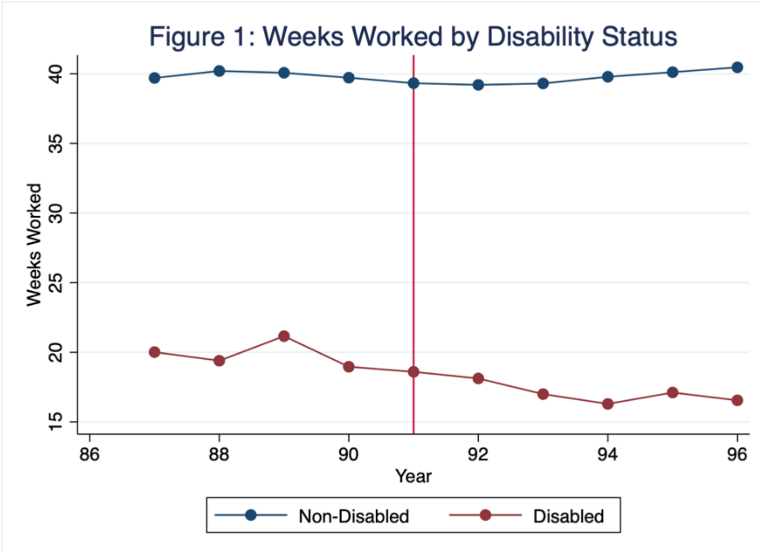
Name: Nabib Ahmed
Email: nahmed@college.harvard.edu
Subject: ECON1016 Assignment 1
Date: Wednesday July 1, 2020

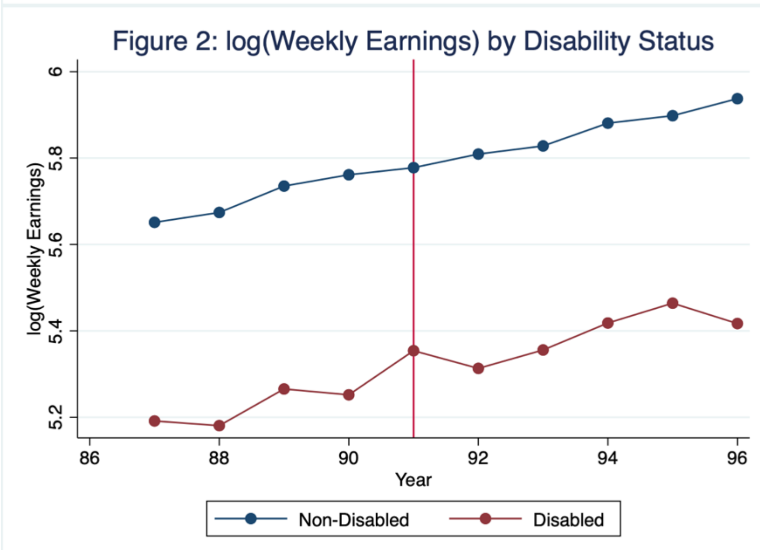
1.

a. In a Difference in Differences research design, we're comparing the difference in the treatment group vs the control group. In this case, the treatment group is workers classified as having a disability and the control group is workers classified as not having a disability.
   The main assumption for such a design is parallel trends, which is that the control and treatment groups 'grow/change' the same way (having similar trends) so that the difference between the two groups are constant over time. In this case, we'd need to assume that the difference in weeks worked and (log of) weekly earnings of workers with and without disabilities is not changing over time (so both are moving at the same rate and direction).
   Violations to the parallel trend assumption may arise due to innovations in assistive technologies. As new assistive technologies emerge, workers with disabilities are more positively impacted (hence more growth in terms of weeks worked and earnings) compared to those who don't have disabilities (they would not be affected as they wouldn't use the technology) causing different trends between the two groups.

b. Graphs made with binscatter on STATA. Code and graph below:

```
* Set Working Directory to the location of the data and qlr.ado:
* Easiest to use drop down menu: file --> change working directory
cd "/Users/nabibahmed/Desktop/ECON1016/Assignment 2"

*Start a log file
capture log close
log using assignment2_log.log, replace

* Import Data
use marcps_w.dta, clear

* Install binscatter command
ssc install binscatter

* Adjusting the variables to reflect work year, not survey year
replace year = year -1

* Limit analysis to age group 21-39
keep if age>=21 & age<=39

* New variable for log(weekly earnings)
generate lnwkwage=log(wsal_val/wkswork)


* Question 1b
* Figure 1
binscatter wkswork year if inrange(year, 86, 96), by(disabl1)
linetype(connect) title("Figure 1: Weeks Worked by Disability
Status") xtitle("Year") ytitle("Weeks Worked") legend(rows(1)
order(1 “Non-disabled” 2 “Disabled”) label(1 "Non-Disabled") label
(2 "Disabled")) xline(91)
graph export fig1.png, replace

* Figure 2
binscatter lnwkwage year if inrange(year, 86, 96), by(disabl1)
linetype(connect) title("Figure 2: log(Weekly Earnings) by
Disability Status") xtitle("Year") ytitle("log(Weekly Earnings)")
legend(rows(1) order(1 “Non-disabled” 2 “Disabled”) label(1
"Non-Disabled") label(2 "Disabled")) xline(91)
graph export fig2.png, replace
```





