Econometrics 710
Final Exam, Spring 2016
Sample Answers

1.

(a) $m(x) = E(y_i|x_i = x) = x'\beta$
$\widehat{\beta} = (X'X)^{-1} X'Y$
$\widehat{m}(x) = x'\widehat{\beta}$

(b) $\sqrt{n}\left(\widehat{\beta} - \beta\right) \to_d N(0, V)$ where $V = Q^{-1}\Omega Q^{-1}$
$\sqrt{n}(\widehat{m}(x) - m(x)) = \sqrt{n}x\left(\widehat{\beta} - \beta\right) \to_d N(0, x'Vx)$
$C = \widehat{m}(x) \pm 2\sqrt{n^{-1}x'\widehat{V}x}$ where $\widehat{V} = \widehat{Q}^{-1}\widehat{\Omega}\widehat{Q}^{-1}$, $\widehat{Q} = n^{-1}X'X$, $\widehat{\Omega} = n^{-1}\sum_{i=1}^n x_i x_i' \widehat{e}_i^2$, $\widehat{e}_i = y_i - x_i'\widehat{\beta}$

(c) percentile bootstrap confidence interval

- Generate an iid sample $\{y_i^*, x_i^*\}$ by drawing with replacement from the sample
- Set $\widehat{\beta}^* = (X^{*\prime}X^*)^{-1} X^{*\prime}Y^*$, $\widehat{m}^*(x) = x'\widehat{\beta}^*$. Note x is fixed
- Repeat B times. Collect the B values of $\widehat{m}^*(x)$
- Sort the B values of $\widehat{m}^*(x)$ and take the 2.5% and 97.5% quantiles, q_1 and q_2, of its empirical distribution
- $[q_1, q_2]$ is the percentile interval for $m(x)$

(d) percentile-t bootstrap confidence interval

- Generate an iid sample $\{y_i^*, x_i^*\}$ by drawing with replacement from the sample
- Set $\widehat{\beta}^* = (X^{*\prime}X^*)^{-1} X^{*\prime}Y^*$, $\widehat{m}^*(x) = x'\widehat{\beta}^*$, $\widehat{V}^* = \widehat{Q}^{*-1}\widehat{\Omega}^*\widehat{Q}^{*-1}$, $\widehat{Q}^* = n^{-1}X^{*\prime}X^*$, $\widehat{\Omega}^* = n^{-1}\sum_{i=1}^n x_i^* x_i^{*\prime} \widehat{e}_i^{*2}$, $\widehat{e}_i^* = y_i^* - x_i^{*\prime}\widehat{\beta}^*$
- Set $t^* = (\widehat{m}^*(x) - \widehat{m}(x)) / \sqrt{n^{-1}x'\widehat{V}^*x}$
- Repeat B times. Collect the B values of t^*
- Sort the B values of $\widehat{m}^*(x)$ and take the 2.5% and 97.5% quantiles, q_1 and q_2, of its empirical distribution
- $[\widehat{m}(x) - q_2\sqrt{n^{-1}x'\widehat{V}x},\ \widehat{m}(x) - q_2\sqrt{n^{-1}x'\widehat{V}x}]$ is the percentile-t interval for $m(x)$

2. Take

$$y_{it} = x_{it}'\beta + u_i + e_{it}$$

Take individual-specific means

$$\overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}, \text{ etc.}$$

Applied to the equation:

$$\overline{y}_i = \overline{x}_i'\beta + u_i + \overline{e}_i$$

Subtract:

$$y_{it}^* = x_{it}^{*\prime}\beta + e_{it}^*$$

where $y_{it}^* = y_{it} - \overline{y}_i$ etc. Estimate by 2SLS (IV) using z_{it}^* as an instrument

$$\widehat{\beta} = \left(\sum_{t=1}^{T} \sum_{i=1}^{n} z_{it}^* x_{it}^{*\prime} \right)^{-1} \left(\sum_{t=1}^{T} \sum_{i=1}^{n} z_{it}^* y_{it}^* \right)$$

Or, equivalently, use z_{it} as an instrument.

3.

(a) Since $Y = Z\gamma\beta + e$

$$\begin{aligned} \widehat{\beta} &= \left(\widehat{\gamma}' Z' Z \widehat{\gamma}\right)^{-1} \widehat{\gamma}' Z' Y \\ &= \left(\widehat{\gamma}' Z' Z \widehat{\gamma}\right)^{-1} \widehat{\gamma}' Z' Z \gamma\beta + \left(\widehat{\gamma}' Z' Z \widehat{\gamma}\right)^{-1} \widehat{\gamma}' Z' e \\ &= \left(\widehat{\gamma}' \left(\frac{1}{n} Z' Z \right) \widehat{\gamma} \right)^{-1} \widehat{\gamma}' \left(\frac{1}{n} Z' Z \right) \gamma\beta + \left(\widehat{\gamma}' \left(\frac{1}{n} Z' Z \right) \widehat{\gamma} \right)^{-1} \widehat{\gamma}' \left(\frac{1}{n} Z' e \right) \end{aligned}$$

