Economics 670: Econometrics

Department of Economics, Finance and Legal Studies

University of Alabama

Fall 2020

Final Exam — Key

The exam consists of four questions on four pages. Each question is of equal value.

1. Consider the regression model $y = X\beta + e$, where there exists a vector of instruments Z such that $E(e|Z) = 0$. With this information, answer the following:

   (a) Suppose $E(e|X) = 0$, derive an unbiased estimator for β.

   (b) For the estimator derived in part (a), show that the estimator is unbiased.

   (c) Suppose $E(e|X) \neq 0$, derive an unbiased estimator for β.

   (d) For the estimator derived in part (c), show that the estimator is unbiased.

   (e) For the estimator derived in part (c), show that the estimator is consistent.

(a) w/ $E(e|x) = 0$ as is sufficient

$$\min_{\beta} (y - x\beta)'(y - x\beta)$$

$$\Rightarrow -x'(y - x\beta) = 0$$

$$\Rightarrow x'y = x'x\beta$$

$$\hat{\beta} = (x'x)^{-1}x'y$$

(b) $E(\hat{\beta}|x) = E\left[(x'x)^{-1}x'y \,|\, x\right]$

$$= E\left[(x'x)^{-1}x'(x\beta + e) \,|\, x\right]$$

$$= \beta + E\left[(x'x)^{-1}x'e \,|\, x\right]$$

$$= \beta + (x'x)^{-1}x'E(e|x)$$

$$= \beta$$

(c) w/ $E(e|x) \neq 0$ & $E(e|z) = 0$ we can resort to 2SLS

$$y = x\beta + e$$

$$x = z\Gamma + u$$

$$\Rightarrow y = (z\Gamma + u)\beta + e$$

$$= z\Gamma\beta + v$$

$$\Rightarrow \hat{\beta} = (\Gamma' z' z \Gamma)^{-1}(\Gamma' z' y)$$

$$\hat{\Gamma} = (z'z)^{-1} z' x$$

$$\Rightarrow \hat{\beta}_{2SLS} = (\hat{\Gamma}' z' z \hat{\Gamma})^{-1} \hat{\Gamma} z' y$$

$$= [x' z (z'z)^{-1} z' x]^{-1} x' z (z'z)^{-1} z' y$$

(d) similar to part (b)

$$E(\hat{\beta}_{2SLS} | z) = E\{[x'z(z'z)^{-1}z'x]^{-1} x'z(z'z)^{-1}z'y \,|\, z\}$$

$$= \beta + [x'z(z'z)^{-1}z'x]^{-1} x'z(z'z)^{-1} z' E(e|z)$$

$$= \beta$$

(c) under the (7) assumptions from class

$$\hat{\beta}_{2SLS} - \beta = \left[(X'Z)(Z'Z)^{-1}(Z'X)\right]^{-1}(X'Z)(Z'Z)^{-1}(Z'e)$$

$$= \left[\left(\tfrac{1}{n}X'Z\right)\left(\tfrac{1}{n}(Z'Z)\right)^{-1}\left(\tfrac{1}{n}Z'X\right)\right]^{-1}\left(\tfrac{1}{n}X'Z\right)\left(\tfrac{1}{n}Z'Z\right)^{-1}\left(\tfrac{1}{n}Z'e\right)$$

$$\xrightarrow{p} \left(Q_{XZ} Q_{ZZ}^{-1} Q_{ZX}\right)^{-1} Q_{XZ} Q_{ZZ}^{-1} \; E(z_i e_i)$$

$$= 0$$

via WLLN for sample moments, the CMT for Q_{ZZ} & $Q_{XZ} Q_{ZZ}^{-1} Q_{ZX}$ & the assumption $E(z_i e_i) = 0$

2. Consider the regression model $y = X\beta + e$, where $e|X$ i.i.d. $N(0, \sigma^2 I_n)$. Noting that the pdf of a normal distribution is defined as

$$\phi(e) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{1}{2}\left(\frac{e-\mu_e}{\sigma}\right)^2\right]$$

where μ_e is the mean of e and σ^2 is its variance, answer the following:

(a) Derive the log-likelihood function.

