Group Project

ECO 4000 Statistical Analysis for Economics and Finance

06 May, 2022

Maximum Points : 35

Due Date : 18 May, 2022 (Before Final Exam)

You will be using the same data set **terrorism.dta**. You will be required to write this assignment in R Markdown that includes your code, results, plots, and answers. **You are responsible for submitting only one hard copy per group.**

In addition to that, as a group, you are required to submit a summary (approx. 250 words per member) wherein each member describes their individual contribution to this group project.

Variable	Definition
ftmpop	Number of fatalities from terrorist incidents in the country, 1998-2004, per million population (U.S. State Department)
evmpop	Number of terrorist events in the country, 1998-2004, per million population (U.S. State Department)
gdppc	GDP per capita in the country (World Bank)
lackpf	Index of the lack of political freedoms (Freedom House), 1-7 scale, 7 = extremely limited political freedoms
language	Index of linguistic fractionalization (0 to 1 scale, 0 = no fractionalization)
ethnic	Index of ethnic fractionalization (0 to 1 scale, 0 = no fractionalization)
religion	Index of religious fractionalization (0 to 1 scale, 0 = no fractionalization)
mideast, latinam, easteurope, africa, eastasia	= 1 if the country is in the indicated region, = 0 otherwise

Once again,You should restrict your analysis to the observations for which there are data on GDP per capita (i.e. for which gdppc is non-missing) and for which there are terrorist fatalities (i.e. for which ftmpop > 0). It will be helpful to use your modified data from the previous assignment.

1) Create a new binary variable higdppc, which equals one if gdppc is greater than or equal to the median in the data set, and which equals zero otherwise; also create the interaction variables hi_lack = higdppc x lackpf and hi_lack2 = higdppc x lackpf2. (5 Points)

2) Produce a regression table using "stargazer" package to fill in the values in the table below. (10 Points)

	(6)	(7)	(8)	(9)
Dependent variable:	ln(*ftmpop*)	ln(*ftmpop*)	ln(*ftmpop*)	ln(*ftmpop*)
Regressor:				
higdppc	( )	( )	( )	( )
lackpf	( )	( )	( )	( )
$lackpf^2$	—	( )	( )	( )
higdppc × *lackpf*	( )	( )	—	—
higdppc × $lackpf^2$	—	( )	—	—
ethnic	—	—	( )	( )
religion	—	—	( )	( )
higdppc × *ethnic*	—	—	( )	( )
higdppc × *religion*	—	—	( )	( )
Mideast	—	—	—	( )
Other regional dummies (latinam, easteurope, africa, eastasia)?	No	No	No	Yes
Intercept	( )	( )	( )	( )
Regression summary statistics:				
$\bar{R}^2$				
R^2				
SER				
n				

Notes: The "other regional dummies" are *latinam*, *easteurope*, *africa*, and *eastasia* (the omitted case is Western Europe combined with North America).

3) Regression (6) produces two regression lines, one for higdppc = 0 and one for higdppc = 1. Produce a scatterplot of lnftmpop vs. lackpf, showing the two regression lines. This can be done either by producing two scatterplots, one for higdppc = 0 and one for higdppc = 1, or by combining the two scatter plots into a single graph. Use regression (6) to write out the estimated regression lines for the two groups (in slope-intercept form). (5 Points)

4) Use the results in the table to answer the following questions. (10 Points)

a) Is the difference between the two slopes plotted in the scatterplot of Question 3 statistically significantly different from zero at the 5% significance level? Explain. (No Code Needed) .In a sentence or two, interpret the sign of the coefficients in regression (6) and the scatterplot in Question 2; that is, explain in everyday terms the findings shown in that scatterplot.

b) Using regression (7), test the hypothesis that the coefficients on higdppc x lackpf and $higdppc \times lackpf^2$ are zero, against the hypothesis that one or the other (or both) is nonzero. State in words what the hypothesis is that you are testing.(Show the Code)

c) Using regression (7), test the hypothesis that the coefficients on lackpf2 and $higdppc \times lackpf^2$ are zero, against the hypothesis that one or the other (or both) is nonzero. State what the hypothesis is that you are testing.(Show the Code)

d) Are the coefficients on the other regional binary variables in regression (9) (latinam, easteurope, africa, eastasia) jointly statistically significant at the 5% significance level? Explain. (Show the Code)

e) As we learned in the class, there will be many possible sources of omitted variables bias(OVB). In your opinion what do you think are the possible sources of OVB?

5) One theory is that ethnic and religious diversity leads to strife and terrorism when economic resources are poor, but if overall economic conditions are strong then ethnic and religious diversity are more readily tolerated. Do regressions (8) and (9) support this theory? Explain. (2 Points)

6) Write a paragraph summarizing your findings about the relation between terrorist fatalities and economic conditions, political freedoms, and ethnic and religious diversity. Your discussion should be based on your results tables from your individual computer assignment(one that you created earlier) as well as this project. (that is, on the empirical evidence, not your opinions). (3 Points)