Econ 709 Problem Set 2

Sarah Bass *

November 22, 2020

Question 3.2

$$\begin{aligned}
\hat{\beta}_{ols} &= (X'X)^{-1}X'Y \\
\hat{\beta}_Z &= (Z'Z)^{-1}Z'Y \\
&= (C'X'XC)^{-1}C'X'Y \\
&= C^{-1}(X'X)^{-1}C'^{-1}C'X'Y \\
&= C^{-1}(X'X)^{-1}X'Y \\
&= C^{-1}\hat{\beta}_{ols}
\end{aligned}$$

$$\begin{aligned}
\hat{e}_{ols} &= (I - X(X'X)^{-1}X')Y \\
\hat{e}_Z &= (I - Z(Z'Z)^{-1}Z')Y \\
&= (I - XC(C'X'XC)^{-1}C'X')Y \\
&= (I - XCC^{-1}(X'X)^{-1}X')Y \\
&= (I - X(X'X)^{-1}X')Y \\
&= \hat{e}_{ols}
\end{aligned}$$

Question 3.5

$$\begin{aligned}
\hat{e} &= (I - X(X'X)^{-1}X')Y \\
\hat{\beta}_e &= (X'X)^{-1}X'\hat{e} \\
&= (X'X)^{-1}X'(I - X(X'X)^{-1}X')Y \\
&= (X'X)^{-1}X'Y - (X'X)^{-1}X'X(X'X)^{-1}X'Y \\
&= (X'X)^{-1}X'Y - (X'X)^{-1}X'Y \\
&= 0
\end{aligned}$$

*I have discussed this problem set with Emily Case, Michael Nattinger, Alex Von Hafften, and Danny Edgel.

Question 3.6

$$\begin{aligned}\hat{\beta}_{\hat{Y}} &= (X'X)^{-1}X'\hat{Y} \\ &= (X'X)^{-1}X'X(X'X)^{-1}X'Y \\ &= (X'X)^{-1}X'Y \\ &= \hat{\beta}_{ols}\end{aligned}$$

Question 3.7

Let $\Gamma = \begin{pmatrix} I_{n_1} \\ \bar{0} \end{pmatrix}$ where $\bar{0}$ is an $n_2 \times n_1$ vector of zeros. Then $X_1 = X\Gamma$, so

$$\begin{aligned}PX_1 &= X(X'X)^{-1}X'X_1 \\ &= X(X'X)^{-1}X'X\Gamma \\ &= X\Gamma \\ &= X_1\end{aligned}$$

$$\begin{aligned}MX_1 &= (I - X(X'X)^{-1}X')X_1 \\ &= (I - X(X'X)^{-1}X')X\Gamma \\ &= (X - X(X'X)^{-1}X'X)\Gamma \\ &= (X - X)\Gamma \\ &= 0\end{aligned}$$

Question 3.11

Let X contain a constant.

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\hat{Y}_i &= \frac{1}{n}\sum_{i=1}^{n} Y_i - \hat{e}_i \\ &= \frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{1}{n}\sum_{i=1}^{n}\hat{e}_i \\ &= \frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{X'\hat{e}}{n} \\ &= \frac{1}{n}\sum_{i=1}^{n} Y_i\end{aligned}$$

Note, $\frac{1}{n}\sum_{i=1}^{n}\hat{e}_i = \frac{X'\hat{e}}{n} = 0$ since X contains a column of constants.

Question 3.12

First note that equation 3.53 can't be estimated by OLS because $D_1 + D_2 = \vec{1}$ (a vector containing 1 in every element), so the constant column of X is no longer linearly independent, and $X'X$ is not invertible.

Part A

Equations 3.54 and 3.55 contain the same information.

$$\begin{aligned} D_1\alpha_1 + D_2\alpha_2 + e &= D_1\alpha_1 + (\vec{1} - D_1)\alpha_2 \\ &= D_1(\alpha_1 - \alpha_2) + \vec{\alpha_2} \\ &= \mu + D_1\phi \end{aligned}$$

Therefore, the regressions are the same with $\mu = \alpha_2$ and $\phi = \alpha_1 - \alpha_2$.

