$. $\square$

Assumption NE′

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

Note that assumption NE′(i) strengthens assumption NE both by increasing the size requirement from $-1/2$ to -1 and by assuming near epoch dependence *uniformly* on (Θ, ρ). This latter condition ensures that near epoch dependence (and therefore the mixingale property) are preserved regardless of the behavior of $\{\theta_n^o\}$.

We now have sufficient structure to guarantee that $\{B_n^o\}$ is $O(1)$, via McLeish's inequality (theorem 3.11), as well as to ensure that Z_{nt} will exhibit appropriate dependence properties. Uniform positive definiteness of $\{B_n^o\}$ is a condition which must simply be imposed.

Assumption PD (positive definiteness)

(i) For $\{\theta_n^o\}$ and $\{\theta_n^*\}$ as defined in assumption ID, the sequences $\{B_n^o\}$ and $\{B_n^*\}$ are uniformly positive definite. □

Because $\lambda'\lambda = 1$, we now have sufficient conditions to ensure that $Z_{nt} = \lambda' B_n^{o\,-1/2} M_{nt}^o$ is r-integrable uniformly in n, t. By construction, $E(M_{nt}^o) = 0$, so $E(Z_{nt}) = 0$. Given assumptions NE′ and MX′, M_{nt}^o and therefore $Z_{nt} = \lambda' B_n^{o\,-1/2} M_{nt}^o$ satisfy the appropriate dependence conditions.

Now in this case

$$\begin{aligned} v_n^2 &\equiv \operatorname{var}\left(\sum_{t=1}^n Z_{nt}\right) = \operatorname{var}\left(\sum_{t=1}^n \lambda' B_n^{o\,-1/2} M_{nt}^o\right) \\ &= \lambda' B_n^{o\,-1/2} \operatorname{var}\left(\sum_{t=1}^n M_{nt}^o\right) B_n^{o\,-1/2}\lambda = \lambda' B_n^{o\,-1/2}(nB_n^o) B_n^{o\,-1/2}\lambda \\ &= n \end{aligned}$$

so that $v_n^{-2} = n^{-1}$ is indeed $O(n^{-1})$ as theorem 5.3 requires.

Underlying the proof of asymptotic normality is an application of the uniform law of large numbers to $\nabla_\theta \psi_n(\theta)$. To help ensure that this applies, we impose the following smoothness condition on $\nabla_\theta q_t(\theta)$.

Assumption SM

(ii) $\{\nabla_\theta q_t(\theta)\}$ is a.s. Lipschitz-L_1. □

We now have conditions sufficient to apply theorem 5.3, which yields $B_n^{o\,-1/2}\sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o) \overset{A}{\sim} N(0, I_k)$. The conditions imposed also guarantee $\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \tilde{Q}_n^o) = o_p(1)$, which gives the following result.

Theorem 5.4

Given assumptions DG, OP′, MX′, SM(i) and (ii), DM′(i) and (ii), and NE′(i) and (ii), $\{B_n^o\}$ is $O(1)$. If assumption PD(i) also holds, then

$$B_n^{o\,-1/2}\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \bar{Q}_n^o) \overset{A}{\sim} N(0, I_k). \quad \square$$

This establishes (5.2). To obtain (5.1), we set $\Theta_n = \Theta$ so that $\theta_n^o = \theta_n^*$, and ensure that $\nabla_\theta \bar{Q}_n^* = 0$ by requiring that θ_n^* minimize $\bar{Q}_n$ interior to Θ uniformly in n. We modify assumption ID.

Assumption ID′

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

Corollary 5.5

Given assumptions DG, OP′, MX′, SM(i) and (ii), DM′(i) and (ii), and NE′(i) and (ii), $\{B_n^*\}$ is $O(1)$. If assumptions PD(i) and ID′(i) also hold,

$$B_n^{*\,-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime} \overset{A}{\sim} N(0, I_k). \quad \square$$

This result establishes a key condition for the desired asymptotic normality for $\hat{\theta}_n$. A similar result for $\tilde{\theta}_n$ is not available, essentially because the nature of the constraints imposed in obtaining $\tilde{\theta}_n$ usually make it unnatural to assume that θ_n^o is interior to Θ_n. Nevertheless, theorem 5.4 provides the structure needed in chapter 7 to examine the local power properties of statistics based on the constrained estimator.

