Econometrics EC 424/524
George W. Evans, Winter 2019.
Problem Set 2
Due in class Tuesday, January 29.

1. Obtain the standard formulae for b_1 and b_2, the LS coefficient estimates $y_i = b_1 + b_2 x_i + e_i$ for the simple regression model

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i,\ i = 1, \ldots, n,$$

from the general result

$$b = (X'X)^{-1}X'y.$$

To do this write $X = \begin{bmatrix} \iota & x \end{bmatrix}$, where $\iota' = (1, 1, \ldots, 1)$ and $x' = (x_1, x_2, \ldots, x_n)$, and $b' = (b_1, b_2)$.

2. Consider the multiple regression model $y = X\beta + \varepsilon$. Suppose the K independent variables x_i', for each observation i, are linearly transformed to the K variables $z_i' = x_i'A$, where A is a $K \times K$ nonsingular matrix. The $n \times K$ matrix of observations of the transformed regressors is thus $Z = XA$. (i) Show that e_X, the residual vector of a LS regression of y on X, is the same as e_Z, the residual vector of a LS regression of y on Z. (ii) Compare the coefficients b from the LS regression of y on X to the coefficients c from the LS regression of y on Z. [Note: the key results are given in the "Transformed Variables" Theorem in Greene, Ch. 3.]

3. For an $n \times K$ matrix X, let $M = I - X(X'X)^{-1}X'$. Partition X into two submatrices as $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ and let $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$.
(i) Show that $MX = 0$. Explain why this implies $MX_1 = 0$. (ii) What is $M_1 M$? (iii) Let e be the residuals of the LS regression of y on X. Let $\tilde{e}$ be the residuals of the regression of e on X_1. Give the relationship of $\tilde{e}$ to e.

4. Let X be $n \times K$ with rank K and let x_k be the kth column of X. Consider the LS regression of x_k on X, i.e.

$$x_k = Xb + e,$$

where $b' = (b_1, \ldots, b_K)$. By appealing directly to the definition of Least Squares, explain why $b_k = 1$ and $b_i = 0$ for $i \neq k$.

5. Use the data set from Problem Set 1. The variable inc is the log of real disposable income per capita. Compute its percentage growth rate (at annual rates):

$$\text{dinc} = 400 \cdot (\text{inc} - \text{inc(-1)}),$$

where inc(-1) denotes the lagged value of inc.

(i) Use LS to estimate the model

$$\text{dgexp} = \beta_1 + \beta_2 \text{ dpg} + \beta_3 \text{ dinc} + \varepsilon.$$

Report 90% confidence intervals on β_2 and on β_3, computed using the standard errors and the critical point of the t-distribution obtained from the tables posted on Canvas.

(ii) Reestimate the model including also lagged independent variables, i.e.

$$\text{dgexp} = \beta_1 + \beta_2 \text{ dpg} + \beta_3 \text{ dinc} + \beta_4 \text{ dpg(-1)} + \beta_5 \text{ dinc(-1)} + \varepsilon.$$

Test the null hypothesis at the 5% level that there are no lagged effects, i.e.

$$H_0 : \beta_4 = \beta_5 = 0.$$

Conduct this test using the F statistic computed from the R^2 values for two regressions of the form (i) and (ii), and using the F critical value obtained from tables posted on Canvas. To do this test correctly the two regressions have to be estimated using the same sample period. Consequently, (i) has to be reestimated so that the first data point used is the same as in (ii). In doing the test, provide key details as well as the result of the test, i.e. report the F statistic, the degrees of freedom for the test, and give the 5% critical point of the test.

6*.[1] The Stata file Growth.DTA contains data on average growth rates (*Growth*) from 1960 to 1995 for 65 countries with along with variables that are potentially related to growth. (For a description see the file posted on Canvas).

Omitting the data for Malta (which is an outlier), run a regression of *Growth* on *TradeShare, YearsSchool, Rev_ Coups, Assassinations* and *GRDP60*.

(a) What is the value of the coefficient on *Rev_ Coups*? Interpret the value of this coefficient. Is it large or small in a real-world sense?

(b) Use the regression to predict the annual average growth rate for a country that has average values for all regressors.

(c) Repeat (b) but now assume that the country's value for *TradeShare* is one standard deviation above the mean.

(d) Test whether, taken as a group, *YearsSchool, Rev_ Coups, Assassinations and GRDP60*, can be omitted from the regression. What is the p-value of the F-statistic?

[1] Starred problems are required for EC 524 and are extra credit for EC 424.