Economics 1016
Summer 2020

Gregory A. Bruich, Ph.D.
Harvard University

Empirical Exercise #1
Differences-in-Differences Empirical Exercise
Due date: Friday, July 10 at 11:59pm

Introduction

In this assignment, you will implement a difference-in-differences (DD) approach to examine whether the Americans with Disabilities Act of 1990 (ADA) improved employment outcomes for the disabled. You will use the same data from the March CPS used by Acemoglu and Angrist (2001), called "marcps_w.dta."

The goal of this first assignment is to acquire some familiarity working with data in statistical software. You are welcome to use any software you would like. Links to download and installation instructions for Stata and R are below:

- For FAS students, Stata 16 is available from the software download page on the Harvard University Information Technology website. If the link does not work, try changing your internet browser.
- To use R, you should install both R and RStudio on your Computer. You can download R from https://cran.r-project.org and you can download RStudio Desktop from https://www.rstudio.com/products/rstudio/download/

The Americans with Disabilities Act of 1990 (ADA) requires employers to accommodate disabled workers and bans employment discrimination against the disabled. Specifically, the ADA requires employers to offer "reasonable accommodation" to disabled workers and bans discrimination against the disabled in wage determination, hiring, and firing. The ADA was signed into law in 1990 and came into effect in 1992. Although the ADA was meant to improve employment outcomes for disabled workers, our theoretical analysis of mandated benefits with minimum wages shows that the ADA could have unintended consequences.

You will focus on workers ages 21-39. Workers classified as having a disability (`disabl1` = 1) will form the *treatment group* and workers classified as not having a disability (`disabl1` = 0) will form the *control group* for your analysis. You will examine two employment outcomes: weeks worked and log(weekly earnings).

***Note**: Be sure to read the Stata hints before starting each problem.* Programming and statistical software is learned by working with data, not by reading about programming or watching others program. The most important thing as you are getting started is to remember that it is ok if you make a mistake or receive an error message! You will learn by figuring out the right code.

Table 1
Documentation of Selected Variables from marcps_w.dta

Variable	Definition
wkswork	Number of weeks worked last year (i.e., in year = survey year -1)
wsal_val	Wage and salary earnings last year (i.e., in year = survey year -1)
year	Year of survey
disabl1	Does the individual have a disability? 1 = yes, 0 = no

Gregory A. Bruich, Ph.D.
Harvard University

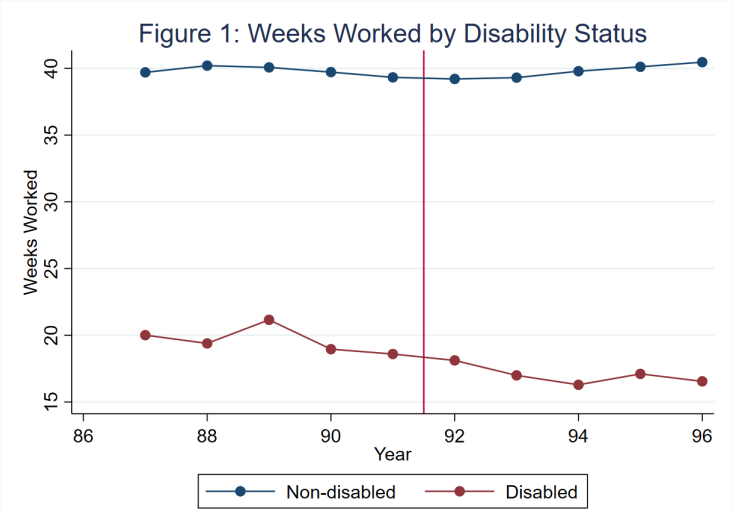
Instructions

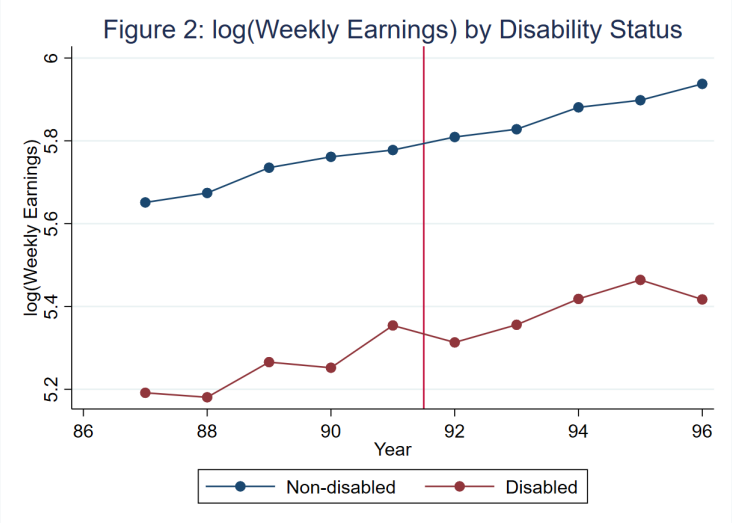
Please submit your Problem Set on Canvas. Your submission should include three files:

1. Your answers to the following questions as a pdf document.
2. A .do file, .R file, or .py/.ipynb file with your code
3. A .pdf of log file of your STATA, R, or Python output

Questions

1. *The parallel trends assumption.*
    a. State the parallel trends assumption in this empirical context, and explain how it could be violated.
    b. Replicate Figure 1 and Figure 2 (shown below).





c. Does it appear that the parallel trends assumption is violated for weeks worked? Does it appear that the parallel trends assumption is violated for log(earnings)? Explain.

