Technical Appendix for "Consistent and Non-Degenerate Model Specification Tests Against Smooth Transition and Neural Network Alternatives"

Jonathan B. Hill*
Dept. of Economics
University of North Carolina - Chapel Hill

May 11, 2008

In this technical appendix we detail a bootstrap/monte-carlo technique for approximating the p-value of test statistics with non-standard limits. See Section 1. Omitted proofs, and supporting lemmata required to prove the main results of the paper, are presented in Section 2.

1. Approximating the Limit Distribution The statistics $\text{ave}_{\Theta_\xi} T_n(\theta)$ and $\sup_{\theta\in\Theta_\xi} T_n(\theta)$ have limit null distributions that depend upon the covariance function $V(\theta_1, \theta_2)$ and therefore upon the underlying distribution of $\{y_t, x_t\}$.

1.1 Asymptotics

Let $\tilde{z}_n$ denote the sample of $pk + 1$-vectors $\{\tilde{z}_1, ..., \tilde{z}_n\}$, and let $\{v_t\}_{t=1}^n$ be *iid* standard normal random variables. Define the following processes, letting $g(\cdot)$ denote any continuous function on Θ_ξ:

$$\hat{z}_n(\hat{\phi}, 0, \theta) \equiv \hat{V}(\theta)^{-1/2}\sqrt{n}\hat{s}_n(\hat{\phi}, 0, \theta)$$
$$\hat{S}_n(\hat{\phi}, 0, \theta) \equiv \frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}_t \tilde{w}(\delta, \tilde{z}_t) F(\tau' \tilde{z}_t) v_t$$
$$\hat{Z}_n(\hat{\phi}, 0, \theta) \equiv \hat{V}(\theta)^{-1/2}\sqrt{n}\hat{S}_n(\hat{\phi}, 0, \theta)$$
$$\hat{T}_n(\theta) \equiv \hat{Z}_n(\hat{\phi}, 0, \theta)' \hat{Z}_n(\hat{\phi}, 0, \theta)$$
$$\hat{g}_n \equiv g(\hat{T}_n(\theta)).$$

It is easy to simulate $\hat{S}_n(\hat{\phi}, 0, \theta)$ and compute a larger number (say J) of statistics $\hat{T}_n(\theta)$ by which an approximate p-value, $\hat{p}_n^J$, can be computed: see Section 1.2,

*Dept. of Economics, University of North Carolina - Chapel Hill; jbhill@email.unc.edu; www.unc.edu/~jbhill.

below. In order to prove $\hat{p}_n^J$ converges to the true p-value under H_0, we must show $\hat{Z}_n(\hat{\phi}, 0, \theta)$ "converges weakly in probability" to the same Gaussian process to which $\hat{z}_n(\hat{\phi}, 0, \theta)$ converges, $z(\theta)$, defined in Theorem 2, cf. Giné and Zinn (1990).

Let $F_0(\cdot)$ denote the asymptotic null distribution of g_0, define $p_0 \equiv 1 - F_0(g_0)$, and define the asymptotic p-value $p_n \equiv 1 - F_0(g_n)$. Similarly, let $\hat{F}_n(\cdot)$ denote the distribution of $\hat{g}_n$ conditioned on $\tilde{z}_n$, and define $\hat{p}_n \equiv 1 - \hat{F}_n(g_n)$.

Denote by $\Rightarrow_p$ weak convergence in probability (see Giné and Zinn, 1990: Sections 2 and 3).

THEOREM C.1 Let $\xi > 0$. Under Assumptions A-C, $\hat{Z}_n(\hat{\phi}, 0, \theta) \Rightarrow_p z(\theta)$. That is (cf. Giné and Zinn, 1990: eq. 3.4),

$$\sup_{x \in \mathbb{R}} \left| P\left(\sup_{\theta \in \Theta_\xi} |\hat{Z}_n(\hat{\phi}, 0, \theta)|_1 \leq x | \tilde{z}_n \right) - P\left(\sup_{\theta \in \Theta_\xi} |z(\theta)|_1 \leq x \right) \right| \to 0. \tag{B.1}$$

Remark 1: Assumption C may be replaced with Assumption D if the null is linearity and $\tilde{w}(\delta, \tilde{z}_t) = \tilde{z}_t$. See Theorem 5.

