

# ECON 753 Homework 3

Jesús Lara Jáuregui

10/24/2020

## Part 1: Exercises from Berndt, chapter 5

### Exercise 1

#### Part (a)

In this exercise work with the Current Population Survey for the years 1978 and 1985, that shows demographic, social and economic information of 550 and 534 individuals respectively.

In this part I calculate the arithmetic and geometric mean of the the wage rate and annual wages for 1978. The results are shown in Table 1 below.

Table 1: Geometric and Arithmetic Mean for 1978 Wages

|        | Geometric Mean | Arithmetic Mean |
|--------|----------------|-----------------|
| Hourly | 5.37           | 6.06            |
| Annual | 10741.87       | 12125.53        |

Table 2: Mean and Standard Deviation of Education and Experience

|            | Mean  | Standard Deviation |
|------------|-------|--------------------|
| Education  | 12.54 | 2.77               |
| Experience | 18.72 | 13.35              |

We have data for the log of the wage rate, so to calculate the arithmetic mean we exponentiate that variable and then get the average of that new variable. To get the mean of the anual wage we just multiply that number by 2000, which are the assumed work hours of a year. In contrast, the geometric mean is calculated by getting the average of the log wage and then exponentiating it. Hence, it is no surprise that the two results are different. The arithmetic mean of the wage rate and annual wages is greater than the geometric mean because the latter reduces large values more than it increases small values (of the wage-rate). Given that logarithm and exponentiation are not linear transformations, the arithmetic and geometric means will, in general, be different.

In Table 2 we show mean and standard deviation of education and experience. As we would expect, the standard deviation of education is much smaller (2.8) than that of experience (13.4). The former means that we would expect to have a very large share of the population within the range 9.5-15.5 of years of education. This is so, because, in general, almost all workers have a minimum years of education which is greater than zero. In contrast, in the labor market there is always a greater dispersion of experience, which is captured by the standard deviation of 13.4

Table 4: Years of Education and Mean Wages by Race

|               | Education |      | lnWage      |      |             |
|---------------|-----------|------|-------------|------|-------------|
|               | Mean      | SD   | Mean Hourly | SD   | Mean Annual |
| <b>Gender</b> |           |      |             |      |             |
| Male          | 12.40     | 3.05 | 6.13        | 3.42 | 12256.64    |
| Female        | 12.76     | 2.22 | 4.32        | 2.50 | 8632.72     |
| <b>Race</b>   |           |      |             |      |             |
| Whites        | 12.81     | 2.49 | 5.55        | 3.38 | 11100.34    |
| NonWhites     | 11.72     | 3.44 | 4.54        | 2.55 | 9084.33     |
| Hispanic      | 10.31     | 3.66 | 4.62        | 1.96 | 9233.10     |

**Part (b)**

In this part we get the frequencies of nonwhites, females and hispanics. The results are shown in Table 3. The share of each category is given by the arithmetic mean, whereas the total count is given by the mean times the sample size or, equivalently, by the sum over each of these dummy variables.

Table 3: Means and Number of Individuals in Demographic Groups

|       | Mean | Count |
|-------|------|-------|
| nonwh | 0.10 | 57    |
| hisp  | 0.07 | 36    |
| fe    | 0.38 | 207   |

Our results show that 10% of the individuals in the sample are nonwhite, 7% hispanic and 38% female.

**Part (c)**

In this part we subdivide the sample by gender and race and analyze the mean and standard deviation of years of education and the wage rate. The results are in Table 4

Regarding education and gender, men and women have on average almost the same years of education, although the standard deviation for men is greater. Regarding race, the results show that whites in the subsample have on average one more year of education than nonwhites, and the latter have, on average, one more year of education than hispanics.

Regarding the wage rate, the average in the sample is greater for men than for women and is greater for whites than for nonwhites and hispanics. The same is true for the mean annual wages.

**Part (d)**

In part d we repeat the comparisons of shares, means and standard deviations of wage and education for 1985. With respect to shares, the results are in Table 5

There we can see that share of nonwhites in the sample increased and that of hispanics decreased, although the changes are small. In contrast, the share of females in the sample increased from 38% in 1978 to 46% in 1985

In Table 6 we show the results for Education and Real Wages, incorporating the percentage change that took place from 1978 to 1985. Real wages were calculated with a price index which is equal to one for 1978 and 1.649 for 1985

Table 5: Comparison of demographics between 1978 and 1985

|       | 1978 |       | 1985 |       |
|-------|------|-------|------|-------|
|       | Mean | Count | Mean | Count |
| nonwh | 0.10 | 57    | 0.13 | 67    |
| hisp  | 0.07 | 36    | 0.05 | 27    |
| fe    | 0.38 | 207   | 0.46 | 245   |