(b) Derive the estimator of β from the log-likelihood function in part (a).

(c) Derive the estimator of σ^2 from the log-likelihood function in part (a).

(d) Using the results from parts (b) and (c), what is the maximized log-likelihood?

(e) Derive the asymptotic distribution of the estimator of β from part (b).

(a) $$f(y|x) = \prod_{i=1}^{n} f(y_i|x_i)$$

$$= \prod_{i=1}^{n} \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} exp\left[-\frac{1}{2\sigma^2}(y_i - x_i'\beta)^2\right]$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_i'\beta)^2\right]$$

$$\equiv \mathcal{L}(\beta, \sigma^2)$$

$$\ln \mathcal{L}(\beta,\sigma^2) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_i'\beta)^2$$

(b) $$\frac{\partial \ln \mathcal{L}(\beta,\sigma^2)}{\partial \beta} = \frac{1}{\hat{\sigma}^2_{MLE}} \sum_{i=1}^{n} x_i (y_i - x_i' \hat{\beta}_{MLE}) = 0$$

$$\Rightarrow \hat{\beta}_{MLE} = \left(\sum_{i=1}^{n} x_i x_i'\right)^{-1} \left(\sum_{i=1}^{n} x_i y_i\right)$$

(c) $$\frac{\partial \ln \mathcal{L}(\beta,\sigma^2)}{\partial \sigma^2} = \frac{-n}{2\hat{\sigma}^2_{MLE}} + \frac{1}{2\hat{\sigma}^4_{MLE}} \sum_{i=1}^{n} (y_i - x_i' \hat{\beta}_{MLE})^2 = 0$$

$$\Rightarrow \hat{\sigma}^2_{MLE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i' \hat{\beta}_{MLE})^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} \hat{e}_i^2$$

(d) $$\ln \mathcal{L}(\hat{\beta}_{MLE}, \hat{\sigma}^2_{MLE}) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \hat{\sigma}^2_{MLE}$$

$$- \frac{1}{2\hat{\sigma}^2_{MLE}} \sum_{i=1}^{n} (y_i - x_i' \hat{\beta}_{MLE})^2$$

$$= -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \hat{\sigma}^2_{MLE} - \frac{1}{2\hat{\sigma}^2_{MLE}} \sum_{i=1}^{n} \hat{e}_i^2$$

$$= -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \hat{\sigma}^2_{MLE} - \frac{n\hat{\sigma}^2_{MLE}}{2\hat{\sigma}^2_{MLE}}$$

$$= -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \hat{\sigma}^2_{MLE} - \frac{n}{2}$$

(c) recall

$$e|x \sim N(0, \sigma^2 I_n)$$

recall

$$\hat{\beta} - \beta = (x'x)^{-1} x'e$$

$$\Rightarrow \hat{\beta} - \beta|_x \sim (x'x)^{-1} x' N(0, \sigma^2 I_n)$$

$$\sim N(0, \sigma^2 (x'x)^{-1} x' x (x'x)^{-1})$$

$$= N(0, \sigma^2 (x'x)^{-1})$$

3. Consider the regression model $y = x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + e = X\beta + e$, where β is a $k(=3) \times 1$ vector, $E(e|x_1, x_2, x_3) = 0$ and $E(e^2|x_1, x_2, x_3) = \sigma^2$. We are interested in testing linear restrictions on the coefficient vector, i.e., $\theta = r(\beta) = R'\beta$. With this information, answer the following:

(a) Consider the null $H_0 : \beta_1 = 0$ versus the alternative $H_1 : \beta_1 \neq 0$, define R.

(b) For the null hypothesis in part (a), define the test statistic and state its asymptotic distribution.

