Problem Set #2

Danny Edgel
Econ 709: Economic Statistics and Econometrics I
Fall 2020

November 22, 2020

Collaborated with Sarah Bass, Emily Case, Michael Nattinger, and Alex Von Hafften

Question 1

Recall that $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$. Then, if $Z = XC$,

$$\begin{aligned}
\hat{\beta}_Z &= (Z'Z)^{-1}Zy = \left[(XC)'XC\right]^{-1}(XC)'y \\
&= \left(C'X'XC\right)^{-1}C'X'y = C^{-1}(X'X)^{-1}C'^{-1}C'X'y \\
&= C^{-1}(X'X)^{-1}X'y = C^{-1}\hat{\beta}_{OLS}
\end{aligned}$$

Also recall that $\hat{\varepsilon}_{OLS} = Y - X\hat{\beta}_{OLS}$. Then,

$$\begin{aligned}
\hat{\varepsilon}_Z &= Y - Z\hat{\beta}_Z = y - Z(Z'Z)^{-1}Zy \\
&= \left(I - XC((XC)'XC)^{-1}XC\right)y = \left(I - XCC^{-1}(X'X)^{-1}C'^{-1}C'X\right)y \\
&= \left(I - X(X'X)^{-1}X\right)y = y - X(X'X)^{-1}Xy \\
&= y - X\hat{\beta}_{OLS} = \hat{\varepsilon}_{OLS}
\end{aligned}$$

Question 2

3.5) Recall from question 1 that $\hat{\varepsilon}_{OLS} = (I - X(X'X)^{-1}X')Y$. Then,

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

3.6) Let $\hat{Y} = X(X'X)^{-1}X'Y$ and $\hat{\beta}_Y$ represent the OLS coefficient from a regression of $\hat{Y}$ on X. Then,

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

3.7) Let $X = [X_1\ X_2]$ be an m by n matrix and recall that $P = X(X'X)^{-1}X'$ and $M = I - P$. Let $n = n_1 + n_2$, where X_1 is an m by n_1 matrix and X_2 is m by n_2. Then, we can define $\Gamma = \begin{pmatrix} I_{n_1} \\ 0 \end{pmatrix}$ such that $X_1 = X\Gamma$. Thus,

$$\begin{aligned} PX_1 &= PX\Gamma = X(X'X)^{-1}X'X\Gamma = X\Gamma = X_1 \\ MX_1 &= (I-P)X_1 = X_1 - PX_1 = X_1 - X_1 = 0 \end{aligned}$$

Question 3

3.11) Let X contain only a non-zero constant, $c \in \mathbb{R}$, such that $X = c\mathbb{1}_n$, where n is the number of elements in Y and $\mathbb{1}_n$ is an $n \times 1$ vector of ones. Then,

$$\begin{aligned} \hat{Y} &= X(X'X)^{-1}X'Y = (c\mathbb{1}_n)\left[(c\mathbb{1}_n)'(c\mathbb{1}_n)\right]^{-1}(c\mathbb{1}_n)'Y \\ &= c\mathbb{1}_n\left(c^2(\mathbb{1}_n'\mathbb{1}_n)\right)^{-1}c(\mathbb{1}_n'Y) = c^2\mathbb{1}_n\left(c^2 n\right)^{-1}n\bar{Y} \\ &= \frac{c^2 n}{c^2 n}\bar{Y}\mathbb{1}_n = \bar{Y}\mathbb{1}_n \end{aligned}$$

Thus, $\hat{Y}$ is a column vector where every entry is $\overline{Y}$

3.12) Equation (3.53) cannot be estimated by OLS. Equation (3.53) can be rewritten as $Y = X\beta + \varepsilon$, where $X = [\mathbb{1}_n\ D_1\ D_2]$ and $\beta = \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{pmatrix}$. $D_1 + D_2 = \mathbb{1}_n$, so $\text{rank}(X) \neq k$, violating the first Gauss-Markov asusumption.

