Econometrica Supplementary Material

SUPPLEMENT TO "TESTING A PARAMETRIC MODEL AGAINST A NONPARAMETRIC ALTERNATIVE WITH IDENTIFICATION THROUGH INSTRUMENTAL VARIABLES"

(*Econometrica*, Vol. 74, No. 2, March 2006, 521–538)

BY JOEL L. HOROWITZ

MATHEMATICAL APPENDIX: PROOFS OF THEOREMS

TO MINIMIZE THE COMPLEXITY of the presentation, it is assumed here that $p=1$ and $r=0$. The proofs for $p>1$ and/or $r>0$ are identical after replacing quantities for $p=1$, $r=0$ with the analogous quantities for the more general case. Let f_{XW} denote the density function of (X, W).

Define

$$S_{n1}(x) = n^{-1/2}\sum_{i=1}^{n} U_i f_{XW}(x, W_i),$$

$$S_{n2}(x) = n^{-1/2}\sum_{i=1}^{n} [g(X_i) - G(X_i, \theta_0)] f_{XW}(x, W_i),$$

$$S_{n3}(x) = n^{-1/2}\sum_{i=1}^{n} [G(X_i, \theta_0) - G(X_i, \hat{\theta}_n)] f_{XW}(x, W_i),$$

$$S_{n4}(x) = n^{-1/2}\sum_{i=1}^{n} U_i[\hat{f}_{XW}^{(-i)}(x, W_i) - f_{XW}(x, W_i)],$$

$$S_{n5}(x) = n^{-1/2}\sum_{i=1}^{n} [g(X_i) - G(X_i, \theta_0)][\hat{f}_{XW}^{(-i)}(x, W_i) - f_{XW}(x, W_i)],$$

and

$$S_{n6}(x) = n^{-1/2}\sum_{i=1}^{n} [G(X_i, \theta_0) - G(X_i, \hat{\theta}_n)][\hat{f}_{XW}^{(-i)}(x, W_i) - f_{XW}(x, W_i)].$$

Then $S_n(x) = \sum_{j=1}^{6} S_{nj}(x)$.

LEMMA 1: *As* $n \to \infty$,

$$\begin{aligned} S_{n3}(x) &= -\Gamma(x)' n^{1/2}(\hat{\theta}_n - \theta_0) + o_p(1) \\ &= -\Gamma(x)' n^{-1/2}\sum_{i=1}^{n} \gamma(U_i, X_i, W_i, \theta_0) + o_p(1) \end{aligned}$$

uniformly over $z \in [0, 1]$.

PROOF: A Taylor series expansion gives

$$S_{n3}(x) = -n^{-1}\sum_{i=1}^{n} G_\theta(X_i, \tilde{\theta}_n) f_{XW}(x, W_i) n^{1/2}(\hat{\theta}_n - \theta_0),$$

where $\tilde{\theta}_n$ is between $\hat{\theta}_n$ and θ_0. Application of Jennrich's (1969) uniform law of large numbers gives the first result of the lemma. The second result follows from the first by applying Assumption 3. *Q.E.D.*

LEMMA 2: *As* $n \to \infty$, $|\partial \hat{f}_{XW}^{(-i)}(x, w)/\partial z - \partial f_{XW}(x, w)/\partial z| = o[(\log n)/(n^{1/2}h^2) + h]$ *almost surely uniformly over* $(z, w) \in [0, 1]^2$.

PROOF: This is a modified version of Theorem 2.2(2) of Bosq (1996) and is proved the same way as that theorem. *Q.E.D.*

LEMMA 3: *As* $n \to \infty$, $S_{n4}(x) = o_p(1)$ *uniformly over* $x \in [0, 1]$.

PROOF: Let $I_1, \ldots, I_m$ be a partition of $[0, 1]$ into m intervals of length $1/m$. For each $j = 1, \ldots, m$, choose a point $x_j \in I_j$. Define $\Delta f_{XW}^{(-i)}(x, w) = \hat{f}_{XW}^{(-i)}(x, w) - f_{XW}(x, w)$. Then for any $\varepsilon > 0$,

$$\begin{aligned} S_{n4}(x) &= n^{-1/2}\sum_{j=1}^{m}\sum_{i=1}^{n} U_i I(x \in I_j)\Delta f_{XW}^{(-i)}(x, W_i) \\ &= n^{-1/2}\sum_{j=1}^{m}\sum_{i=1}^{n} U_i I(x \in I_j)\Delta f_{XW}^{(-i)}(x_j, W_i) \\ &\quad + n^{-1/2}\sum_{j=1}^{m}\sum_{i=1}^{n} U_i I(x \in I_j)[\Delta f_{XW}^{(-i)}(x, W_i) - \Delta f_{XW}^{(-i)}(x_j, W_i)] \\ &\equiv S_{n41}(x) + S_{n42}(x). \end{aligned}$$