In order to obtain the asymptotic normality result for $\hat{\theta}_n$, we still need to ensure the existence of a nonstochastic sequence $\{A_n : \Theta \to \mathbb{R}^{k \times k}\}$ with the desired properties. In view of lemma 5.2, a natural candidate for A_n is

$$A_n = \nabla_\theta^2 \bar{Q}_n = \nabla_\theta \bar{\psi}_n'(\nabla_\psi^2 g_n \circ \bar{\psi}_n)\nabla_\theta \bar{\psi}_n + (\nabla_\psi g_n \circ \bar{\psi}_n \otimes I_k)\nabla_\theta^2 \bar{\psi}_n.$$

The desired uniform convergence will follow by application of lemma 3.4 provided that $\psi_n(\theta) - \bar{\psi}_n(\theta)$, $\nabla_\theta \psi_n(\theta) - \nabla_\theta \bar{\psi}_n(\theta)$, and $\nabla_\theta^2 \psi_n(\theta) - \nabla_\theta^2 \bar{\psi}(\theta)$ converge to zero a.s. uniformly on Θ. Sufficient conditions are already in place to ensure the convergence of $\psi_n(\theta)$. To ensure that of $\nabla_\theta \psi_n(\theta)$ and $\nabla_\theta^2 \psi_n(\theta)$ we add sufficient structure to apply the uniform law of large numbers, theorem 3.18. We use the following additional smoothness, domination, and near epoch dependence conditions.

Assumption SM

(iii) $\{\nabla_\theta^2 q_t(\theta)\}$ is a.s. Lipschitz-L_1. □

Assumption DM′

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

Assumption NE

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

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

Note that assumption NE(ii) is implied by assumption NE′(ii) which we earlier imposed. The next result does not require the strength of the latter condition, so we do not impose it. Also, for ease of reference we identify assumption NE(iii) with a condition which we now label assumption NE′(iii).

Assumption NE′

(iii) Assumption NE(iii) holds. □

These conditions allow us to establish uniform convergence of $Q_n(\theta)$, $\nabla_\theta Q_n(\theta)$, and $\nabla_\theta^2 Q_n(\theta)$. Our interest here is on $\nabla_\theta^2 Q_n(\theta)$. We present the other results for convenience. Parts (a) and (b) of the next result are used in chapter 7.

Theorem 5.6

(a) Given assumptions DG, OP, MX, SM(i), DM, and NE(i),

$$\bar{Q}_n \equiv g_n \circ \bar{\psi}_n : \Theta \to \mathbb{R}$$

is continuous on Θ uniformly in n and $Q_n(\theta) - \bar{Q}_n(\theta) \to 0$ a.s. uniformly on Θ;

(b) Given assumptions DG, OP′(i) and (ii), MX, SM(i) and (ii), DM′(i) and (ii), and NE(i) and (ii),

$$\nabla_\theta \bar{Q}_n = (\nabla_\psi g_n \circ \bar{\psi}_n) \nabla_\theta \bar{\psi}_n : \Theta \to \mathbb{R}^k$$

is continuous on Θ uniformly in n and $\nabla_\theta Q_n(\theta) - \nabla_\theta \bar{Q}_n(\theta) \to 0$ a.s. uniformly on Θ;

(c) Given assumptions DG, OP′, MX, SM, DM′, and NE,

$$A_n \equiv \nabla_\theta^2 \bar{Q}_n = \nabla_\theta \bar{\psi}_n' (\nabla_\psi^2 g_n \circ \bar{\psi}_n) \nabla_\theta \bar{\psi}_n + (\nabla_\psi g_n \circ \bar{\psi}_n \otimes I_k) \nabla_\theta^2 \bar{\psi}_n : \Theta \to \mathbb{R}^{k \times k}$$

where $\nabla_\theta^2 \bar{\psi}_n(\theta) \equiv n^{-1} \sum_{t=1}^n \nabla_\theta^2 E(q_t(\theta)) = n^{-1} \sum_{t=1}^n E(\nabla_\theta^2 q_t(\theta))$ is continuous on Θ uniformly in n and $\nabla_\theta^2 Q_n(\theta) - A_n(\theta) \to 0$ a.s. uniformly on Θ. □

Result (c) is the one relevant for our immediate purposes.

The only condition of theorem 5.1 which remains to be considered is that requiring $\{A_n^*\}$ to be $O(1)$ and uniformly positive definite. Given theorem 5.6(c), the domination conditions imposed ensure that $\{A_n^*\}$ is $O(1)$. It is convenient simply to impose the requirement that $\{A_n^*\}$ is uniformly positive definite.

Assumption PD

(ii) $\{A_n^*\}$ and $\{A_n^o\}$ are uniformly positive definite. □

The condition imposed on $\{A_n^o\}$ is used in chapter 7.

We now can state the desired asymptotic normality result for our unconstrained optimization estimator.