Remark 2: By appealing to (B.1) and the continuous mapping theorem, we deduce $\hat{T}_n(\theta) \Rightarrow_p T(\theta)$ and $\hat{g}_n \Rightarrow_p g(T(\theta)) = g_0$. We immediately conclude $\hat{F}_n(x) \xrightarrow{p} F_0(x)$ uniformly in $x \in \mathbb{R}$, and therefore $\hat{p}_n - p_n = o_p(1)$. Because F_0 is a continuous function on the real line, $p_n \Rightarrow p_0$ under H_0 by Theorem 3 and the continuous mapping theorem, where p_0 is uniformly distributed on [0,1]. Moreover, $p_n \xrightarrow{p} 0$ under H_1 by Theorem 3 ($g_n \to \infty$ with probability one). Hence the asymptotic p-value satisfies $\hat{p}_n \Rightarrow p_0$ under H_0 and $\hat{p}_n \to 0$ under H_1.

Proof of Theorem C.1. From Theorem 3, $\hat{z}_n(\hat{\phi}, 0, \theta)$ converges weakly to some Gaussian element $z(\theta)$ on $\mathbb{C}[\Theta_\xi]$ with covariance function $E[z(\theta_1)z(\theta_2)']$ $= V(\theta_1)^{-1/2}V(\theta_1, \theta_2)V(\theta_2)^{-1/2}$. In order to show $\hat{Z}_n(\hat{\phi}, 0, \theta)$, conditioned on the sample $\tilde{z}_n$, converges weakly to the same Gaussian element $z(\theta)$, we must demonstrate $\hat{Z}_n(\hat{\phi}, 0, \theta)$, conditioned on $\tilde{z}_n$, and $\hat{z}(\hat{\phi}, 0, \theta)$ converge pointwise to the same multivariate normal distribution; and $\hat{Z}_n(\hat{\phi}, 0, \theta)$, conditioned on $\tilde{z}_n$, is stochastically equicontinuous in $\mathbb{C}[\Theta_\xi]$, cf. Giné and Zinn (1990: Theorem 3.1).

Step 1 ($\hat{Z}_n(\hat{\phi}, 0, \theta) \xrightarrow{d} N(0, I_{pk+1})$): Operate conditionally on the sample $\tilde{z}_n$, and define

$$S_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \epsilon_t g_t(\theta) v_t, \quad Z_n(\theta) = V(\theta)^{-1/2} \sqrt{n} S_n(\phi, 0, \theta).$$

Using steps identical to the line of proof of Lemma A.1, recalling $\hat{\phi}$ and $\hat{V}(\theta)$ are consistent under either hypothesis, and v_t is $(0,1)$-*iid*, it is easy to show

$$\sup_{\theta \in \Theta_\xi} \left| \hat{Z}_n(\hat{\phi}, 0, \theta) - Z_n(\theta) \right|_1 = o_p(1).$$

Thus, it suffices to consider only $Z_n(\theta)$ in the following.

Because v_t is $(0,1)$-*iid*, the covariance function of $Z_n(\theta)$, conditioned on $\tilde{z}_n$, is[1]

$$E[Z_n(\theta_1)Z_n(\theta_2)'|\tilde{z}_n] = V(\theta_1)^{-1/2}\frac{1}{n}\sum_{t=1}^{n}\epsilon_t^2 g_t(\theta_1)g_t(\theta_2)'V(\theta_2)^{-1/2}.$$

Using Theorem 17 of Bierens (1991) and Lemma B.2, it is straightforward to show

$$\sup_{(\theta_1,\theta_2)\in\Theta_\xi} \left| \frac{1}{n}\sum_{t=1}^{n}\epsilon_t^2 g_t(\theta_1)g_t(\theta_2)' - V(\theta_1,\theta_2) \right|_1 = o_p(1),$$

where $V(\theta_1,\theta_2) \equiv E[\epsilon_t^2 g_t(\theta_1)g_t(\theta_2)']$. Mimicking the line of proof of Lemma A.2, under H_0 and $v_t \overset{iid}{\sim} N(0,1)$ it follows

$$Z_n(\theta) \xrightarrow{d} N(0, I_{pk+1})$$

in distribution pointwise in Θ_ξ. This implies $Z_n(\theta)$ and $z_n(\phi,0,\theta)$ have the same pointwise limit distribution because mean-zero multivariate normal distribution are fully characterized by their covariance functions. Hence, from Lemma A.1 we conclude $\hat{Z}_n(\hat{\phi},0,\theta)$ and $\hat{z}(\hat{\phi},0,\theta)$ have the same pointwise limit distribution.