(a) Neither (3.54) nor (3.55) is more general. The two specifications have the same explanatory power. In (3.54), the average of Y for men is given by α_1, and the average for women is given by α_2. In (3.55), the averages are $\mu + \phi$ and μ, respectively. Thus, given the parameters for one specification, you could calculate the parameters of the other with:

$$\begin{aligned} \mu + \phi &= \alpha_1 & \qquad \phi &= \alpha_2 - \alpha_1 \\ \mu &= \alpha_2 & \qquad \alpha_2 &= \mu \end{aligned}$$

(b) $\mathbb{1}_n'D_1 = n_1$, $\mathbb{1}_n'D_2 = n_2$

3.13) (a) Letting $X = [D_1\ D_2]$ and $\hat{\beta} = \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix}$, we can solve:

$$\begin{aligned} \hat{\beta} &= (X'X)^{-1}X'Y \\ &= \begin{pmatrix} \mathbb{1}_n'D_1 & 0 \\ 0 & \mathbb{1}_n D_2 \end{pmatrix}^{-1}\begin{pmatrix} D_1'Y \\ D_2'Y \end{pmatrix} = \frac{1}{n_1 n_2}\begin{pmatrix} n_2 & 0 \\ 0 & n_1 \end{pmatrix}\begin{pmatrix} \sum_{i=1}^n D_{1i}Y_i \\ \sum_{i=1}^n D_{2i}Y_i \end{pmatrix} \\ &= \frac{1}{n_1 n_2}\begin{pmatrix} n_2\sum_{i=1}^n D_{1i}Y_i \\ n_1\sum_{i=1}^n D_{2i}Y_i \end{pmatrix} = \begin{pmatrix} \frac{1}{n_1}\sum_{i=1}^n D_{1i}Y_i \\ \frac{1}{n_2}\sum_{i=1}^n D_{2i}Y_i \end{pmatrix} = \begin{pmatrix} \overline{Y}_1 \\ \overline{Y}_2 \end{pmatrix} \end{aligned}$$

(b) In plain English, Y^* is the demeaned value of Y, using the means for men and women separately. Econometrically, as shown below, Y^* is the residualized value of Y, or, rather, the value of Y that cannot be explained by gender alone:

$$\begin{aligned} Y^* &= Y - D_1\overline{Y}_1 - D_2\overline{Y}_2 = Y - \left(D_1\overline{Y}_1 + D_2\overline{Y}_2\right) \\ &= Y - [D_1 \ D_2]\begin{pmatrix}\hat{\gamma}_1 \\ \hat{\gamma}_2\end{pmatrix} = \hat{\mu} \end{aligned}$$

It logically follows, then, that X^* is the residualized value of X, from a regression of X on D_1 and D_2.

(c) Let $D = [D_1 \ D_2]$, $\hat{\alpha} = \begin{pmatrix}\hat{\alpha}_1 \\ \hat{\alpha}_2\end{pmatrix}$, and $\hat{\gamma} = \begin{pmatrix}\hat{\gamma}_1 \\ \hat{\gamma}_2\end{pmatrix}$. From part (b), we can rewrite:

$$Y^* = Y - D_1\hat{\gamma}_1 - D_2\hat{\gamma}_2 = Y - D\hat{\gamma} = (I_n - D(D'D)^{-1}D')Y = M_DY$$

Where M_D is the orthogonal projection matrix for D. Similarly, $X^* = M_DX$. Thus, we can derive:

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

Since The second recgression is a partition of D and X, then by Theorem 3.4,

$$\hat{\beta} = (X'M_DX)^{-1}X'M_DY$$

Thus, $\hat{\beta} = \widetilde{\beta}$.

Question 4

Let $\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$ and $X = [X_1 \ X_2]$. By the definition of R^2,

$$\begin{aligned} R_1^2 &= 1 - \frac{\hat{\varepsilon}'\hat{\varepsilon}}{(Y-\overline{Y})'(Y-\overline{Y})} = 1 - \frac{(Y-X\hat{\beta})'(Y-X\hat{\beta})}{(Y-\overline{Y})'(Y-\overline{Y})} \\ &= 1 - \frac{(Y-X_1\hat{\beta}_1 - X_2\hat{\beta}_2)'(Y-X_1\hat{\beta}_1 - X_2\hat{\beta}_2)}{(Y-\overline{Y})'(Y-\overline{Y})} \end{aligned}$$