Theorem 5.7

Given assumptions DG, OP′, MX′, SM, DM′, NE′, ID′, and PD,

$$B_n^{*-1/2} A_n^* \sqrt{(n)}(\hat{\theta}_n - \theta_n^*) \overset{A}{\sim} N(0, I_k).$$ □

This result provides a fundamental basis for testing hypotheses. In practical situations, one needs estimates of A_n^* and B_n^*. These are the focus of the next chapter.

MATHEMATICAL APPENDIX

Proof of theorem 5.1

Because $\hat{\theta}_n - \theta_n^* \to 0$ a.s. where $\{\theta_n^*\}$ is interior to Θ uniformly in n, there exists a sequence $\{\theta_n^\ddagger : \Omega \to \Theta\}$ measurable-$F/B(\Theta)$ and tail equivalent to $\hat{\theta}_n$ (i.e. $\theta_n^\ddagger = \hat{\theta}_n$ a.a. n a.s.) such that $\theta_n^\ddagger$ takes its values in a convex compact neighborhood of θ_n^* for all n. Because $\hat{\theta}_n$ minimizes $Q_n(\theta)$ a.s. and because $\theta_n^\ddagger$ is tail equivalent to $\hat{\theta}_n$ and interior to Θ for a.a. n, it follows that

$$\nabla_\theta Q_n(\theta_n^\ddagger) = 0 \quad \text{a.a. } n \quad \text{a.s.}$$

By the mean value theorem for random functions (Jennrich 1969, lemma 3) there exist mean values $\ddot{\theta}_n^{(i)}$, $i = 1, \ldots, k$ lying on the segment between $\theta_n^\ddagger$ and θ_n^* such that

$$\nabla_\theta Q_n(\theta_n^\ddagger)' = \nabla_\theta Q_n^{*\prime} + \nabla_\theta^2 \ddot{Q}_n(\theta_n^\ddagger - \theta_n^*),$$

where $\nabla_\theta^2 \ddot{Q}_n$ is the $k \times k$ hessian of Q_n with row i evaluated at $\ddot{\theta}_n^{(i)}$, $i = 1, \ldots, k$. Multiplying by $\sqrt{(n)}$ and using the first order condition above gives

$$\sqrt{(n)}\nabla_\theta Q_n^{*\prime} + \nabla_\theta^2 \ddot{Q}_n \sqrt{(n)}(\theta_n^\ddagger - \theta_n^*) = 0 \quad \text{a.a. } n \quad \text{a.s.}$$

Because $\nabla_\theta^2 Q_n(\theta) - A_n(\theta) \to 0$ a.s. and $\ddot{\theta}_n^{(i)} - \theta_n^* \to 0$ a.s. as a consequence of $\theta_n^\ddagger - \theta_n^* \to 0$ a.s. it follows from theorem 2.3 of Domowitz and White (1982) that $\nabla_\theta^2 \ddot{Q}_n - A_n^* \to 0$ a.s. Because A_n^* is uniformly positive definite, it follows that $\nabla_\theta^2 \ddot{Q}_n$ is nonsingular a.a. n a.s. Thus

$$\sqrt{(n)}(\theta_n^\ddagger - \theta_n^*) = -\nabla_\theta^2 \ddot{Q}_n^{-1} \sqrt{(n)}\nabla_\theta Q_n^* \quad \text{a.a. } n \quad \text{a.s.}$$

Premultiplying by $B_n^{*-1/2} A_n^*$ gives

$$\begin{aligned} B_n^{*-1/2} A_n^* \sqrt{(n)}(\theta_n^\ddagger - \theta_n^*) &= -B_n^{*-1/2} A_n^* \nabla_\theta^2 \ddot{Q}_n^{-1} \sqrt{(n)}\nabla_\theta Q_n^* \quad \text{a.a. } n \quad \text{a.s.} \\ &= -B_n^{*-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime} - B_n^{*-1/2} A_n^*(\nabla_\theta^2 \ddot{Q}_n^{-1} - A_n^{*-1}) \\ &\quad \times B_n^{*1/2} B_n^{*-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime} \quad \text{a.a. } n \quad \text{a.s.} \\ &= -B_n^{*-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime} + o_p(1). \end{aligned}$$

The last equality follows because $B_n^{*-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime}$ is $O_p(1)$ by asymptotic normality, $\{B_n^{*1/2}\}$, $\{B_n^{*-1/2}\}$, and $\{A_n^*\}$ are each $O(1)$ given the assumptions, and $\nabla_\theta^2 \ddot{Q}_n^{-1} - A_n^{*-1} \to 0$ a.s. Because $B_n^{*-1/2}\sqrt{(n)}\nabla_\theta Q_n^* \overset{A}{\sim} N(0, I_k)$, it follows that $-B_n^{*-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime} \overset{A}{\sim} N(0, I_k)$ also. Thus, by lemma 4.7 of White (1984),

$$B_n^{*-1/2} A_n^* \sqrt{(n)}(\theta_n^\ddagger - \theta_n^*) \overset{A}{\sim} N(0, I_k).$$

