Economics 403
Michaelmas Test

Oliver Linton

December 5th, 2002

Answer question one and one other. Explain all your answers.

1. (a) FALSE: the square of the OLS estimate b is not unbiased since,

$$\begin{aligned} E(b^2) &= E[(b - E(b) + E(b))^2] \\ &= E[(b - E(b))^2] + E(b)^2 = \\ &= V(b) + \beta^2 > \beta^2 \end{aligned}$$

   (b) FALSE: In case of autocorrelation of error terms, OLS estimator of conditonal variance can be biased in either direction and when the bias is upward, you tend to accept too often.

   (c) TRUE: When n is large,

$$\sqrt{n}(\overline{X} - 1) \xrightarrow{d} N(0, 1),$$

   by CLT. Since the function $f(x) = 1/x$ is continuously differentiable function at $x = 1$ and $E(\overline{X}) = 1$, applying the Delta method,

$$\sqrt{n}\left(\frac{1}{\overline{X}} - 1\right) \xrightarrow{d} N(0, 1),$$

   i.e. $\frac{1}{\overline{X}} \xrightarrow{a} N(0, \frac{1}{n})$.

   (d) FALSE: Asymptotically $LM = W$, i.e. they have the same power. It is however true that in finite samples, for linear regression models, $LM \leq W$, since

$$LM = \frac{W}{1 - \frac{W}{n}}$$

   and we use the (same) asymptotic critical values, which means that in finite samples W rejects more than LM does.

(e) TRUE: But only when the smaller variance is at least 3 times smaller than the larger one. Taking the average of the two estimates, assuming that the two populations are indipendent, we have:

$$\begin{aligned}
\widehat{\mu}^* &= \frac{1}{2}(\widehat{\mu}_{even} + \widehat{\mu}_{odd}) \\
E(\widehat{\mu}^*) &= \mu = E(\mu_{even}) = E(\mu_{odd}) \\
V(\widehat{\mu}^*) &= \frac{1}{4}V(\widehat{\mu}_{even}) + \frac{1}{4}V(\widehat{\mu}_{odd})
\end{aligned}$$

and $V(\widehat{\mu}_{even}) < V(\widehat{\mu}^*)$ only if $V(\widehat{\mu}_{odd}) > 3 * V(\widehat{\mu}_{even})$.

2. (a)

$$\begin{aligned}
\mathrm{p}\lim_{\mathrm{n}\to\infty}(\widehat{\beta}) &= \mathrm{p}\lim_{\mathrm{n}\to\infty}\left(\frac{\sum x_i y_i}{\sum x_i^2}\right) = \mathrm{p}\lim_{\mathrm{n}\to\infty}\left(\frac{\sum(x_i^* + u_i)(\beta x_i^* + \beta u_i - \beta u_i + \varepsilon_i)}{\sum(x_i^* + u_i)^2}\right) \\
&= \beta + \left(\frac{\mathrm{p}\lim_{\mathrm{n}\to\infty}\frac{1}{n}\sum(x_i^* + u_i)(\varepsilon_i - \beta u_i)}{\mathrm{p}\lim_{\mathrm{n}\to\infty}\frac{1}{n}\sum(x_i^* + u_i)^2}\right) \\
&= \beta - \beta\frac{var(u_i)}{\mathrm{E}(x_i^* + u_i)^2} \neq \beta
\end{aligned}$$

b

$$\begin{aligned}
\widetilde{\beta} &= \frac{\sum(y_i^* + \varepsilon_i)}{\sum(x_i^* + u_i)} = \frac{\sum(\beta x_i^* + \varepsilon_i)}{\sum(x_i^* + u_i)} = \\
&= \beta\frac{\sum x_i^*/n}{\left(\sum x_i^* + u_i\right)/n} + \frac{\sum \varepsilon_i/n}{\left(\sum x_i^* + u_i\right)/n}
\end{aligned}$$

by Slutzky. This quantity is consistent if $E(|x_i^*|) < \infty$ and $E(x_i^*) \neq 0$, by LLN:

$$\frac{\sum x_i^*}{n} \xrightarrow{p} E(x_i^*) \neq 0$$

since by hypothesis, $E(u_i) = E(\varepsilon_i) = 0$, by LLN:

$$\begin{aligned}
\frac{\sum u_i}{n} &\xrightarrow{p} 0 \\
\frac{\sum \varepsilon_i}{n} &\xrightarrow{p} 0
\end{aligned}$$

Suppose that $E(x_i) \neq 0$. Then

$$\widetilde{\beta} - \beta = \frac{\frac{1}{n}\sum_{i=1}^n \varepsilon_i - \beta\frac{1}{n}\sum_{i=1}^n u_i}{\frac{1}{n}\sum_{i=1}^n x_i} \to^p 0$$

by the Slutsky theorem and what shown in part (b), we have

$$\sqrt{n}(\widetilde{\beta}-\beta)=\frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\varepsilon_i-\beta u_i)}{E(x_i)}+o_p(1),$$

which, by Cramer rule, is asymptotically normal with mean zero and variance

$$\frac{\mathrm{var}(\varepsilon_i)+\beta^2\mathrm{var}(u_i)}{E^2(x_i)}.$$

(b) i Draw an xy axis with a line through the intercept and then a point (x, y) that is to the northwest of the line. Now drop a perpendicular from that point to the x axis, this line will cut through the $y=\beta x$ line at the OLS fit. Draw in the orthogonal projection from (y, x) to the line $y = bx$ - this should cut the line at a right angle. You now should be staring at two triangles, which happen to be 'similar', which means they are scale modules of one another - since they are both right triangles and the angle between the hypothenuse and the short side are the same. You want the orthogonal distance from (y, x) to the line $y = \beta x$ and what you know is the OLS distance, the hypotenuse of the upper triangle, which is $(y-\beta x)$. Regarding the lower triangle, the base is x, the height is βx and the hypotenuse is $\sqrt{x^2+(\beta x)^2}$, and it is similar to a right triangle with the same angle at origin, base $x = 1$ and height β whose hypothenuse is $\sqrt{1+\beta^2}$. You then ratio up this latter triangle and the upper one to get the orthogonal distance you wanted in the first place, Q. In fact lenght of the longer side in the upper triangle is $\sqrt{Q}$ and by proportionality $Q=\frac{1}{1+\beta^2}(y-\beta x)^2$.

