

1.

- (a) This is easiest solved using matrix notation. Write the model as $y = X_1\beta_1 + X_2\beta_2e$ and the short regression as $y = X_1\hat{\beta}_1 + \hat{e}$. Let $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$. By the properties of least-squares and the fact that $M_1X_1 = 0$,

$$\begin{aligned}\hat{e} &= M_1y \\ &= M_1(X_1\beta_1 + X_2\beta_2 + e) \\ &= M_1(X_2\beta_2 + e)\end{aligned}$$

Thus since M_1 is idempotent

$$\begin{aligned}(n - k_1)s^2 &= (X_2\beta_2 + e)'M_1M_1(X_2\beta_2 + e) \\ &= (X_2\beta_2 + e)'M_1(X_2\beta_2 + e) \\ &= e'M_1e + \beta_2'X_2M_1X_2\beta_2 + 2\beta_2'X_2M_1e.\end{aligned}$$

Since X_2 and M_1 are functions of X , and $E(e | X) = 0$

$$\begin{aligned}(n - k_1)E(s^2 | X) &= E(e'M_1e | X) + \beta_2'X_2M_1X_2\beta_2 \\ &= (n - k_1)\sigma^2 + \beta_2'X_2M_1X_2\beta_2.\end{aligned}$$

the second equality since $E(ee' | X) = I\sigma^2$ and $\text{tr}[M_1] = \text{rank}(M_1) = n - k_1$ imply that

$$E(e'M_1e | X) = \text{tr}[M_1E(ee' | X)] = \text{tr}[M_1]\sigma^2 = (n - k_1)\sigma^2$$

Therefore we find

$$E(s^2 | X) = \sigma^2 + \frac{1}{n - k_1}\beta_2'X_2M_1X_2\beta_2.$$

Common errors:

- i. Assuming (implicitly or explicitly) that $\beta_2 = 0$
 - ii. Pretending that $\hat{e}'\hat{e} = e'e$
 - iii. Assuming that s^2 must be unbiased because it is in a correctly-specified model.
- (b) Note that

$$s^2 = \left(\frac{n}{n - k_1}\right) \left(\frac{1}{n}e'M_1e + \frac{1}{n}\beta_2'X_2M_1X_2\beta_2 + 2\frac{1}{n}\beta_2'X_2M_1e\right)$$

and $\frac{n}{n - k_1} \rightarrow 1$. We learned in class that

$$\frac{1}{n}e'M_1e \rightarrow_p \sigma^2.$$

Indeed,

$$\frac{1}{n}e'M_1e = \frac{1}{n}e'e - \frac{1}{n}e'X_1\left(\frac{1}{n}X_1'X_1\right)^{-1}\frac{1}{n}X_1'e \rightarrow_p \sigma^2$$

since $\frac{1}{n}X_1'X_1 \rightarrow_p Q_{11}$ and $\frac{1}{n}X_1'e \rightarrow_p 0$.

Next,

$$\begin{aligned} \frac{1}{n}\beta_2'X_2'M_1e &= \beta_2' \left(\frac{1}{n}X_2'e - \frac{1}{n}X_2'X_1 \left(\frac{1}{n}X_1'X_1 \right)^{-1} \frac{1}{n}X_1'e \right) \\ &\xrightarrow{p} \beta_2' \left(0 - Q_{21} (Q_{11})^{-1} 0 \right) \\ &= 0 \end{aligned}$$

Finally,

$$\begin{aligned} \frac{1}{n}\beta_2'X_2'M_1X_2\beta_2 &= \beta_2' \left[\frac{1}{n}X_2'X_2 - \frac{1}{n}X_2'X_1 \left(\frac{1}{n}X_1'X_1 \right)^{-1} \frac{1}{n}X_1'X_2 \right] \beta_2 \\ &\xrightarrow{p} \beta_2' \left[Q_{22} - Q_{21} (Q_{11})^{-1} Q_{12} \right] \beta_2 \end{aligned}$$

In sum

$$s^2 \xrightarrow{p} \sigma^2 + \beta_2' \left[Q_{22} - Q_{21} (Q_{11})^{-1} Q_{12} \right] \beta_2.$$

Common errors:

- i. Confusing probability limits and expectations
- ii. Assuming that s^2 must be consistent because it is under correct specification.

2.