Because $\{\theta_n^\ddagger\}$ and $\{\hat{\theta}_n\}$ are tail equivalent, it follows immediately that

$$B_n^{*-1/2} A_n^* \sqrt{(n)}(\hat{\theta}_n - \theta_n^*) \overset{A}{\sim} N(0, I_k). \qquad \square$$

Proof of lemma 5.2

Given assumption OP′(ii), $q_t(\theta)$ is measurable-$F/B(\mathbb{R}^l)$ for each θ in Θ so that $\psi_n(\theta) \equiv n^{-1}\sum_{t=1}^n q_t(\theta)$ is measurable-F/B for each θ in Θ. Because for each θ in Θ, $\nabla_\theta \psi_n(\theta)$ is the limit of a sequence of measurable difference quotients, it follows that $\nabla_\theta \psi_n(\theta)$ is measurable-$F/B(\mathbb{R}^k)$ for each θ in Θ. A similar argument establishes that $\nabla_\theta^2 \psi_n(\theta)$ is measurable-$F/B(\mathbb{R}^{k \times k})$ for each θ in Θ. Because $\nabla_\psi g_n$ is continuous by assumption OP′(i), it is measurable (theorem 13.2 of Billingsley 1979, p. 154). For each θ in Θ, the composition $\nabla_\psi g_n(\psi_n(\theta))$ is measurable by theorem 13.3 of Billingsley (1979, p. 154). A similar argument establishes the measurability of $\nabla_\psi^2 g_n(\psi_n(\theta))$. Because products and sums of measurable functions are measurable (lemma 2.6 of Bartle 1966, p. 9), it follows that the elements of $\nabla_\theta Q_n(\theta)$ and $\nabla_\theta^2 Q_n(\theta)$ are measurable-F/B for each θ in Θ.

Next, let $F_t \in F$ be the set of all $\omega \in \Omega$ such that $q_t(\omega, \cdot)$ is continuously differentiable of order 2 on Θ. By assumption OP′(ii), $P(F_t) = 1$. It follows that $\psi_n(\omega, \cdot) \equiv n^{-1}\sum_{t=1}^n q_t(\omega, \cdot)$ is continuously differentiable of order 2 on Θ for all ω in $F^n \equiv \bigcap_{t=1}^n F_t$, $P(F^n) = 1$. For fixed $\omega \in F^n$, the chain rule implies that $Q_n(\omega, \cdot)$ is differentiable on Θ with gradient

$$\underset{l \times k}{\nabla_\theta Q_n(\omega, \cdot)} = \underset{1 \times l}{\nabla_\psi g_n(\psi_n(\omega, \cdot))} \, \underset{l \times k}{\nabla_\theta \psi_n(\omega, \cdot)},$$

Continuity follows because compositions, products, and sums of continuous functions are continuous. Application of the product rule and chain rule implies that $\nabla_\theta Q_n(\omega, \cdot)$ is differentiable so that $Q_n(\omega, \cdot)$ has hessian

$$\begin{aligned} \underset{k \times k}{\nabla_\theta^2 Q_n(\omega, \cdot)} &= \underset{k \times l}{\nabla_\theta \psi_n(\omega, \cdot)'} \, \underset{l \times l}{\nabla_\psi^2 g_n(\psi_n(\omega, \cdot))} \, \underset{l \times k}{\nabla_\theta \psi_n(\omega, \cdot)} \\ &\quad + \underset{k \times lk}{(\nabla_\psi g_n(\psi_n(\omega, \cdot)) \otimes I_k)} \, \underset{lk \times k}{\nabla_\theta^2 \psi_n(\omega, \cdot)}. \end{aligned}$$

Continuity follows because compositions, products, and sums of continuous functions are continuous. As the result holds for all $\omega \in F^n$, $P(F) = 1$, the proof is complete. $\square$

Proof of theorem 5.3

See Wooldridge (1986, corollary 4.4). $\square$

Proof of theorem 5.4

First we establish that $\bar{Q}_n$ is differentiable on Θ. Recall that

$$\bar{\psi}_n(\theta) = n^{-1} \sum_{t=1}^{n} E(q_t(\theta)).$$

Because $q_t(\theta)$ and $\nabla_\theta q_t(\theta)$ are r-dominated uniformly in t for $r > 2$ by assumption DM′(i) and (ii), it follows from corollary 5.9 of Bartle (1966) that $\nabla_\theta E(q_t(\theta)) = E(\nabla_\theta q_t(\theta))$, which implies that

$$\nabla_\theta \bar{\psi}_n(\theta) = n^{-1} \sum_{t=1}^{n} E(\nabla_\theta q_t(\theta)).$$