Table 6: Comparison of education and real wages between 1978 and 1985

|               | Education |      |       |      |         | lnWage      |      |             |      |         |
|---------------|-----------|------|-------|------|---------|-------------|------|-------------|------|---------|
|               | 1978      |      | 1985  |      | %Change | 1978        |      | 1985        |      | %Change |
|               | Mean      | SD   | Mean  | SD   |         | Mean Hourly | SD   | Mean Hourly | SD   |         |
| <b>Gender</b> |           |      |       |      |         |             |      |             |      |         |
| Male          | 12.40     | 3.05 | 13.01 | 2.77 | 0.05    | 6.13        | 3.42 | 5.29        | 3.21 | -0.14   |
| Female        | 12.76     | 2.22 | 13.02 | 2.43 | 0.02    | 4.32        | 2.50 | 4.19        | 2.86 | -0.03   |
| <b>Race</b>   |           |      |       |      |         |             |      |             |      |         |
| Whites        | 12.81     | 2.49 | 13.17 | 2.48 | 0.03    | 5.55        | 3.38 | 4.89        | 3.17 | -0.12   |
| NonWhites     | 11.72     | 3.44 | 12.64 | 2.60 | 0.08    | 4.54        | 2.55 | 4.33        | 2.47 | -0.05   |
| Hispanic      | 10.31     | 3.66 | 11.52 | 4.05 | 0.12    | 4.62        | 1.96 | 3.74        | 3.35 | -0.19   |
| <b>Total</b>  |           |      |       |      |         |             |      |             |      |         |
| Whole Sample  | 12.54     | 2.77 | 13.02 | 2.62 | 0.04    | 5.37        | 3.26 | 4.75        | 3.12 | -0.11   |

Just by looking at the signs, we can see the main trends regarding these two variables: the average years of education increased for all groups, whereas the average wage rate fell also for all groups. For the whole sample, the average years of education increased by 4% and the real wage rate fell by 11%.

However, this trends, although general, are not exactly the same for all groups. The average years of education increased more for men than for women. However, the average real wage rate for men fell by 14% whereas for women the decline was 3%. Regarding race, Hispanics had the greatest increase in the average years of education. However, they also experienced the largest decline in real wage (19%).

### Part (e)

In this part we want to analyze the distribution of wages and log wages. In particular, we want to test whether they are normally distributed. To do that, we follow two approaches. In the first one, we calculate the proportion of observations that are within 6 different groups defined in terms of the sample mean and sample standard deviation. We compare them with the proportions of a normal distribution for these six groups. The second approach, more formal, consists of chi-squared goodness-of-fit tests for normal distribution.

From Table 7 we can observe that the proportions of log wages match almost perfectly with those of the normal distributions, whereas the proportions of Wage for each group are substantially different. Now we conduct the chi-squared test. The null-hypothesis  $H_0$  is that wages and log wages are normally distributed. The results of the test are shown below.

As we would expect, for lnwage the p-value is 0.59 so we failed to reject  $H_0$ . In contrast, the p-value for wage is close to zero, so we reject the null-hypothesis  $H_0$  at the 1% significance that wages are normally distributed.

Table 7: Counts and Proportions of lnwage and wage Distributions

|              | lnwage |            | wage   |            | Normal Distribution<br>Proportions |
|--------------|--------|------------|--------|------------|------------------------------------|
|              | Counts | Proportion | Counts | Proportion |                                    |
| wi<=w-2s     | 9      | 0.02       | 0      | 0.00       | 0.02                               |
| w-2s<wi<=w-s | 80     | 0.15       | 45     | 0.08       | 0.14                               |
| w-s<wi<=w    | 194    | 0.35       | 278    | 0.51       | 0.34                               |
| w<wi<=w+s    | 193    | 0.35       | 164    | 0.30       | 0.34                               |
| w+s<wi<=w+2s | 63     | 0.11       | 45     | 0.08       | 0.14                               |
| w+2s<wi      | 11     | 0.02       | 18     | 0.03       | 0.02                               |

Table 8: Chi-Squared Goodness-of-Fit Tests for Normal Distribution

|        | Statistic | pValue |
|--------|-----------|--------|
| lnwage | 3.74      | 0.59   |
| wage   | 84.99     | 0.00   |

## Problem 6

Part (a)

Part (b)

Part c

Part d

## Part 2: Replication of Card & Krueger (1994)

In this part I replicate the first three rows of Table 3 of Card & Krueger(1994), in which they present the basic Diff-in-Diff results of the effect of the increase in minimum wage in New Jersey in employment in the fast-food sector. I present the results of the three panels in three different tables. The variable analyzed is Full-Time Equivalent employment (FTE), which, at the store level is defined by the equation:

$$FTE_i = FullTimeEmployees_i + Managers_i + \frac{1}{2}PartTimeEmployees_i$$

I create that variable and then group the data by state and period (after en before the treatment) to build Panel 1 of Table 3

