Economics 1016
Summer 2020

Gregory A. Bruich, Ph.D.
Harvard University

Stata Hints for Empirical Project #1

Notes and commands that may be useful to you but are not necessarily required to answer the questions.

Set-up:

- Visit the software download page on the Harvard University Information Technology website (http://downloads.fas.harvard.edu/download). If the link does not work, try changing your internet browser.

  Download and install Stata to your computer. If you installed stata last year, you should un-install stata before installing the new version.

- Install the `binscatter` command by running the following:

  ```
  ssc install binscatter
  ```

  This is a command created by Michael Stepner for producing binned scatter graphs. The documentation including a helpful introductory slide deck is here: https://michaelstepner.com/binscatter/.

  You can also type `help binscatter` to see the syntax for the command.

- Open the "marcps_w.dta" data file. You can use the drop down menu: file -> open

- The variable year is the survey year, but the labor supply variables refer to the previous year.

  ```
  replace year = year -1
  ```

- Limit your analysis to observations on people ages 21-39, for example by using the command:

  ```
  keep if age>=21 & age<=39
  ```

- Create a new variable that is log(weekly earnings) by running the command:

  ```
  generate lnwkwage=log(wsal_val/wkswork)
  ```

Question 1:

Part (a) and (b)

Use `binscatter` to replicate the graph. The documentation on Michael Stepner's website may be helpful. You can also open the `binscatter` help file (or for any other command) by typing:

```
help binscatter
```

Question 2:

The following commands will create the variable `post92`, which takes the value '1' for observations after the ADA was implemented and '0' otherwise and the variable `disabl1_post92`, the interaction term:

```
generate post92 = (year >= 92)

generate disabl1_post92 =  disabl1*post92
```

When running the regressions, use the command "regress" with the option "`robust`" to report heterskedasticity robust standard errors. To regress a variable y on x1 and x2, the command would be:

```
regress y x1 x2, robust
```

Question 3:

Set-up:

Install `coefplot` by running the following:

```
ssc install coefplot
```

Part (a)

You can generate year indicator variables for years 1988-1996 as follows using this example for how to do it for the observations from 1988:

```
generate y88 = (year == 88)
```

You can generate interaction terms between disability status and year as follows using this example for how to do it for the observations from 1988:

```
generate disabl1_y88 = disabl1*y88
```

Part (b)

Here is how to produce the coefficient plot after you have run your regression:

```
coefplot, recast(connected) ciopts(recast(rline) lpattern(dash)) vertical ///
keep(disabl1_y*) xline(4.5) yline(0) ///
coeflabel(disabl1_y88 = "1988" disabl1_y89  = "1989" ///
disabl1_y90  = "1990" disabl1_y91  = "1991" disabl1_y92  = "1992" ///
disabl1_y93  = "1993" disabl1_y94  = "1994" disabl1_y95  = "1995" ///
disabl1_y96  = "1996") xtitle(Year) ///
ytitle("Coefficient and 95% CI") ///
title("Figure 3: Weeks Worked by Disabled vs. Non-disabled by Year")

graph export "Figure3.png", replace
```