Because $\bar{Q}_n = g_n \circ \bar{\psi}_n$ and because g_n and $\bar{\psi}_n$ are differentiable given assumptions OP′ and DM′(i), it follows from the chain rule that $\bar{Q}_n$ is differentiable and that

$$\nabla_\theta \bar{Q}_n = (\nabla_\psi g_n \circ \bar{\psi}_n)\nabla_\theta \bar{\psi}_n, \quad n = 1, 2, \ldots .$$

We establish the desired result by first showing that

$$\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \tilde{Q}_n^o) \xrightarrow{P} 0$$

and then establishing that

$$B_n^{o\,-1/2}\sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o)' \stackrel{A}{\sim} N(0, I_k).$$

Now by lemma 5.2

$$\nabla_\theta Q_n^o = \nabla_\psi g_n(\psi_n^o)\nabla_\theta \psi_n^o.$$

Taking a mean value expansion of $\nabla_\psi g_n(\psi_n^o)$ around $\bar{\psi}_n^o$ gives

$$\nabla_\psi g_n(\psi_n^o) = \nabla_\psi g_n(\bar{\psi}_n^o) + (\psi_n^o - \bar{\psi}_n^o)'\nabla_\psi^2 \ddot{g}_n,$$

where each column of $\nabla_\psi^2 \ddot{g}_n$ is evaluated at a mean value lying on the segment connecting ψ_n^o and $\bar{\psi}_n^o$. Thus

$$\nabla_\theta Q_n^o = \nabla_\psi g_n(\bar{\psi}_n^o)\nabla_\theta \psi_n^o + (\psi_n^o - \bar{\psi}_n^o)'\nabla_\psi^2 \ddot{g}_n \nabla_\theta \psi_n^o.$$

Applying the chain and product rules to $\tilde{Q}_n$ gives

$$\nabla_\theta \tilde{Q}_n = \nabla_\theta \bar{Q}_n + (\psi_n^o - \bar{\psi}_n^o)'(\nabla_\psi^2 g_n \circ \bar{\psi}_n)\nabla \bar{\psi}_n + \nabla_\psi g_n(\bar{\psi}_n^o)(\nabla_\theta \psi_n - \nabla_\theta \bar{\psi}_n).$$

Setting $\theta = \theta_n^o$ gives

$$\nabla_\theta \tilde{Q}_n^o = \nabla_\psi g_n(\bar{\psi}_n^o)\nabla_\theta \bar{\psi}_n^o + (\psi_n^o - \bar{\psi}_n^o)'\nabla_\psi^2 g_n(\bar{\psi}_n^o)\nabla_\theta \bar{\psi}_n^o + \nabla_\psi g_n(\bar{\psi}_n^o)(\nabla_\theta \psi_n^o - \nabla_\theta \bar{\psi}_n^o).$$

Thus

$$\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \tilde{Q}_n^o) = \sqrt{(n)}(\psi_n^o - \bar{\psi}_n^o)[\nabla_\psi^2 \ddot{g}_n \nabla_\theta \psi_n^o - \nabla_\psi^2 g_n(\bar{\psi}_n^o)\nabla_\theta \bar{\psi}_n^o].$$

We show that this is $o_p(1)$ by showing that $\sqrt{(n)}(\psi_n^o - \bar{\psi}_n^o)$ is $O_p(1)$ and that

$$\nabla_\psi^2 \ddot{g}_n \nabla_\theta \psi_n^o - \nabla_\psi^2 g_n(\bar{\psi}_n^o)\nabla_\theta \bar{\psi}_n^o \xrightarrow{P} 0.$$

To establish the latter, we apply proposition 2.16 of White (1984). Given assumptions OP′(i), MX′, DM′(i) and (ii), SM(i) and (ii), and NE′(i) and (ii), it follows from the uniform law of large numbers, theorem 3.17, that

$$\psi_n(\theta) - \bar{\psi}_n(\theta) \to 0 \quad \text{a.s.}$$

and

$$\nabla_\theta \psi_n(\theta) - \nabla_\theta \bar{\psi}_n(\theta) \to 0 \quad \text{a.s.}$$

