**SUPPLEMENT TO**

# GMM ESTIMATION AND UNIFORM SUBVECTOR INFERENCE WITH POSSIBLE IDENTIFICATION FAILURE

**By**

**Donald W. K. Andrews and Xu Cheng**

**October 2011**

**COWLES FOUNDATION DISCUSSION PAPER NO. 1828**

**COWLES FOUNDATION FOR RESEARCH IN ECONOMICS**
**YALE UNIVERSITY**
**Box 208281**
**New Haven, Connecticut 06520-8281**

**http://cowles.econ.yale.edu/**

# Supplemental Appendices

for

# GMM Estimation and Uniform Subvector Inference with Possible Identification Failure

Donald W. K. Andrews
Cowles Foundation for Research in Economics
Yale University

Xu Cheng
Department of Economics
University of Pennsylvania

First Draft: August, 2007
Revised: October 24, 2011

# 11. Outline

This Supplement includes two Supplemental Appendices (denoted C and D) to the paper "GMM Estimation and Uniform Subvector Inference with Possible Identification Failure." Supplemental Appendix C gives some results that are used in the verification of the assumptions for the two examples in this paper. Supplemental Appendix D provides additional numerical results to those provided in the paper for the nonlinear regression model with endogeneity.

# 12. Supplemental Appendix C: Verification of Assumptions

In this Supplemental Appendix, we provide some results that are used in the main paper when verifying the assumptions in the two examples considered.

## 12.1. Law of Large Numbers and Central Limit Theorem

Here we state some results that are useful in the verification of Assumptions GMM1-GMM5. Specifically, Lemma 12.1 is a uniform convergence result for non-stochastic functions, Lemma 12.2 is a uniform LLN, and Lemma 12.3 is a CLT. The latter two results are for strong mixing triangular arrays. These are standard sorts of results. The proofs of these Lemmas are given in Appendix A of AC2.

**Lemma 12.1.** *Let* $\{q_n(\theta) : n \geq 1\}$ *be non-stochastic functions on* $\Theta$. *Suppose* (i) $q_n(\theta) \to 0 \ \forall \theta \in \Theta$, (ii) $||q_n(\theta_1) - q_n(\theta_2)|| \leq C\delta \ \forall \theta_1, \theta_2 \in \Theta$ *with* $||\theta_1 - \theta_2|| \leq \delta$, $\forall n \geq 1$, *for some* $C < \infty$ *and all* $\delta > 0$, *and* (iii) $\Theta$ *is compact. Then,* $\sup_{\theta \in \Theta} ||q_n(\theta)|| \to 0$.

**Assumption S1.** Under any $\gamma_0 \in \Gamma$, $\{W_i : i \geq 1\}$ is a strictly stationary and strong mixing sequence with mixing coefficients $\alpha_m \leq Cm^{-A}$ for some $A > d_\theta q/(q - d_\theta)$ and some $q > d_\theta \geq 2$, or $\{W_i : i \geq 1\}$ is an i.i.d. sequence and the constant $q$ equals $2 + \delta$ for some $\delta > 0$.

**Lemma 12.2.** *Suppose* (i) *Assumption* S1 *holds,* (ii) *for some function* $M_1(w) : \mathcal{W} \to R^+$ *and all* $\delta > 0$, $||s(w, \theta_1) - s(w, \theta_2)|| \leq M_1(w)\delta$, $\forall \theta_1, \theta_2 \in \Theta$ *with* $||\theta_1 - \theta_2|| \leq \delta$, $\forall w \in \mathcal{W}$, (iii) $E_\gamma \sup_{\theta \in \Theta} ||s(W_i, \theta)||^{1+\varepsilon} + E_\gamma M_1(W_i) \leq C \ \forall \gamma \in \Gamma$ *for some* $C < \infty$ *and*

$\varepsilon > 0$, *and* (iv) $\Theta$ *is compact. Then,* $\sup_{\theta\in\Theta} ||n^{-1}\sum_{i=1}^n s(W_i,\theta) - E_{\gamma_0}s(W_i,\theta)|| \to_p 0$ *under* $\{\gamma_n\} \in \Gamma(\gamma_0)$ *and* $E_{\gamma_0}s(W_i,\theta)$ *is uniformly continuous on* $\Theta$ $\forall\gamma_0 \in \Gamma$.