Part B

$$\begin{aligned} \vec{1}'D_1 &= \sum_{i=1}^{n} 1\{\text{person } i \text{ is a man}\} = n_1 \\ \vec{1}'D_2 &= \sum_{i=1}^{n} 1\{\text{person } i \text{ is a woman}\} = n_2 \end{aligned}$$

Question 3.13

Part A

Let $X = [D_1 D_2]$. Order our observations such that the first n_1 observations are men and the remaining n_2 observations are women. Then $X'X = \begin{pmatrix} \vec{1}'_{n_1}\vec{1}_{n_1} & \bar{0} \\ \bar{0} & \vec{1}'_{n_2}\vec{1}_{n_2} \end{pmatrix}$

$$\begin{aligned} \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix} &= \begin{pmatrix} \vec{1}'_{n_1}\vec{1}_{n_1} & \bar{0} \\ \bar{0} & \vec{1}'_{n_2}\vec{1}_{n_2} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{n_1} y_i \\ \sum_{i=n_1+1}^{n_1+n_2} y_i \end{pmatrix} \\ &= \begin{pmatrix} n_1 & \bar{0} \\ \bar{0} & n_2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{n_1} y_i \\ \sum_{i=n_1+1}^{n_1+n_2} y_i \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{n_1}\sum_{i=1}^{n_1} y_i \\ \frac{1}{n_2}\sum_{i=n_1+1}^{n_1+n_2} y_i \end{pmatrix} \\ &= \begin{pmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{pmatrix} \end{aligned}$$

Part B

$Y^* = Y - D_1\bar{Y}_1 - D_2\bar{Y}_2$ simplifies to $Y^* = \hat{u}$, which is the deviation from average for men and women.

$X^* = Y - D_1\bar{X}_1{}' - D_2\bar{X}_2{}'$ is the residual of the regression $X = D_1b_1 + D_2b_2$. From part A, we know $b_1 = \bar{X}_1, b_2 = \bar{X}_2$. So X^* is a matrix of regressors transformed to be in deviations from the average for men or women, depending on the gender identity of the individual.

Part C

$$
\begin{aligned}
\hat{\beta} &= (XM_DX)^{-1}X'M_DY \\
\tilde{\beta} &= (X'^*X^*)^{-1}X'^*Y^* \\
&= (XM_DX)^{-1}X'M_DY \\
&= \hat{\beta}
\end{aligned}
$$

where we solved for $\hat{\beta}$ via theorem 3.4.

Question 3.16

Let $X = [X_1X_2], \hat{\beta} = [\hat{\beta}_1'\hat{\beta}_2']', \hat{\beta}^* = [\tilde{\beta}_1'\vec{0}_{n_2}']'$ where $\vec{0}_{n_2}$ is an n_2 sized matrix of zeros.

$$
\begin{aligned}
R_2^2 &= 1 - \frac{\sum_{i=1}^n \hat{e}_i^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= 1 - \frac{\hat{e}'\hat{e}}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= 1 - \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&\geq 1 - \frac{(Y - X\hat{\beta}^*)'(Y - X\hat{\beta}^*)}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= R_1^2
\end{aligned}
$$

where the inequality comes from OLS minimizing the sum of squared residuals.

If X_2 is orthogonal to Y then

$$
\begin{aligned}
X_2'Y = 0 &\Rightarrow \hat{\beta}_2 = 0 \\
&\Rightarrow \tilde{\beta} = \hat{\beta} \\
&\Rightarrow R_2^2 = R_1^2
\end{aligned}
$$

3.21

If X_1 or X_2 is orthogonal to Y, then the orthogonal regressor will have estimated coefficients of 0 in both equations. If only one regressor is orthogonal, then the equation with both regressors will reduce to the same as the equation with only one regressor (that is not orthogonal). If both regressors are orthogonal to Y, then all of the coefficient estimates will be 0.

If X_1 and X_2 are orthogonal to each other, then by theorem 3.4 we have that:

$$\begin{aligned}\hat{\beta}_1 &= (X_1' M_2 X_1)^{-1}(X_1' M_2 Y) \\ &= ((M_2 X_1)'(M_2 X_1))^{-1}((M_2 X_1)'Y) \\ &= (X_1' X_1)^{-1}(X_1' Y) \\ &= \tilde{\beta}_1\end{aligned}$$

By symmetry, the same condition holds for $\hat{\beta}_2 = \tilde{\beta}_2$

Question 3.22

$$\begin{aligned}\tilde{\beta} &= (X_1' X_1)^{-1} X_1' Y \\ \tilde{u} &= Y - X_1 \tilde{\beta} \\ \tilde{\beta}_2 &= (X_2' X_2)^{-1} X_2' \tilde{u} \\ &= (X_2' X_2)^{-1} X_2'(Y - X_1 \tilde{\beta}_1) \\ \hat{\beta}_2 &= (X_2' X_2)^{-1} X_2'(Y - X_1 \hat{\beta}_1)\end{aligned}$$

So $\tilde{\beta}_2 = \hat{\beta}_2$ only when $\tilde{\beta}_1 = \hat{\beta}_1$. As we showed in the previous problem, this occurs when X_1 andX_2 are orthogonal or when one (or both) of the regressors is orthogonal to Y.

