Econometrics 710
Midterm Exam Sample Answers
March 17, 2005

1. OLS is estimating the linear projection of y_i on x_i. This is an approximation to the (nonlinear) conditional mean of y_i given x_i.

   The slope coefficient is
$$\beta_1 = \frac{E\left(\left(x_i - Ex_i\right) y_i\right)}{E\left(\left(x_i - Ex_i\right)^2\right)}$$
   Since $Ex_i = 0$ this simplifies to
$$\frac{E\left(x_i\left(x_i^2 + \varepsilon_i\right)\right)}{E\left(x_i^2\right)} = \frac{E\left(x_i^3\right) + E\left(x_i \varepsilon_i\right)}{E\left(x_i^2\right)} = \frac{\mu_3}{\mu_2}$$
   The intercept is
$$\begin{aligned} \beta_0 &= Ey_i - Ex_i\beta_1 \\ &= E\left(x_i^2 + \varepsilon_i\right) \\ &= \mu_2. \end{aligned}$$
   Thus $\beta_0 = \mu_2$ and $\beta_1 = \mu_3/\mu_2$.

2. This is a projection model. From the theory of asymptotic efficiency, we know that the pooled OLS estimator $\hat{\beta}$ is asymptotically efficient, in the sense that no other estimator can have lower asymptotic variance. However, this does not address the question of whether or not the two estimators can have equal asymptotic efficiency. It turns out that generally the pooled estimator is more efficient, with equal efficiency if and only if $n_1 = n_2$.

   To assess the relative efficiency, we need to calculate the asymptotic variances of the two estimators.

   First, standard asymptotic theory shows that
$$\sqrt{n_1 + n_2}\left(\hat{\beta} - \beta\right) \rightarrow_d N(0, V)$$
   where $V = Q^{-1}\Omega Q^{-1}$ with $Q = Ex_i x_i'$ and $\Omega = Ex_i x_i' e_i^2$. Thus
$$Var(\hat{\beta}) \simeq \left(\frac{1}{n_1 + n_2}\right) V. \tag{1}$$
   Second
$$\begin{aligned} Var\left(\tilde{\beta}\right) &= Var\left(\frac{1}{2}\left(\hat{\beta}_1 + \hat{\beta}_2\right)\right) \\ &= \frac{1}{4}\left(Var(\hat{\beta}_1) + Var(\hat{\beta}_2) + Cov(\hat{\beta}_1, \hat{\beta}_2) + Cov(\hat{\beta}_2, \hat{\beta}_1)\right) \\ &= \frac{1}{4}\left(Var(\hat{\beta}_1) + Var(\hat{\beta}_2)\right) \end{aligned}$$
   the final equality since $\hat{\beta}_1$ and $\hat{\beta}_2$ are independent (as the samples are independent). From standard

theory

$$\sqrt{n_1}\left(\hat{\beta}_1-\beta\right) \rightarrow_d N(0,V)$$

$$\sqrt{n_2}\left(\hat{\beta}_2-\beta\right) \rightarrow_d N(0,V)$$

The covariance matrices V are the same as before since all the observations are drawn from the same distribution. The sample sizes n_1 and n_2, however, are different, so the variances of the estimators $\hat{\beta}$, $\hat{\beta}_1$ and $\hat{\beta}_2$ are different. From the asymptotic distribution we calculate

$$\begin{aligned} Var(\hat{\beta}_1) &\simeq n_1^{-1}V \\ Var(\hat{\beta}_2) &\simeq n_2^{-1}V \end{aligned}$$

Therefore

$$\begin{aligned} Var\left(\tilde{\beta}\right) &= \frac{1}{4}\left(n_1^{-1}V+n_2^{-1}V\right) \\ &= \left(\frac{1}{4n_1}+\frac{1}{4n_2}\right)V \qquad (2) \end{aligned}$$

We have calculated the asymptotic variances for $\hat{\beta}$ and $\tilde{\beta}$ in (1) and (2). They are both proportional to the matrix V, so efficiency is decided by the scale term.

Which is more efficient? Interestingly, they are equally efficient if $n_1 = n_2$, for in this case it is easy to see that (1)=(2). If $n_1 \neq n_2$ this is not true, however. For example, if $n_2 = 3n_1$, then

$$Var(\hat{\beta}) = \frac{1}{4n_1}V < \frac{1}{3n_1}V = Var\left(\tilde{\beta}\right)$$

To see that the pooled estimator $\hat{\beta}$ is more efficient, observe that

$$\frac{1}{n_1+n_2} \leq \left(\frac{1}{4n_1}+\frac{1}{4n_2}\right) \qquad (3)$$

with equality only if $n_1 = n_2$. To show (3), by rearranging, it is equivalent to

$$\sqrt{n_1 n_2} \leq \frac{n_1+n_2}{2}$$

which is true by the property of geometric means.

A more formal method of doing the asymptotic argument is to write $n = n_1 + n_2$, $n_1 = \alpha_1 n$ and $n_2 = \alpha_2 n$, and then assume that α_1 and α_2 are constant as $n \rightarrow \infty$. The details are essentially unchanged.

3. Since the true equation is a homoskedastic linear regression, we know that the asymptotic variance of the long regression $\hat{\beta}_1$ is

$$\begin{aligned} Var(\hat{\beta}_1) &= \frac{1}{n}\left(M_{11}-M_{12}M_{22}^{-1}M_{21}\right)^{-1}Var(\varepsilon_i) \\ &= \frac{1}{n}M_{11}^{-1}\sigma^2 \qquad (4) \end{aligned}$$

the second equality since $M_{21} = 0$.