c. The parallel trends assumption seems to be met for weeks worked (figure 1) because the two groups have a relatively flat trend. Although we do see a slight negative trend for disabled workers, whereas for non-disabled, its flatter (so the difference seems to be getting wider), it is marginal. Similarly, the parallel trends assumption seems to be met for log(earnings) because both groups seem to be trending upwards at the same rate. Albeit, it does seem to change towards 1995, however that maybe because of other factors (like new policies or changes in the workforce). So parallel trends seem to be satisfied for both outcome variables since the two groups trend together.

2.

a. The regression result from STATA

```
. * Regression A
. regress wkswork disabl1 post92 disabl1_post92, robust

Linear regression                               Number of obs     =    410,572
                                                F(3, 410568)      =    5667.88
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0492
                                                Root MSE          =       19.3
```

wkswork	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
disabl1	-20.19257	.2324933	-86.85	0.000	-20.64825	-19.73689
post92	-.0656345	.0614244	-1.07	0.285	-.1860244	.0547554
disabl1_po~92	-2.49399	.329498	-7.57	0.000	-3.139796	-1.848184
_cons	39.80184	.0416529	955.56	0.000	39.7202	39.88347

The coefficient for β_3 (the difference-in-differences estimator) is -2.49. Now for a person with a disability (i.e. *disabl1* = 1), the model says:

$$wksworked_{it} = \beta_0 + \beta_1 + \beta_2 post92_t + \beta_3 post92_t + \varepsilon_{it}$$

And for a person without a disability (i.e. *disabl1* = 0), the model says:

$$wksworked_{it} = \beta_0 + \beta_2 post92_t + \varepsilon_{it}$$

The difference then between the two equations (weeks worked for person with disability minus weeks worked for person without disability) results in:

$$wks\widehat{work}ed_{it} = \beta_1 + \beta_3 post92_t$$

From this above equation, we see that when the ADA passed in 1992 (and $post92_t$ goes from 0 to 1), the difference in the two groups (disabled and non-disabled workers) changes by β_3 (from -20.19 to -22.68) . This means the ADA increased the gap between weeks worked by disabled and non-disabled workers by 2.49 weeks.

The coefficient has a p-value of 0.0, which means it is statistically significant at the 5% level (since the p-value is less than 0.05). Since non-disabled people continued working 40 hours, and the ADA seems to widen the gap between disabled and non-disabled workers, this effectively means that disabled workers work less weeks after ADA passed.

b. The regression results from STATA

```
. * Regression B
. regress lnwkwage disabl1 post92 disabl1_post92, robust

Linear regression                               Number of obs     =    333,406
                                                F(3, 333402)      =    1412.40
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0159
                                                Root MSE          =      .8277
```

lnwkwage	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
disabl1	-.4702201	.0162366	-28.96	0.000	-.5020434	-.4383969
post92	.147005	.0028853	50.95	0.000	.1413498	.1526602
disabl1_po~92	-.0091121	.0236234	-0.39	0.700	-.0554133	.0371891
_cons	5.720342	.001958	2921.59	0.000	5.716505	5.72418

The coefficient for β_3 (the difference-in-differences estimator) is -0.009. By the rationale from part a (taking the difference between groups), this means the ADA increased the wage gap between disabled and non-disabled workers by 0.09% (use percentage because we're working with log-linear regression).

The coefficient has a p-value of 0.7, which means it is not statistically significant at the 5% level (since the p-value is more than 0.05). Since the coefficient is not statistically significant, we can conclude the ADA did not have impact on the weekly earnings for the disabled.