Now let $\widetilde{\beta}_2 = 0^*\hat{\beta}_2$. Then, for $\widetilde{\beta} = \begin{pmatrix} \widetilde{\beta}_1 \\ \widetilde{\beta}_2 \end{pmatrix}$, $X\widetilde{\beta} = X_1\widetilde{\beta}_1$ and $\widetilde{\beta}_1 = \hat{\beta}_1 = (X_1'M_2X_1)^{-1}X_1'M_2Y$:

$$\begin{aligned} R_2^2 &= 1 - \frac{(Y-X\widetilde{\beta})'(Y-X\widetilde{\beta})}{(Y-\overline{Y})'(Y-\overline{Y})} = 1 - \frac{(Y-X_1\widetilde{\beta}_1 - X_2\widetilde{\beta}_2)'(Y-X_1\widetilde{\beta}_1 - X_2\widetilde{\beta}_2)}{(Y-\overline{Y})'(Y-\overline{Y})} \\ &= 1 - \frac{(Y-X_1\hat{\beta}_1)'(Y-X_1\hat{\beta}_1)}{(Y-\overline{Y})'(Y-\overline{Y})} \\ &\leq 1 - \frac{(Y-X_1\hat{\beta}_1 - X_2\hat{\beta}_2)'(Y-X_1\hat{\beta}_1 - X_2\hat{\beta}_2)}{(Y-\overline{Y})'(Y-\overline{Y})} = R_1^2 \end{aligned}$$

It is possible for $R_1^2 = R_2^2$, if $\hat{\beta}_2 = 0$, which occurs if X_2 is orthogonal to Y.

Question 5

3.21) As a standard OLS coefficient in a non-partitioned regression, $\widetilde{\beta}_1 = (X_1'X_1)^{-1}X_1'Y$. By Theorem 3.4, $\hat{\beta}_1 = (X_1'M_2X_1)^{-1}X_1'M_2Y$. Thus, $\widetilde{\beta}_1 = \hat{\beta}_1$ if $X_1'M_2 = X_1'$. This will be true if:

$$\begin{aligned} X_1'M_2 &= X_1' \\ X_1'(I - X_2(X_2'X_2)^{-1}X_2') &= X_1' \\ X_1' - X_1'X_2(X_2'X_2)^{-1}X_2') &= X_1' \end{aligned}$$

Which holds if $X_1'X_2 = 0$, in which case X_1 and X_2 are orthogonal. The same is true for $\widetilde{\beta}_2$ and $\hat{\beta}_2$, by the same mathematical logic. The coefficients will also be equal if either X_1 or X_2 (or both) are orthogonal to Y, since this will lead to OLS coefficients of zero.

3.22) In a partitioned regression, $\hat{\beta}_2 = (X_2'X_2)^{-1}X_2'(y - X_1)\hat{\beta}_1$. Then,

$$\widetilde{u} = Y - X_1\widetilde{\beta}_1 = Y - X_1(X_1'X_1)^{-1}X_1'Y = (I - X_1(X_1'X_1)^{-1}X_1')Y = M_1Y$$
$$\widetilde{u} = X_2\widetilde{\beta}_2$$
$$\widetilde{\beta}_2 = (X_2'X_2)^{-1}X_2'\widetilde{u} = (X_2'X_2)^{-1}X_2'(Y - X_1\widetilde{\beta}_1)$$

Thus, $\widetilde{\beta}_2 = \hat{\beta}_2$ only if $\widetilde{\beta}_1 = \hat{\beta}_1$, which, as we solved in 3.21, is only true if X_1 and X_2 are orthogonal. This is rarely the case, but weirder things have happened in 2020 alone.

3.23) The equation for the regression of Y on X is $Y = X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + \hat{\varepsilon}$, and the equation for the regression of Y on Z can be simplified as follows:

$$Y = Z_1\widetilde{\beta}_1 + Z_2\widetilde{\beta}_2 + \widetilde{\varepsilon} = X_1\widetilde{\beta}_1 + (X_2 - X_1)\widetilde{\beta}_2 + \widetilde{\varepsilon} \quad = X_1(\widetilde{\beta}_1 - \widetilde{\beta}_2) + X_1\widetilde{\beta}_2 + \widetilde{\varepsilon}$$

The two regressions capture the same variation in Y, where $\hat{\beta} = \widetilde{\beta}_1 - \widetilde{\beta}2$ and $\hat{\beta}_2 = \widetilde{\beta}_2$. Most importantly, $\hat{\varepsilon} = \widetilde{\varepsilon}$, so $\hat{\sigma}^2 = \widetilde{\sigma}^2$.