uniformly on Θ and that $\bar{\psi}_n$ and $\nabla_\theta \bar{\psi}_n$ are continuous on Θ uniformly in n. It follows from theorem 2.3 of Domowitz and White (1982) that $\psi_n^o - \bar{\psi}_n^o \to 0$ a.s. and $\nabla_\theta \psi_n^o - \nabla_\theta \bar{\psi}_n^o \to 0$ a.s. Assumption DM′(i) and (ii) ensures that $\{\bar{\psi}_n^o\}$ and $\{\nabla_\theta \bar{\psi}_n^o\}$ are $O(1)$. Further, the continuity of $\nabla_\psi^2 g_n$ uniformly in n and the fact that the mean values used in evaluating $\nabla_\psi^2 \ddot{g}_n$ lie between ψ_n^o and $\bar{\psi}_n^o$ ensure that $\nabla_\psi^2 \ddot{g}_n - \nabla_\psi^2 g_n(\bar{\psi}_n^o) \to 0$ a.s., where $\{\nabla_\psi^2 g_n(\bar{\psi}_n^o)\}$ is $O(1)$ given assumptions DM′(i) and OP′(i). It follows from proposition 2.16 of White (1984) that

$$\nabla_\psi^2 \ddot{g}_n \nabla_\theta \psi_n^o - \nabla_\psi^2 g_n(\bar{\psi}_n^o)\nabla_\theta \bar{\psi}_n^o \to 0 \quad \text{a.s.}$$

Convergence in probability follows because the first term is measurable.

Next,

$$\sqrt{(n)}(\psi_n^o - \bar{\psi}_n^o) = n^{-1/2} \sum_{t=1}^{n} q_{nt}^o - E(q_{nt}^o).$$

Define $Z_{nt} \equiv \lambda'(q_{nt}^o - E(q_{nt}^o))$ and consider

$$\sqrt{(n)}\lambda'(\psi_n^o - \bar{\psi}_n^o) = n^{-1/2} \sum_{t=1}^{n} Z_{nt}$$

for arbitrary $\lambda \in \mathbb{R}^l$, $\lambda'\lambda = 1$. Given assumptions DG, OP′(ii), DM′(i), NE′(i), and MX′, it follows from lemma 3.14 that $\{Z_{nt}\}$ is a mixingale of size -1 and $c_{nt} = \max(\|Z_{nt}\|_r, 1) \leq \Delta < \infty$ given assumption DM′(i). By Chebyshev's inequality

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

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

By McLeish's inequality

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

$$\leq nK\Delta^2,$$

so that

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

as $\Delta_o \to \infty$. Hence, $\sqrt{(n)}\lambda'(\psi_n^o - \bar{\psi}_n^o)$ is $O_p(1)$ for arbitrary $\lambda \in \mathbb{R}^l$, $\lambda'\lambda = 1$, implying that $\sqrt{(n)}(\psi_n^o - \bar{\psi}_n^o)$ is $O_p(1)$. Therefore, by corollary 2.36 of White (1984)

$$\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \tilde{Q}_n^o) = \sqrt{(n)}(\psi_n^o - \bar{\psi}_n^o)[\nabla_\psi^2 \bar{g}_n \nabla_\theta \psi_n^o - \nabla^2 g_n(\bar{\psi}_n^o)\nabla_\theta \bar{\psi}_n^o] \xrightarrow{P} 0.$$

Next we show that $\{B_n^{o\,-1/2}\}$ is $O(1)$. For this it suffices that $\{B_n^o\}$ is $O(1)$ and uniformly positive definite. Unifom positive definiteness is guaranteed by assumption PD(i), and it remains to show that $\{B_n^o\}$ is $O(1)$.

Define $Z_{nt} \equiv \lambda' M_{nt}^o$, where $\lambda \in \mathbb{R}^k$, $\lambda'\lambda = 1$, and

$$M_{nt}^o \equiv \nabla_\theta \bar{\psi}_n^{o\prime} \nabla_\psi^2 g_n(\bar{\psi}_n^o)(q_{nt}^o - E(q_{nt}^o)) + [\nabla_\theta q_{nt}^o - E(\nabla_\theta q_{nt}^o)]'\nabla_\psi g_n(\bar{\psi}_n^o)',$$

so that $\text{var}\,(n^{-1/2}\sum_{t=1}^{n} Z_{nt}) = \lambda' B_n^o \lambda$. Now

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

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

Given assumptions DG, OP′, DM′(i) and (ii), NE′(i) and (ii), and MX′, it follows from lemma 3.13 that Z_{nt} is a mixingale of size -1 (and *a fortiori* of size $-1/2$) with $c_{nt} \leq \Delta < \infty$ for all n, t. Applying McLeish's inequality, we have

$$\text{var}\, n^{-1/2}\sum_{t=1}^{n} Z_{nt} \leq n^{-1}(nK\Delta^2) = K\Delta^2 < \infty.$$

Thus $\{\lambda' B_n^o \lambda\}$ is $O(1)$ for arbitrary $\lambda \in \mathbb{R}^k$, $\lambda'\lambda = 1$, implying that $\{B_n^o$ … $O(1)$.