**Comment.** Note that the centering term in Lemma 12.2 is $E_{\gamma_0}s(W_i,\theta)$, rather than $E_{\gamma_n}s(W_i,\theta)$.

**Lemma 12.3.** *Suppose* (i) *Assumption* S1 *holds,* (ii) $s(w) \in R$ *and* $E_\gamma|s(W_i)|^q \le C$ $\forall\gamma \in \Gamma$ *for some* $C < \infty$ *and* $q$ *as in Assumption* S1. *Then,* $n^{-1/2}\sum_{i=1}^n(s(W_i) - E_{\gamma_n}s(W_i)) \to_d N(0, V_s(\gamma_0))$ *under* $\{\gamma_n\} \in \Gamma(\gamma_0)$ $\forall\gamma_0 \in \Gamma$, *where* $V_s(\gamma_0) = \sum_{m=-\infty}^{\infty} Cov_{\gamma_0}(s(W_i), s(W_{i+m}))$.

## 12.2. Probit Model with Endogeneity

Here we establish some results that are used when verifying Assumptions GMM1-GMM5 for the probit model with endogeneity.

### 12.2.1. Moment Conditions in (2.17)

First, we show that $E_{\gamma_0}(\ell(\theta)|X_i,Z_i)$ is maximized at $(a,\zeta_1) = (a_0,\zeta_{1,0})$. Note that

$$E_{\gamma_0}(\ell(\theta)|X_i,Z_i) = L_i(\theta_0)\log L_i(\theta) + (1 - L_i(\theta_0))\log(1 - L_i(\theta))$$

because $E_{\gamma_0}(y_i|X_i,Z_i) = L_i(\theta_0)$. Now we view $E_{\gamma_0}(\ell(\theta)|X_i,Z_i)$ as a function of $L_i(\theta)$. The first- and second-order derivatives of $E_{\gamma_0}(\ell(\theta)|X_i,Z_i)$ wrt $L_i(\theta)$ are

$$\begin{aligned}\frac{\partial}{\partial L_i(\theta)}E_{\gamma_0}(\ell(\theta)|X_i,Z_i) &= \frac{L_i(\theta_0) - L_i(\theta)}{L_i(\theta)(1 - L_i(\theta))} \text{ and}\\ \frac{\partial^2}{\partial L_i^2(\theta)}E_{\gamma_0}(\ell(\theta)|X_i,Z_i) &= -\frac{L_i(\theta_0) + L_i^2(\theta) - 2L_i(\theta)L_i(\theta_0)}{L_i^2(\theta)(1 - L_i(\theta))^2}. \qquad (12.1)\end{aligned}$$

The second-order derivative is negative for all $\theta \in \Theta$. When $L_i(\theta) = L_i(\theta_0)$, the first-order derivative is 0. Hence, $E_{\gamma_0}(\ell(\theta)|X_i,Z_i)$, viewed as a function of $L_i(\theta)$, has a unique global maxima at $L_i(\theta_0)$. Because the df of the standard normal distribution is strictly increasing, $E_{\gamma_0}(\ell(\theta)|X_i,Z_i)$ is maximized at $\theta$ if and only if $P_\phi(Z_i'(\beta\pi - \beta_0\pi_0) + X_i'(\zeta_1 - \zeta_{10}) = 0) = 1$. This implies that $E_{\gamma_0}(\ell(\theta)|X_i,Z_i)$ is maximized if and only if $\beta\pi = \beta_0\pi_0$ and $\zeta_1 = \zeta_{10}$ because $P_\phi(\overline{Z}_i'c = 0) < 1$ for $c \neq 0$.

