

Problem Set 4

May 9, 2021

Rules

- Solutions to problem sets must be completed and submitted individually.
- The deadline for submission is **23.59 on 16 May 2020**. Submissions after the deadline will not be accepted.
- The datasets for the STATA exercises can be found at SUCourse.
- **Submissions must include two files: 1) one word file including your answers 2) one STATA log file.** Results in the log file and interpreted results must match. Otherwise, you will get zero points for those exercises. **Do not upload zipped files.**
- You are expected to answer all questions very clearly. Therefore, please be careful about use of language and writing. **For instance; if you are asked to interpret a coefficient, you need to interpret it in a way that someone who cannot see the data and does not know econometrics can understand what you mean.**
- Double check that you write your name/surname and student ID number.
- Failure in fulfilling any of these will result in a FAIL grade for that homework.

Question 1- Minimum Wages and Employment (60 points)

In this problem set, you are expected to replicate some of the main results of an influential paper by David Card and Alan Krueger:

David Card and Alan Krueger. *"Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania."* American Economic Review 84 (September 1994). (<http://davidcard.berkeley.edu/papers/njmin-aer.pdf>)

Check this youtube video for the discussion on minimum wage: <https://youtu.be/mWwXmH-n5Bo?t=318>

The data set used in the paper is at SUCourse+ under Problem set 4: "Problem set 4 > card_krueger_data.dta".

Some Tips for replication:

- First you need to generate full time employment variables (as described in the paper) for both time periods.
- In Table 3, in order to obtain the standard errors for the 3rd column, you can use "ttest" command in STATA. Check the help file for "ttest" to see how to use it.
- In Table 4, note that the dependent variable is the difference in full time employment in both periods. Therefore, you need to generate a new variable indicating that difference.
- In Table 4 and Table 7, use if condition at the end of your regression commands in order to limit the sample as described in the paper (note that the authors are not using the observations for which the wage information is missing).
- In Table 7, first you need to generate the price variables (as described in the paper) indicating the prices in both periods. These price variables are used to calculate the dependent variable in Table 7. Note that, in Table 7, the authors use the sample for which the wage and fte information is not missing.

1) What is the main research question of the paper? Discuss the importance (policy implications) of this research question. Briefly describe the data and the methodology used to answer this research question? (10 points)

2) Using the data (card_krueger_data.dta), calculate the **following statistics** which are reported in Table 3 in the paper (Read the related text and the table notes carefully in order understand how the authors obtained these statistics. Hint: In order to calculate the standard error for the mean differences, you can use "ttest" command in STATA.). Discuss what these statistics indicate. Based on these statistics, what is the estimated impact of minimum wage increase on full time employment in NJ? (15 points)

TABLE 3—AVERAGE EMPLOYMENT PER
IN NEW JERSEY

Variable	Stores by state		
	PA (i)	NJ (ii)	Difference, NJ – PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	– 2.89 (1.44)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	– 0.14 (1.07)

3) Using the data, replicate the **following statistics** which are reported in Table 4 in the paper (Read the related text and the table notes carefully in order understand how the authors obtained these statistics. There might be small differences between your results and the paper. It is acceptable if you obtain coefficient estimates of around 2.28 instead of 2.33 and 2.30.). Write the empirical model that is estimated in column (ii). Discuss the meaning of the coefficient of "New Jersey dummy" in column (ii). What is the main advantage of specification in column (ii), as compared to the simple calculations in question (2). (20 points)

TABLE 4—REDUCED-FORM MODELS FOR CHA

Independent variable	(i)	(ii)
1. New Jersey dummy	2.33 (1.19)	2.30 (1.20)
2. Initial wage gap ^a	—	—
3. Controls for chain and ownership ^b	no	yes
4. Controls for region ^c	no	no
5. Standard error of regression	8.79	8.78
6. Probability value for controls ^d	—	0.34

4) Replicate the **following statistics** which are reported in Table 7 in the paper (Read the related text and the table notes carefully in order understand how the authors obtained these statistics.). What is the research question in this analysis? Discuss the meaning of the coefficient of "New Jersey dummy" in column (ii). (15 points)

TABLE 7—REDUCED-FORM MODELS FOR CHANGE 1

Independent variable	Dependent variable	
	(i)	(ii)
1. New Jersey dummy	0.033 (0.014)	0.037 (0.014)
2. Initial wage gap ^a	—	—
3. Controls for chain and ^b ownership	no	yes
4. Controls for region ^c	no	no
5. Standard error of regression	0.101	0.097

Question 2 - Impact of Traffic Laws (15 points)

Use the data in TRAFFIC2.DTA for this exercise.

1) Run an OLS regression of *prcfat* on a linear time trend, monthly dummy variables, and the variables *wkends*, *unem*, *spdlaw*, and *beltlaw*. Discuss your findings (Interpret the size and significance of the estimated coefficients. While discussing the results for month dummies, mention only the min. and max. coefficients of month dummies.)

Question 3 - Policy Analysis (25 points)

Use the data in INJURY.DTA for this exercise.

1) Using the data for Kentucky, reestimate equation (13.12) in the book, adding as explanatory variables male, married, and a full set of industry and injury type dummy variables. Interpret the estimated coefficient for "*afchnge*" , "*highearn*" and "*afchnge * highearn*" variables. What do your results imply

about the impact of policy?(10 points)

2) What do you think about the small R-squared from part (1)? Does this mean the equation is useless? (5 points)

3) Estimate equation (13.12) using the data for Michigan. Compare the estimates on the interaction term for Michigan and Kentucky. Is the Michigan estimate statistically significant? How do you explain the different results (in terms of significance) for Kentucky and Michigan? (10 points)