Step 2 (Stochastic Equicontinuity in Θ_ξ**):** From Theorem 3.1 of Giné and Zinn (1990), and appealing to the Cauchy-Schwartz inequality, it suffices to show

$$\lim_{\varpi\to 0}\limsup_{n\to\infty} P\left(E_v\left[\sup_{\theta_i\in\Theta_\xi, |\theta_1-\theta_2|<\varpi} |Z_n(\theta_1) - Z_n(\theta_2)|_1^2\right] > \varepsilon\right) = 0. \quad \text{(B.2)}$$

for all $\varepsilon > 0$, where E_v denotes the expectation only with respect to $(v_t)_{t=1}^n$, and

$$Z_n(\theta_1) - Z_n(\theta_2) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\epsilon_t v_t \psi_t(\theta_1,\theta_2),$$

say, where $\psi_t(\theta_1,\theta_2) = V(\theta_1)^{-1/2}g_t(\theta_1) - V(\theta_2)^{-1/2}g_t(\theta_2)$. In order to prove (B.2) it suffices first to bound

$$E_v\left[\sup_{\theta_i\in\Theta_\xi, |\theta_1-\theta_2|<\varpi} |Z_n(\theta_1) - Z_n(\theta_2)|_1^2\right],$$

[1] We exploit the symmetry of $V(\theta)^{-1/2}$.

and then bound

$$E\left(E_v\left[\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}|Z_n(\theta_1)-Z_n(\theta_2)|_1^2\right]\right)$$

by some $C\times\varpi$, $0<C<\infty$.

Step 2.1 ($E_v[\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}|Z_n(\theta_1)-Z_n(\theta_2)|_1^2]$): FrByom Liaponov's inequality there exists some finite $B>0$ such that

$$\begin{aligned}
&E_v\left[\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}|Z_n(\theta_1)-Z_n(\theta_2)|_1^2\right]\\
&\quad\le B\times E_v\left[\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}|Z_n(\theta_1)-Z_n(\theta_2)|_2^2\right]\\
&\quad= B\times E_v\left[\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}\sum_{l=1}^{pk+1}\frac{1}{n}\sum_{s,t}^{n}\epsilon_s\epsilon_t v_s v_t\psi_{s,l}(\theta_1,\theta_2)\psi_{t,l}(\theta_1,\theta_2)\right]\\
&\quad\le B\frac{1}{n}\sum_{t=1}^{n}\epsilon_t^2\sum_{l=1}^{pk+1}\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}\psi_{t,l}(\theta_1,\theta_2)^2\\
&\quad\le B\frac{1}{n}\sum_{t=1}^{n}\epsilon_t^2\left(\sum_{l=1}^{pk+1}\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}\left|\psi_{t,l}(\theta_1,\theta_2)\right|\right)^2\\
&\quad= B\frac{1}{n}\sum_{t=1}^{n}\epsilon_t^2\left(\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}|\psi_t(\theta_1,\theta_2)|\right)^2
\end{aligned}\tag{B.3}$$

where the second inequality follows from $E_v[v_sv_t]=0$, $\forall s\ne t$, and 1 otherwise.

Step 2.2 ($E(E_v[\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}|Z_n(\theta_1)-Z_n(\theta_2)|^2])<\infty$): The Cauchy-Schwartz inequality and (B.3) imply

$$\begin{aligned}
&E\left(E_v\left[\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}|Z_n(\theta_1)-Z_n(\theta_2)|_1^2\right]\right)\\
&\quad\le B\frac{1}{n}\sum_{t=1}^{n}E\left[\epsilon_t^2\left(\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}|\psi_t(\theta_1,\theta_2)|\right)^2\right]\\
&\quad\le B\|\epsilon_t\|_4^2\left\|\sup_{\theta_i\in\Theta_\xi,|\theta_1-\theta_2|<\varpi}|\psi_t(\theta_1,\theta_2)|\right\|_4^2.
\end{aligned}\tag{B.4}$$

Now, by the mean-value-theorem and the definition of $\psi_t(\theta_1, \theta_2)$, for each ϖ

$$\sup_{\theta_i \in \Theta_\xi, |\theta_1 - \theta_2| < \varpi} |\psi_t(\theta_1, \theta_2)| \tag{B.5}$$
$$\leq \sum_{l=1}^{s} \sup_{\theta_i \in \Theta_\xi, |\theta_1 - \theta_2| < \varpi} \left| \frac{\partial}{\partial \theta_l} \left[V(\theta)^{-1/2} g_t(\theta) \right] \right|_1 \times |\theta_{1,l} - \theta_{2,l}|$$
$$\leq \sum_{l=1}^{s} \sup_{\theta \in \Theta_\xi} \left| \frac{\partial}{\partial \theta_l} \left[V(\theta)^{-1/2} g_t(\theta) \right] \right|_1 \times \varpi,$$

where $s = pk + 1 + (pk+1)q$. Using the Minkowski inequality, from (B.4) and (B.5) it therefore suffices to prove

$$\lim_{\varpi \to 0} \sum_{l=1}^{s} \left\| \sup_{\theta \in \Theta_\xi} \left| \frac{\partial}{\partial \theta_l} \left[V(\theta)^{-1/2} g_t(\theta) \right] \right|_1 \right\|_4 \times \varpi = 0.$$