We also have $\nabla \bar{Q}_n^o < \infty$, so that by corollary 2.36 of White (1984)

$$B_n^{-1/2}\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \bar{Q}_n^o)' - B_n^{o\,-1/2}\sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o)'$$
$$= B_n^{o\,-1/2}\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \tilde{Q}_n^o)' \xrightarrow{P} 0.$$

We complete the proof by showing that

$$B_t^{-1/2}\sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o)' \stackrel{A}{\sim} N(0, I_k).$$

Now

$$B_t^{-1/2}\sqrt{(n)}(\nabla_\theta \tilde{Q}_n^o - \nabla_\theta \bar{Q}_n^o)' = n^{-1/2}\sum_{t=1}^{n} B_n^{o\,-1/2} M_{nt}^o.$$

Define $Z_{nt} = \lambda' B_n^{o\,-1/2} M_{nt}^o$ where $\lambda \in \mathbb{R}^k$, $\lambda'\lambda = 1$. We apply theorem … Now $\|Z_{nt}\|_r \leq \Delta < \infty$, $r > 2$ by the Minkowski inequality, becau… $\{\nabla_\theta \bar{\psi}_n^o\}$, $\{\nabla_\psi^2 g_n(\bar{\psi}_n^o)\}$, and $\{\nabla_\psi g_n(\bar{\psi}_n^o)\}$ are $O(1)$ and assumption D… ensures the r-integrability of q_{nt}^o and ∇q_{nt}^o uniformly in n, t. By constru… tion, $E(Z_n) = 0$. Given assumptions NE′(i) and (ii) and MX′, $\{Z_{nt}\}$ … near epoch dependent on $\{V_t\}$ of size -1, where $\{V_t\}$ is mixing with ϕ … of size $-r/(r-1)$ or α_m of size $-2r/(r-2)$. In this case,

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

$$= \lambda' B_n^{o\,-1/2}\,\text{var}\left(\sum_{t=1}^{n} M_{nt}^o\right) B_n^{o\,-1/2}\lambda$$

$$= n\lambda' B_n^{o\,-1/2} B_n^o B_n^{o\,-1/2}\lambda$$

$$= n$$

so that $v_n^{-2} = n^{-1}$ is $O(n^{-1})$. Thus, by theorem 5.3

$$v_n^{-1}\sum_{t=1}^{n} Z_{nt} \stackrel{A}{\sim} N(0, 1)$$

for arbitrary $\lambda \in \mathbb{R}^k$, $\lambda'\lambda = 1$. By the Cramér-Wold device (e.g. theo… 5.1 of White 1984), we have

$$B_n^{o\,-1/2}\sqrt{(n)}(\nabla_\theta Q_n^o - \nabla_\theta \bar{Q}_n^o)' \stackrel{A}{\sim}$$

Proof of corollary 5.5

By setting $\Theta_n = \Theta$, it follows immediately from theorem 5.4 that $\{B_n^*\}$ is $O(1)$ and

$$B_n^{*\,-1/2}\sqrt{(n)}(\nabla_\theta Q_n^* - \nabla_\theta \bar{Q}_n^*)' \overset{A}{\sim} N(0, I_k).$$

Now $\nabla_\theta \bar{Q}_n = (\nabla_\psi g_n \circ \bar{\psi}_n)\nabla\bar{\psi}_n$ is continuous on Θ because of the continuity of $\nabla_\psi g_n$, $\bar{\psi}_n$, and $\nabla_\theta \bar{\psi}_n$. Because θ_n^* minimizes $\bar{Q}_n$ interior to Θ by assumption ID′(i), it follows that $\nabla_\theta \bar{Q}_n^* = 0$. Thus

$$B_n^{*\,-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime} \overset{A}{\sim} N(0, I_k). \qquad \square$$

Proof of theorem 5.6(*a*)

Given assumptions DG, OP, MX, SM(i), DM, and NE, $\psi_n(\theta) - \bar{\psi}_n(\theta) \to 0$ a.s. uniformly on Θ and $\bar{\psi}_n$ is continuous on Θ uniformly in n by theorem 3.18. It follows from lemma 3.4 that

$$\bar{Q}_n \equiv g_n \circ \bar{\psi}_n$$

is continuous on Θ uniformly in n and $Q_n(\theta) - \bar{Q}_n(\theta) \to 0$ a.s. uniformly on Θ.

