Econ 7010 - Midterm 1

Ryan T. Godwin

Winter 2022

The exam is 70 minutes long, and has **50 marks**. Upload your answers to the UM Learn dropbox, within 80 minutes of exam start. You may quickly submit a low-quality version of your exam, and then upload a higher quality version after the 80 minutes.

Consider the model:

$$\boldsymbol{y} = \beta_1 + \beta_2 x + \boldsymbol{\epsilon}$$

which in matrix form is written:

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1}$$

where the X matrix contains a column of 1s and the x variable, and for example, looks like:

$$X = \begin{bmatrix} 1 & 3.1 \\ 1 & 1.2 \\ \vdots & \vdots \\ 1 & 3.3 \end{bmatrix}$$

For model 1, denote the associated LS estimator, predicted values, and residuals as: $\boldsymbol{b}$, $\hat{\boldsymbol{y}}$, and $\boldsymbol{e}$, respectively. Now, suppose that the 2nd column of X is *multiplied* by 2. For example, the X matrix now looks like:

$$X^{\star} = \begin{bmatrix} 1 & 6.2 \\ 1 & 2.4 \\ \vdots & \vdots \\ 1 & 6.6 \end{bmatrix}$$

Consider a similar model to the one above:

$$\boldsymbol{y} = X^{\star}\boldsymbol{\beta}^{\star} + \boldsymbol{\epsilon}^{\star} \tag{2}$$

Denote the LS estimator, predicted values, and residuals associated with model 2 as $\boldsymbol{b}^{\star}$, $\hat{\boldsymbol{y}}^{\star}$, and $\boldsymbol{e}^{\star}$. Note that:

$$X^{\star} = XA, \text{ where } A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

and finally, note that if A, B, and C are all $(k \times k)$ matrices then a property of the matrix inverse is:

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$$

a) What is the relation between $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{y}}^\star$? Prove. [**5 marks**]

Answer. A property of the LS predicted values is that they are invariant to non-singular linear transformations of the X matrix. To see this, write:

$$\begin{aligned}\hat{\boldsymbol{y}}^\star = P_{X^\star}\boldsymbol{y} = P_{XA}\boldsymbol{y} &= XA\left[(XA)'XA\right]^{-1}(XA)'\boldsymbol{y} \\ &= XA\left[A'X'XA\right]^{-1}A'X'\boldsymbol{y}\end{aligned}$$

and since A', $X'X$, and A are all $(k \times k)$ matrices, the above property of the matrix inverse applies:

$$\begin{aligned}\hat{\boldsymbol{y}}^\star &= XAA^{-1}(X'X)^{-1}(A')^{-1}A'X'\boldsymbol{y} \\ &= X(X'X)^{-1}X'\boldsymbol{y} \\ &= P_X\boldsymbol{y} = \hat{\boldsymbol{y}}\end{aligned}$$

b) Is the $\bar{R}^2$ from model 1 larger than, smaller than, or equal to the $\bar{R}^2$ from model 2? Explain. [**5 marks**]

Answer. The adjusted R-square from model 1 is:

$$\bar{R}^2 = \left[1 - \frac{\boldsymbol{e}'\boldsymbol{e}/(n-k)}{\boldsymbol{y}'M_i\boldsymbol{y}/(n-1)}\right]$$

and the adjusted R-square from model 2 is:

$$\bar{R}^2 = \left[1 - \frac{\boldsymbol{e}^{\star\prime}\boldsymbol{e}^\star/(n-k)}{\boldsymbol{y}'M_i\boldsymbol{y}/(n-1)}\right]$$

Similar to part (a), the residuals are *invariant to non-singular linear transformations* of the X matrix. Since $\boldsymbol{e}^\star = \boldsymbol{e}$, the adjusted R-square from the two models are identical.

c) How exactly does $\boldsymbol{b}^\star$ compare to $\boldsymbol{b}$? Prove. (For example, is b_1 the same as $b_1^\star$? Is $b_2^\star$ twice the value, half the value, or the same as b_2?) [**15 marks**]

Answer. The formula for $\boldsymbol{b}^\star$ can be written as:

$$\begin{aligned}\boldsymbol{b}^\star &= \left[(XA)'(XA)\right]^{-1}(XA)'\boldsymbol{y} \\ &= \left[A'X'XA\right]^{-1}A'X'\boldsymbol{y} \\ &= A^{-1}\left(X'X\right)^{-1}\left(A'\right)^{-1}A'X'\boldsymbol{y} \\ &= A^{-1}\left(X'X\right)^{-1}X'\boldsymbol{y} \\ &= A^{-1}\boldsymbol{b}\end{aligned}$$

Now, A^{-1} is:

$$A^{-1} = \frac{1}{2}\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \tag{3}$$

so:

$$\boldsymbol{b}^{\star} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ \frac{1}{2}b_2 \end{bmatrix} \tag{4}$$

The slope estimator from model 2 is *half* the value of the slope estimator from model 1. That is, if the x variable is doubled, the slope estimator is halved.