A Taylor series expansion gives

$$S_{n42}(x) = n^{-1/2}\sum_{j=1}^{m}\sum_{i=1}^{n} U_i I(x \in I_j)[\partial \Delta f_{XW}^{(-i)}(\tilde{x}_j, W_i)/\partial x](x - x_j),$$

where $\tilde{x}_j$ is between x_j and x. Therefore, it follows from Lemma 2 that

$$\begin{aligned} |S_{n42}(x)| &\le n^{-1/2}m^{-1}\sum_{j=1}^{m}\sum_{i=1}^{n} |U_i| I(x \in I_j)|\partial \Delta f_{XW}^{(-i)}(\tilde{x}_j, W_i)/\partial x| \\ &\le n^{-1/2}m^{-1} o_p[(\log n)/(n^{1/2}h^2) + h] \end{aligned}$$

$$\times \sum_{j=1}^{m} \sum_{i=1}^{n} |U_i| I(x \in I_j)$$

$$= O_p[(\log n)/(mh^2) + n^{1/2}h/m]$$

uniformly over $x \in [0, 1]$. In addition, for any $\varepsilon > 0$,

$$\mathbf{P}\Big[\sup_{x\in[0,1]} |S_{n41}(x)| > \varepsilon\Big] = \mathbf{P}\left[\max_{1\le j\le m}\left|n^{-1/2}\sum_{i=1}^{n} U_i \Delta f_{XW}^{(-i)}(x_j, W_i)\right| > \varepsilon\right]$$
$$\le \sum_{j=1}^{m} \mathbf{P}\left[\left|n^{-1/2}\sum_{i=1}^{n} U_i \Delta f_{XW}^{(-i)}(x_j, W_i)\right| > \varepsilon\right].$$

However, $\mathbf{E}[U_i \Delta f_{XW}^{(-i)}(x_j, W_i)] = 0$ and standard calculations for kernel estimators show that

$$\operatorname{var}\left[n^{-1/2}\sum_{i=1}^{n} U_i \Delta f_{XW}^{(-i)}(x, W_i)\right] = O[(nh^2)^{-1} + h^4]$$

for any $x \in [0, 1]$. Therefore, it follows from Chebyshev's inequality that

$$\sum_{j=1}^{m} \mathbf{P}\left[\left|n^{-1/2}\sum_{i=1}^{n} U_i \Delta f_{XW}^{(-i)}(x_j, W_i)\right| > \varepsilon\right] = O[m/(nh^2\varepsilon^2) + mh^4/\varepsilon^2],$$

which implies that

$$\mathbf{P}\Big[\sup_{x\in[0,1]} |S_{n41}(x)| > \varepsilon\Big] = O[m/(nh^2\varepsilon^2) + mh^4/\varepsilon^2].$$

The lemma now follows by choosing m so that $n^{-1/2}m \to C_3$ as $n \to \infty$, where C_3 is a constant such that $0 < C_3 < \infty$. *Q.E.D.*

LEMMA 4: *As* $n \to \infty$, $S_{n6}(x) = o_p(1)$ *uniformly over* $x \in [0, 1]$.

PROOF: A Taylor series expansion gives

$$S_{n6}(x) = n^{-1}\sum_{i=1}^{n} G_\theta(X_i, \tilde{\theta}_n)[\hat{f}_{XW}^{(-i)}(x, W_i) - f_{XW}(x, W_i)]n^{1/2}(\hat{\theta}_n - \theta_0),$$

where $\tilde{\theta}_n$ is between $\hat{\theta}_n$ and θ_0. The result follows from boundedness of G_θ, $n^{1/2}(\hat{\theta}_n - \theta_0) = O_p(1)$, and $[\hat{f}_{XW}^{(-i)}(x, W_i) - f_{XW}(x, W_i)] = O[h^2 + (\log n)/(nh^2)^{1/2}]$ almost surely uniformly over $x \in [0, 1]$. *Q.E.D.*

LEMMA 5: *Under* H_0, $S_n(x) = B_n(x) + o_p(1)$ *uniformly over* $x \in [0, 1]$.