The standard errors are shown in parenthesis. This table shows the main results of the study, where can be seen that:

- (1) Before the treatment stores in PA were, on average, bigger in terms of FTE than those of New Jersey
- (2) That difference became smaller after the treatment
- (3) FTE fell in PA after the treatment, while that of NJ remained constant or slightly increased
- (4) Most importantly, the diff-in-diff estimate, which under the Conditional Independence Assumption (or parallel trends) is a estimation of the causal effect of minimum wage on FTE, is non-negative (2.75 in our results)

In this part, I am getting different standard errors in the first two cells of bottom-row (that compare the mean within each state). In all this exercise I obtained the standard errors of the difference in means by conducting

Table 9: Simple Regression of log Wages to Determine Wage Discrimination

|                         | <i>Dependent variable:</i>  |
|-------------------------|-----------------------------|
|                         | lnwage                      |
| Constant                | 0.488***<br>(0.098)         |
| fe                      | -0.306***<br>(0.034)        |
| union                   | 0.207***<br>(0.037)         |
| nonwh                   | -0.157***<br>(0.055)        |
| hisp                    | -0.027<br>(0.069)           |
| ed                      | 0.075***<br>(0.007)         |
| ex                      | 0.026***<br>(0.005)         |
| exsq                    | -0.0003***<br>(0.0001)      |
| Observations            | 550                         |
| R <sup>2</sup>          | 0.392                       |
| Adjusted R <sup>2</sup> | 0.385                       |
| Residual Std. Error     | 0.385 (df = 542)            |
| F Statistic             | 50.012*** (df = 7; 542)     |
| <i>Note:</i>            | *p<0.1; **p<0.05; ***p<0.01 |

Table 10: Means and Number of Individuals in Demographic Groups

|       | Mean | Count |
|-------|------|-------|
| other | 0.83 | 457   |
| nonwh | 0.10 | 57    |
| hisp  | 0.07 | 36    |

Table 11: Comparison of Demographics Variables

|           | Education |       | Experience |        | Female |        | Unionized |        | Wages  |        |
|-----------|-----------|-------|------------|--------|--------|--------|-----------|--------|--------|--------|
|           | ed        | diff  | ex         | diff.1 | fe     | diff.2 | union     | diff.3 | lnwage | diff.4 |
| Others    | 12.81     | 0.00  | 18.11      | 0.00   | 0.37   | 0.00   | 0.30      | 0.00   | 1.71   | 0.00   |
| NonWhites | 11.72     | -1.09 | 22.16      | 4.05   | 0.49   | 0.13   | 0.39      | 0.08   | 1.51   | -0.20  |
| Hispanic  | 10.31     | -2.51 | 21.06      | 2.95   | 0.33   | -0.03  | 0.19      | -0.11  | 1.53   | -0.18  |

Table 12: Regressions of lnwage by Racial Category

|                         | <i>Dependent variable:</i> |                            |                       |
|-------------------------|----------------------------|----------------------------|-----------------------|
|                         | Others<br>(1)              | lnwage<br>NonWhites<br>(2) | Hispanics<br>(3)      |
| Constant                | 0.456***<br>(0.111)        | 0.755*<br>(0.389)          | 0.682***<br>(0.234)   |
| fe                      | -0.338***<br>(0.037)       | -0.059<br>(0.144)          | -0.235**<br>(0.104)   |
| union                   | 0.197***<br>(0.040)        | 0.205<br>(0.139)           | 0.417***<br>(0.123)   |
| ed                      | 0.077***<br>(0.008)        | 0.054**<br>(0.026)         | 0.055***<br>(0.015)   |
| ex                      | 0.028***<br>(0.005)        | 0.003<br>(0.020)           | 0.021<br>(0.014)      |
| exsq                    | -0.0003***<br>(0.0001)     | -0.00001<br>(0.0004)       | -0.0003<br>(0.0003)   |
| Observations            | 457                        | 57                         | 36                    |
| R <sup>2</sup>          | 0.420                      | 0.149                      | 0.489                 |
| Adjusted R <sup>2</sup> | 0.414                      | 0.066                      | 0.404                 |
| Residual Std. Error     | 0.376 (df = 451)           | 0.484 (df = 51)            | 0.287 (df = 30)       |
| F Statistic             | 65.423*** (df = 5; 451)    | 1.789 (df = 5; 51)         | 5.752*** (df = 5; 30) |

*Note:*

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table 13: RSS comparison

| Sample   | RSS   |
|----------|-------|
| Complete | 81.35 |
| White    | 63.83 |
| Hisp     | 2.46  |
| Other    | 11.94 |

Table 14: Chow test information

| DF 1 | DF 2 | F statistic | P-Value |
|------|------|-------------|---------|
| 5    | 65   | 0.52        | 0.24    |

