ECON 523: Program Evaluation for International Development

Replication Project: Card and Krueger (1994)

Due March 15

Overview of the Assignment

This exercise makes use of the data set `Card-Krueger-1994-data.dta`, which was first used in the paper "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania" by David Card and Alan Krueger, published in the *American Economic Review* in 1994. The authors examine the impacts of a change in New Jersey's minimum wage law that took effect in April of 1992.

For this assignment, you will produce a replication and extension of Card and Krueger's analysis. Specifically, you will be replicating Table 2 as well as portions of Tables 3 and 4, as discussed below. In addition, you will need to produce one table of additional analysis — either a replication of a specification (or multiple specifications) that appears later in the paper, or new analysis of your own devising. More original and/or challenging extensions will receive more credit.

When you submit this assignment, you will be asked to upload: (1) a write-up of your analysis that should be at most 5–8 pages long, including tables, and (2) a do file that produces your results from the raw data.

Honor code related text. This is an independent project. You are welcome (in fact, encouraged) to discuss your approach with other students in the class. You are welcome to discuss which variables you want to drop, which specifications you plan to include, etc. However, you should not share you Stata code with anyone, and both your write-up and your do file should represent your own independent work. You may be asked to talk me through your do file in order to receive full credit on this assignment. If any part of your written work is identical to that of another student or to any document available on the internet, you will receive a zero on the assignment.

This assignment is hard. Replicating published papers is often substantially harder than one would expect. Dealing with the data will be frustrating. You **will not** be able to replicate every coefficient and standard error exactly (at least I can't). Struggling to replicate published papers is good practice for unraveling the mysteries in your own data. I'll be evaluating your efforts to match the published coefficients, but I do not expect anyone to successfully replicate everything. That is why this is a project (i.e. preparation for real life, research-wise) and not a problem set.

Specific Guidance on What to Replicate, and How to Do It

You must replicate the following pieces of Card and Krueger's paper:

1. Table 2 – all of the coefficients and hypothesis tests contained therein

2. Part of Table 3, specifically the contents of the red box (note that you may be able to replicate every coefficient in the red box, but you will not be able to replicate every standard error):

TABLE 3—AVERAGE EMPLOYMENT PER STORE BEFORE AND AFTER THE RISE IN NEW JERSEY MINIMUM WAGE

Variable	Stores by state			Stores in New Jersey[a]			Differences within NJ[b]	
	PA (i)	NJ (ii)	Difference, NJ − PA (iii)	Wage = $4.25 (iv)	Wage = $4.26–$4.99 (v)	Wage ≥ $5.00 (vi)	Low– high (vii)	Midrange– high (viii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	− 2.89 (1.44)	19.56 (0.77)	20.08 (0.84)	22.25 (1.14)	− 2.69 (1.37)	− 2.17 (1.41)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	− 0.14 (1.07)	20.88 (1.01)	20.96 (0.76)	20.21 (1.03)	0.67 (1.44)	0.75 (1.27)
3. Change in mean FTE employment	− 2.16 (1.25)	0.59 (0.54)	2.76 (1.36)	1.32 (0.95)	0.87 (0.84)	− 2.04 (1.14)	3.36 (1.48)	2.91 (1.41)
4. Change in mean FTE employment, balanced sample of stores[c]	− 2.28 (1.25)	0.47 (0.48)	2.75 (1.34)	1.21 (0.82)	0.71 (0.69)	− 2.16 (1.01)	3.36 (1.30)	2.87 (1.22)
5. Change in mean FTE employment, setting FTE at temporarily closed stores to 0[d]	− 2.28 (1.25)	0.23 (0.49)	2.51 (1.35)	0.90 (0.87)	0.49 (0.69)	− 2.39 (1.02)	3.29 (1.34)	2.88 (1.23)

Notes: Standard errors are shown in parentheses. The sample consists of all stores with available data on employment. FTE (full-time-equivalent) employment counts each part-time worker as half a full-time worker. Employment at six closed stores is set to zero. Employment at four temporarily closed stores is treated as missing.
[a]Stores in New Jersey were classified by whether starting wage in wave 1 equals $4.25 per hour ($N = 101$), is between $4.26 and $4.99 per hour ($N = 140$), or is $5.00 per hour or higher ($N = 73$).
[b]Difference in employment between low-wage ($4.25 per hour) and high-wage (≥ $5.00 per hour) stores; and difference in employment between midrange ($4.26–$4.99 per hour) and high-wage stores.
[c]Subset of stores with available employment data in wave 1 and wave 2.
[d]In this row only, wave-2 employment at four temporarily closed stores is set to 0. Employment changes are based on the subset of stores with available employment data in wave 1 and wave 2.