(c) Consider the null $H_0 : \beta_1 = \beta_2$ versus the alternative $H_1 : \beta_1 \neq \beta_2$, define R.

(d) For the null hypothesis in part (c), define the test statistic and state its asymptotic distribution.

(e) Consider the null $H_0 : \beta_1 = \beta_2 = 0$ versus the alternative H_0 is not true, define R and the corresponding test statistic.

(a) $H_0: \beta_1 = 0 \Rightarrow \beta_1 = \theta = r(\beta) = R'\beta = [1 \ 0 \ 0]\begin{bmatrix}\beta_1\\ \beta_2\\ \beta_3\end{bmatrix}$

(b) $t = \dfrac{\hat{\beta}_1 - \beta_{10}}{se(\hat{\beta}_1)} = \dfrac{\hat{\beta}_1 - 0}{\sqrt{s^2(X'X)^{-1}_{[11]}}} \sim t_{n-3}$

(c) $H_0: \beta_1 = \beta_2 \Leftrightarrow H_0: \beta_1 - \beta_2 = 0$

$\Rightarrow \beta_1 - \beta_2 = \theta = r(\beta) = R'\beta = [1 \ -1 \ 0]\begin{bmatrix}\beta_1\\ \beta_2\\ \beta_3\end{bmatrix}$

(d) $t = \dfrac{\hat{\beta}_1 - \hat{\beta}_2 - 0}{se(\hat{\beta}_1 - \hat{\beta}_2)} \sim t_{n-3}$

(e) $H_0: \beta_1 = 0 \wedge \beta_2 = 0 \Leftrightarrow H_0: \begin{matrix}\beta_1 = 0\\ \beta_2 = 0\end{matrix}$

$\Rightarrow \theta = r(\beta) = R'\beta = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}\beta_1\\ \beta_2\\ \beta_3\end{bmatrix}$

$W = (\hat{\theta} - \theta_0)'\hat{V}_{\hat{\theta}}^{-1}(\hat{\theta} - \theta_0) = r(\hat{\beta})'(\hat{R}'\hat{V}_{\beta}\hat{R})^{-1} r(\hat{\beta}) \sim \chi^2_2$

4. Suppose we are interested in the role of education (educ) and experience (exper) on wages (wage). However, the level of education (educ) depends upon unobserved ability (abil). We hope to resolve this endogeneity via the instrument, mother's education (mothereduc). Considering the lines of R code below, answer the following:

```
exper2 <- exper^2
reg1 <- lm(educ~exper+exper2+mothereduc)
h.educ <- fitted(reg1)
reg2 <- lm(log(wage)~h.educ+exper+exper2)
```

(a) Define each variable in this scenario: outcome (Y), endogenous (X), and exdogenous (both included and excluded, $Z = (Z_1, Z_2)$).

(b) Write down the equations being estimated in the code in terms of Y, X and Z with the corresponding parameters and errors.

(c) State the assumptions of the model in part (b).

(d) Use a directed acyclic graph (DAG) to draw the model in part (b).

(e) What can be said about the standard errors obtained directly from the fourth line of code?

(a) $Y = \text{wage}$ $X = \text{educ}$ $Z_1 = (\text{exper}, \text{exper}^2)$

$Z_2 = \text{mothereduc}$

(b) $X = Z\Gamma + U$ $Y = X\beta + e$

(c) $E(e|X) \neq 0$ $E(e|Z) = 0$

$E(ZZ')$ is pd $E(ZX')$ has full rank

(d)
$$
\begin{tikzcd}
Z_1 \arrow[rr] \arrow[dr] & & Y \\
Z_2 \arrow[r] & X \arrow[ur] & \\
& U \arrow[u] \arrow[r, leftrightarrow] & e \arrow[uu]
\end{tikzcd}
$$

(e) they do not recognize that $\hat{X}$ is predicted & therefore are not correct