# Problem Set 3

ECON 306 - Introduction to Econometrics
Spring 2021
Delina E Agnosteva

## Due Date: April 1 at 11:59PM on Canvas.

**INSTRUCTIONS:** Solve the following questions to the best of your ability. Ask me if you do not know how to solve any of these questions before the due date. I will work with you if you are having trouble solving these.

*To receive full credit for this assignment, the problem set needs to be submitted to* **Canvas in a single PDF document** *containing your 1) Stata log file, 2) any figures (scatterplots, histograms, etc.), and 3) any written explanations and answers. All of these components need to be attached together in that order. Late submissions will NOT be accepted. DO NOT email! No assignments will be accepted via email.*

First of all, for this problem set, you will have to submit the Stata log file. Stata can record your session into a file called a log file but does not start a log automatically; you must tell Stata to record your session. By default, the resulting log file contains what you type and what Stata produces in response, recorded in a format called Stata Markup and Control Language (SMCL). The file can be printed or converted to plain text for incorporation into documents you create with your word processor. You can find more information here: https://www.stata.com/manuals13/u15.pdf.

So, in the beginning of your Stata .do file write the following command: *log using session* (or a different file name). Then, at the very end of your .do file, include *translate session.smcl session.pdf*. This would translate your Stata SMCL log files directly into PDF files and then use Adobe Acrobat to merge PDF files together. You will *need to* turn in this log file to receive full credit for this assignment.

I would strongly suggest compiling the log file in Stata **after you have completed all of your code and can run it smoothly without any errors**. In that way, your log file would not contain any lines of code that does not produce any results or any duplicate results. Please do your best to include comments in your code (using the ∗ sign in your Stata .do file) and to make the solutions to the different problems as clearly marked as possible. Otherwise, the graders might have to penalize you, if they cannot follow your work. And then I will have to re-grade your work and the whole process becomes highly inefficient.

# Growth.xlsx Data.

The data Growth contains data on average growth rates over 1960-1995 for 65 countries, along with variables that are potentially related to growth. These data were provided by Professor Ross Levine of Brown University and were used in his paper with, Thorsten Beck and Norman Loayza "Finance and the Sources of Growth" Journal of Financial Economics, 2000, Vol. 58, pp. 261-300. (Side note: The Journal of Financial Economics is one of the top 3 Finance academic journals and publishing a paper in that journal is quite prestigious. The other 2 top Finance journals are the Journal of Finance and the Review of Financial Studies.) There is an additional PDF file on Canvas that describes the variables.

Use the data to complete the following:

**Problem 1.** The data is currently in an Excel format. You need to **import** it into Stata and convert it into a Stata dataset (.dta). Once you have the data imported in Stata, save it as 'Growth.dta'.

**Problem 2.** Construct a scatterplot of average annual growth rate (*Growth*) on the average trade share (*TradeShare*).
**Does there appear to be a relationship between the variables?**

**Problem 3.** One country, Malta, has a trade share much larger than the other countries. Find Malta on the scatterplot. (If you are feeling adventurous you could even label the point representing Malta on the scatter plot. But you are not required to.)
**Does Malta look like an outlier?**

**Problem 4.** Using all observations, run a regression of *Growth* on *TradeShare*. **What is the estimated slope? What is the estimated intercept? Use the regression to predict the growth rate for a country with a trade share of 0.5 and for another with a trade share equal to 1.0.**

**Problem 5.** Estimate the same regression, excluding the data for Malta. (Hint: But do not completely drop Malta from the data.) **What is the estimated slope? What is the estimated intercept? Use the regression to predict the growth rate for a country with a trade share of 0.5 and for another with a trade share equal to 1.0.**

**Problem 6.** Perform a formal test for heteroskedasticity in the data. (In class, we performed a Breusch-Pagan / Cook-Weisberg test for heteroskedasticity in one of our in-class demos.)

**Problem 7.** Excluding the data for Malta, run a regression of *Growth* on *TradeShare*, but now controlling for heteroskedasticity in the data. **What is the value of the coefficient on *TradeShare*? What happened to the regression results after controlling for heteroskedasticity.**

**Problem 8.** Plot the estimated regression functions from Problem 4 and Problem 5 into the same scatterplot. You need to create a graph that contains 1) a scatterplot of the data; 2) the regression line from Problem 4; 3) the regression line from Problem 5. I have created and shown you a similar scatter plot in some of our in-class Stata demos.

**Problem 9.** After having excluded Malta from the regression and controlling for heteroskedasticity, is the estimated regression slope statistically significant? That is, can you reject the null hypothesis $H_0 : \beta_1 = 0$ vs. a two-sided alternative hypothesis at the 10%, 5%, or 1% significance level? (Stata has already generated all of this information for you.)

**Problem 10.** What is the $p$-value associated with the coefficient's $t$-statistic? (Stata has already generated all of this information for you.)

**Problem 11.** After having excluded the data for Malta, construct a table that shows the sample mean, standard deviation, and minimum and maximum values for the series *Growth*, *TradeShare*, *YearsSchool*, *Oil*, *Rev_Coups*, *Assassinations*, and *RGDP60*.

**Problem 12.** After having excluded the data for Malta, run a regression of *Growth* on *TradeShare*, *YearsSchool*, *Rev_Coups*, *Assassinations*, and *RGDP60*. **What is the value of the coefficient on *Rev_Coups*?**

**Problem 13.** Use the regression to predict the average annual growth rate for a country that has average values for all regressors.

**Problem 14.** Why is *Oil* omitted from the regression? What would happen if it were included?