Econometrics 710
Final Exam, Spring 2009
Sample Answers

1.

(a) Estimator:

$$\begin{aligned}\hat{\beta} &= \left(\sum_{i=1}^{n} x_i x_i'\right)^{-1}\left(\sum_{i=1}^{n} x_i y_i\right) \\ \hat{e}_i &= y_i - x_i'\hat{\beta} \\ \hat{\mu}_3 &= \frac{1}{n}\sum_{i=1}^{n} \hat{e}_i^3\end{aligned}$$

(b) Percentile Bootstrap

i. Draw an observation (y_i^*, x_i^*) randomly from the observed sample $\{y_i, x_i\}$

ii. Repeat this n times, to obtain a sample $\{y_i^*, x_i^*\}$, $i = 1, ..., n$

iii. Compute the estimator $\hat{\mu}_3$ on this bootstrap sample. This is

$$\begin{aligned}\hat{\beta}^* &= \left(\sum_{i=1}^{n} x_i^* x_i^{*\prime}\right)^{-1}\left(\sum_{i=1}^{n} x_i^* y_i^*\right) \\ \hat{e}_i^* &= y_i^* - x_i^{*\prime}\hat{\beta}^* \\ \hat{\mu}_3^* &= \frac{1}{n}\sum_{i=1}^{n} \hat{e}_i^{*3}\end{aligned}$$

iv. Repeat this B times, to obtain $\{\hat{\mu}_{3b}^*\}$, $b = 1, ..., B$.

v. Let $\hat{q}_{.05}$ and $\hat{q}_{.95}$ be the 5% and 95% empirical quantiles of $\hat{\mu}_{3b}^*$. These can be computed as the $[.05(B+1)]$'th and $[.95(B+1)]$'th order statistics of $\hat{\mu}_{3b}^*$

vi. The Efron percentile confidence interval for μ_3 is $[\hat{q}_{.05}, \hat{q}_{.95}]$

2.

(a) The estimated covariance matrix for $\hat{\beta}_1, \hat{\beta}_2$ is

$$\hat{V}_\beta = \begin{bmatrix} s(\hat{\beta}_1)^2 & \hat{\rho} s(\hat{\beta}_1) s(\hat{\beta}_2) \\ \hat{\rho} s(\hat{\beta}_1) s(\hat{\beta}_2) & s(\hat{\beta}_2)^2 \end{bmatrix}$$

(recall that the standard errors are the square roots of the diagaonal elements of the estimated covariance matrix).

The the variance estimate for $\hat{\theta} = \hat{\beta}_1 - \hat{\beta}_2$ is

$$\hat{V}_\theta = \begin{pmatrix} 1 & -1 \end{pmatrix} \hat{V}_\beta \begin{pmatrix} 1 \\ -1 \end{pmatrix} = s(\hat{\beta}_1)^2 - 2\rho s(\hat{\beta}_1)^2 s(\hat{\beta}_2)^2 + s(\hat{\beta}_2)^2$$

the standard error is

$$s(\hat{\theta}) = \sqrt{s(\hat{\beta}_1)^2 - 2\rho s(\hat{\beta}_1)^2 s(\hat{\beta}_2)^2 + s(\hat{\beta}_2)^2}$$

An asymptotic 95% confidence interval for θ is

$$\hat{C} = \hat{\theta} \pm 2s(\hat{\theta}) = \hat{\beta}_1 - \hat{\beta}_2 \pm 2\sqrt{s(\hat{\beta}_1)^2 - 2\hat{\rho} s(\hat{\beta}_1)^2 s(\hat{\beta}_2)^2 + s(\hat{\beta}_2)^2}$$

(b) No. All we know is $-1 < \hat{\rho} < 1$

(c) At the observed values, the confidence interval for θ is

$$\hat{C} = 0.2 \pm 2\,(.07)\,\sqrt{2\,(1-\hat{\rho})} \simeq 0.2 \pm 0.2\sqrt{(1-\hat{\rho})}$$

We don't know $\hat{\rho}$, but the question is whether or not we can reject $\theta = 0$, or equivalently if $\hat{C}$ includes 0. Looking at the above interval, this occurs if and only if $\hat{\rho} > 0$. Therefore the reported evidence, by itself, does not support the conclusion that β_1 exceeds β_2. While it may be true that $\hat{\rho} > 0$ (and then we can reject $\beta_1 = \beta_2$), this is unknown from the reported information, so the author's claim is (at best) unsupported.