Proof of theorem 5.6(*b*)

Given assumptions DG, OP′, MX, SM(i) and (ii), DM′(i) and (ii), and NE(i) and (ii), $\psi_n(\theta) - \bar{\psi}_n(\theta) \to 0$ a.s. uniformly on Θ, $\nabla_\theta \psi_n(\theta) - \nabla_\theta \bar{\psi}_n(\theta) \to 0$ a.s. uniformly on Θ, and $\bar{\psi}_n$ and $\nabla_\theta \bar{\psi}_n$ are continuous on Θ uniformly in n. It follows from lemma 3.4 that

$$\nabla_\theta \bar{Q}_n = (\nabla_\psi g_n \circ \bar{\psi}_n)\nabla_\theta \bar{\psi}_n$$

is continuous on Θ uniformly in n and $\nabla_\theta Q_n(\theta) - \nabla_\theta \bar{Q}_n(\theta) \to 0$ a.s. uniformly on Θ.

Proof of theorem 5.6(*c*)

Given assumptions DG, OP′, MX, SM, DM′, and NE, $\psi_n(\theta) - \bar{\psi}_n(\theta) \to 0$ a.s. uniformly on Θ, $\nabla_\theta \psi_n(\theta) - \nabla_\theta \bar{\psi}_n(\theta) \to 0$ a.s. uniformly on Θ, $\nabla_\theta^2 \psi_n(\theta) - \nabla_\theta^2 \bar{\psi}_n(\theta) \to 0$ a.s. uniformly on Θ, and $\bar{\psi}_n$, $\nabla_\theta \bar{\psi}_n$, and $\nabla_\theta^2 \bar{\psi}_n$ are continuous on Θ uniformly in n. It follows from lemma 3.4 that

$$A_n \equiv \nabla_\theta^2 \bar{Q}_n = \nabla_\theta \bar{\psi}_n'(\nabla_\psi^2 g_n \circ \bar{\psi}_n)\nabla_\theta \bar{\psi}_n + (\nabla_\psi g_n \circ \bar{\psi}_n \otimes I_k)\nabla_\theta^2 \bar{\psi}_n$$

is continuous on Θ uniformly in n and $\nabla^2 Q_n(\theta) - A_n(\theta) \to 0$ a.s. uniformly on Θ. $\square$

Proof of theorem 5.7

Given assumptions DG and OP, by lemma 5.2 $Q_n \equiv g_n \circ \psi_n$ satisfies the measurability and continuity requirements of theorem 5.1. Given assumptions DG, OP′, MX′, SM, DM′(i), NE′(i), and ID′(i), it follows from theorem 3.18 that $\hat{\theta}_n - \theta_n^* \to 0$ a.s. where $\hat{\theta}_n$ solves $\min_\Theta Q_n(\theta)$ a.s. Given assumption ID′(i), $\{\theta_n^*\}$ is interior to Θ uniformly in n. By corollary 5.5, given assumptions DG, OP′, MX′, SM, DM′(i) and (ii), NE′(i) and (ii), ID′(i), and PD(i), there exists $\{B_n^*\}$ $O(1)$ and uniformly positive definite given assumption PD(i) such that

$$B_n^{*\,-1/2}\sqrt{(n)}\nabla_\theta Q_n^{*\prime} \overset{A}{\sim} N(0, I_k).$$

By theorem 5.6(c), $\nabla^2 Q_n(\theta) - A_n(\theta) \to 0$ a.s. uniformly on Θ where A_n is continuous on Θ uniformly in n. Further, $\{A_n^*\}$ is $O(1)$ given the domination condition, assumption DM′, and the uniform continuity of $\nabla_\psi g_n$ and $\nabla_\psi^2 g_n$ ensured by assumption OP′(i). Assumption PD(ii) ensures that $\{A_i^*\}$ is uniformly positive definite. The conditions of theorem 5.1 are therefore satisfied, and it follows that

$$B_n^{*\,-1/2} A_i^* \sqrt{(n)}(\hat{\theta}_n - \theta_n^*) \overset{A}{\sim} N(0, I_k). \qquad \square$$

REFERENCES

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

Bates, C. 1984: Nonlinear parametric models with dependent observations, University of Rochester, Department of Economics PhD dissertation.

Billingsley, P. 1979: *Probability and Measure*. New York: John Wiley and Sons.

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

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

Serfling, R. J. 1968: Contributions to central limit theory for dependent processes, *Annals of Mathematical Statistics* 39, 1158–75.

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

Withers, C. S. 1981: Central limit theorems for dependent variables I, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandete Gebiete* 57, 509–34.

Wooldridge, J. 1986: Asymptotic properties of econometric estimators, University of California San Dego, Department of Economics PhD dissertation.