### 12.2.2. Weight Matrix

In this section, we derive the elements of $\mathcal{W}_{e,i}(\theta;\gamma_0)$ in (8.2) and show that it is positive definite a.s. $\forall\theta\in\Theta$. Note that

$$P_{\gamma_0}(y_i=1|\overline{Z}_i)=L_i(\theta_0) \text{ and } P_{\gamma_0}(y_i=0|\overline{Z}_i)=1-L_i(\theta_0). \tag{12.2}$$

The upper left element of $\mathcal{W}_{e,i}(\theta;\gamma_0)$ is

$$\mathcal{W}_{11,i}(\theta)=E_{\gamma_0}(w_{1,i}(\theta)^2(y_i-L_i(\theta))^2|\overline{Z}_i)=w_{1,i}(\theta)^2(L_i(\theta_0)-2L_i(\theta_0)L_i(\theta)+L_i(\theta)^2). \tag{12.3}$$

To calculate the off-diagonal term of $\mathcal{W}_{e,i}(\theta;\gamma_0)$, note that

$$\begin{aligned}E_{\gamma_0}(V_i|\overline{Z}_i,y_i=1)&=E_{\gamma_0}(V_i|\overline{Z}_i,U_i>-(Z_i'\beta_0\pi_0+X_i'\zeta_{1,0}))=\sigma_v\rho\frac{L_i'(\theta_0)}{L_i(\theta_0)} \text{ and}\\ E_{\gamma_0}(V_i|\overline{Z}_i,y_i=0)&=E_{\gamma_0}(V_i|\overline{Z}_i,-U_i>Z_i'\beta_0\pi_0+X_i'\zeta_{1,0})=-\sigma_v\rho\frac{L_i'(\theta_0)}{1-L_i(\theta_0)}.\end{aligned} \tag{12.4}$$

The off-diagonal term of $\mathcal{W}_{e,i}(\theta;\gamma_0)$ is

$$\begin{aligned}&\mathcal{W}_{12,i}(\theta)\\ =\;&E_{\gamma_0}(w_{1,i}(\theta)(y_i-L_i(\theta))(Y_i-Z_i'\beta-X_i'\zeta_2)|\overline{Z}_i)\\ =\;&w_{1,i}(\theta)\sum_{k=0,1}(k-L_i(\theta))\left[E_{\gamma_0}(V_i|\overline{Z}_i,y_i=k)+Z_i'(\beta_0-\beta)+X_i'(\zeta_{2,0}-\zeta_2)\right]P_{\gamma_0}(y_i=k|\overline{Z}_i)\\ =\;&w_{1,i}(\theta)\left[(1-L_i(\theta))\sigma_v\rho\frac{L_i'(\theta_0)}{L_i(\theta_0)}L_i(\theta_0)+L_i(\theta)\sigma_v\rho\frac{L_i'(\theta_0)}{1-L_i(\theta_0)}(1-L_i(\theta_0)\right]+\\ &w_{1,i}(\theta)\left[(1-L_i(\theta))L_i(\theta_0)-L_i(\theta)(1-L_i(\theta_0))\right]\left[Z_i'(\beta_0-\beta)+X_i'(\zeta_{2,0}-\zeta_2)\right]\\ =\;&w_{1,i}(\theta)\left[\sigma_v\rho L_i'(\theta_0)+(L_i(\theta_0)-L_i(\theta))\left(Z_i'(\beta_0-\beta)+X_i'(\zeta_{2,0}-\zeta_2)\right)\right].\end{aligned} \tag{12.5}$$

The lower-right element of $\mathcal{W}_{e,i}(\theta;\gamma_0)$ is

$$\mathcal{W}_{22,i}(\theta)=E_{\gamma_0}((Y_i-Z_i'\beta-X_i'\zeta_2)^2|\overline{Z}_i)=\sigma_v^2+(Z_i'(\beta_0-\beta)+X_i'(\zeta_{2,0}-\zeta_2))^2. \tag{12.6}$$

Next we show that $\mathcal{W}_{e,i}(\theta;\gamma_0)$ is positive definite a.s. when $\theta=(\psi_0,\pi)$. This holds if

$$\begin{aligned}&\mathcal{W}_{11,i}(\theta)\mathcal{W}_{22,i}(\theta)-\mathcal{W}_{12,i}(\theta)^2\\ =\;&\sigma_v^2w_{1,i}(\theta)^2\left[L_i(\theta_0)-2L_i(\theta_0)L_i(\theta)+L_i(\theta)^2-\rho^2L_i'(\theta_0)^2\right]>0 \text{ a.s.}\end{aligned} \tag{12.7}$$

