ECO 5350
Intro. Econometrics

T. Fomby
Fall 2016

EXERCISE 1

Purpose: To learn such concepts as causal effect, experimental data, observational data, data types in economics – time series data, cross-section data, pooled cross-section/time series data, panel (longitudinal) data – and a little bit about operating the STATA computer program on Apps.SMU.edu. **This exercise is due on Thursday, September 8.**

Work problems 1, 3, and 4 on page 15 of your textbook and computer exercises C6 and C7 on pages 16 and 17. I will choose 3 of these problems/exercises for grading so you need to answer each of the problems and exercises on a separate page or pages so that I can easily grade them. Be sure and put your name on each of the problems/exercises.

Problem #4

computer Exercise C6

Keys to

Chapter 1 Exercises

Problem 4. States (and provinces) that have control over taxation sometimes reduce taxes in an attempt to spur economic growth. Suppose that you are hired by a state to estimate the effect of corporate tax rates on, say, the growth in per capita gross state product (GSP)

(i) What kind of data would you need to collect to undertake a statistical analysis?

Answer: You would need a **panel data set**. That is, you would need to have several years' worth of Data on each of the 50 states in the United States and, in particular, each state's annual growth rate of per capita GSP (as the dependent variables) along with some state-specific explanatory variables observed in corresponding years. These explanatory variables would include the tax rate within the state as it varies over time. But additional explanatory variables might include the amount of natural resources within each state, the number of urban areas in the state, the level of education of the citizens of the state, and the number of Fortune 500 corporate headquarters in the state.

(ii) Is it feasible to do a controlled experiment? What would be required?

Answer: Really, it is not feasible to do a controlled experiment in this case. That would require that the U.S. have a dictator who would randomly divide the states into control group states and treatment group states. Then the dictator would tell the legislators of the control group states that they have to have the same tax rate and cannot change it for the next 20 years while the dictator would tell the state legislatures in the treatment group to charge different tax rates of varying degrees. But in our democracy this is not feasible. We must work with observational data and use as many relevant control (explanatory) variables as we can to tease out the lone effect of tax rate changes on the growth rates of states.

(iii) Is a correlation analysis between GSP growth and tax rates likely to be convincing? Explain.

Answer: As argued in the answer to question (ii) above, we need a lot of more control variables than just tax rate changes to discern the lone effect of tax rate changes on growth rates of the individual states. A simple correlation between tax rate changes and state growth rates will not be sufficient. After controlling for other factors the partial correlation between tax rates and growth rates could easily change and be zero. One just has to analyze the full set of data in a panel data setting.

C6. Here is the STATA program and output that are necessary to answer the various parts of this question.

(i) 2197 counties in 1996 data set
1051 counties had zero murders in 1996
2166 counties had zero executions in 1996
(2166/2197)*100 = 98.5% of counties had zero executions.

(ii) Largest number of murders in a county = 1403
Largest number of executions in a county = 3
Average number of executions per county = 0.0159. There are a huge number of counties, 2166, that had no executions in 1996, thus the small average number of executions per county.

(iii) The correlation between murders and executions is positive and equal to 0.2095 with a p-value of 0.0000. This means the correlation is highly significant.

(iv) Surely, more executions do not cause more murders. By definition, the more murders there are, the more executions there eventually will be. If anything, more murders cause, with substantial delay, more executions.

```
1   * Only use 1996 data
2   keep if year == 1996
3   * Getting a count of the number of counties
4   codebook countyid
5   * Give us the number of counties that had zero murders during 1996
6   codebook countyid if murders == 0
7   * Give us the number of counties that had zero executions during 1996
8   codebook countyid if execrate == 0
9   * Here we get the minimum and maximum number of murders across all of the
10  * states in 1996
11  summarize murders
12  * Herre we get the minimum and maximum number of executions across all of the
13  * states in 1996
14  summarize execs
15  * Calculate the Pearson Correlation Coefficient and its p-value between
16  * murders and executions
17  pwcorr murders execs, sig
18
```

(R)
Statistics/Data Analysis

(R)
Statistics/Data Analysis

14.0

StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC http://www.stata.com
979-696-4600 stata@stata.com
979-696-4601 (fax)

Notes:
1. Unicode is supported; see help unicode_advice .
2. New update available; type - update all -

```
1 . use "E:\E5350\E5350 f16\Exercises\countymurders.dta", clear

2 . do "C:\Users\00008904\AppData\Local\Temp\STD01000000.tmp"

3 . * Only use 1996 data
4 . keep if year == 1996
  (35,152 observations deleted)

5 . * Getting a count of the number of counties
6 . codebook countyid
```

countyid — **county identifier: 1000*statefips + countyf**

type:	numeric (long)		
range:	[1001,56045]	units:	1
unique values:	2,197	missing .:	0/2,197
mean:	32921.9		
std. dev:	15531.7		

percentiles:	10%	25%	50%	75%	90%
	8045	20105	36065	48049	51171

```
7 . * Give us the number of counties that had zero murders during 1996
8 . codebook countyid if murders == 0
```

countyid — **county identifier: 1000*statefips + countyf**

type:	numeric (long)		
range:	[1011,56045]	units:	1
unique values:	1,051	missing .:	0/1,051
mean:	31793.8		
std. dev:	14776.7		

percentiles:	10%	25%	50%	75%	90%
	16053	20003	30093	47085	51065

```
 9 . * Give us the number of counties that had zero executions during 1996
10 . codebook countyid if execrate == 0
```

countyid — **county identifier: 1000*statefips + countyf**

```
                  type:  numeric ( long)

                 range:  [ 1001,56045]                units:     1
           unique values:  2,166                      missing .:  0/2,166

                  mean:    32895.2
             std. dev:     15505.9

           percentiles:        10%       25%       50%       75%       90%
                              8047     20105     36060     48041     51169
```

```
11 . * Here we get the minimum and maximum number of murders across all of the
12 . * states in 1996
13 . summarize murders
```

Variable	Obs	Mean	Std. Dev.	Min	Max
murders	2,197	6.390077	39.76102	0	1403

```
14 . * Herre we get the minimum and maximum number of executions across all of the
15 . * states in 1996
16 . summarize execs
```

Variable	Obs	Mean	Std. Dev.	Min	Max
execs	2,197	.0159308	.14226	0	3

```
17 . * Calculate the Pearson Correlation Coefficient and its p-value between
18 . * murders and executions
19 . pwcorr murders execs, sig
```

	murders	execs
murders	1.0000	
execs	0.2095 0.0000	1.0000

```
20 .
end of do-file

21 .
```