Now, suppose that a 3rd column is added to the original X matrix, for example:

$$X^{\star\star} = \begin{bmatrix} 1 & 3.1 & 4.7 \\ 1 & 1.2 & 0.8 \\ \vdots & \vdots & \vdots \\ 1 & 3.3 & 9.2 \end{bmatrix}$$

Let the LS estimator associated with the $X^{\star\star}$ matrix be denoted $\boldsymbol{b}^{\star\star}$.

d) Do you think that $b_1^{\star\star}$ and $b_2^{\star\star}$ will be numerically different from b_1 and b_2 (from the original model 1)? Explain why or why not. [**5 marks**]

Answer. Unless the newly added regressor is *orthogonal* to the original regressors, the LS estimates will change. In general, when regressors are added or dropped from the model, all of the LS estimates change, because the x variables are correlated with each other.

To see this, partition the model into two blocks. Denote the 1st 2 columns of the X matrix as the block X_1 and the 3rd column as X_2. Similarly, partition b_1 and b_2 into $\boldsymbol{b}_1$ and partition b_3 into $\boldsymbol{b}_2$. The formula for b_1 may be written as:

$$\boldsymbol{b}_1 = \left(X_1{}'X_1\right)^{-1}\left[X_1'\boldsymbol{y} - X_1'X_2\boldsymbol{b}_2\right] \tag{5}$$

Only if (i) X_1 and X_2 are orthogonal ($X_1'X_2 = 0$); or (ii) $\boldsymbol{b}_2 = 0$, does $\boldsymbol{b}_1 = \left(X_1{}'X_1\right)^{-1}X_1'\boldsymbol{y}$, which is the LS estimator for model 1.

e) Suppose that the 3rd column of $X^{\star\star}$ is correlated with $\boldsymbol{\epsilon}$. Do you think that $b_1^{\star\star}$, $b_2^{\star\star}$ and $b_3^{\star\star}$ are biased estimators? Prove. [**15 marks**]

Answer. Correlation between X and $\boldsymbol{\epsilon}$ is a violation of assumption A.5, leading the LS estimator to be biased. However, by using the partitioned formulas, we will be able to see that only the slope estimator associated with the 3rd column of X (b_3) is biased.

Follow the same partitioning notation as in part (c). Characterize the correlation between the 2nd block of regressors and $\boldsymbol{\epsilon}$ as: $\mathrm{E}[\boldsymbol{\epsilon}] = X_2\gamma$ (for example). The LS estimator for the 2nd block of regressors is:

$$\boldsymbol{b}_2 = \left(X_2'M_1X\right)^{-1}X_2'M_1\boldsymbol{y},$$

Substituting in the population model for $\boldsymbol{y}$, and expanding, we have:

$$\begin{aligned}\boldsymbol{b}_2 &= \left(X_2'M_1X_2\right)^{-1}X_2'M_1\left(X_1\boldsymbol{\beta}_1 + X_2\boldsymbol{\beta}_2 + \epsilon\right) \\ &= \left(X_2'M_1X_2\right)^{-1}X_2'M_1X_1\boldsymbol{\beta}_1 + \left(X_2'M_1X_2\right)^{-1}X_2'M_1X_2\boldsymbol{\beta}_2 \\ &\quad + \left(X_2'M_1X_2\right)^{-1}X_2'M_1\epsilon\end{aligned}$$

Using $M_1 X_1 = 0$, and simplifying, we get:

$$\boldsymbol{b}_2 = 0 + \boldsymbol{\beta}_2 + \left(X_2' M_1 X_2\right)^{-1} X_2' M_1 \epsilon$$

Finally, taking the expectation of $\boldsymbol{b}_2$ and using $\mathrm{E}\left[\epsilon\right] = X_2 \gamma$ we get:

$$\mathrm{E}\left[\boldsymbol{b}_2\right] = \mathbf{0} + \boldsymbol{\beta}_2 + \gamma$$

We will get a similar result for $\mathrm{E}\left[\boldsymbol{b}_2\right]$, except that the third term will cancel since $M_2 X_2 = 0$:

$$\mathrm{E}\left[\boldsymbol{b}_1\right] = \boldsymbol{\beta}_1 + 0 + \left(X_1 M_2' X_1\right)^{-1} X_1' M_2 X_2 \gamma = \boldsymbol{\beta}_2$$

So, $\boldsymbol{b}_1$ is unbiased, while $\boldsymbol{b}_2$ is biased.

f) Suppose that you instead estimate $\boldsymbol{\beta}$ using $\hat{\boldsymbol{\beta}}$, where $\hat{\boldsymbol{\beta}}$ is a linear estimator, and where $E[\hat{\boldsymbol{\beta}}] = \boldsymbol{\beta}$. What can you say about $V[\hat{\boldsymbol{\beta}}]$? [**5 marks**]

Answer. According to the Gauss-Markov theorem, any other linear and unbiased estimator besides the LS estimator must have higher variance. That is, it must be the case that $\mathrm{V}[\hat{\boldsymbol{\beta}}] - \mathrm{V}[\boldsymbol{b}]$ is a positive semi definite matrix.

END.