AREC 345: Global Poverty and Economic Development

Problem Set 2

Department of Agricultural and Resource Economics
University of Maryland
Fall 2016

Problem Set 1 is due at the start of section on September 23. Problem sets turned in more than 5 minutes after the start of section will be marked as late. All problem sets must be turned in as hard copies; points will be deducted if multiple pages are not stapled together.

To complete this problem set, you will need to download the data set `arec345ps2data.xlsx` from the course website (it is the same data set you used for Problem Set 1). The data set contains 7 variables: `Country code`, `Country`, `Log GDP per capita in 2010`, `Absolute latitude`, `Under 5 mortality`, and dummies for being located in South Asia and Sub-Saharan Africa.

1. Use the `LINEST` command in Excel to estimate a bivariate regression of `Log GDP per capita` on `Absolute latitude` (in other words, a regression in which log GDP per capita is the dependent variable and absolute latitude is the independent variable).

   The `LINEST` command estimates the best linear fit for a given data set, allowing us to explore the relationships between the independent and dependent variables. To run the `LINEST` command, start with an Excel worksheet in which your dependent variable is in the first column and your independent variable is just to the right of the dependent variable. For example, if you wanted to estimate the linear regression equation

   $$y = a + b \cdot x + e$$

   to explore the relationship between independent variable x and dependent variable y, you'd have the variable y in Column A and x in Column B.

   In this case, you'll probably want to create a new worksheet with log GDP in the first column and `Absolute latitude` in the second column. The data set we're using contains 156 observations, so the dependent variable, for example, will be contained in cells `A2:A157` and the independent variable will be contained in cells `B2:B157`.

   We are estimating two parameters: the constant and the coefficient on `Absolute latitude`. To use the `LINEST` command, highlight a block of cells two cells across and two down. Then type the following:

   ```
   =LINEST(A2:A157,B2:157,TRUE,TRUE)
   ```

   and hit `CONTROL-SHIFT-ENTER`. Your block of four cells now contains the following **in this order**: the coefficient estimate on `Absolute latitude` with the estimated standard error below it and then, to the right, the estimated constant and the associated standard error.

   Please see the TA for assistance if you are having trouble using Excel to run the regression. You are welcome to use Stata instead of Excel if you are already familiar with it.

Report your coefficient estimates. What do they suggest about the relationship between proximity to the equator and economic development? Use the coefficient estimate and the standard error to calculate the t-statistic. Is the coefficient on latitude statistically significant? At what confidence level?

2. Now confirm that you have the correct coefficient estimate by using the following formula: the regression coefficient estimate (in a bivariate regression) should be equal to the correlation coefficient multiplied by the ratio of the standard deviation of the dependent variable to the standard deviation of the independent variable:

$$\hat{b} = \text{corr}\left(\texttt{Log GDP per capita}, \texttt{Latitude}\right) \cdot \frac{\text{SD}(\texttt{Log GDP per capita})}{\text{SD}(\texttt{Latitude})} \tag{1}$$

   (a) Calculate the correlation between `Log GDP per capita` and `Absolute latitude` using the `CORREL` command in Excel.

   (b) Calculate the standard deviation of `Log GDP per capita` using the `STDEV` command.

   (c) Calculate the standard deviation of `Absolute latitude` using the `STDEV` command.

   (d) Calculate $\hat{b}$ using the formula provided above (in Equation 1). Confirm that it matches the estimated slope coefficient from your answer to (1).

3. Generate a new variable, `Tropical`, that is equal to one if `Absolute latitude` is less than or equal to 23.5 degrees and equal to zero otherwise. Then generate a second dummy variable, `Tropical Africa`, that is equal to one if both `Tropical` and `Sub-Saharan Africa` are equal to one and equal to zero otherwise. Calculate the average of `Log GDP per capita` in the four groups defined by these two variables:

   (a) What is the average `Log GDP per capita` in tropical Africa (i.e. when `Tropical Africa`=1)?

   (b) What is the average `Log GDP per capita` in non-tropical Africa (i.e. when `Sub-Saharan Africa`=1 and `Tropical`=0)?

   (c) What is the average `Log GDP per capita` in tropical non-Africa?

   (d) What is the average `Log GDP per capita` in non-tropical non-Africa?

4. Regress `Log GDP per capita` on three dummy variables: `Sub-Saharan Africa`, `Tropical`, and `Tropical Africa`. To do this, you will want to create a new worksheet with `Log GDP per capita` in Column A and the three dummy variables in Columns B, C, and D. Report your coefficient estimates. What do the coefficient estimates suggest about the relationship between tropical geography and income in Africa vs. the rest of the world?

5. How does the estimated intercept (i.e. constant term) from your answer to (4) compare to your answer to (3d)?

6. Show that the estimated intercept plus the coefficient on `Tropical` sum to your answer to (3c).

7. Show that the estimated intercept plus the all three estimated slope coefficients sum to your answer to (3a).