Econometrics 710
Midterm Exam
March 11, 2014
Sample Answers

1. The answer could be expressed either with the conditional variance $\text{var}(\widehat{\beta} \mid X)$ or unconditional $\text{var}(\widehat{\beta})$. The former is easier. To evaluate the latter you can use the fact that the regression model implies $E(\widehat{\beta} \mid X) = \beta$ and thus $\text{var}(\widehat{\beta}) = E\,\text{var}(\widehat{\beta} \mid X)$.

   (a) We know that if we set $D = diag\{\sigma_1^2, ..., \sigma_n^2\}$ where $\sigma_i^2 = E(e_i^2|x_i)$, then $\text{var}(\widehat{\beta} \mid X) = (X'X)^{-1}(X'DX)(X'X)^{-1}$. Since $n^{-1}X'X = I_k$ it follows that

   $$\begin{aligned}\text{var}(\widehat{\beta} \quad &| \quad X) = (nI_k)^{-1}(X'DX)(nI_k)^{-1} \\ &= n^{-2}X'DX \\ &= n^{-2}\sum_{i=1}^{n} x_i x_i' \sigma_i^2\end{aligned}$$

   It follows that

   $$\text{var}(\widehat{\beta}) = E\left(n^{-2}\sum_{i=1}^{n} x_i x_i' \sigma_i^2\right) = n^{-1}E\left(x_i x_i' \sigma_i^2\right).$$

   (b) The (conditional) covariances take the form

   $$\text{cov}\left(\widehat{\beta}_j, \widehat{\beta}_\ell \mid X\right) = n^{-2}\sum_{i=1}^{n} x_{ji}x_{\ell i}\sigma_i^2$$

   which can be anything. For example, if $\sigma_i^2 = x_{ji}x_{\ell i}$ then $\text{cov}\left(\widehat{\beta}_j, \widehat{\beta}_\ell \mid X\right) = n^{-2}\sum_{i=1}^{n} x_{ji}^2 x_{\ell i}^2 > 0$.

   (c) Under conditional homoskedasticity, $E(e_i^2|x_i) = 0$, then

   $$\text{var}(\widehat{\beta} \mid X) = n^{-2}\sum_{i=1}^{n} x_i x_i' \sigma^2 = \frac{\sigma^2}{n} I_k$$

   so the off-diagonals are all zero. Since the latter is independe of X then

   $$\text{var}(\widehat{\beta}) = \frac{\sigma^2}{n} I_k$$

   as well. In either case (conditional or unconditional), $E(e_i^2|x_i) = 0$ is a sufficient condition for $\widehat{\beta}_j$ and $\widehat{\beta}_\ell$ to be mutually uncorrelated.

2.

   (a) $\widehat{\gamma} = (Y'Y)^{-1}(Y'X) = \left(\frac{1}{n}\sum_{i=1}^{n} y_i^2\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} y_i x_i\right)^{-1}$

   (b) $\widehat{\theta} = 1/\widehat{\gamma}$

   (c) There is nothing different from the standard projection model, except the notation has been switched from y to x and vice-versa. Thus

   $$\sqrt{n}\,(\widehat{\gamma} - \gamma) \to_d N(0, V_\gamma)$$

   where

   $$V_\gamma = \frac{E\left(y_i^2 u_i^2\right)}{\left(Ey_i^2\right)^2}$$

   Let $g(\gamma) = 1/\gamma$ and note $\frac{d}{d\gamma}g(\gamma) = -1/\gamma^2$. Thus by the Delta Method, assuming $\gamma \neq 0$,

   $$\sqrt{n}\left(\widehat{\theta} - \theta\right) \to_d N(0, V_\theta)$$

where

$$V_\theta = \left(\frac{1}{\gamma^2}\right)^2 V_\gamma = \frac{E\left(y_i^2 u_i^2\right)}{\gamma^4 \left(E y_i^2\right)^2}$$

(d) A moment estimator of V_θ is

$$\widehat{V}_\theta = \frac{\frac{1}{n}\sum_{i=1}^n y_i^2 \widehat{u}_{ii}^2}{\widehat{\gamma}^4 \left(\frac{1}{n}\sum_{i=1}^n y_i^2\right)^2} = \frac{n\sum_{i=1}^n y_i^2 \widehat{u}_{ii}^2}{\widehat{\gamma}^4 \left(\sum_{i=1}^n y_i^2\right)^2}$$

where $\widehat{u}_i = x_i - y_i \widehat{\gamma}$. A standard error is thus

$$se\left(\widehat{\theta}\right) = \sqrt{\frac{n\sum_{i=1}^n y_i^2 \widehat{u}_{ii}^2}{\widehat{\gamma}^4 \left(\sum_{i=1}^n y_i^2\right)^2}}$$

3.

