Cornell University
Department of Economics

Econ 620
Instructor: Prof. Kiefer

Solution to Problem set # 2

1)
a) $y = \frac{x}{\alpha}$
Hence, $\frac{1}{y} = \alpha - \frac{\beta}{x}$. Therefore, let y'=$\frac{1}{y}$ and let $x' = -\frac{1}{x}.Then,$ y'= $\alpha + \beta x'$.
b) $y = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$
Hence, $\frac{1}{y} = 1 + e^{-\alpha-\beta x}$. Then, $\ln(\frac{1}{y} - 1) = \ln(\frac{1-y}{y}) = -\alpha - \beta x$. Therefore, let y'= $ln(\frac{y}{1-y})$ and we get that y'=$\alpha + \beta x$.

2)

Unbiasedness implies that $c_1 + c_2 = 1$ since $E(b) = (c_1 + c_2)\beta$.
Therefore, the problem consists of minimizing $Var(b) = c_1^2 \mathrm{v}_1 + c_2^2 \mathrm{v}_2$ subject to $c_1 + c_2 = 1$.
Second order conditions will be met, since the objective funcion is quadratic and the restriction is linear and from the first order conditions we can conclude that $\mathrm{c}_1 = \frac{\mathrm{v}_2}{\mathrm{v}_1+\mathrm{v}_2}$ and $\mathrm{c}_2 = \frac{\mathrm{v}_1}{\mathrm{v}_1+\mathrm{v}_2}$.

3)
First of all, note that

$$Z'Z = \begin{bmatrix} y' \\ X' \end{bmatrix} \begin{bmatrix} y & X \end{bmatrix} = \begin{bmatrix} y'y & y'X \\ X'y & X'X \end{bmatrix}$$

Now, the OLS estimate is given by;

$$\widehat{\beta} = (X'X)^{-1} X'y = \begin{bmatrix} 20 & 0 \\ 0 & 75 \end{bmatrix}^{-1} \begin{bmatrix} 10 \\ 25 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{3} \end{bmatrix}$$

To find s^2, we have to find $e'e$;

$$e'e = \left(y - X\widehat{\beta}\right)' \left(y - X\widehat{\beta}\right) = y'y - \widehat{\beta}'X'X\widehat{\beta}$$
$$= [100] - \begin{bmatrix} \frac{1}{2} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 20 & 0 \\ 0 & 75 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{3} \end{bmatrix} = \frac{260}{3} = 86.667$$

Then, from the formula for s^2;

$$s^2 = \frac{e'e}{N-k} = \frac{\frac{260}{3}}{(20-2)} = \frac{130}{27} = 4.8148$$

Finally, we have

$$R^2 = 1 - \frac{e'e}{y'Ay}$$

Note that

$$y'Ay = y'\left[I - \mathbf{1}\left(\mathbf{1}'\mathbf{1}\right)^{-1}\mathbf{1}'\right]y = y'y - \frac{1}{N}y'\mathbf{1}\mathbf{1}'y$$
$$= y'y - N\overline{y}^2 = 100 - 20 \times \left(\frac{10}{20}\right)^2 = 95$$

Then,

$$R^2 = 1 - \frac{\frac{260}{3}}{95} = 0.0877$$

The full data set is a (20×3) matrix whose first column is the vector of observations on the dependent variable and the remaining columns are the vector of ones corresponding to the constant term and vector of observations on the independent variable. On the other hand, the cross product matrix given in the question is simply a (3×3) matrix. Still, we can extract every bit of information we need from the cross product matrix only. We do not lose anything just by looking at the cross product matrix. In other words, the cross product matrix is a sufficient statistic for the regression model.

4)

$$(\text{Corr}(\hat{y},\text{y}))^2 = \frac{\left[\sum_i (y_i - \overline{y})\left(\hat{y_i} - \hat{\overline{y}}\right)\right]^2}{\sum_i (y_i - \overline{y})^2 \sum_i \left(\hat{y_i} - \hat{\overline{y}}\right)^2} = \frac{y'A\hat{y}}{y'Ay}\frac{y'A\hat{y}}{\hat{y}'A\hat{y}} = \frac{y'A\hat{y}}{y'Ay} = R^2$$

The second equality comes from the fact that $A = I - \mathbf{1}\left(\mathbf{1}'\mathbf{1}\right)^{-1}\mathbf{1}'$. The third equality follows because $y'A\hat{y} = (\hat{y} + e)'A\hat{y} = \hat{y}'A\hat{y} + e'A\hat{y} = \hat{y}'A\hat{y} + e'\hat{y} = \hat{y}'A\hat{y}$ since $e'\hat{y} = 0$.