(a) By definition,

$$V = (Ex_i^2)^{-1} (E(x_i^2 e_i^2)) (Ex_i^2)^{-1} = E(x_i^2 e_i^2)$$

and

$$V^0 = (Ex_i^2)^{-1} Ee_i^2 = \sigma^2$$

where $\sigma^2 = Ee_i^2$.

(b) By the definition of covariance and the above equations,

$$\begin{aligned} C &= cov(x_i^2, e_i^2) \\ &= E(x_i^2 e_i^2) - E(x_i^2) E(e_i^2) \\ &= V - V^0 \end{aligned}$$

Thus $C = V - V^0$ (or $V = C + V^0$).

Common errors:

- i. Assuming that e_i is homoskedastic (e.g., stating that the assumptions imply homoskedasticity)
- ii. Assuming that $C = 0$

3. A point forecast of y_{n+1} takes the form $x'\hat{\beta}$ for some estimate $\hat{\beta}$ of β . A complete answer requires describing the choice of estimator $\hat{\beta}$, and it is best if this choice is justified.

(a) One option is least-squares $\hat{\beta} = (X'X)^{-1} X'y$. While this estimator is not semiparametrically efficient in the model, it can be justified as simple and robust to misspecification.

(b) Another option is FGLS. $\hat{\beta} = (X'\hat{D}^{-1}X)^{-1} X'\hat{D}^{-1}y$ where $\hat{D} = \text{diag}(\hat{\sigma}_i^2)$, $\hat{\sigma}_i^2 = z_i'\hat{\gamma}$ and $\hat{\gamma} = (Z'Z)^{-1} Z'\hat{\eta}$ where $\hat{\eta}$ is the vector with i 'th entry \hat{e}_i^2 where $\hat{e}_i = y_i - x_i'\hat{\beta}$ and $\hat{\beta}$ is the OLS estimator. Given that the model is specified as a regression with a parametric variance equation, the FGLS estimator is semiparametrically efficient.

A standard forecast interval takes the form

$$x' \hat{\beta} \pm 2 \sqrt{z' \hat{\gamma} + \frac{1}{n} x' \hat{V} x'}$$

where $\hat{\gamma}$ is an estimate of γ and \hat{V} is an estimate of the asymptotic variance of the estimator $\hat{\beta}$. The natural estimator for $\hat{\gamma}$ is described above. The estimate \hat{V} depends on the choice for $\hat{\beta}$. If $\hat{\beta}$ is OLS, then either $\hat{V} = \left(\frac{1}{n} X' X\right)^{-1} \left(\frac{1}{n} \sum_i x_i x_i' \hat{e}_i^2\right) \left(\frac{1}{n} X' X\right)^{-1}$ or $\hat{V} = \left(\frac{1}{n} X' X\right)^{-1} \left(\frac{1}{n} \sum_i x_i x_i' (z_i' \hat{\gamma})\right) \left(\frac{1}{n} X' X\right)^{-1}$. If $\hat{\beta}$ is FGLS, a good choice is $\hat{V} = \left(\frac{1}{n} X' \hat{D}^{-1} X\right)^{-1}$.

On a cautionary note, it may be observed that this forecast interval is correct only when the error e_i is normally distributed.

Common errors:

- i. Assuming that the error is homoskedastic (even though the question explicitly assumes a heteroskedastic variance equation).
 - ii. Assuming that γ is known
 - iii. Assuming that β is known
 - iv. Stating that the forecast is $\hat{\beta}' x$ without describing $\hat{\beta}$.
 - v. Picking the least-squares estimator but not describing why this choice is made.
 - vi. Specifying the forecast interval as $x' \hat{\beta} \pm 2 \sqrt{\hat{\sigma}^2 + \frac{1}{n} x' \hat{V} x'}$ or $x' \hat{\beta} \pm 2 \sqrt{\frac{1}{n} x' \hat{V} x'}$
4. It was not stated explicitly, but implicit in the notation we can see that γ is real valued. A convenient way to write the estimator $\hat{\gamma}$ is

$$\hat{\gamma} = \left(\hat{\beta}' X' X \hat{\beta}\right)^{-1} \hat{\beta}' X' Z$$