Note that

$$L_i(\theta_0) - 2L_i(\theta_0)L_i(\theta) + L_i(\theta)^2 = (L_i(\theta) - L_i(\theta_0))^2 + L_i(\theta_0) - L_i(\theta_0)^2 \tag{12.8}$$
$$\geq L_i(\theta_0)(1 - L_i(\theta_0)) = \lambda(-Z_i'\beta_0\pi - X_i'\zeta_{1,0})\lambda(Z_i'\beta_0\pi + X_i'\zeta_{1,0})L_i'(\theta_0)^2 > \rho^2 L_i'(\theta_0)^2 \text{ a.s.},$$

where $\lambda(x) = (1 - L(x))/L'(x)$ for $x \in R$. The last inequality holds because $\log \lambda(x)$ is strictly convex (see Baricz (2008)), which implies that $\lambda(-Z_i'\beta_0\pi - X_i'\zeta_{1,0})\lambda(Z_i'\beta_0\pi + X_i'\zeta_{1,0}) > \lambda(0) > 1 \geq \rho^2$ a.s. Moreover, $\mathcal{W}_{11,i}(\theta), \mathcal{W}_{22,i}(\theta) > 0\ \forall\theta \in \Theta$. Hence, $\mathcal{W}_{e,i}(\theta; \gamma_0)$ is positive definite a.s. when $\theta = (\psi_0, \pi)$.

## 13. Supplemental Appendix D: Numerical Results

Here we report some additional numerical results for the nonlinear regression model with endogeneity.

Figures S-1 and S-2 report asymptotic and finite-sample ($n = 500$) densities of the estimators for $\beta$ and $\pi$ when $\pi_0 = 3.0$. Figures S-3 to S-6 report asymptotic and finite-sample ($n = 500$) densities of the $t$ and QLR statistics for $\beta$ and $\pi$ when $\pi_0 = 1.5$. Figures S-7 and S-8 report CP's of nominal 0.95 standard and robust $|t|$ and QLR CI's for $\beta$ and $\pi$ when $\pi_0 = 3.0$.

## Reference

Baricz, A. (2008) Mills' ratio: monotonicity patterns and functional inequalities. *Journal of Mathematical Analysis and Applications* 340, 1362–1370.

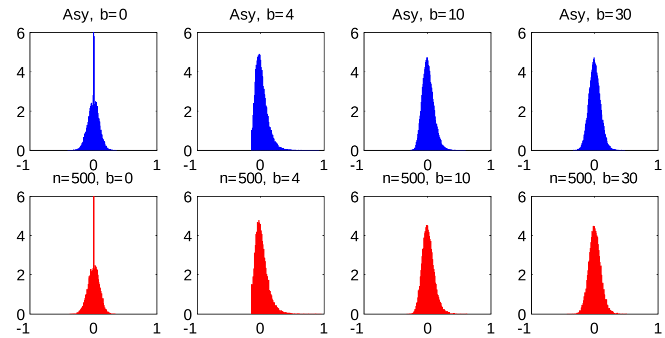


Figure S-1. Asymptotic and Finite-Sample ($n = 500$) Densities of the Estimator of $\beta$ in the Nonlinear Regression Model with Endogeneity when $\pi_0 = 3.0$.

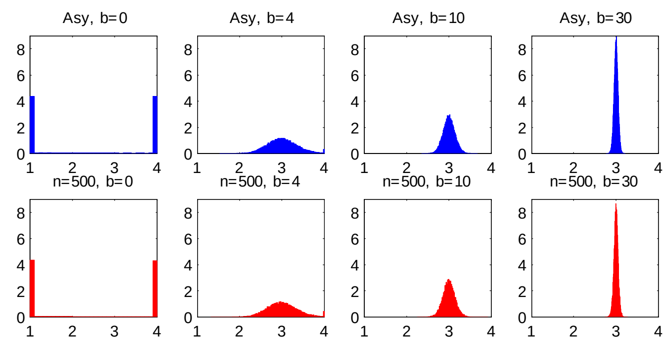


Figure S-2. Asymptotic and Finite-Sample ($n = 500$) Densities of the Estimator of $\pi$ in the Nonlinear Regression Model with Endogeneity when $\pi_0 = 3.0$.