This will be true if $|| \sup_{\theta \in \Theta_\xi} |(\partial / \partial \theta_l)[V(\theta)^{-1/2} g_t(\theta)]|_1 ||_4$ is bounded for each l, which follows from Lemmas A.5 and A.6. ■

1.2 Algorithm

The p-value algorithm is identical to that of Hansen (1991, 1996), although Giné and Zinn (1984, 1990) detail a generic procedure. Generate a double array of *iid* standard normal variables, $[v_{t,j}]_{t=1,j=1}^{n,J}$. For each $j = 1...J$, compute $\hat{S}_{n,j}(\hat{\phi}, 0, \theta) = 1/n \sum_{t=1}^{n} \hat{\epsilon}_t \tilde{z}_{2t}(\delta) F_t(\tau) v_{t,j}$, $\hat{T}_{n,j}(\theta) = n \hat{S}_{n,j}(\hat{\phi}, 0, \theta)' \hat{V}(\theta)^{-1} \hat{S}_{n,j}(\hat{\phi}, 0, \theta)$, and $\hat{g}_{n,j} = g(\hat{T}_{n,j}(\theta))$. The approximate asymptotic p-value $\hat{p}_n^J$ is simply the sample frequency $J^{-1} \sum_{j=1}^{J} (g(\hat{T}_{n,j}(\theta)) > g(T_n(\theta)))$. Because the J-samples $[v_{t,j}]_{t=1}^{n}$ are independent of each other, an appeal to the Glivenko-Cantelli Theorem guarantees $\hat{p}_n^J \xrightarrow{p} \hat{p}_n$ as $J \to \infty$, and Theorem 7 guarantees $\hat{p}_n \Rightarrow p_0$ under H_0, etc.

2. Omitted Proofs (Lemmas A.2.1, A.5-A.6) Recall:

LEMMA A.2.1 *Under the conditions of Lemma A.2, for each* $\theta \in \Theta_\xi$

$$\operatorname*{plim}_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} w_t(\theta)^2 = \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} E[w_t(\theta)^2] = 1,$$

and for some $\kappa > 0$

$$\operatorname*{plim}_{n \to \infty} \sum_{t=1}^{n} E|w_t(\theta)/\sqrt{n}|_1^{2+\kappa} = 0.$$

LEMMA A.5 *Under Assumptions A-B and under both H_0 and H_1,*

i. $\left|\hat{A}(\hat{\phi}) - A(\phi)\right|_1 = o_p(1)$,

ii. $\sup_{\theta\in\Theta}\left|\hat{b}(\hat{\phi},\theta) - b(\phi,\theta)\right|_1 = o_p(1)$,

iii. $\sup_{\theta\in\Theta}\sup_{\phi\in\Phi}\left|\hat{b}(\phi,\theta) - b(\phi,\theta)\right|_1 = o_p(1)$

iv. $\sup_{\theta\in\Theta}\left|\hat{V}(\theta) - V(\theta)\right|_1 = o_p(1)$.

LEMMA A.6 *For some positive constants $C, K < \infty$,*

i. $|A(\phi)^{-1}|_1 \leq K$,

ii. $\sup_{\theta\in\Theta_\xi}|b(\theta,\phi)|_1 \leq K$,

iii. $\sup_{\theta\in\Theta_\xi}\left|\frac{\partial}{\partial\theta}b(\phi,\theta)\right|_1 \leq K$,

iv. $\left\|\sup_{\theta\in\Theta_\xi}|g_t(\theta)|_1\right\|_{4+\kappa} \leq K$,

v. $\left\|\sup_{\theta\in\Theta_\xi}\left|\frac{\partial}{\partial\theta}g_t(\theta)\right|_1\right\|_{4+\kappa} \leq K$,

vi. $\sup_{\theta\in\Theta_\xi}\left|V(\theta)^{-1/2}\right|_1^2 \leq C(pk+1)\left[\inf_{\theta\in\Theta_\xi}\lambda_{\min}(V(\theta))\right]^{-1} \leq K$

vii. $\sup_{\theta\in\Theta_\xi}\left|\frac{\partial}{\partial\theta_l}V(\theta)^{-1/2}\right|_1 \leq K, \quad \text{for } l = 1...pk+1.$

Proof of Lemma A.2.1. By the normalization $r'r = 1$, for each t

$$\begin{aligned}
E[w_t(\theta)^2] &= E\left[\epsilon_t r'V(\theta)^{-1/2}g_t(\theta)\right]^2 \\
&= r'V(\theta)^{-1/2}E\left[\epsilon_t^2 g_t(\theta)g_t(\theta)'\right]V(\theta)^{-1/2}r \\
&= r'V(\theta)^{-1/2}V(\theta)V(\theta)^{-1/2}r = r'r = 1.
\end{aligned}$$

The weak limit $\text{plim}_{n\to\infty} 1/n\sum_{t=1}^n w_t(\theta)^2 - 1 = 0$ then follows from

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^n \epsilon_t^2 g_t(\theta)g_t(\theta)' - V(\theta)\right|_1 = o_p(1),$$

proved in Step 3 of the line of proof of Lemma A.5, below.