Table 15: Replication of Rows 1-3 of Table 1 of Card and Krueger

| Variable           | Stores By State |                 |                 | Stores in NJ    |                 |                 | Differences in NJ |                 |
|--------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-------------------|-----------------|
|                    | PA              | NJ              | NJ-PA           | Wage=4.25       | Wage=4.26-4.99  | Wage>4.99       | Low-high          | Midrange-high   |
| FTE before         | 23.33<br>(1.35) | 20.44<br>(0.51) | -2.89<br>(1.44) | 19.56<br>(0.77) | 20.08<br>(0.84) | 22.25<br>(1.14) | -2.69<br>(1.37)   | -2.17<br>(1.41) |
| FTE after          | 21.17<br>(0.94) | 21.03<br>(0.52) | -0.14<br>(1.08) | 20.88<br>(1.01) | 20.96<br>(0.76) | 20.21<br>(1.03) | 0.66<br>(1.44)    | 0.74<br>(1.27)  |
| Change in mean FTE | -2.17<br>(1.65) | 0.59<br>(0.73)  | 2.75<br>(1.34)  | 1.32<br>(1.13)  | 0.87<br>(1.53)  | -2.04<br>(1.27) | 3.36<br>(1.3)     | 2.91<br>(1.22)  |

a t-test (where the null hypothesis is that the means are equal) and then extracting the associated standard errors. The difference may come from the way t.test manages missing values. I conclude that because I get the correct results for the standard errors of FTE for each group and also in mean comparisons between states.

Then I group NJ data by wage category before the treatment: Low (Minimum wage=\$4.25), Midrange (\$4.26-\$4.99), and High (>\$4.99). Below is the replication of Panel 2:

In the panel we can see that, in general, stores with a high starting wage are bigger in terms of FTE than the rest before the treatment. However, this changes after the increase in minimum wage since stores with a high starting wage before the treatment are the only ones whose average FTE falls after the increase in minimum wage.

Again, my results are different in the standard errors that compare the difference in the means within state and wage category over time. I finally present the results of Panel 3, that compares FTE between wage categories and over time

There, in the bottom row, we can see that the difference between Low wage and High wage stores

### Part 3: Advances in replication

The data consists of four main variables: trade flows, tariffs, rules origin and input-output (IO) information. As I said in problem set 3, IO is used for robustness check and I will make that if I have time. I gathered all the data and am in the cleaning process. Up to this point, I successfully cleaned the data on trade flows, which consists of the imports of Mexico from NAFTA partners and non-NAFTA countries. Below I show a replication of panels 1 and 2 of Table 2 of Conconi et al. (2018)

I am almost done with the data of tariffs and rules of origin. Once I have them I will be able to replicate Tables 1 and 2 of descriptive statistics. With that, the replication of the main regressions should be very straightforward.

Table 16: Trade: Panels 1 and two of Table 2

| HS Code and Description           | Mexican imports non-NAFTA |            | Mexican imports NAFTA |             |
|-----------------------------------|---------------------------|------------|-----------------------|-------------|
|                                   | 1991                      | 2003       | 1991                  | 2003        |
| 01-05: animal products            | 105015.80                 | 396708.2   | 974826.06             | 2300023.24  |
| 06-15: vegetables                 | 163746.87                 | 245054.6   | 1470589.97            | 4893730.51  |
| 16-24: foodstuffs                 | 81746.57                  | 133052.3   | 630233.92             | 2241186.71  |
| 25-27: mineral products           | 122199.61                 | 718771.1   | 1295175.35            | 4932581.91  |
| 28-38: chemicals                  | 166600.88                 | 1194191.4  | 1844428.29            | 7416817.92  |
| 39-40: plastics/rubber            | 164092.39                 | 1365552.9  | 1364731.88            | 11502199.59 |
| 41-43: raw hides, skins, leathers | 22956.68                  | 222964.6   | 183434.99             | 796171.69   |
| 44-49: wood products              | 39232.65                  | 359856.6   | 1291545.32            | 4276956.29  |
| 50-63: textiles                   | 325965.40                 | 1468048.9  | 662750.51             | 6778142.79  |
| 64-67: footwear/headgear          | 82539.92                  | 260928.9   | 72100.04              | 60278.68    |
| 68-71: stone/glass                | 39363.39                  | 525586.3   | 244871.21             | 1492556.97  |
| 72-84: metals                     | 192869.11                 | 1585282.5  | 2045470.12            | 9408639.86  |
| 84-85: machinery/electrical       | 1224048.48                | 21999538.6 | 6540453.92            | 36770636.94 |
| 86-89: transportation             | 135931.48                 | 1444667.6  | 1230502.75            | 11583126.23 |
| 90-97: miscellaneous              | 324508.53                 | 1839537.2  | 1361375.09            | 5104762.24  |