5)

We know that $\widehat{\beta}_{yx} = \frac{\sum_{i=1}^n (y_i - \overline{y})(x_i - \overline{x})}{\sum_{i=1}^n (x_i - \overline{x})^2}$ and $\widehat{\beta}_{xy} = \frac{\sum_{i=1}^n (y_i - \overline{y})(x_i - \overline{x})}{\sum_{i=1}^n (y_i - \overline{y})^2}$.

a) $\widehat{\beta}_{yx} = 1/\widehat{\beta}_{xy}$ if $\sum_{i=1}^n (x_i - \overline{x})^2 . \sum_{i=1}^n (y_i - \overline{y})^2 = [\sum_{i=1}^n (y_i - \overline{y})(x_i - \overline{x})]^2$. This will happen only when $\mathrm{R}^2 = 1$. To see why, go to part c).

b) The sign is determined by the numerator and hence both will have the same sign.

c) $\widehat{\beta}_{yx}\widehat{\beta}_{xy} = \frac{\sum_{i=1}^n (y_i - \overline{y})(x_i - \overline{x})}{\sum_{i=1}^n (x_i - \overline{x})^2}\frac{\sum_{i=1}^n (y_i - \overline{y})(x_i - \overline{x})}{\sum_{i=1}^n (y_i - \overline{y})^2} = \frac{\left[\sum_{i=1}^n (y_i - \overline{y})(x_i - \overline{x})\right]^2}{\sum_{i=1}^n (x_i - \overline{x})^2 \sum_{i=1}^n (y_i - \overline{y})^2} = r^2$

d) Since both estimators have the same sign, it follows from c) that $\mid \widehat{\beta}_{yx} \mid\mid \widehat{\beta}_{xy} \mid \leq 1$. Hence, dividing by $\mid \widehat{\beta}_{xy} \mid$, we have that $\mid \widehat{\beta}_{yx} \mid \leq \frac{1}{|\widehat{\beta}_{xy}|}$. That is, the absolute value of the estimator of the slope in the regression of y on x is smaller than that of the regression of x on y.

For the numerical part, $\widehat{\beta}_{yx} = \frac{22.13-200(20.72/200)(11.34/200)}{12.16-200(11.34/200)^2} = \frac{20.955}{11.51} = 1.81948$

$\widehat{\beta}_{xy} = \frac{20.955}{84.96-200(20.72/200)^2} = \frac{20.955}{82.8134} = 0.253$ and $\mathrm{r}^2 = \frac{[20.955]^2}{11.51 \quad 82.8134} = 0.4603 = \widehat{\beta}_{yx}\widehat{\beta}_{xy}$ and $1.81948 = \widehat{\beta}_{yx} \leq 1/\widehat{\beta}_{xy} = 3.9526$

6)

a)

E[b| $X] = \beta + (X'X)^{-1}X'E[\varepsilon \mid X] = \beta$ since $E[\varepsilon \mid X] = 0$.

Therefore, E[b]=E(E[b| X]) =E(β) = β.

Therefore, if the regressors are stochastic, as long as tthey are uncorrelated with the error term (and all other assumptions hold), the OLS estimator of β is still unbiased.

b)

$Var(b) = E_X[Var(b \mid X)] + Var_X E[b \mid X]$

Since $E_X[Var(b \mid X)] = E_X[Var((X'X)^{-1}X'\varepsilon \mid X)] = E_X[(X'X)^{-1}X'Var(\varepsilon \mid X)X(X'X)^{-1}] =$

$= E_X[(X'X)^{-1}X'\sigma^2 IX(X'X)^{-1}] = \sigma^2 E_X[(X'X)^{-1}]$ and also, $Var_X E[b \mid X] = Var_X(\beta) = 0$, it follows,

$Var(b) = \sigma^2 E[(X'X)^{-1}]$