The second stated limit follows from the following bound. By l_1-norm properties, the envelope inequality (e.g. Theorem 21.3 of Davidson 1994), Assumption A, and Lemma A.6, for some small $\kappa > 0$ and some finite $M > 0$

$$\begin{aligned} &E|w_t(\theta)|_1^{2+\kappa} \\ &\quad \le |r|_1^{2+\kappa} \left|V(\theta)^{-1/2}\right|_1^{2+\kappa} E|g_t(\theta)|_1^{2+\kappa} \\ &\quad \le |r|_1^{2+\kappa} \sup_{\theta\in\Theta_\xi} \left|V(\theta)^{-1/2}\right|_1^{2+\kappa} \left\|\sup_{\theta\in\Theta_\xi} |g_t(\theta)|_1\right\|_{2+\kappa}^{2+\kappa} \le M, \end{aligned}$$

where $|r|^{2+\kappa} < \infty$ is trivial[2]. Thus, $\sum_{t=1}^{r} E|w_t(\theta)|^{2+\kappa}/n^{1+\kappa/2} = o(1/n^{\kappa/2})$. ■

Proof of Lemma A.5.

Claim (i). Under either hypothesis Assumption A and Theorem 8.2.2 of Bierens (1994) ensure $\hat{\phi} \xrightarrow{p} \phi$. $|\hat{A}(\hat{\phi}) - A(\phi)|_1 = o_p(1)$ follows from

$$\sup_{\phi\in\Theta} \left| \frac{1}{n}\sum_{t=1}^{n} \partial f(\phi, \tilde{z}_t)\partial' f(\phi, \tilde{z}_t) - E\left[\partial f(\phi, \tilde{z}_t)\partial' f(\phi, \tilde{z}_t)\right] \right|_1 = o_p(1),$$

which follows easily from Assumption A and Theorem 17 of Bierens (1991).

Claim (ii). By the consistency of $\hat{\phi}$, $\sup_{\theta\in\Theta} |\hat{b}(\hat{\phi},\theta) - b(\phi,\theta)| = o_p(1)$ follows if

$$\sup_{\theta\in\Theta}\sup_{\phi\in\Phi} \left|\hat{b}(\phi,\theta) - b(\phi,\theta)\right|_1 = o_p(1). \tag{B.6}$$

Rate (B.6) follows from Theorem 17 of Bierens (1991) by writing for each $i = 1...pk+1$, each $j = 1...m$, and any $\phi \in \Phi$,

$$\frac{1}{n}\sum_{t=1}^{n} \psi^{(i,j)}(\Gamma^{(i,j)}(\tilde{z}_t, \phi, \theta)) = \frac{1}{n}\sum_{t=1}^{n} F(\tau'\tilde{z}_t)w_i(\delta, \tilde{z}_t)\partial_j f(\phi, \tilde{z}_t)$$

where $\psi^{(i,j)}(\xi_1, \xi_2) = \xi_1 \times \xi_2$ for each i and j, and

$$\begin{aligned} \Gamma^{(i,j)}(\tilde{z}_t, \phi, \theta) &= \left(\gamma_1^{(i,j)}(\tilde{z}_t, \phi, \theta), \gamma_2^{(i,j)}(\tilde{z}_t, \phi, \theta)\right)' \\ &= \left(F(\tau'\tilde{z}_t)w_i(\delta, \tilde{z}_t),\ \partial_j f(\phi, \tilde{z}_t)\right)'. \end{aligned}$$

By Assumption A we know $\gamma_l^{(i,j)}(\tilde{z}_t, \phi, \theta)$ is for each $\tilde{z}_t$ a continuous real function on $\Theta \times \Phi$, and for each (ϕ, θ) Borel measurable functions on $\mathbb{R}^{pk+1}$. We require

[2] It is here that the restriction $\xi > 0$ is imperative. Using $\xi = 0$ (i.e. $\Theta_0 = \Theta$), from Lemma B.2 we obtain for some finite $B > 0$, $\sup_{\theta\in\Theta} |V(\theta)^{-1/2}|_1 \le B \sup_{\theta\in\Theta}[\lambda_0(V(\theta))^{-1}]$. If $0 \in \Theta$, either the right-hand-side supremum does not exist, or $\sup_{\theta\in\Theta}[\lambda_{\min}(V(\theta))^{-1}] = +\infty$ if we extend the real line. In either case, a finite bound would still have to be established. Depending on the specification of $\tilde{z}_t(\delta)$, $\lambda_0(V(\theta)) = 0$ for any $\theta = (\delta', 0')'$ is certainly possible.