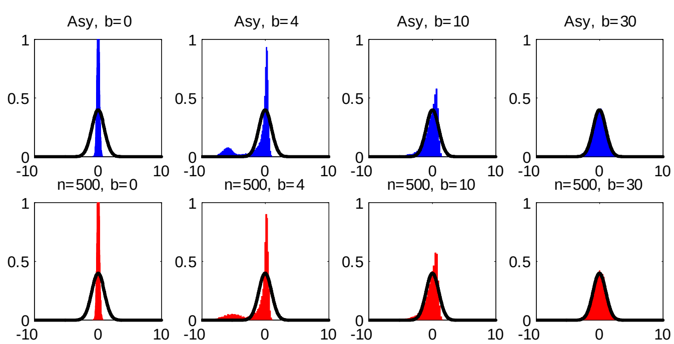


Figure S-3. Asymptotic and Finite-Sample ($n = 500$) Densities of the $t$ Statistic for $\beta$ in the Nonlinear Regression Model with Endogeneity when $\pi_0 = 1.5$ and the Standard Normal Density (Black Line).

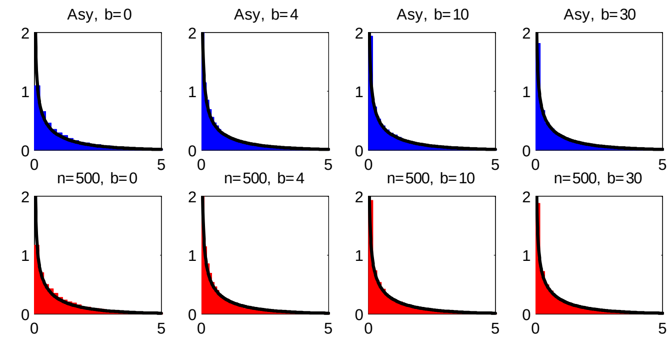


Figure S-4. Asymptotic and Finite-Sample (n=500) Densities of the QLR Statistic for $\beta$ in the Nonlinear Regression Model with Endogeneity when $\pi_0 = 1.5$ and the $\chi^2_1$ Density (Black Line).

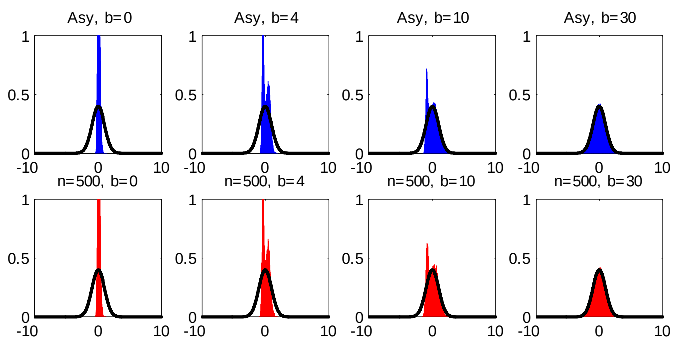


Figure S-5. Asymptotic and Finite-Sample ($n = 500$) Densities of the $t$ Statistic for $\pi$ in the Nonlinear Regression Model with Endogeneity when $\pi_0 = 1.5$ and the Standard Normal Density (Black Line).