Since $M_{21} = 0$, the short regression is unbiased. Thus $\tilde{\beta}_1$ is estimating the equation

$$y_i = x_{1i}'\beta_1 + u_i$$

where

$$u_i = x_{2i}'\beta_2 + \varepsilon_i$$

The asymptotic variance of this estimator is

$$Var(\tilde{\beta}_1) = \frac{1}{n} M_{11}^{-1} E\left(x_{1i}x_{1i}'u_i^2\right) M_{11}^{-1}$$

Note that

$$u_i^2 = \left(\beta_2' x_{2i}\right)^2 + 2\beta_2' x_{2i}\varepsilon_i + \varepsilon_i^2$$

so

$$\begin{aligned} E\left(x_{1i}x_{1i}'u_i^2\right) &= E\left(x_{1i}x_{1i}'\left(\beta_2' x_{2i}\right)^2\right) + 2E\left(x_{1i}x_{1i}'\beta_2' x_{2i}\varepsilon_i\right) + E\left(x_{1i}x_{1i}'\varepsilon_i^2\right) \\ &= Q + M_{11}\sigma^2 \end{aligned}$$

where

$$Q = E\left(x_{1i}x_{1i}'\left(\beta_2' x_{2i}\right)^2\right)$$

Therefore

$$Var(\tilde{\beta}_1) = \frac{1}{n} M_{11}^{-1} Q M_{11}^{-1} + \frac{1}{n} M_{11}^{-1}\sigma^2. \tag{5}$$

The asymptotic variances of $\hat{\beta}_1$ and $\tilde{\beta}_1$ are given in (4) and (5).

Comparing the two expressions, observe that if $\beta_2 \neq 0$ then $Q > 0$. Therefore

$$Var(\hat{\beta}_1) < Var(\tilde{\beta}_1)$$

with equality only if $\beta_2 = 0$. Thus in general, $\hat{\beta}_1$ is more efficient than $\tilde{\beta}_1$. They are equally efficient only when $\beta_2 = 0$

4.

(a) We know that the projection coefficient is

$$\beta_1^* = \beta_1 + M_{11}^{-1} M_{12}\beta_2$$

The bias is

$$Bias(\tilde{\beta}_1) = M_{11}^{-1} M_{12}\beta_2.$$

(b) OLS is estimating the equation

$$y_i = x_{1i}'\beta_1^* + u_i^*$$

where

$$\begin{aligned} u_i^* &= y_i - x_{1i}'\beta_1^* \\ &= \left(x_{2i} - M_{21}M_{11}^{-1}x_{1i}\right)'\beta_2 + \varepsilon_i \\ &= x_{2i}^{*\prime}\beta_2 + \varepsilon_i \end{aligned}$$

where

$$x_{2i}^* = x_{2i} - M_{21}M_{11}^{-1}x_{1i},$$

the residual from the regression of x_{2i} on x_{1i}.
The asymptotic variance of $\tilde{\beta}_1$ is

$$Var(\tilde{\beta}_1) = \frac{1}{n}M_{11}^{-1}E\left(x_{1i}x_{1i}'u_i^2\right)M_{11}^{-1}$$

We calculate that

$$u_i^{*2} = \left(\beta_2'x_{2i}^*\right)^2 + 2\beta_2'x_{2i}^*\varepsilon_i + \varepsilon_i^2$$

so

$$E\left(x_{1i}x_{1i}'u_i^{*2}\right) = Q^* + M_{11}\sigma^2.$$

where

$$Q^* = E\left(x_{1i}x_{1i}'\left(\beta_2'x_{2i}^*\right)^2\right).$$

In this case,

$$Var(\tilde{\beta}_1) = \frac{1}{n}M_{11}^{-1}Q^*M_{11}^{-1} + \frac{1}{n}M_{11}^{-1}\sigma^2.$$

(c) The asymptotic mean-squared error of $\tilde{\beta}_1$ is

$$\begin{aligned} MSE(\tilde{\beta}_1) &= Bias(\tilde{\beta}_1)Bias(\tilde{\beta}_1)' + Var(\tilde{\beta}_1) \\ &= M_{11}^{-1}M_{12}\beta_2\beta_2'M_{21}M_{11}^{-1} + \frac{1}{n}M_{11}^{-1}Q^*M_{11}^{-1} + \frac{1}{n}M_{11}^{-1}\sigma^2 \end{aligned}$$

(d) The asymptotic mean-squared error of $\hat{\beta}_1$ is

$$\begin{aligned} MSE(\hat{\beta}_1) &= Var(\hat{\beta}_1) \\ &= \frac{1}{n}\left(M_{11} - M_{12}M_{22}^{-1}M_{21}\right)^{-1}\sigma^2 \end{aligned}$$

In general, $MSE(\hat{\beta}_1)$ and $MSE(\tilde{\beta}_1)$ are difficult to compare. However, observe that when $\beta_2 = 0$ and $M_{21} \neq 0$, then $MSE(\tilde{\beta}_1) = \frac{1}{n}M_{11}^{-1}\sigma^2 < MSE(\hat{\beta}_1)$ so $\tilde{\beta}_1$ is more efficient. On the other hand, we showed in problem 3 that when $\beta_2 \neq 0$ and $M_{21} = 0$ that $\hat{\beta}_1$ is more efficient. It follows that neither estimator strictly dominates the other, and either can be more efficient than the other, depending on the specific values.