$\sup_{|\xi|\leq d}\left|(\partial/\partial\xi)\psi^{(i,j)}(\xi)\right| = O(d^{\mu})$ as $d\to\infty$, which easily follows by Assumption A for $\mu = 1$. Moreover, for each i and j we need

$$\max_k\left\{\sup_{\theta\in\Theta}\sup_{\phi\in\Phi}\left|\gamma_k^{(i,j)}(\tilde{z}_t,\phi,\theta)\right|\right\}\leq\rho\bar{d}(\tilde{z}_t)$$

for some non-negative Borel function $\bar{d}(\tilde{z}_t)$ on $\mathbb{R}^{pk+1}$ such that $\sup_t E\left[\bar{d}(\tilde{z}_t)^{1+\mu+\kappa}\right] < \infty$ for some $\kappa > 0$, where μ is defined above, and $\rho = 1$[3]. Simply define

$$\bar{d}(\tilde{z}_t) = \sup_{\theta\in\Theta}|F(\tau'\tilde{z}_t)w(\delta,\tilde{z}_t)| + \sup_{\phi\in\Phi}|\partial f(\phi,\tilde{z}_t)|\,,$$

and $\sup_t E\left[\bar{d}(\tilde{z}_t)^{1+\mu+\kappa}\right] < \infty$ follows from Assumption A with $\mu = 1$.

Claim (iii). Using the fact that each $|\hat{\phi}-\phi|_1$, $\sup_{\theta\in\Theta}|\hat{b}(\hat{\phi},\theta) - b(\phi,\theta)|_1$ and $|\hat{A}(\hat{\phi}) - A(\phi)|_1$ is $o_p(1)$, it is straightforward to show under either hypothesis $\sup_{\theta\in\Theta}|\hat{V}(\theta) - n^{-1}\sum_{t=1}^n\epsilon_t^2 g_t(\theta)g_t(\theta)'|_1 = o_p(1)$. In order to show $\sup_{\theta\in\Theta}|n^{-1}\sum_{t=1}^n\epsilon_t^2 g_t(\theta)g_t(\theta)' - V(\theta)|_1 = o_p(1)$, apply Theorem 17 of Bierens (1991) to each element $\epsilon_t^2 g_{t,i}(\theta)g_{t,j}(\theta)$: define for each i and j, $\gamma_1^{(i,j)}(\tilde{z}_t,\theta) = \epsilon_t^2 g_{t,i}(\theta)$ and $\gamma_2^{(i,j)}(\tilde{z}_t,\theta) = g_{t,j}(\theta)$, set $\psi^{(i,j)}(\xi_1,\xi_2) = \xi_1\times\xi_2$ and $\bar{d}(\tilde{z}_t) = \epsilon_t^2\sup_{\theta\in\Theta}|g_t(\theta)| + \sup_{\theta\in\Theta}|g_t(\theta)|_1$, and put $\mu = \rho = 1$. Now use the Cauchy-Schwartz inequality, and the moment bounds in Assumption A and Lemma A.6. The result $\sup_{\theta\in\Theta}|\hat{V}(\theta) - V(\theta)|_1 = o_p(1)$ then follows from the triangular inequality.

Claim (iv). Immediate from Claims (i)-(iii). ■

Proof of Lemma A.6.

Claim (i). Each $0 < \lambda_i(A) < \infty$ because A is positive definite. By Liaponov's inequality for some finite $B > 0$

$$\begin{aligned}|A^{-1}|_1 &\leq |A^{-1/2}|_1^2 \leq B|A^{-1/2}|_2^2\\ &= B\times Tr\left(A^{-1/2\prime}A^{-1/2}\right) = B\times Tr\left(A^{-1}\right)\\ &= B\sum\nolimits_i\lambda_i(A^{-1}) = B\sum\nolimits_i 1/\lambda_i(A)\leq C.\end{aligned}$$

[3] The identity $\rho = 1$ is trivial here because we do not include moving average terms: see Bierens (1991).