3.

(a) Using matrix notation

$$\hat{\beta} = \left(X'ZWZ'X\right)^{-1}\left(X'ZWZ'Y\right)$$

(b) By the law of iterated expecatations

$$\begin{aligned} E\left(z_i e_i\right) &= E\left(E\left(z_i e_i \mid z_i\right)\right) \\ &= E\left(z_i E\left(e_i \mid z_i\right)\right) \\ &= E\left(z_i E\left(\delta n^{-1/2} + u_i \mid z_i\right)\right) \\ &= E\left(z_i \delta n^{-1/2}\right) \\ &= \mu_z \delta n^{-1/2} \\ &\neq 0 \end{aligned}$$

The third step using $e_i = \delta n^{-1/2} + u_i$ and the fourth using $E\left(u_i \mid z_i\right) = 0$. This contradicts (1).

(c) Since $Y = X\beta + e$

$$\sqrt{n}\left(\hat{\beta} - \beta\right) = \sqrt{n}\left(X'ZWZ'X\right)^{-1}\left(X'ZWZ'e\right)$$

Then using $e = \delta n^{-1/2} + u$

$$\begin{aligned} \sqrt{n}\left(\hat{\beta} - \beta\right) &= \sqrt{n}\left(X'ZWZ'X\right)^{-1}\left(X'ZW\left(Z'\delta n^{-1/2} + Z'u\right)\right) \\ &= \left(\frac{1}{n}X'ZW\frac{1}{n}Z'X\right)^{-1}\left(\frac{1}{n}X'ZW\left(\frac{1}{n}Z'\delta + \frac{1}{\sqrt{n}}Z'u\right)\right) \end{aligned}$$

(d) By the WLLN,

$$\frac{1}{n}\sum_{i=1}^{n} z_i x_i' \rightarrow_p E\left(z_i x_i'\right) = Q$$

$$\frac{1}{n}\sum_{i=1}^{n} z_i \rightarrow_p E\left(z_i\right) = \mu_z$$

and by the CLT

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i \rightarrow_d N(0, \Omega)$$

where

$$\Omega = E\left(z_i z_i' u_i^2\right)$$

Thus

$$\frac{1}{n}Z'\delta + \frac{1}{\sqrt{n}}Z'u \rightarrow_d \mu_z\delta + N(0,\Omega) = N(\mu_z\delta, \Omega)$$

and

$$\sqrt{n}\left(\hat{\beta}-\beta\right) = \left(\frac{1}{n}X'ZW\frac{1}{n}Z'X\right)^{-1}\left(\frac{1}{n}X'ZW\left(\frac{1}{n}Z'\delta+\frac{1}{\sqrt{n}}Z'u\right)\right)$$
$$\rightarrow_d \left(Q'WQ\right)^{-1}Q'WN(\mu_z\delta,\Omega)$$
$$N((Q'WQ)^{-1}Q'W\mu_z\delta, (Q'WQ)^{-1}Q'W\Omega WQ\left(Q'WQ\right)^{-1})$$

The asymptotic distribution is non-central normal with a classic covariance matrix. Thus the GMM estimator is asymptotically biased.

4.

(a) No statistical model was specified. Therefore least-squares regression is estimating projections. The coefficient in (3) is the projection

$$\beta = \left(Ex_ix_i'\right)^{-1}Ex_iy_i$$

which defines the error

$$e_i = y_i - x_i'\beta$$

The coefficient being estimated in (4) is the projection of e_i on z_i

$$\begin{aligned}\gamma &= \left(Ez_iz_i'\right)^{-1}Ez_ie_i \\ &= \left(Ez_iz_i'\right)^{-1}Ez_i\left(y_i - x_i'\beta\right) \\ &= \left(Ez_iz_i'\right)^{-1}\left(Ez_iy_i - Ez_ix_i'\left(Ex_ix_i'\right)^{-1}Ex_iy_i\right)\end{aligned}$$