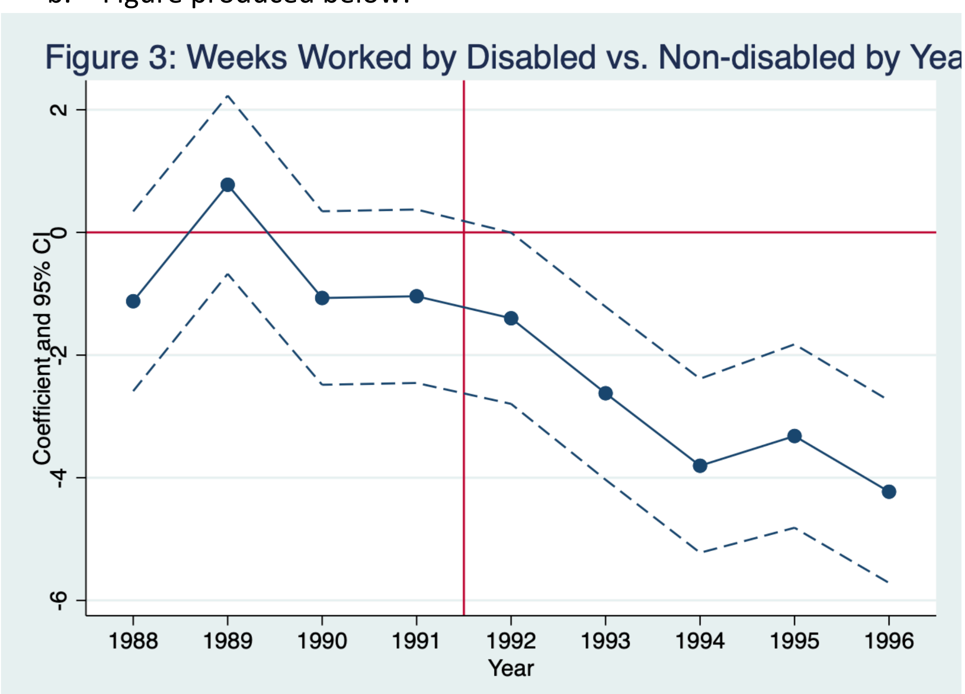
c. From the analysis in part (a) and (b), I have to conclude the ADA was harmful for the employment outcomes for the disabled. In both cases, the gap widens between the disabled and non-disabled labor force (less equality between the two groups), and the disabled worked less weeks after ADA passed.

3.

a. The regression results from STATA

```
. regress wkswork disabl1 y88 y89 y90 y91 y92 y93 y94 y95 y96 disabl1_y88 disabl1_y89 d
> isabl1_y90 disabl1_y91 disabl1_y92 disabl1_y93 disabl1_y94 disabl1_y95 disabl1_y96, r
> obust

Linear regression                               Number of obs     =    410,572
                                                F(19, 410552)     =     907.84
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0497
                                                Root MSE          =     19.296

------------------------------------------------------------------------------
             |               Robust
     wkswork |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     disabl1 |  -19.69036   .5152861   -38.21   0.000     -20.7003   -18.68041
         y88 |   .5023846    .132676     3.79   0.000     .2423436    .7624256
         y89 |   .3703711   .1301568     2.85   0.004     .1152678    .6254745
         y90 |   .0180991    .131201     0.14   0.890    -.2390508    .2752491
         y91 |  -.3746421   .1322421    -2.83   0.005    -.6338327   -.1154515
         y92 |  -.4946092   .1339358    -3.69   0.000    -.7571194    -.232099
         y93 |  -.3940256   .1358811    -2.90   0.004    -.6603484   -.1277028
         y94 |   .0851376   .1355252     0.63   0.530    -.1804877    .3507628
         y95 |   .4158822   .1410101     2.95   0.003     .1395067    .6922577
         y96 |   .7642089    .140064     5.46   0.000     .4896878     1.03873
 disabl1_y88 |  -1.122271   .7474665    -1.50   0.133    -2.587283    .3427406
 disabl1_y89 |   .7766638   .7411266     1.05   0.295    -.6759219    2.229249
 disabl1_y90 |  -1.069676   .7216558    -1.48   0.138      -2.4841    .3447474
 disabl1_y91 |  -1.040961   .7216694    -1.44   0.149    -2.455412    .3734888
 disabl1_y92 |   -1.39992   .7121371    -1.97   0.049    -2.795687   -.0041524
 disabl1_y93 |  -2.622487   .7194216    -3.65   0.000    -4.032532   -1.212443
 disabl1_y94 |  -3.804497   .7236397    -5.26   0.000    -5.222809   -2.386185
 disabl1_y95 |  -3.319968   .7631029    -4.35   0.000    -4.815627    -1.82431
 disabl1_y96 |  -4.227817   .7585819    -5.57   0.000    -5.714614   -2.741019
       _cons |   39.70219   .0928464   427.61   0.000     39.52021    39.88416
------------------------------------------------------------------------------
```

b. Figure produced below:



c. Before the policy in 1992, we see the trend for differences between the two groups is flat, as from 1990 to 1991, there appeared to be no change in the difference of weeks worked by the two groups. However, after ADA passed in 1992, we see a negative trend in the difference in the weeks worked by disabled vs. non-disabled workers, meaning that disabled workers worked less weeks. A potential reason for this is that ADA may have increased the income effect for disabled worked. With the ADA banning discrimination against disabled workers, disabled workers may have gotten access to higher paying jobs, thus becoming richer, and then via the income effect, working less. Thus, the above graph could be the result of this income effect boost from the ADA (this would make sense with figure 2, which shows a positive trend in *log(Weekly Earnings)* for disabled workers).

d. This type of graph is helpful in assessing the parallel trends assumption because it directly measures the differences in the outcome variable between the two groups. In the Figure 1 and 2, we first need to find the difference of each group from before and after the event/time, and then compare those differences. Whereas in figure 3, we have the differences between the two groups already, and can take one difference using the before and after event values. We can check for the parallel trends assumption by seeing if the line stays at constant value while in the before-event period and while in the after-event period (because it would imply the two groups trend together since their differences don't change in either period).