Claim (ii). By Assumption A and the envelope and Cauchy-Schwartz inequalities:

$$\begin{aligned}
\sup_{\theta\in\Theta_\xi}|b(\theta,\phi)|_1 &\le E\left[\sum\nolimits_{i,j}\sup_{\theta\in\Theta_\xi}|F(\tau'\tilde{z}_t)w_i(\delta,\tilde{z}_t)|_1\sup_{\phi\in\Phi}|\partial_j f(\phi,\tilde{z}_t)|_1\right]\\
&\le \left(E\left[\sum\nolimits_i\sup_{\theta\in\Theta_\xi}|F(\tau'\tilde{z}_t)w_i(\delta,\tilde{z}_t)|_1\right]^2\right)^{1/2}\\
&\qquad\times\left(E\left[\sum\nolimits_j\sup_{\phi\in\Phi}|\partial_j f(\phi,\tilde{z}_t)|_1\right]^2\right)^{1/2}\\
&= \left\|\sup_{\theta\in\Theta_\xi}|F(\tau'\tilde{z}_t)w(\delta,\tilde{z}_t)|_1\right\|_2\left\|\sup_{\theta\in\Theta_\xi}|\partial f(\phi,\tilde{z}_t)|_1\right\|_2\le C.
\end{aligned}$$

Claim (iii). Similarly:

$$\begin{aligned}
&\sup_{\theta\in\Theta_\xi}\left|\frac{\partial}{\partial\theta}b(\theta,\phi)\right|_1\\
&\quad\le\left\|\sup_{\theta\in\Theta_\xi}\left|\frac{\partial}{\partial\theta}F(\tau'\tilde{z}_t)w(\delta,\tilde{z}_t)\right|_1\right\|_2\left\|\sup_{\theta\in\Theta_\xi}|\partial f(\phi,\tilde{z}_t)|_1\right\|_2\le C.
\end{aligned}$$

Claim (iv). Next, by Assumption A, Minkowski's inequality, and properties of the l_1-norm, and (i) and (ii):

$$\begin{aligned}
&\left\|\sup_{\theta\in\Theta_\xi}|g_t(\theta)|_1\right\|_{4+\kappa}\\
&\quad\le\left\|\sum_{i=1}^{pk+1}\sup_{\theta\in\Theta_\xi}|F(\tau'\tilde{z}_t)w_i(\delta,\tilde{z}_t)|_1\right\|_{4+\kappa}\\
&\qquad\quad+\sup_{\theta\in\Theta_\xi}|b(\phi,\theta)|_1\times|A^{-1}|_1\times\left\|\sum_{i=1}^{pk+1}\sup_{\phi\in\Phi}|\partial_i f(\phi,\tilde{z}_t)|_1\right\|_{4+\kappa}\\
&\quad=\left\|\sup_{\theta\in\Theta_\xi}|F(\tau'\tilde{z}_t)w(\delta,\tilde{z}_t)|_1\right\|_{4+\kappa}\\
&\qquad\quad+\sup_{\theta\in\Theta_\xi}|b(\phi,\theta)|_1\times|A^{-1}|_1\times\left\|\sup_{\theta\in\Theta_\xi}|\partial f(\phi,\tilde{z}_t)|_1\right\|_{4+\kappa}\le C
\end{aligned}$$

Claim (v). Similarly, by Assumption A, (i) and (iii):

$$
\begin{aligned}
&\left\| \sup_{\theta\in\Theta_\xi} \left| \frac{\partial}{\partial\theta} g_t(\theta) \right|_1 \right\|_{4+\kappa} \\
&\quad \leq \left\| \sup_{\theta\in\Theta_\xi} \left|\frac{\partial}{\partial\theta} F_t(\tau) w(\delta, \tilde{z}_t)\right|_1 \right\|_{4+\kappa} \\
&\qquad + \sup_{\theta\in\Theta_\xi} \left| \frac{\partial}{\partial\theta} b(\phi,\theta) \right|_1 \times |A^{-1}|_1 \times \left\| \sup_{\phi\in\Phi} |\partial f(\phi, \tilde{z}_t)|_1 \right\|_{4+\kappa} \leq C.
\end{aligned}
$$

Claim (vi). Next, by Liaponov's inequality for some finite $B > 0$:

$$
\begin{aligned}
|V(\theta)^{-1/2}|_1^2 &\leq B|V(\theta)^{-1/2}|_2^2 = B \times Tr\left(V(\theta)^{-1/2\prime} V(\theta)^{-1/2}\right) \\
&= B \times Tr\left(V(\theta)^{-1}\right) = B \times \sum_{i=0}^{pk} \lambda_i(V(\theta)^{-1}) \\
&= B \times \sum_{i=0}^{pk} \frac{1}{\lambda_i(V(\theta))} \\
&\leq B \times (pk+1)\, \lambda_{\min}(V(\theta))^{-1}
\end{aligned}
$$

hence

$$
\sup_{\theta\in\Theta_\xi} |V(\theta)^{-1/2}|_1^2 \leq B \times (pk+1) \left[\inf_{\theta\in\Theta_\xi} \lambda_{\min}(V(\theta))\right]^{-1} \leq C,
$$

which is guaranteed for some finite C by Assumption C: $\inf_{\theta\in\Theta_\xi} \lambda_{\min}(V(\theta)) > 0$.