(a) Since the samples are independent, the estimators are independent and thus their joint asymptotic covariance matrix (and estimate) is block diagonal: $\left[\begin{array}{cc} \widehat{V}_{\beta_1} & 0 \\ 0 & \widehat{V}_{\beta_2} \end{array}\right]$. The minimum-distance criterion takes the form

$$\begin{aligned} J_n(\beta) &= n \left(\begin{array}{c} \widehat{\beta}_1 - \beta \\ \widehat{\beta}_2 - \beta \end{array}\right)' \left[\begin{array}{cc} \widehat{V}_{\beta_1} & 0 \\ 0 & \widehat{V}_{\beta_2} \end{array}\right]^{-1} \left(\begin{array}{c} \widehat{\beta}_1 - \beta \\ \widehat{\beta}_2 - \beta \end{array}\right) \\ &= n\left(\widehat{\beta}_1 - \beta\right)' \widehat{V}_{\beta_1}^{-1} \left(\widehat{\beta}_1 - \beta\right) + n\left(\widehat{\beta}_2 - \beta\right)' \widehat{V}_{\beta_2}^{-1} \left(\widehat{\beta}_2 - \beta\right) \end{aligned}$$

The FOC for minimization are

$$0 = -2n\widehat{V}_{\beta_1}^{-1}\left(\widehat{\beta}_1 - \widetilde{\beta}\right) - 2n\widehat{V}_{\beta_2}^{-1}\left(\widehat{\beta}_2 - \widetilde{\beta}\right)$$

with solution

$$\widetilde{\beta} = \left(\widehat{V}_{\beta_1}^{-1} + \widehat{V}_{\beta_2}^{-1}\right)\left(\widehat{V}_{\beta_1}^{-1}\widehat{\beta}_1 + \widehat{V}_{\beta_2}^{-1}\widehat{\beta}_2\right).$$

This is a weighted average of the estimators $\widehat{\beta}_1$ and $\widehat{\beta}_2$, with weights depending on the covariance matrices.

(b) We know that since $\beta_1 = \beta_2 = \beta$

$$\sqrt{n}\left(\widehat{\beta}_1 - \beta\right) \to_d Z_1 \sim N(0, V_{\beta_1})$$

$$\sqrt{n}\left(\widehat{\beta}_2 - \beta\right) \to_d Z_2 \sim N(0, V_{\beta_2})$$

where Z_1 and Z_2 are independent. The convergence is also joint convergence. Furthermore, $\widehat{V}_{\beta_1} \to_p V_{\beta_1}$ and $\widehat{V}_{\beta_2} \to_p V_{\beta_2}$. It follows that

$$\begin{aligned} \sqrt{n}\left(\widetilde{\beta} - \beta\right) &= \left(\widehat{V}_{\beta_1}^{-1} + \widehat{V}_{\beta_2}^{-1}\right)^{-1}\left(\widehat{V}_{\beta_1}^{-1}\sqrt{n}\left(\widehat{\beta}_1 - \beta\right) + \widehat{V}_{\beta_2}^{-1}\sqrt{n}\left(\widehat{\beta}_2 - \beta\right)\right) \\ &\to_d \left(V_{\beta_1}^{-1} + V_{\beta_2}^{-1}\right)^{-1}\left(V_{\beta_1}^{-1} Z_1 + V_{\beta}^{-1} Z_2\right) \\ &\sim \left(V_{\beta_1}^{-1} + V_{\beta_2}^{-1}\right)^{-1} N(0, V_{\beta_1}^{-1} + V_{\beta_2}^{-1}) \\ &= N(0, \left(V_{\beta_1}^{-1} + V_{\beta_2}^{-1}\right)^{-1}) \end{aligned}$$

