Economics 423
Intermediate Econometrics

Computer Project #2 (100 points)

Instructions:
-You must submit this to soriano@umd.edu by 11:59 PM by the due date specified on the most recent version of the syllabus.
-A proper submission includes a log file (or whatever the equivalent is in R) in a text or PDF format. I should not have to open Stata, R, or Python to see your code and output.
-For each line of code, include a brief comment that explains what you are intending to do with that line. (For example: sum y x z *Summarize variables I am using). This is your best way to receive partial credit if you make a mistake. This is especially true for people using R or Python, since I do not use R or Python myself.
-Comment your answers to the questions at the beginning or the end of the log. Keep them all in one place so I do not need to search the output for each answer.
-Please see my solution Computer Project #1 for guidance on formatting.
-At the beginning of the file, specify the one person that you worked with (if anyone).
-Please note: I reserve the right to ask anyone to stay in office hours to explain their answers if I have reason to believe their work in not their own.

Background and Purpose

The purpose of this project is to estimate a fixed effects model using panel data and to contrast the results with (a) using OLS to pool the data; (b) estimating a random effects model. The project will also raise issues such as how to treat time in pooled data models. The equations you will estimate are Kuznets curves for traffic fatalities—that is, you will see how the death rate due to traffic fatalities in a country varies with per capita income in the country.

A standard Kuznets curve describes an inverted U-shaped relationship between some externality and per capita income. In the case of traffic fatalities one would initially expect fatalities to worsen as income grows since motor vehicles per capita are growing rapidly and deaths per motor vehicle are high. (Many of these deaths are pedestrians, killed by buses and trucks.) As incomes grow, fatalities per vehicle begin to decline. This may happen in part due to people taking motorized transit rather then walking, due to safer vehicles and/or safer roads.

The data you will use are described in the paper *Traffic Fatalities and Economic Growth* by Elizabeth Kopits and Maureen Cropper.[1] The data set consists of the following variables for 88 countries:

Name
Country code
Year
Traffic fatalities

[1]Elizabeth Kopits and Maureen Cropper, "Traffic Fatalities and Economic Growth," Accident Analysis and Prevention, 37 (2005) 169-178. See also World Bank Policy Research Working Paper 3035, April 2003.

Population
Per capita GDP, measured in 1985 international dollars
HD1 dummy (=1 if UN Human Development Index > 0.8)

The number of years of data varies from country to country so that you are dealing with an unbalanced panel.

You will estimate equations relating the natural logarithm of traffic fatalities/population (F/P) to the logarithm of per capita income (Y) and the log of per capita income squared:

$$(1) \quad \ln(F/P)_{it} = b_0 + b_1 \ln Y_{it} + b_2 (\ln Y_{it})^2 + b_3 t + \alpha_i + u_{it}$$

where t is time. If variables were not in natural logarithms, the equation might predict negative levels of F/P at income levels that one would observe in the near future.

Tasks

1. Read the paper Traffic Fatalities and Economic Growth. You will wind up estimating models similar to the quadratic models (eq. (2)) in the Working Paper version of the paper. Begin by running (1) using OLS. Estimate the model with and without clustered standard errors. Why do the standard errors differ? Which are correct? **[20 points for output showing transformation of variables and OLS results.]**

2. Estimate a Kuznets curve for ln(F/P) using fixed effects and an exponential time trend (equation (1) above). Are your results statistically significant? At what value of per capita income does F/P peak? [Calculate the FE estimates with and without clustered standard errors.] Contrast the value at which F/P peaks with the value in 1. **[20 points for output; 5 for correct peak.]**

3. We learned in Ch. 10 of WR to de-trend our variables. Is it likely that the variables for all 88 countries follow a common time trend? Create a separate time trend for the HD1 countries and re-estimate the model with the two time trends. How does this change your results from step 2.? **[20 points for output.]**

4. Contrast the results in 3. with what happens if you estimate a random effects model. Write a paragraph analyzing the results. Are the results similar to of different from the FE model? Would you expect this? **[20 points for output.]**

5. Explain why economists would favor estimating the Fixed Effects model in 3. rather than a random effects model or the OLS version of equation (1). What do the country-specific intercepts possibly capture in this equation? **[20 points]**

You must hand in your STATA output and write out answers to the above questions. Points will be assigned as indicated above.