PROOF: Under H_0, $S_{n2}(x) = S_{n5}(x) = 0$ for all x. Now apply Lemmas 1, 3, and 4. *Q.E.D.*

PROOF OF THEOREM 1: Under H_0, $S_n(x) = B_n(x) + o_p(1)$ uniformly over $x \in [0, 1]$ by Lemma 5. Therefore,

$$\tau_n = \int_0^1 B_n^2(x)\,dx + o_p(1).$$

The result follows by writing $\int_0^1 [B_n^2(x) - \mathbf{E}B_n(x)^2]\,dx$ as a degenerate U statistic of order 2. See, for example, Serfling (1980, pp. 193–194). *Q.E.D.*

PROOF OF THEOREM 2: By Theorem 5.1a of Bhatia, Davis, and McIntosh (1983), $|\hat{\omega}_j - \tilde{\omega}_j| = O(\|\hat{\Omega} - \tilde{\Omega}\|)$. Moreover, standard calculations for kernel density estimators show that $\|\hat{\Omega} - \tilde{\Omega}\| = O[(\log n)/(nh^2)^{1/2}]$. Part (i) of the theorem follows by combining these two results. Part (ii) is an immediate consequence of part (i). *Q.E.D.*

PROOF OF THEOREM 3: Let $\tilde{z}_\alpha$ denote the $1 - \alpha$ quantile of the distribution of $\sum_{j=1}^{\infty} \tilde{\omega}_j \chi_{1j}^2$. Because of Theorem 2, it suffices to show that if H_1 holds, then under sampling from $Y = g(X) + U$,

$$\lim_{n\to\infty} \mathbf{P}(\tau_n > \tilde{z}_\alpha) = 1.$$

This will be done by proving that

$$\operatorname*{plim}_{n\to\infty} n^{-1}\tau_n = \int_0^1 [(Tq)(x)]^2\,dx > 0.$$

To do this, observe that by Jennrich's (1969) uniform law of large numbers, $n^{-1/2}S_{n2}(x) = (Tq)(x) + o_p(1)$ uniformly over $x \in [0, 1]$. Moreover, $S_{n5}(x) = o(h^{-1}\log n) = o(n^{1/6}\log n)$ a.s. uniformly over $x, w \in [0, 1]$ because $\hat{f}_{XW}^{(-i)}(x, w) - f_{XW}(x, w) = o[(\log n)/(nh^2)^{1/2}]$ a.s. uniformly over $x \in [0, 1]$. Combining these results with Lemma 5 yields

$$n^{-1/2}S_n(x) = n^{-1/2}B_n(x) + (Tq)(x) + o_p(1).$$

A further application of Jennrich's (1969) uniform law of large numbers shows that $n^{-1/2}S_n(x) \to^p (Tq)(x)$, so $n^{-1}\tau_n \to^p \int_0^1 [(Tq)(x)]^2\,dx$. *Q.E.D.*

PROOF OF THEOREM 4: Arguments like those leading to Lemma 5 show that

$$S_n(x) = B_n(x) + S_{n2}(x) + S_{n5}(x) - \mathbf{E}(W\Delta)'\tilde{\gamma}'(TG_\theta)(x) + o_p(1)$$

uniformly over $x \in [0, 1]$. Moreover,

$$S_{n5}(x) = n^{-1}\sum_{i=1}^{n} \Delta(X_i)[\hat{f}_{XW}^{(-i)}(x, W_i) - f_{XW}(x, W_i)]$$
$$= O[(\log n)/(nh^2)^{1/2}]$$

almost surely uniformly over x. In addition,

$$S_{n2}(x) = n^{-1}\sum_{i=1}^{n} \Delta(X_i) f_{XW}(x, W_i)$$
$$= (T\Delta)(x) + o(1)$$

almost surely uniformly over x. Therefore, $S_n(x) = B_n(z) + \mu(x) + o_p(1)$ uniformly over x. However,

$$B_n(x) + \mu(x) = \sum_{j=1}^{\infty} \tilde{b}_j \psi_j(x),$$

where $\tilde{b}_j = b_j + \mu_j$ and b_j is defined as in the proof of Theorem 1. The b_j's are asymptotically distributed as independent $N(\mu_j, \omega_j)$ variates. Now proceed as in Serfling's (1980, pp. 195–199) derivation of the asymptotic distribution of a second-order degenerate U statistic. *Q.E.D.*

PROOF OF THEOREM 5: Let $z_{g\alpha}$ denote the critical value under the model $Y = g(X) + U$, $g \in \mathcal{F}_{nc}$. Let $\hat{z}_{\varepsilon\alpha g}$ denote the corresponding estimated approximate critical value. Observe that $z_{g\alpha}$ is bounded and $\hat{z}_{\varepsilon\alpha g}$ is bounded in probability uniformly over $g \in \mathcal{F}_{nc}$.