ii Write

$$(y_i-\beta x_i)^2=(\beta_0-\beta)^2x_i^2+(\varepsilon_i-\beta_0u_i)^2+2(\varepsilon_i-\beta_0u_i)(\beta_0-\beta)x_i$$

and hence

$$\begin{aligned}
Q_n(\beta) &= \frac{(\beta_0-\beta)^2}{1+\beta^2}\frac{1}{n}\sum_{i=1}^{n}x_i^2\\
&\quad+\frac{\beta_0-\beta}{1+\beta^2}\frac{2}{n}\sum_{i=1}^{n}x_i(\varepsilon_i-\beta_0u_i)\\
&\quad+\frac{1}{1+\beta^2}\frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i-\beta_0u_i)^2\\
&\to {}_p\frac{(\beta_0-\beta)^2}{1+\beta^2}E[x_i^2]\\
&\quad-2\frac{(\beta_0-\beta)}{1+\beta^2}\beta_0E[u_i^2]\\
&\quad+\frac{1}{1+\beta^2}\left(E[\varepsilon_i^2]+\beta_0^2E[u_i^2]\right)\\
&\equiv Q(\beta).
\end{aligned}$$

Suppose that $E[\varepsilon_i^2] = E[u_i^2] = \sigma^2$ and $E[x_i^2] = m_{xx}$ then

$$\begin{aligned} Q(\beta) &= \frac{(\beta_0 - \beta)^2 m_{xx} - 2(\beta_0 - \beta)\beta_0\sigma^2 + (1+\beta_0^2)\sigma^2}{1+\beta^2} \\ &= \sigma^2 \frac{(\beta_0 - \beta)^2[1 + (m_{xx}/\sigma^2) - 1] - 2(\beta_0 - \beta)\beta_0 + (1+\beta_0^2)}{1+\beta^2} \\ &= \sigma^2 \left[1 + \frac{(\beta_0 - \beta)^2[(m_{xx}/\sigma^2) - 1]}{1+\beta^2}\right]. \end{aligned}$$

Therefore,

$$Q(\beta_0) = \sigma^2$$

and

$$Q(\beta) > \sigma^2$$

for any $\beta \neq \beta_0$ provided $m_{xx} \neq \sigma^2$. Thus we have consistency so long as $E[\varepsilon_i^2] = E[u_i^2] = \sigma^2$ and $m_{xx} \neq \sigma^2$. Be very generous with this question.

3. (a)

$$\begin{aligned} Q &= y' M_x y \\ &= y_1' M_1 y_1 + y_2' M_2 y_2 \\ &= Q_1 + Q_2 \end{aligned}$$

since

$$\begin{aligned} y' P_X y &= \\ &= \begin{bmatrix} y_1' & y_2' \end{bmatrix} \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_2)^{-1} \end{bmatrix} \begin{bmatrix} X_1' & 0 \\ 0 & X_2' \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= y_1' X_1 (X_1'X_1)^{-1} X_1' y_1 + y_2' X_2 (X_2'X_2)^{-1} X_2 y_2 \\ &= y_1' P_1 y_1 + y_2' P_2 y_2 \end{aligned}$$

The restricted model, under $H_0 : \beta_1 = \beta_2$ is

$$y = Z\beta + u$$

from which we get the residuals:

$$Q_R = y' M_Z y$$

The F statistic, with a number of restriction equal to the number of columns in X_2, is:

$$F = \frac{(Q_R - Q)/k}{Q/n - 2k} \sim F(k, n_1 + n_2 - 2k)$$

The F statistics is than compared with the critical value $F^{\alpha}_{(k,n-2k)}$ and re reject H_0 if $F > F^{\alpha}_{(k,n-2k)}$.

(b)

$$\begin{aligned} y &= Z\beta + W\gamma + u \\ &= X_1\beta + X_2\beta + X_2\gamma + u \\ &= X_1\beta + X_2(\beta + \gamma) + u \end{aligned}$$

denoting $\beta = \beta_1$ and $\beta+\gamma = \beta_2$, the test $\beta_1 = \beta_2$ is equivalent to $\beta = \beta+\gamma$, i.e. $\gamma = 0$.

(c) Given the rank of the matrix X_2, only n_2 elements of the vector β will be identifiable and the residuals corresponding to all observations that belong to group 2 will be zero. The number of degrees of freedom for the numerator of the F statistic must therefore be at most n_2. The number of degrees of freedom for the denominator willl be the number of observations for which $y'M_Xy \neq 0$, which will normally be n_1, minus the number of regressors that affect those observations which will be k, for a total of $n_1 - k$.

(d) Imposing $\gamma_i = 0$, with $i = 1, 2, ..., k - rank(X_2)$, defining γ' with $rank(X_2)$ non zero element, and than testing $H_0 : \gamma^{'} = 0$.

(e) If the two variances are known and $\frac{\sigma_1^2}{\sigma_2^2} = k < \infty$, rescale X_2 multiplying by $\sqrt{k}$,so that $var(u_2/\sqrt{k}) = k\sigma_2^2 = \sigma_1^2$ and precede as in homoskedastic case.