Since $Z = X\beta\gamma + u$, we see

$$\begin{aligned} \hat{\gamma} &= \left(\hat{\beta}' X' X \hat{\beta}\right)^{-1} \hat{\beta}' X' (X\beta\gamma + u) \\ &= \left(\hat{\beta}' \frac{1}{n} X' X \hat{\beta}\right)^{-1} \hat{\beta}' \frac{1}{n} X' X \beta\gamma + \left(\hat{\beta}' \frac{1}{n} X' X \hat{\beta}\right)^{-1} \hat{\beta}' \frac{1}{n} X' u \end{aligned}$$

Then since $\hat{\beta} \xrightarrow{p} \beta$ and $\frac{1}{n} X' u \xrightarrow{p} 0$.

$$\hat{\gamma} \xrightarrow{p} (\beta' Q \beta)^{-1} \beta' Q \beta \gamma + (\beta' Q \beta)^{-1} \beta' 0 = \gamma$$

(Technically, this result requires that $\beta' Q \beta > 0$, otherwise γ is not identified.)

Another way to solve this is to write $\hat{\beta} = (X' X)^{-1} X' y$ and then

$$\begin{aligned} \hat{\gamma} &= \left(y' X (X' X)^{-1} (X' X) (X' X)^{-1} X' y\right)^{-1} \left(y' X (X' X)^{-1} X' Z\right) \\ &= \left(y' X (X' X)^{-1} X' y\right)^{-1} \left(y' X (X' X)^{-1} X' Z\right) \\ &= \left(\left(\frac{1}{n} y' X\right) \left(\frac{1}{n} X' X\right)^{-1} \left(\frac{1}{n} X' y\right)\right)^{-1} \left(\left(\frac{1}{n} y' X\right) \left(\frac{1}{n} X' X\right)^{-1} \left(\frac{1}{n} X' Z\right)\right) \\ &\xrightarrow{p} \left(E(y_i x_i') (E(x_i x_i'))^{-1} E(x_i y_i)\right)^{-1} \left(E(y_i x_i') (E(x_i x_i'))^{-1} E(x_i z_i)\right) \quad (1) \end{aligned}$$

We want to show that this equals γ . Since $y = x'_i\beta + e_i$ and $E x_i e_i = 0$,

$$E(x_i y_i) = E(x_i (x'_i\beta + e_i)) = E(x_i x'_i) \beta$$

and since $z_i = x'_i\beta\gamma + u_i$ and $E x_i u_i = 0$,

$$E(x_i z_i) = E(x_i (x'_i\beta\gamma + u_i)) = E(x_i x'_i) \beta\gamma$$

Therefore the right-hand-side of (1) equals

$$\left(\beta' E(x_i x'_i) (E(x_i x'_i))^{-1} E(x_i x'_i) \beta \right)^{-1} \left(\beta' E(x_i x'_i) (E(x_i x'_i))^{-1} E(x_i x'_i) \beta \gamma \right) = \gamma$$

The extra credit problem asked for the asymptotic distribution of $\hat{\gamma}$. In general this is tricky as you have to handle the joint distribution of $\hat{\beta}$ and $\hat{\gamma}$. But when $\gamma = 0$ the problem simplifies. Note that from the above equation when $\gamma = 0$

$$\begin{aligned} \sqrt{n}\hat{\gamma} &= \left(\hat{\beta}' \frac{1}{n} X' X \hat{\beta} \right)^{-1} \hat{\beta}' \frac{1}{\sqrt{n}} X' u \\ &\xrightarrow{d} (\beta' Q \beta)^{-1} \beta' N(0, \Omega_u) \\ &= N \left(0, \frac{\beta' E(x_i x'_i u_i^2) \beta}{(\beta' E(x_i x'_i) \beta)^2} \right) \end{aligned}$$

Common errors:

- (a) Attempting to demonstrate consistency by taking expectations
- (b) Treating $\hat{\beta}$ as a constant rather than a random variable
- (c) Treating $\hat{\beta}$ as an invertible matrix
- (d) Treating $\hat{\beta}$ as if it is a function of X . e.g. $E(\hat{\beta}' x_i u_i | X) = \hat{\beta}' x_i E(u_i | X)$ (this is incorrect since $\hat{\beta}$ is a function of X and y and the problem does not make an assumption about the relationship between e_i and u_i)
- (e) Saying that the WLLN asserts that $n^{-1} \sum_{i=1}^n (\hat{\beta}' x_i)^2 \xrightarrow{p} E(\hat{\beta}' x_i)^2$ ($\hat{\beta}' x_i$ is not iid, as $\hat{\beta}$ depends on the full sample).