Claim (vii). By standard properties of matrix differentiation

$$
\frac{\partial}{\partial\theta_l} V(\theta)^{-1/2} = -(1/2)[V(\theta)^{-1/2} \times \frac{\partial}{\partial\theta_l} V(\theta) \times V(\theta)^{-1}].
$$

Hence, for some finite $B > 0$, by Liaponov's inequality and (vi),

$$
\begin{aligned}
&\sup_{\theta\in\Theta_\xi} \left| \frac{\partial}{\partial\theta_l} V(\theta)^{-1/2} \right|_1 \\
&\quad \leq \sup_{\theta\in\Theta_\xi} \left| V(\theta)^{-1/2} \times \frac{\partial}{\partial\theta_l} V(\theta) \times V(\theta)^{-1} \right|_1 \\
&\quad \leq \sup_{\theta\in\Theta_\xi} \left| V(\theta)^{-1/2} \right|^3 \left| \frac{\partial}{\partial\theta_l} V(\theta) \right|_1 \\
&\quad \leq B(pk+1)^{3/2} \left[\inf_{\theta\in\Theta_\xi} \lambda_{\min}(V(\theta))\right]^{-3/2} \sup_{\theta\in\Theta_\xi} \left|\frac{\partial}{\partial\theta_l} V(\theta)\right|
\end{aligned}
$$

where $\inf_{\theta\in\Theta_\xi}\lambda_{\min}(V(\theta)) > 0$ by Assumption C. The proof is complete when we show the l_1-normed $|(\partial/\partial\theta_l)V(\theta)|$ is uniformly bounded by some finite $M > 0$.

The covariance matrix derivative $(\partial/\partial\theta_l)V(\theta)$ is computed as

$$\begin{aligned}
\frac{\partial}{\partial\theta_l}V(\theta) &= \frac{\partial}{\partial\theta_l}E\left[\epsilon_t^2 g_t(\theta)g_t(\theta)'\right] \\
&= \left(E\left[\epsilon_t^2\frac{\partial}{\partial\theta_l}g_{t,i}(\theta)g_{t,j}(\theta)\right]\right)_{i,j} \\
&\quad + \left(E\left[\epsilon_t^2 g_{t,i}(\theta)\frac{\partial}{\partial\theta_l}g_{t,j}(\theta)\right]\right)_{i,j}.
\end{aligned}$$

By the envelope and repeated Cauchy-Schwartz inequalities,

$$\begin{aligned}
&\sup_{\theta\in\Theta_\xi}\left|\frac{\partial}{\partial\theta_l}V(\theta)\right|_1 \\
&\leq 2\sum_{i,j=1}^{pk+1}\sup_{\theta\in\Theta_\xi}\left|E\left[\epsilon_t^2\frac{\partial}{\partial\theta_l}g_{t,i}(\theta)g_{t,j}(\theta)\right]\right| \\
&\leq 2E\left[\epsilon_t^2\sum_{i=1}^{pk+1}\sup_{\theta\in\Theta_\xi}\left|\frac{\partial}{\partial\theta_l}g_{t,i}(\theta)\right|_1\times\sum_{i=1}^{pk+1}\sup_{\theta\in\Theta_\xi}|g_{t,i}(\theta)|_1\right] \\
&\leq 2\left\|\epsilon_t\right\|_4^2\left\|\sum\nolimits_{i=1}^{pk+1}\sup_{\theta\in\Theta_\xi}\left|\frac{\partial}{\partial\theta_l}g_{t,i}(\theta)\right|_1\right\|_4\left\|\sum\nolimits_{i=1}^{pk+1}\sup_{\theta\in\Theta_\xi}|g_{t,i}(\theta)|_1\right\|_4 \\
&= 2\left\|\epsilon_t\right\|_4^2\left\|\sup_{\theta\in\Theta_\xi}\left|\frac{\partial}{\partial\theta_l}g_t(\theta)\right|_1\right\|_4\left\|\sup_{\theta\in\Theta_\xi}|g_t(\theta)|_1\right\|_4 \leq M.
\end{aligned}$$

■