Since $E(x_i | z_i) = \gamma' z_i$, by standard regression results, $\widehat{\gamma} \to_p \gamma$ as $n \to \infty$. Also, $\frac{1}{n} Z' Z \to_p Q = E(z_i z_i')$ and $\frac{1}{n} Z' e \to_p E(z_i e_i) = 0$, the latter since the assumptions state $E(e_i | z_i) = 0$. Applied to the above equation we find

$$\widehat{\beta} \to_p \left(\gamma' Q \gamma\right)^{-1} \gamma' Q \gamma \beta + \left(\gamma' Q \gamma\right)^{-1} \gamma' 0 = \beta$$

and hence $\widehat{\beta}$ is consistent for β

(b) From the above expression

$$\begin{aligned} \sqrt{n}\left(\widehat{\beta} - \beta\right) &= \left(\widehat{\gamma}' \left(\frac{1}{n} Z' Z \right) \widehat{\gamma} \right)^{-1} \widehat{\gamma}' \left(\frac{1}{n} Z' Z \right) \gamma\beta - \beta & (1) \\ &+ \left(\widehat{\gamma}' \left(\frac{1}{n} Z' Z \right) \widehat{\gamma} \right)^{-1} \widehat{\gamma}' \left(\frac{1}{\sqrt{n}} Z' e \right) & (2) \end{aligned}$$

Under the assumption that $\beta = 0$, (1) equals 0 so only (2) remains.Thus

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) = \left(\widehat{\gamma}' \left(\frac{1}{n} Z' Z \right) \widehat{\gamma} \right)^{-1} \widehat{\gamma}' \left(\frac{1}{\sqrt{n}} Z' e \right).$$

Under the CLT, $\frac{1}{\sqrt{n}} Z' e \to_d N(0, \Omega)$ where $\Omega = E\left(z_i z_i' e_i^2\right)$. Hence

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) \to_d \left(\gamma' Q \gamma\right)^{-1} \gamma' N(0, \Omega) = N\left(0, \left(\gamma' Q \gamma\right)^{-1} \left(\gamma' \Omega \gamma\right) \left(\gamma' Q \gamma\right)^{-1}\right)$$

(c) Without $\beta = 0$ the component (1) is not eliminated. Since $\widehat{\gamma}$ converges at rate $n^{1/2}$ this term affects the asymptotic distribution.

(d) The t-ratio is

$$
\begin{aligned}
t &= \frac{\widehat{\beta}}{\sqrt{\frac{1}{n}\left(\widehat{\gamma}'\widehat{Q}\widehat{\gamma}\right)^{-1}\left(\gamma'\widehat{\Omega}\gamma\right)\left(\widehat{\gamma}'\widehat{Q}\widehat{\gamma}\right)^{-1}}} \\
\widehat{Q} &= n^{-1}\sum_{i=1}^{n}\widehat{\pi}_i\widehat{\pi}_i' \\
\widehat{\Omega} &= n^{-1}\sum_{i=1}^{n}\widehat{\pi}_i\widehat{\pi}_i'\widehat{e}_i^2 \\
\widehat{e}_i &= y_i - \widehat{\pi}_i\widehat{\beta}
\end{aligned}
$$

The t-ratio is asymptotically normal if $\widehat{\gamma} \to_p \gamma$, $\widehat{Q} \to_p Q$ and $\widehat{\Omega} \to_p \Omega$. The first two were shown in part (a). Showing that $\widehat{\Omega}$ converges is too detailed for an exam.

An asympotic 5% test rejects H_0 if $|t| > 2$, otherwise fails to reject the hypothesis.

The test is appropriate for the hypothesis $H_0 : \beta = 0$ because asymptotic normality holds for $\beta = 0$ as shown in part (b). However, the test is not appropriate for other values of β due to the issue raised in part (c). Thus the standard errors should not be used to form a confidence interval for β.

4.

(a) Rejection of a test is a Bernoulli random variable, e.g. iid {0,1} with $P(1) = p$. The null is $p = 0.05$. The variance of the estimated size from B simulation replications is $p(1-p)/B$. So under $H_0 : p = 0.05$ and $B = 100$ the theoretical standard deviation is $\sqrt{(.05)(.95)/100} = 0.022$. Thus 0.07 is about 1 standard deviation away from the null hypothesis of 0.05. This is not statistically significant based on an asymptotic (large B) test. You could alternative calculate the exact probability using the Binomial distribution.

I would tell my colleague that the numerical finding does NOT suggest that the test is over-rejecting. The evidence is consistent with random noise. They should increase the number of replications to determine a more precise estimate.