We prove (2.12); the proof of (2.13) is similar. Define $D_n(x) = S_{n3}(x) + S_{n6}(x) + \mathbf{E}[S_{n2}(x) + S_{n5}(x)]$ and $\tilde{S}_n(x) = S_n(x) - D_n(x)$. Then $\tau_n = \|\tilde{S}_n + D_n\|^2$. Use the inequality

(A1) $$a^2 \geq 0.5b^2 - (b-a)^2$$

with $a = S_n$ and $b = D_n$ to obtain

$$\mathbf{P}(\tau_n > z_{g\alpha}) \geq \mathbf{P}(0.5\|D_n\|^2 - \|\tilde{S}_n\|^2 > z_{g\alpha}).$$

For any finite $M > 0$,

$$\mathbf{P}(0.5\|D_n\|^2 - \|\tilde{S}_n\|^2 \leq z_{g\alpha})$$
$$= \mathbf{P}(0.5\|D_n\|^2 \leq z_{g\alpha} + \|\tilde{S}_n\|^2, \|\tilde{S}_n\|^2 \leq M)$$

$$+ \mathbf{P}(0.5\|D_n\|^2 \leq z_{g\alpha} + \|\tilde{S}_n\|^2, \|\tilde{S}_n\|^2 > M)$$
$$\leq \mathbf{P}(0.5\|D_n\|^2 \leq z_{g\alpha} + M) + \mathbf{P}(\|\tilde{S}_n\|^2 > M),$$

where $\|\tilde{S}_n\|$ is bounded in probability uniformly over $g \in \mathcal{F}_{nc}$. Therefore, for each $\varepsilon > 0$ there is $M_\varepsilon < \infty$ such that, for all $M > M_\varepsilon$,

$$\mathbf{P}(0.5\|D_n\|^2 - \|\tilde{S}_n\|^2 \leq z_{g\alpha}) \leq \mathbf{P}(0.5\|D_n\|^2 \leq z_{g\alpha} + M) + \varepsilon.$$

Equivalently,

$$\mathbf{P}(0.5\|D_n\|^2 - \|\tilde{S}_n\|^2 > z_{g\alpha}) \geq \mathbf{P}(0.5\|D_n\|^2 > z_{g\alpha} + M) - \varepsilon$$

and

$$\text{(A2)} \qquad \mathbf{P}(\tau_n > z_{g\alpha}) \geq \mathbf{P}(0.5\|D_n\|^2 > z_{g\alpha} + M) - \varepsilon.$$

Now

$$S_{n2}(x) + S_{n5}(x) = n^{-1/2} \sum_{i=1}^{n} [g(X_i) - G(X_i, \theta_g)] \hat{f}_{XW}^{(-i)}(x, W_i).$$

Therefore,

$$\mathbf{E}[S_{n2}(x) + S_{n5}(x)]$$
$$= n^{-1/2} \mathbf{E} \sum_{i=1}^{n} [g(X_i) - G(X_i, \theta_g)][f_{XW}(x, W_i) + h^2 R_n(x)],$$

where $R_n(x)$ is nonstochastic, does not depend on g, and is bounded uniformly over $x \in [0, 1]$. It follows that

$$\mathbf{E}[S_{n2}(x) + S_{n5}(x)] = n^{1/2}(Tq_g)(x) + O[n^{1/2} h^2 \|q_g\|]$$

and

$$\mathbf{E}[S_{n2}(x) + S_{n5}(x)] \geq 0.5 n^{1/2}(Tq_g)(x)$$

uniformly over $g \in \mathcal{F}_{nc}$ for all sufficiently large n.

Now

$$|S_{n3}(x) + S_{n6}(x)|$$
$$\leq \sup_{\xi \in [0,1], g \in \mathcal{F}_{nc}} n^{1/2} |G(\xi, \hat{\theta}_n) - G(\xi, \theta_g)| n^{-1} \sum_{i=1}^{n} \hat{f}_{XW}^{(-i)}(x, W_i).$$

Therefore, it follows from the definition $\mathcal{F}_{nc}$ and uniform convergence of $\hat{f}_{XW}^{(-i)}$ to f_{XW} that $\|S_{n3} + S_{n6}\| = O_p(1)$ uniformly over $g \in \mathcal{F}_{nc}$. A further application of (A1) with $a = D_n(x)$ and $b = \mathbf{E}[S_{n2}(x) + S_{n5}(x)]$ gives

$$\text{(A3)} \qquad \|D_n\|^2 \geq 0.125n\|Tq_g\|^2 + O_p(1)$$

uniformly over $g \in \mathcal{F}_{nc}$ as $n \to \infty$. Inequality (2.12) follows by substituting (A3) into (A2) and choosing C to be sufficiently large. *Q.E.D.*

Dept. of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, U.S.A.; joel-horowitz@northwestern.edu.

REFERENCES

BHATIA, R., C. DAVIS, AND A. MCINTOSH (1983): "Perturbation of Spectral Subspaces and Solution of Linear Operator Equations," *Linear Algebra and Its Applications*, 52/53, 45–67.

BOSQ, D. (1996): *Nonparametric Statistics for Stochastic Processes*. New York: Springer-Verlag.

JENNRICH, R. I. (1969): "Asymptotic Properties of Non-Linear Least Squares Estimators," *The Annals of Mathematical Statistics*, 40, 633–643.

SERFLING, R. J. (1980): *Approximation Theorems of Mathematical Statistics*. New York: John Wiley & Sons.