Question 6

(due w/ PS3)

Question 7

(a) Recall the estimator $\hat{\beta}_1$ for a partition regression: $\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'(Y - X_2\hat{\beta}_2)$. Then,

$$\begin{aligned}
\mathbb{E}(\hat{\beta}_1|X) &= \mathbb{E}\left[(X_1'X_1)^{-1}X_1'(Y - X_2\hat{\beta}_2)|X\right] \\
&= (X_1'X_1)^{-1}X_1'\,\mathbb{E}\left[Y - X_2\hat{\beta}_2|X\right] \\
&= (X_1'X_1)^{-1}X_1'\,\mathbb{E}\left[Y|X\right] - (X_1'X_1)^{-1}X_1'X_2\,\mathbb{E}\left[\hat{\beta}_2|X\right]
\end{aligned}$$

Where $\hat{\beta}_2 = (X_2'M_1X_2)^{-1}X_2'M_1Y$ and $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$, so:

$$\begin{aligned}
\mathbb{E}(\hat{\beta}_1|X) &= (X_1'X_1)^{-1}X_1'\,\mathbb{E}\left[Y|X\right] - (X_1'X_1)^{-1}X_1'X_2(X_2'M_1X_2)^{-1}X_2'M_1\,\mathbb{E}\left[Y|X\right] \\
&= (X_1'X_1)^{-1}X_1'\left[I - X_2(X_2'M_1X_2)^{-1}X_2'M_1\right]\mathbb{E}\left[Y|X\right]
\end{aligned}$$

(b)

$$\hat{\hat{\beta}}_1 = (X_1'X_1)^{-1}X_1'\hat{y} = (X_1'X_1)^{-1}X_1'X\hat{\beta}$$
$$\mathbb{E}(\hat{\hat{\beta}}_1|X) = \mathbb{E}\left[(X_1'X_1)^{-1}X_1'X\hat{\beta}|X\right] = (X_1'X_1)^{-1}X_1'X\,\mathbb{E}\left[\hat{\beta}|X\right]$$

Since the Gauss-Markov assumptions are satisfied, $\mathbb{E}(\hat{\beta}|X) = \beta$, so:

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

Thus, $\hat{\hat{\beta}}_1$ is only an unbiased estimator of β_1 if $X_1'X_2 = 0$, i.e. if X_1 and X_2 are orthogonal.

(c) Note that $X_1 = X\Gamma$, where $\Gamma = [I_{k_1}\ \ 0 * X_2]$

$$\begin{aligned}\widetilde{\beta}_1 &= (X_1'X_1)^{-1}X_1'y \\ \widetilde{y} &= X_1\widetilde{\beta}_1 \\ \widetilde{\widetilde{\beta}} &= (X'X)^{-1}X'\widetilde{y} = (X'X)^{-1}X'X_1\widetilde{\beta}_1 = \Gamma\widetilde{\beta}_1 \\ \widetilde{\widetilde{\beta}} &= \begin{pmatrix}\widetilde{\beta}_1 \\ 0\end{pmatrix}\end{aligned}$$

(d) $R^2 = 1$, as shown below.

$$\begin{aligned}R^2 &= 1 - \frac{\widetilde{\varepsilon}'\widetilde{\varepsilon}}{(Y-\overline{Y})'(Y-\overline{Y})} \\ \widetilde{\varepsilon} &= \widetilde{y} - X\widetilde{\widetilde{\beta}} = X_1\widetilde{\beta}_1 - X\begin{pmatrix}\widetilde{\beta}_1 \\ 0\end{pmatrix} = X_1\widetilde{\beta}_1 - X_1\widetilde{\beta}_1 = 0\end{aligned}$$

(e) Note that

$$\begin{aligned}Var(\widetilde{\beta}_1|X) &= Var((X_1'X_1)^{-1}X_1'y|X) \\ &= (X_1'X_1)^{-1}X_1'Var(y|X)X_1(X_1'X_1)^{-1} \\ &= (X_1'X_1)^{-1}X_1'\sigma^2 I X_1(X_1'X_1)^{-1} \\ &= \sigma^2(X_1'X_1)^{-1}X_1'X_1(X_1'X_1)^{-1} \\ &= \sigma^2(X_1'X_1)^{-1}\end{aligned}$$

Where $\widetilde{\widetilde{\beta}} = \begin{pmatrix}\widetilde{\beta}_1 \\ 0\end{pmatrix}$, such that:

$$Var(\widetilde{\widetilde{\beta}}|X) = \begin{pmatrix}\sigma^2(X_1'X_1)^{-1} & 0 \\ 0 & 0\end{pmatrix}$$