3. Part of Table 4, specifically the contents of the red box (note that you will probably not be able to replicate every coefficient and standard error exactly, but I am interested to see how close you get):

	Model				
Independent variable	(i)	(ii)	(iii)	(iv)	(v)
1. New Jersey dummy	2.33 (1.19)	2.30 (1.20)	—	—	—
2. Initial wage gap[a]	—	—	15.65 (6.08)	14.92 (6.21)	11.91 (7.39)
3. Controls for chain and ownership[b]	no	yes	no	yes	yes
4. Controls for region[c]	no	no	no	no	yes
5. Standard error of regression	8.79	8.78	8.76	8.76	8.75
6. Probability value for controls[d]	—	0.34	—	0.44	0.40

TABLE 4—REDUCED-FORM MODELS FOR CHANGE IN EMPLOYMENT

Notes: Standard errors are given in parentheses. The sample consists of 357 stores with available data on employment and starting wages in waves 1 and 2. The dependent variable in all models is change in FTE employment. The mean and standard deviation of the dependent variable are -0.237 and 8.825, respectively. All models include an unrestricted constant (not reported).

[a]Proportional increase in starting wage necessary to raise starting wage to new minimum rate. For stores in Pennsylvania the wage gap is 0.

[b]Three dummy variables for chain type and whether or not the store is company-owned are included.

[c]Dummy variables for two regions of New Jersey and two regions of eastern Pennsylvania are included.

[d]Probability value of joint F test for exclusion of all control variables.

4. An additional analysis table of your choosing, either a subsequent specification (or specifications) from the paper, or additional analysis of outcome variables, or alternative specifications exploring the robustness of Card and Krueger's results

Structuring Your Write-Up

You should structure your write-up following the paper template posted on the course website.[1] Your write-up should include the following:

- Your **introduction** should explain why empirical research on the impact of minimum wage laws is important, and then provide an overview of Card and Krueger's empirical approach and their main findings. In the second or third paragraph of the introduction, you should explain which parts of Card and Krueger's analysis you will be replicating. Though your introduction will reflect the ideas in Card and Krueger's original paper, the writing must be your own.

- Your **data** section can be short. Your first paragraph should describe the context of Card and Krueger's data collection in your own words, and the second section should explain what data you actually use in your analysis.

- Your **result**s section should contain five paragraphs:

  - A paragraph describing the empirical strategy and the assumptions required for valid causal inference.

  - A paragraph describing the results of your replication of Table 2 (again, in your own words). You should explain what statistics are reported and how we should interpret the results.

---

[1]You are welcome to use latex instead of word to format your write-up, but you should adhere to the structure of the paper template.

– A paragraph describing the results of your replication of Table 3 (again, in your own words). You should explain what statistics are reported and how we should interpret the results.

– A paragraph describing the results of your replication of Table 4 (again, in your own words). You should explain what statistics are reported and how we should interpret the results.

– A paragraph describing your additional piece of replication: what did you choose to do, why is this an interesting or important extension of the analysis, and what are your findings?

• Your **conclusion** should briefly address the following questions: What have we learned from this exercise? How does the minimum wage impact employment outcomes in the fast food industry? Why is this result important?

Saving Your Empirical Results

You will need to include summary statistics and regression tables in your write-up. They need to be formatted professionally. As we discussed in class, you can do this using Stata's `putexcel` command to format your own bespoke tables, or you can do this using Stata's `esttab` command (together with `estpost` and `eststo`) to make professional-quality tables in excel or word with minimal effort. Detailed guidance on how to use `esttab` to format high-quality, easily replicable tables is available on the Williams ECON department's stata help website. You are also welcome to use `outreg2` or any other command of your choosing. However, to receive full credit, your replication files must generate your analysis tables as a word/rtf, excel/csv, or latex document.