(b) We know that $\hat{\beta}\rightarrow_p\beta$. Then

$$\begin{aligned}\tilde{\gamma} &= \left(Z'Z\right)^{-1}\left(Z'\hat{e}\right) \\ &= \left(\frac{1}{n}Z'Z\right)^{-1}\left(\frac{1}{n}Z'e\right) - \left(\frac{1}{n}Z'Z\right)^{-1}\left(\frac{1}{n}Z'X\right)\left(\hat{\beta}-\beta\right)\end{aligned}$$

$$\rightarrow_p \left(Ez_iz_i'\right)^{-1}Ez_ie_i = \gamma$$

(c) This is

$$W_n = n\tilde{\gamma}'\tilde{V}_\gamma^{-1}\tilde{\gamma}$$

where

$$\begin{aligned}\tilde{V}_\gamma &= \left(\frac{1}{n}Z'Z\right)^{-1}\tilde{\Omega}\left(\frac{1}{n}Z'Z\right)^{-1} \\ \tilde{\Omega} &= \frac{1}{n}\sum_{i=1}^{n}z_iz_i'\tilde{u}_i^2\end{aligned}$$

Since this is a test on all coefficients, we can simplify this to

$$W_n = \left(\frac{1}{\sqrt{n}}\tilde{e}'Z\right)\tilde{\Omega}^{-1}\left(\frac{1}{\sqrt{n}}Z'\hat{e}\right)$$

(d) When $\gamma = 0$, $Ez_ie_i = 0$, so

$$\frac{1}{\sqrt{n}}Z'e \rightarrow_d N(0,\Omega)$$

where

$$\Omega = Ez_iz_i'e_i^2$$

Then

$$\frac{1}{\sqrt{n}}Z'\hat{e} = \frac{1}{\sqrt{n}}Z'e - \frac{1}{n}Z'X\sqrt{n}\left(\hat{\beta}-\beta\right)$$
$$\rightarrow_d N(0,\Omega) + E(z_i x_i')O_p(1)$$

which simplifies to $N(0,\Omega)$ when $E(z_i x_i') = 0$. We find that

$$W_n = \left(\frac{1}{\sqrt{n}}\tilde{e}'Z\right)\tilde{\Omega}^{-1}\left(\frac{1}{\sqrt{n}}Z'\hat{e}\right)$$
$$\rightarrow_d N(0,\Omega)'\Omega^{-1}N(0,\Omega) = \chi^2_\ell$$

A complete proof would show that $\tilde{\Omega} \rightarrow_p \Omega$.

Even though the statistic W_n has ignored the two-step estimation, the asymptotic distribution is still standard. The reason is that the two steps are asymptotically independent when z_i and x_i are uncorrelated.

(e) If $E(z_i x_i') \neq 0$ then the distribution is no long χ^2_ℓ. The reason is that the sampling distribution of $\sqrt{n}\left(\hat{\beta}-\beta\right)$ no longer drops out. For simplicity, suppose that the error is homoskedastic. Then

$$\frac{1}{\sqrt{n}}Z'\hat{e} = \frac{1}{\sqrt{n}}Z'e - \frac{1}{n}Z'X\left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{\sqrt{n}}X'e\right)$$
$$\rightarrow_d \begin{pmatrix} I & -Q_{zx}Q_{xx}^{-1} \end{pmatrix} N\left(0, \begin{bmatrix} Q_{zz} & Q_{zx} \\ Q_{xz} & Q_{xx} \end{bmatrix}\sigma^2\right)$$
$$= N\left(0, \left(Q_{zz} - Q_{zx}Q_{xx}^{-1}Q_{xz}\right)\sigma^2\right)$$

and $\Omega = Q_{zz}\sigma^2$. Hence

$$W_n \rightarrow_d N(0, Q_{zz} - Q_{zx}Q_{xx}^{-1}Q_{xz})Q_{zz}^{-1}N(0, Q_{zz} - Q_{zx}Q_{xx}^{-1}Q_{xz})$$

which is not χ^2_ℓ. Thus in the general case (where z_i and x_i are correlated), the statistic W_n has incorrectly ignored the two-step estimation problem.