2. *The basic difference-in-differences specification.*
   a. Run the following regression. Report and interpret the coefficient estimate for β_3 (the difference-in-differences estimator) and state whether it is statistically significant at the 5%-level. Specifically, what effect did the ADA have on weeks worked for the disabled?

$$wkswork_{it} = \beta_o + \beta_1 disabl1_i + \beta_2 post92_t + \beta_3 disabl1_i \times post92_t + \varepsilon_{it} \qquad (1)$$

   b. Run the following regression. Report and interpret the coefficient estimate of (the difference-in-differences estimator) and state whether it is statistically significant at the 5%-level. Specifically, what effect did the ADA have on weekly earnings for the disabled?

$$lnwkwage_{it} = \beta_o + \beta_1 disabl1_i + \beta_2 post92_t + \beta_3 disabl1_i \times post92_t + \varepsilon_{it} \qquad (2)$$

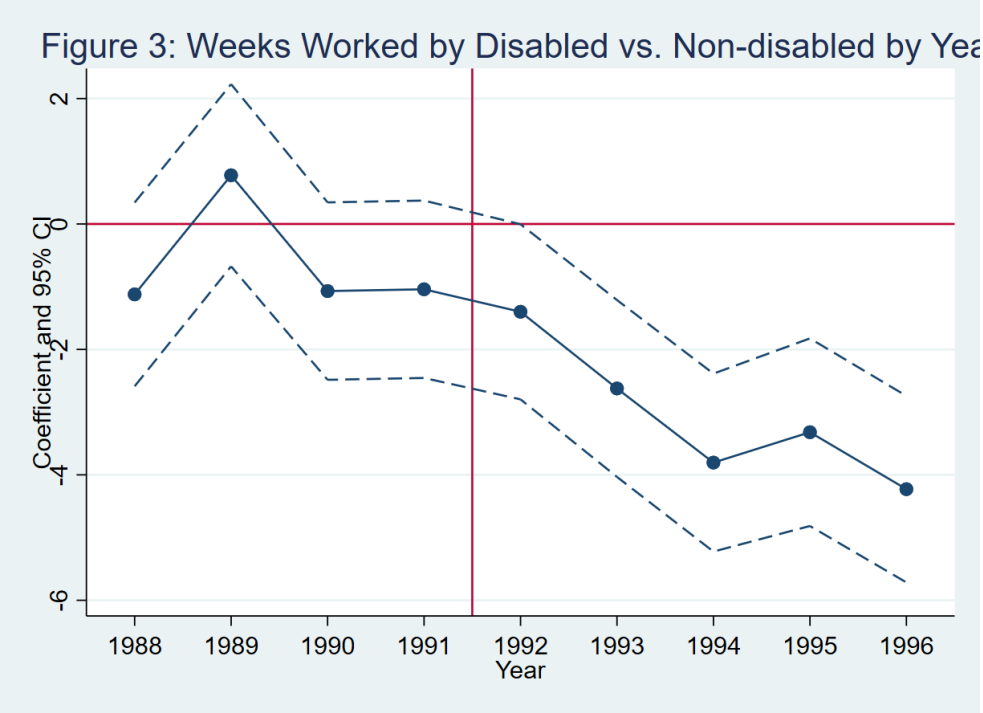
   c. Assume that the regressions you ran in parts (a) and (b) have uncovered the causal effects of the ADA. Do your findings suggest that the ADA was helpful or harmful to employment outcomes for the disabled? Explain.

3. *Non-parametric difference-in-differences specification with separate interaction terms for each year.*

   a. First, run a regression that takes equation (1) in question 2a and makes the following changes: 1) instead of including the variable *post92*, include a series of indicator variables (*y88*, *y89*, *y90* etc.) for the years 1988-1996, and 2) instead of including the variable $disabl1_i * post92_t$, include a series of interactions between disabled status and year (*disabl1_y88*, *disabl1_y89* etc.) for the years 1988-1996.

      Note: 1987 is now the "excluded year," so all the interaction term coefficient estimates represent the change in weeks worked for disabled people relative to non-disabled people in a particular year relative to 1987.

   b. Now, replicate Figure 3, which shows the coefficient estimates on the interaction terms (*disabl1_y88, disabl1_y89,* etc.*)* and their 95% confidence intervals.



c. Does there appear to be an immediate effect of the policy when it was implemented in 1992? What could explain what you observe?

d. As a general matter, how can one use this graph to help assess the parallel trends assumption? What is the advantage of this type of graph over the raw averages plotted in Figure 1 and 2?