

ECO 348K, Problem Set #6, due 4/22/22. Please submit answers + Stata code + .log files on Canvas.

Q1. Open the data BraceroData.dta that I posted on Canvas (note: BraceroData_s13.dta can be opened in Stata 13 and is posted under the [data] folder). This data is from the paper, “Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion” (posted on Canvas as BraceroPaper.pdf) and modified slightly for ease-of-use with this problem set.

The *bracero* program was a series of bilateral agreements between the US and Mexico that permitted temporary immigration of agricultural workers from Mexico to the US in the 1940s-1960s. At its peak, these agreements were responsible for the supply of half a million seasonal workers each year to US farms, with work contracts for six weeks to six months. This program was restricted in 1962 and fully terminated on December 31, 1964 with the stated goal of “improving wages and domestic employment for seasonal farm workers in the US.” In other words, the people who advocated for terminating this program did so because they believed that by restricting the immigration of Mexican workers in this sector, the farms in the US would increase their demand for US workers and increase wages for those workers.

However, because farms can substitute between labor and capital, it is not clear what the effect of this policy will be. When farms can no longer employ immigrant workers through the *bracero* program, will they employ more domestic workers (and increase wages) or will they increase their use of technology so that the same work can be done with fewer workers? Using this data and a difference-in-differences research design, you can test whether ending the *bracero* program increased employment of domestic seasonal workers and increased wages. In the data, I have already defined **treated** = 1 as states that were heavily reliant on *bracero* workers prior to the ending of the program and **treated** = 0 as states that were not heavily reliant on *bracero* workers prior to the ending of the program.

Note: for all regressions in this problem set, cluster your standard errors at the state level.

a. Given the description above, you know that the *bracero* was restricted in 1962 and fully terminated in December 31, 1964. Define a variable named **post** that is equal to one in any year equal to or after the program is first restricted and equal to zero in any year before the program is first restricted. Fill in the difference-in-differences 2 x 2 table below with the appropriate means for **yearly_mex_frac** (you can round the number to the thousandths decimal if you want). Then, using those numbers, provide an estimate of the effect of restricting and eventually terminating the *bracero* program (we will call this “ending the *bracero* program” from here on) on **yearly_mex_frac** in the treated states.

mean(yearly_mex_frac)	Treated = 0	Treated = 1
Post = 0		
Post = 1		

b. Estimate the effect of ending the *bracero* program on **yearly_mex_frac** in the treated states by running a regression, do not include state or year fixed effects in this regression. Report the regression here and compare the estimated effect from that regression to the estimate from Q1a that you derived from the 2 x 2 comparison.

c. Use the [binscatter] command (which can be installed with *ssc install binscatter*) with the option [discrete] to plot the mean of **yearly_mex_frac** for each year. Use [help binscatter] to figure out how to connect each point with a line and to figure out how to make sure there is one set of points plotted for treated states and one set of points plotted for non-treated states. Use the [xline()] option to plot vertical lines at years 1962 and 1965 since those were two events that restricted and then fully terminated the *bracero* program. Insert the final plot here.

d. Estimate a difference-in-differences event study design with **yearly_mex_frac**. Include state and year fixed effects, using 1961 as the omitted year. To interact the treated indicator with year fixed effects and omit 1961, include the following variables in your regression model [ib1961.year##treated]. Report the full regression here.

Use the following code to make an event study figure and include it in your answers. You may need to install the command [coefplot] by running [ssc install coefplot]. Note: the “///” at the end of each line will only work as line breaks if you run the code from the .do file program editor window. If you are running code directly from the command window, then you will need to enter the code all as one line.

```
coefplot ,vertical keep(*.year#1.treated) omit base rename(*.year#1.treated="") ///  
    xlabel(angle(45)) nolabels color(black) plotregion(color(white)) graphregion(color(white)) ///  
    ytitle("Mexican Fraction of Seasonal Workers") xtitle("Year") xline(8 11, lpattern(dash)) ///  
    level(95) ylabel(-0.6(.2)0.6) yscale(r(-.6,.6))
```

e. Based on your results in Q1c and Q1d, discuss the parallel trends assumption in this setting and whether it is likely to hold. Specifically, evaluate whether the coefficients are statistically different from zero for many periods before 1962.

f. Repeat questions (a)-(e) for the following outcomes: **realwage_hourly** and **domestic_seasonal**. Change the labels and scaling of the figures as needed. Based on this evidence on the outcomes **realwage_hourly** and **domestic_seasonal**, did ending the *bracero* program and preventing the immigration of agricultural workers achieve the stated goal of, “improving wages and domestic employment for seasonal farm workers in the US”?

Repeat the regression part of question (b) for: **tomatoes_total** and **asparagus_total**. Next, assume that tomatoes were a crop that could be easily switched to a mechanized production process at the time but that asparagus were a crop where labor could not be easily replaced with machines. Based on this understanding of these crops and the evidence on all of the four outcomes listed above, is it more likely that farms replaced immigrant labor with domestic labor in response to the end of the *bracero* program or that they replaced immigrant labor with capital?

g. Open the dataset titled **Tomatoes.dta**. This dataset contains data on the fraction of tomato farming that was mechanized in each year in California and Ohio. California was a state that lost a lot of seasonal workers when the *bracero* program ended and Ohio is a state that did not lose many workers when the *bracero* program ended. Use the [binscatter] command to plot the fraction of tomato farming that was mechanized in each year, separately for California and Ohio (the command should end up being similar to the one you used in question Q1c). How does this evidence compare with the results from part Q1f?

Stata guide:

These are the Stata commands & functions that I used to get the answers for all of the questions above. Although this is how I did it, you can use any commands you see fit unless stated otherwise in the problem.

gen – used to generate a new variable

sum x if y==z – used to summarize the variable x when the variable y is equal to the value z.

reg y a##b, cluster(c) – used to run a regression of y on binary variables a, b, and the interaction between a and b, using standard errors clustered on the variable c.

ssc install binscatter – used to install the binscatter command

binscatter x y, discrete – used to make a binned scatter plot of x and y, plotting the average value of x at each discrete value of y.

areg y x, a(z) – used to run a regression of y on x controlling for fixed effects of z.

ssc install coefplot – used to install the coefplot command

coefplot - used to execute the *coefplot* command (see code that I gave you above).