

ASSIGNMENT 3

651 Empirical Economics, Fall 2018

Deadline: October 10th, 20.00.

The assignment gives a maximum of 5 points. Answers should be provided in English. The key to a good score is to be clear, precise and concise in your answers. We subtract points for solutions which are hard to follow. Submit a pdf-file of your solutions via the portal on the course web page no later than the stated deadline. Make sure to include your question answers (including any tables and figures) in a single document, and to also attach your do-file for the STATA assignment toward the end of your document. It is fine to copy-paste results from Stata into your answers, but do not simply use screen shots (as they are hard to read). To make results and do-files from Stata easy to read in Word, use Courier with font size 9 and single spacing. Ask the teaching assistants in case you are uncertain about the format for the assignment solutions.

A huge literature in economics and political science investigates the importance of well-functioning institutions for economic development. The next course in the Economics sequence (Economics of Organization) will acquaint you more with this literature and the theoretical arguments as to how and why institutions matter. This and the upcoming assignment will allow you to get a taste of empirical work in this area. Compared to the two previous assignments, you have more freedom how to exactly solve the assignment, though we do provide guidelines for you to follow. If you are uncertain how to interpret any of the questions, ask any of the teaching staff.

Start by download “The QoG Basic Time-Series Data” from the Quality of government Project’s website: <https://qog.pol.gu.se/data/datadownloads/qogbasicdata>. Also make sure to access the Codebook with definitions of the variables in this data set (available under Downloads/Assignments on the course webpage). You will use this data set also in Assignment 4. Then match the quality-of-government with the data set “colonialorigins.dta” which is already available under Downloads/Assignments. The code below should help you with the merging:

```
/*Specify the working directory (so that Stata finds the files). Here I have placed both files in  
C:\Users\erik\Downloads*/  
cd C:\Users\erik\Downloads  
/*Load the quality-of-government data, here named “Quality” */  
use "Quality", clear  
/*Merge the colonial origins data*/  
merge m:1 ccodealp using "colonialorigins.dta"
```

The point of merging these datasets is that “colonialorigins-dta” include a set of time-independent variables not available in the quality-of-government data set which focuses more on time-dependent variables. Ask Andrew in case you encounter problems merging the two datasets.

Your task in this assignment is to analyze the effect of the quality of government on economic prosperity (GDP per capita), using the data at hand. In other words, we want to know whether

an increase in the quality of government in expectation leads to higher prosperity. While answering this question, you should use the methods used in Lecture 11 and 12-13 as well as the previous course content you find relevant. You don't need to exactly follow the order given in the guidelines below, but all the questions should be addressed in your solutions.

Guidelines

1. Choose a measure for the quality of government. As you might already have guessed, the quality of government is neither easy to define nor to quantify, and the QoG data includes a host of potential measures. We expect you to pick one measure which you work with during this assignment. Make sure to (concisely) motivate your choice.
2. Define (and briefly motivate) the set of countries and period of time you intend to study. A key concern here is whether the key variables of interest are available for the countries and time period you consider. Make sure to include at least 20 countries and 10 years (ensuring the data has a panel structure).
3. Inspect the key variables of interest, i.e., the dependent variable (GDP per capita) and the quality-of-government measure you have chosen. Plot their distributions in histograms, their evolution over time, their correlation using scatterplots, and investigate whether there are systematic patterns of missing values. Check whether the min, max, median, mean and standard deviation of the variables make sense. Feel free to transform the variables to (or away from) log form in case you find this called for. Note that since you have panel data here, you may have to transform the data for the figures and statistics to be easy to interpret. For example, you may want to calculate the mean value of your quality-of-government measure across all countries in each year to show the evolution of quality of government over time in a parsimonious way or calculate the average for each country over a period of time when showing the variable distributions in a histogram. The bottom line is that you cannot show everything but must summarize the data in a way that is both informative and relatively easy to digest.
4. Describe the sources of variation in your QoG-measure. How much of the variation is within countries over time and how much is between countries at a particular point in time? What do these descriptive statistics tell you about which types of analyses are most adequate?
5. Answer the question whether quality of government affects GDP per capita by analyzing the data. You should estimate models with pooled OLS, fixed effects and random effects. It is up to you to choose and motivate the control variables that you include in these analyses. Consider at least two different sets of control variables and discuss how the results change. Also note the following:
 - a. Ignore the IV-method discussed in Lecture 14-15) in this assignment.
 - b. The set of appropriate control variables may depend on the methods used. Make sure you use an appropriate set of controls.
 - c. Write out the equations for the models you are estimating and comment on the key identifying assumptions. A good score requires you to discuss these assumptions in the actual context you consider, i.e., not just rephrasing mathematical expressions from the lecture notes. (If you perform several similar

versions of a given model – say with different sets of control variables – you don't have to write out the equation every single time).

- d. Compare the results from different methods and how you interpret them. The point of the exercise is to understand the choices between different methods. It is not a problem per se if different methods give different results.
- e. Do not only comment on how the estimated effect of quality of government changes across specifications, but also on how the standard errors differ.
- f. Present your results in an accessible way. Crafting nice and clear tables is both more difficult and a more important part of empirical work than most students realize.

Best of luck!