

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  $\{y_i^*, x_i^*\}, i = 1, \dots, n$
- iii. Compute the estimator  $\hat{\mu}_3$  on this bootstrap sample. This is

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

- iv. Repeat this  $B$  times, to obtain  $\{\hat{\mu}_{3b}^*\}, b = 1, \dots, B$ .
- 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  $[\cdot 05(B+1)]$ 'th and  $[\cdot 95(B+1)]$ 'th order statistics of  $\hat{\mu}_{3b}^*$
- vi. The Efron percentile confidence interval for  $\mu_3$  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 diagonal elements of the estimated covariance matrix).

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

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

the standard error is

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

An asymptotic 95% confidence interval for  $\theta$  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  $\theta$  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  $\beta_1$  exceeds  $\beta_2$ . 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

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

(b) By the law of iterated expectations

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

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

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

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

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

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

(d) By the WLLN,

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n z_i x_i' &\rightarrow_p E(z_i x_i') = Q \\ \frac{1}{n} \sum_{i=1}^n z_i &\rightarrow_p E(z_i) = \mu_z \end{aligned}$$

and by the CLT

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

where

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

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

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

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

$$\beta = (Ex_i x_i')^{-1} Ex_i y_i$$

which defines the error

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

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

$$\begin{aligned}\gamma &= (Ez_i z_i')^{-1} Ez_i e_i \\ &= (Ez_i z_i')^{-1} Ez_i (y_i - x_i' \beta) \\ &= (Ez_i z_i')^{-1} (Ez_i y_i - Ez_i x_i' (Ex_i x_i')^{-1} Ex_i y_i)\end{aligned}$$

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

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

- (c) This is

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

where

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

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

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

- (d) When  $\gamma = 0$ ,  $Ez_i e_i = 0$ , so

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

where

$$\Omega = Ez_i z_i' e_i^2$$

Then

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

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

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

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  $\chi_\ell^2$ . The reason is that the sampling distribution of  $\sqrt{n}(\hat{\beta} - \beta)$  no longer drops out. For simplicity, suppose that the error is homoskedastic. Then

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

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  $\chi_\ell^2$ . 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.