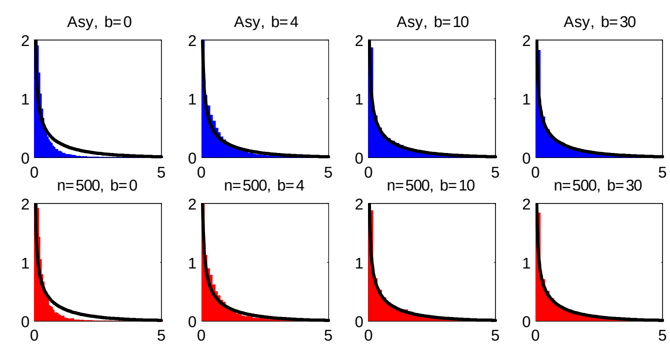


Figure S-6. Asymptotic and Finite-Sample (n=500) Densities of the QLR Statistic for $\pi$ in the Nonlinear Regression Model with Endogeneity when $\pi_0 = 1.5$ and the $\chi^2_1$ Density (Black Line).

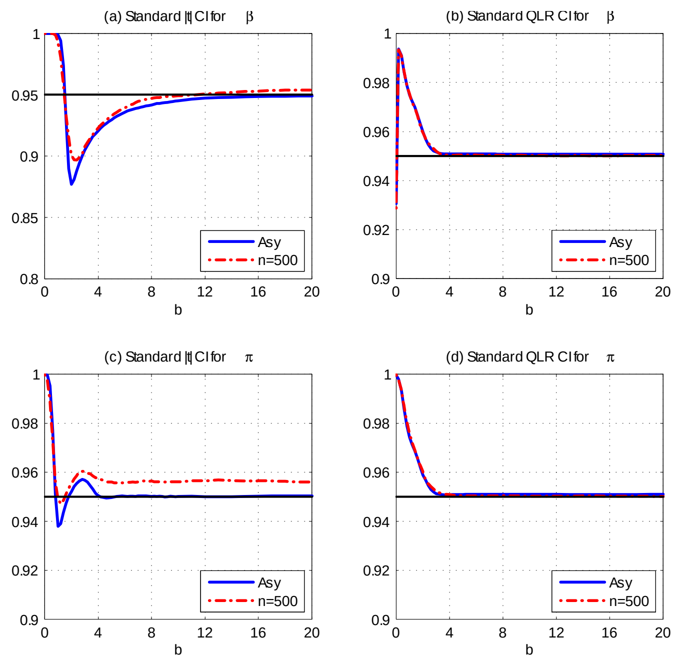


Figure S-7. Coverage Probabilities of Standard $|t|$ and QLR CI's for $\beta$ and $\pi$ in the Nonlinear Regression Model with Endogeneity when $\pi_0 = 3.0$.

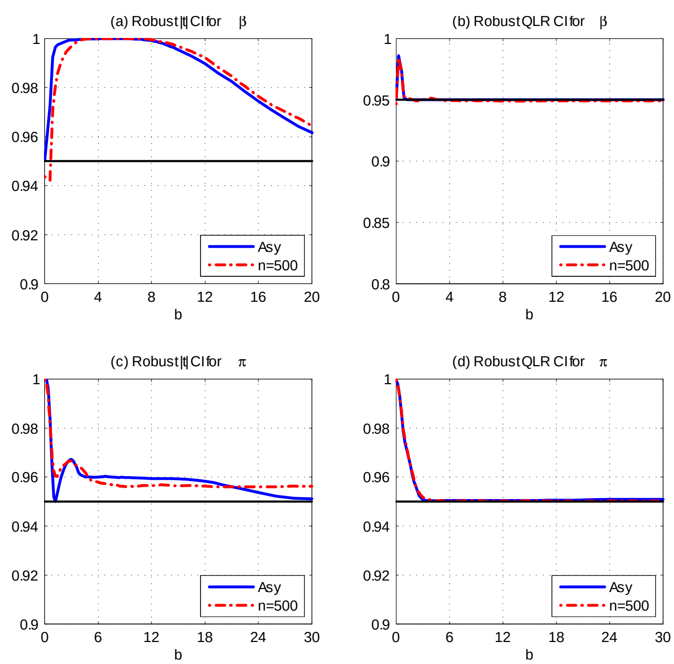


Figure S-8. Coverage Probabilities of Robust $|t|$ and QLR CI's for $\beta$ and $\pi$ in the Nonlinear Regression Model with Endogeneity when $\pi_0 = 3.0$, $\kappa = 1.5, D = 1$, and $s(x) = \exp(-2x)$.