(c) The (approximate) variance of $\widehat{\beta}_1$ is $n_1^{-1}\widehat{V}_{\beta_1}$ and that of $\widehat{\beta}_2$ is $n_2^{-1}\widehat{V}_{\beta_2}$. Thus a minimum-distance criterion can be written as

$$J_n(\beta) = \begin{pmatrix} \widehat{\beta}_1 - \beta \\ \widehat{\beta}_2 - \beta \end{pmatrix}' \begin{bmatrix} n_1^{-1}\widehat{V}_{\beta_1} & 0 \\ 0 & n_2^{-1}\widehat{V}_{\beta_2} \end{bmatrix}^{-1} \begin{pmatrix} \widehat{\beta}_1 - \beta \\ \widehat{\beta}_2 - \beta \end{pmatrix}$$
$$= n_1 \left(\widehat{\beta}_1 - \beta\right)' \widehat{V}_{\beta_1}^{-1} \left(\widehat{\beta}_1 - \beta\right) + n_2 \left(\widehat{\beta}_2 - \beta\right)' \widehat{V}_{\beta_2}^{-1} \left(\widehat{\beta}_2 - \beta\right).$$

Minimizing, we find the solution

$$\widetilde{\beta} = \left(n_1 \widehat{V}_{\beta_1}^{-1} + n_2 \widehat{V}_{\beta_2}^{-1}\right) \left(n_1 \widehat{V}_{\beta_1}^{-1} \widehat{\beta}_1 + n_2 \widehat{V}_{\beta_2}^{-1} \widehat{\beta}_2\right)$$

This is also a weighted average, but now the weights depend on the sample size as well.

To develop an asymptotic theory we need to describe what it means for n_1, n_2 to diverge to infinity. A convenient solution is to assume that both diverge, but $n_1/n_2 \to c$, a constant which can differ from one. In practice, we simply think of c as the observed ratio n_1/n_2. Then we can treat $n_1 = cn_2$, and conduct the asymptotics as $n_2 \to \infty$. Then

$$\sqrt{n_1}\left(\widehat{\beta}_1 - \beta\right) \to_d Z_1 \sim N(0, V_{\beta_1})$$

$$\sqrt{n_2}\left(\widehat{\beta}_2 - \beta\right) \to_d Z_2 \sim N(0, V_{\beta_2})$$

and

$$\sqrt{n_2}\left(\widehat{\beta}_1 - \beta\right) = \sqrt{\frac{n_2}{n_1}}\sqrt{n_1}\left(\widehat{\beta}_1 - \beta\right) \to_d c^{-1/2} Z_1$$

We find

$$\widetilde{\beta} = \left(\frac{n_1}{n_2}\widehat{V}_{\beta_1}^{-1} + \widehat{V}_{\beta_2}^{-1}\right)\left(\frac{n_1}{n_2}\widehat{V}_{\beta_1}^{-1}\widehat{\beta}_1 + \widehat{V}_{\beta_2}^{-1}\widehat{\beta}_2\right)$$
$$\simeq \left(c\widehat{V}_{\beta_1}^{-1} + \widehat{V}_{\beta_2}^{-1}\right)\left(c\widehat{V}_{\beta_1}^{-1}\widehat{\beta}_1 + \widehat{V}_{\beta_2}^{-1}\widehat{\beta}_2\right)$$

and

$$\sqrt{n_2}\left(\widetilde{\beta} - \beta\right) = \left(c\widehat{V}_{\beta_1}^{-1} + \widehat{V}_{\beta_2}^{-1}\right)^{-1} \left(c\widehat{V}_{\beta_1}^{-1}\sqrt{n_2}\left(\widehat{\beta}_1 - \beta\right) + \widehat{V}_{\beta_2}^{-1}\sqrt{n_2}\left(\widehat{\beta}_2 - \beta\right)\right)$$
$$\to_d \left(cV_{\beta_1}^{-1} + V_{\beta_2}^{-1}\right)^{-1}\left(cV_{\beta_1}^{-1} c^{-1/2} Z_1 + V_{\beta}^{-1} Z_2\right)$$

$$\sim \left(cV_{\beta_1}^{-1} + V_{\beta_2}^{-1}\right)^{-1} N(0, cV_{\beta_1}^{-1} + V_{\beta_2}^{-1})$$
$$= N(0, \left(cV_{\beta_1}^{-1} + V_{\beta_2}^{-1}\right)^{-1})$$

If you want to write in terms of n_1 you have

$$\sqrt{n_1}\left(\widetilde{\beta} - \beta\right) = \sqrt{\frac{n_1}{n_2}}\sqrt{n_2}\left(\widetilde{\beta} - \beta\right) \to_d c^{1/2} N(0, \left(cV_{\beta_1}^{-1} + V_{\beta_2}^{-1}\right)^{-1}) = (0, \left(V_{\beta_1}^{-1} + c^{-1}V_{\beta_2}^{-1}\right)^{-1}).$$

The two are equivalent.