Question 3.23

First note that the residuals are the same from both equations:

$$\begin{aligned}\tilde{\beta}_2 &= ((X_2 - X_1)' M_1 (X_2 - X_1))^{-1}((X_2 - X_1)' M_1 Y) \\ &= (X_2' X_2)^{-1} X_2' Y \\ &= \hat{\beta}_2. \\ \tilde{\beta}_1 &= (X_1' X_1)^{-1} X_1'(Y - (X_2 - X_1)\tilde{\beta}_2) \\ &= (X_1' X_1)^{-1} X_1' Y - (X_1' X_1)^{-1} X_1'(X_2 - X_1)\tilde{\beta}_2 \\ &= (X_1' X_1)^{-1} X_1'(Y - X_2 \hat{\beta}_2) + (X_1' X_1)^{-1} X_1' X_1 \hat{\beta}_2 \\ &= \hat{\beta}_1 + \hat{\beta}_2. \\ \Rightarrow \tilde{e} &= X_1 \tilde{\beta}_1 + (X_2 - X_1)\tilde{\beta}_2 \\ &= X_1(\hat{\beta}_1 + \hat{\beta}_2) + (X_2 - X_1)\hat{\beta}_2 \\ &= X_1 \hat{\beta}_1 + X_2 \hat{\beta}_2 \\ &= \hat{e}.\end{aligned}$$

Since the residual variance estimates are a function of the estimated residuals, which are the same from both regressions, $\hat{\sigma}^2 = \tilde{\sigma}^2$.

Question 7

Part A

$$
\begin{aligned}
E[\hat{\beta}|X] &= E[(X'X)^{-1}X'Y|X] \\
&= (X'X)^{-1}X'E[Y|X] \\
&= (X'X)^{-1}X'X\beta \\
&= \beta \\
\Rightarrow E[\hat{\beta}_1|X] &= \beta_1
\end{aligned}
$$

Part B

$$
\begin{aligned}
E[\hat{\hat{\beta}}_1|X] &= E[(X_1'X_1)^{-1}X_1'\hat{Y}|X] \\
&= E[(X_1'X_1)^{-1}X_1'X\hat{\beta}|X] \\
&= E[(X_1'X_1)^{-1}X_1'X(X'X)^{-1}X'Y|X] \\
&= (X_1'X_1)^{-1}X_1'X(X'X)^{-1}X'E[Y|X] \\
&= (X_1'X_1)^{-1}X_1'X(X'X)^{-1}X'X\beta \\
&= (X_1'X_1)^{-1}X_1'X\beta \\
&= (X_1'X_1)^{-1}X_1'(X_1\beta_1 + X_2\beta_2) \\
&= \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2
\end{aligned}
$$

This is equal to β_1 if either $\beta_2 = 0$ or X_1 andX_2 are orthogonal.

Part C

$$
\begin{aligned}
\tilde{\tilde{\beta}} &= (X'X)^{-1}X'\tilde{Y} \\
&= (X'X)^{-1}X'X_1\tilde{\beta}_1 \\
&= \begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix} \tilde{\beta}_1
\end{aligned}
$$

Part D

Let $\tilde{\tilde{Y}} = X\tilde{\tilde{\beta}}, \tilde{e} = \tilde{Y} - \tilde{\tilde{Y}}$.

$$\begin{aligned}
\tilde{\tilde{Y}} &= X\tilde{\tilde{\beta}} \\
&= X \begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix} \tilde{\beta}_1 \\
&= X_1 \tilde{\beta}_1 \\
&= \tilde{Y} \\
\Rightarrow \tilde{e} &= 0
\end{aligned}$$

$$\begin{aligned}
\Rightarrow R^2 &= 1 - \frac{\tilde{e}'\tilde{e}}{\sum_{i=1}^{n} (\tilde{Y}_i - \bar{\tilde{Y}})^2} \\
&= 1 - \frac{0}{\sum_{i=1}^{n} (\tilde{Y}_i - \bar{\tilde{Y}})^2} \\
&= 1
\end{aligned}$$

Part E

$$\begin{aligned}
Var(\tilde{\tilde{\beta}}|X) &= Var(\begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix} (X_1'X_1)^{-1} X_1' Y | X) \\
&= \begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix} (X_1'X_1)^{-1} X_1' Var[Y|X] \left(\begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix} (X_1'X_1)^{-1} X_1' \right)' \\
&= \begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix} (X_1'X_1)^{-1} X_1' \sigma^2 I X_1 (X_1'X_1)^{-1} \begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix}' \\
&= \sigma^2 \begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix} (X_1'X_1)^{-1} X_1' X_1 (X_1'X_1)^{-1} \begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix}' \\
&= \sigma^2 \begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix} (X_1'X_1)^{-1} \begin{pmatrix} I_{k_1} \\ \bar{0} \end{pmatrix}' \\
&= \begin{pmatrix} \sigma^2 (X_1'X_1)^{-1} & \bar{0} \\ \bar{0} & \bar{0} \end{pmatrix}
\end{aligned}$$