(b) If $B = 1000$ then the standard deviation is $\sqrt{(.05)(.95)/1000} = 0.007$. So $(.07 - .05)/.007 = 2.9$ exceeds significance at any conventional level. With this evidence I would agree with my colleauge that the test over-rejects the 5% goal.

However, I may also point out that 7% rejection instead of the nominal 5% is not excessive. This is mild case of over-rejection.

5.

(a) The GMM criterion is proportional to

$$J_n(\beta) = \frac{1}{2}\left(Y - X\beta\right)' Z\Omega^{-1}Z'\left(Y - X\beta\right).$$

The FOC for mimimization is

$$0 = \frac{\partial}{\partial\beta}J_n(\widehat{\beta}) = -X'Z\Omega^{-1}Z'\left(Y - X\widehat{\beta}\right)$$

which has solution

$$\widehat{\beta} = \left(X'Z\Omega^{-1}Z'X\right)^{-1} \left(X'Z\Omega^{-1}Z'Y\right)$$

(b) The Lagrangian for the problem is

$$J_n(\beta, \lambda) = \frac{1}{2} \left(Y - X\beta\right)' Z\Omega^{-1}Z' \left(Y - X\beta\right) + \lambda' R'\beta$$

where λ is a $q \times 1$ Lagrange multiplier. The FOC for minimization is

$$0 = \frac{\partial}{\partial \beta} J_n(\widetilde{\beta}, \widetilde{\lambda}) = -X'Z\Omega^{-1}Z' \left(Y - X\widetilde{\beta}\right) + R\widetilde{\lambda}. \tag{3}$$

Premultiply by $R' \left(X'Z\Omega^{-1}Z'X\right)^{-1}$ and find

$$\begin{aligned}
0 &= -R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} X'Z\Omega^{-1}Z'Y + R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} X'Z\Omega^{-1}Z'X\widetilde{\beta} + R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} \\
&= -R'\widehat{\beta} + R'\widetilde{\beta} + R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\widetilde{\lambda} \\
&= -R'\widehat{\beta} + R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\widetilde{\lambda}
\end{aligned}$$

the final equality since the constrained estimator satisfies $R'\widetilde{\beta} = 0$. Solving for $\widetilde{\lambda}$ we find

$$\widetilde{\lambda} = \left(R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\right)^{-1} R'\widehat{\beta}$$

Substitute this into (3) and find

$$0 = -X'Z\Omega^{-1}Z'Y + X'Z\Omega^{-1}Z'X\widetilde{\beta} + R\left(R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\right)^{-1} R'\widehat{\beta}.$$

Sovling for $\widetilde{\beta}$

$$\begin{aligned}
\widetilde{\beta} &= \left(X'Z\Omega^{-1}Z'X\right)^{-1} \left(X'Z\Omega^{-1}Z'Y\right) - \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\left(R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\right)^{-1} R'\widehat{\beta} \\
&= \widehat{\beta} - \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\left(R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\right)^{-1} R'\widehat{\beta}
\end{aligned}$$

(c) We know

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) \rightarrow_d Z \sim N(0, \left(Q'\Omega^{-1}Q\right)^{-1})$$

where $Q = E(z_i x_i')$. Then using the assumption $R'\beta = 0$

$$\begin{aligned}
\sqrt{n}\left(\widetilde{\beta} - \beta\right) &= \sqrt{n}\left(\widehat{\beta} - \beta\right) - \sqrt{n}\left(X'Z\Omega^{-1}Z'X\right)^{-1} R\left(R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\right)^{-1} R'\widehat{\beta} \\
&= \sqrt{n}\left(\widehat{\beta} - \beta\right) - \sqrt{n}\left(X'Z\Omega^{-1}Z'X\right)^{-1} R\left(R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\right)^{-1} \left(R'\widehat{\beta} - R'\beta\right) \\
&= \sqrt{n}\left(\widehat{\beta} - \beta\right) - \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\left(R' \left(X'Z\Omega^{-1}Z'X\right)^{-1} R\right)^{-1} R'\sqrt{n}\left(\widehat{\beta} - \beta\right) \\
&\rightarrow {}_d Z - \left(Q'\Omega^{-1}Q\right)^{-1} R\left(R' \left(Q'\Omega^{-1}Q\right)^{-1} R\right)^{-1} R'Z \\
&= \left[I_k - \left(Q'\Omega^{-1}Q\right)^{-1} R\left(R' \left(Q'\Omega^{-1}Q\right)^{-1} R\right)^{-1} R'\right] Z
\end{aligned}$$

which is normal with mean zero and variance matrix

$$V_R = \left(Q'\Omega^{-1}Q\right)^{-1} - \left(Q'\Omega^{-1}Q\right)^{-1} R \left(R' \left(Q'\Omega^{-1}Q\right)^{-1} R\right)^{-1} R' \left(Q'\Omega^{-1}Q\right)^{-1}$$

Thus

$$\sqrt{n}\left(\widetilde{\beta} - \beta\right) \rightarrow_d N(0, V_R)$$