- Fastfood.dta and do

ECONOMETRIA APLICADA AVANZADA

Diferencias en Diferencias

Junio 10-11, 2023

Cristina Tello-Trillo, Census Bureau & University of Maryland

Lecture 2 - Introduccion

INTRODUCCION

Bienvenidos!

- Objetivo:

- En esta clase vamos a aprender uno de los métodos mas importantes para el análisis empírico: Diferencias-en-Diferencias

Como vamos a aprender?

- La clase de hoy y manana sera divida entre teoria/conceptos y aplicaciones empiricas del ‘mundo real’.

- Todos deberian tener su laptop con STATA.

Conociéndonos

- Economista empirical aplicada
- Trabajo en el Centro de Estudios Economicos en el U.S. Census Bureau – Ministerio de Comercio
- Profesora Adjunta en la University of Maryland.
- Pre-grado en la PUCP
- PhD en Economia en Yale University
- Trabajo en temas de comercio internacional y mercado de trabajo:
  – Como el comercio con China afecta los salarios, quienes son los mas afectados.
  – Como politicas 'favorable a la familia' puede afectar a las decisiones laborales de la firma.
  – Determinantes de la productividad de la firma.

Evaluacion DID

- Control de lectura
- Problem set DID (Fecha de entrega: Viernes Junio 23 6pm)

This Class

- Natural Experiments
- Difference-in-Differences
- Applications
- Class exercise

Lecture 2 – Natural Experiments & DD

NATURAL (QUASI) EXPERIMENTS

Why use Natural (Quasi) Experiments?

- To solve for the endogeneity problem.

OLS assumptions

- **A1** Linearity in parameters: $Y_i = X_i\ \beta + \epsilon_i$
- **A2** No multi-collinearity
- **A3** Exogeneity of the independent variables
  - $E[\epsilon_i|X] = 0$ -> $E[\epsilon_i]$ -> $\mathrm{E}[\epsilon' X] = 0$
  - No information on X tell us something about the expected value of the ϵ
  - The causal interpretation of the coefficients are valid.
- **A4** Random sampling of observations
- **A5** Spherical errors: Homoscedasticity & no autocorrelation:
  - $Var[\epsilon_i|X] = \sigma^2$
  - $Cov[\epsilon_i, \epsilon_j|X] = 0$
- **A6 (optional):** Error terms should be normally distributed.

Exogeneity

- $Y = X\beta + e$
- The exogeneity assumption $E(e|X) = 0$ which implies $corr(e, X) = 0$ means that the regressor X is exogenous.

- A violation of this condition leads to an **endogeneity problem** which precludes our ability to make causal inferences (OLS is not BLUE).

$$E[\beta^{OLS}|X] = \beta + (X'X)^{-1}X'E[\epsilon|X]$$

Reasons for Endogeneity

- **Omitted Variables**
- **Measurement error**
- **Simultaneity bias**

Example (omitted variables):

- $\ln(w) = \beta_1 educ_i + \beta_2 X_i + \delta U_i + \epsilon_i$
  - X: observed factors that determine earnings, as work experience, gender, family background
  - U : unobserved factors that determine earnings, as individual ability
- $\ln(w) = \beta_1 educ_i + \beta_2 X_i + \underbrace{\mu_i}_{\delta U_i + \epsilon_i}$
  - A person level of education is (at least partially) determined by a person's ability
  - High ability individuals do better in school and therefore choose to attain a higher level of education, and their high ability is the fundamental reason for their high wages.
  - β_1^{OLS} wil not identify the causal effect of one additional year on education on earnings

OLS Assumptions: Exogeneity

Violation of this assp: Endogeneity; Causal effect not properly identify

Solution to Endogeneity: Instrumental Variables, instrument for X

Solution to Endogeneity: Panel Data Fixed Effects, include fixed effects dummies

Solutions to Endogeneity: Differences-in-Differences (changing the framework of analysis)

- Another way to think about **endogeneity** is that it indicates that the regressor X is **non-random/non-independent.**
  – If a regressor X was completely random it shouldn't be correlated with the error term.
  – Something that is **random** is not cause by anything and has not predictability of correlation.

Randomized experiments

- In many of the “hard” sciences, the researcher can simply design experiment to achieve the necessary randomness.
  - E.g. To determine the effect of a new vaccine, you randomly give it to certain patients.
  - E.g To determine the effect of modifying certain gene, you modify it in a random sample of mice.

But, it’s hard to do this in Economics

- Ethical, financial and practical reasons.
  - E.g. we can’t randomly assign education to a random group of people.
- Therefore, we need to rely on what we call “Natural (or quasi) Experiments”
  - **Natural experiments is when some event causes a random assignment of (or change in) a variable of interest X.**
    - **Policy change, gov. randomization, climate event (e.g. earthquake, wildfires)**
  - **Quasi experiment: is when the intervention its on a target population:**
    - **An eligibility cutoff mark (students above a threshold in a exam receive a grant).**

The Impact of the Mariel Boatlift on the Miami Labor Market

Example:

- David Card (1990) use a natural-experiment in which a larger number of Cuban immigrants (125k) entered the Miami, FL labor market in the "Mariel boatlift". Increasing the Miami labor force by 7%.
- This resulted from a temporary lifting of restrictions on emigration from Cuba in 1980.
- Half of the immigrants settled in Miami.
- Card estimate the causal effect on wages of an exogenous increase in labor supply (caused by immigration) by comparing the change in wages of low-skilled workers in Miami to the change in wages of similar workers in other comparable U.S. cities over the same period.
- He conclude that the influx of immigrants had a negligible effect on wages of low skilled workers.

See current academic debate about the impact of immigration on labor market:

- https://www.nber.org/reporter/2011number3/labor-market-effects-immigrants

Natural (Quasi) Experiment

- We can use such “natural/quasi” experiments to ensure randomness of the regressor X and make causal inference.
  - We use the randomness introduced into X by the natural experiment (climate disasters, sudden supply shock, independent change in policy) to uncover the causal effect of X on Y.

Natural Experiments

- Natural experiments can be used in many ways:
  - Use them to construct IV (quarter of birth, proximity to college, seat-belt laws)
  - Use the to construct regression discontinuity (Policy evaluation class: cutoff for home loans at credit score 620 and above)
  - But admittedly, when most people refer to a natural experiment, they are talking about DID (Difference-in-Differences) regression.

Lecture 2 – Natural Experiments & DD

DIFFERENCES-IN-DIFFERENCES

Difference-in-Differences

- DID basically compares outcome Y for the "treated" group to the outcome Y for the "untreated" group where treatment is **randomly** assigned by the natural experiment.

Difference-in-Differences

Like the Mariel boatlift or wawa wasi

- Let's think about a simple evaluation policy.
- If we have data on a bunch of people right before and right after a policy is enacted we can try to identify the effect of a policy.
- Suppose we have two years of data, y0 and y1, and that the policy is enacted in between.
- We could try to identify the effect by simply looking at before and after the policy

$$\bar{Y}_1 - \bar{Y}_0$$

Difference-in-Differences

$$\overline{Y_1} - \overline{Y_0}$$

- The problem is that this attributes any changes in time to the effect policy.
- Suppose that something else happened at time $\tau \in \{0,1\}$ other than just the program.
- We will attribute whatever that is (i.e increasing trend) to the program.

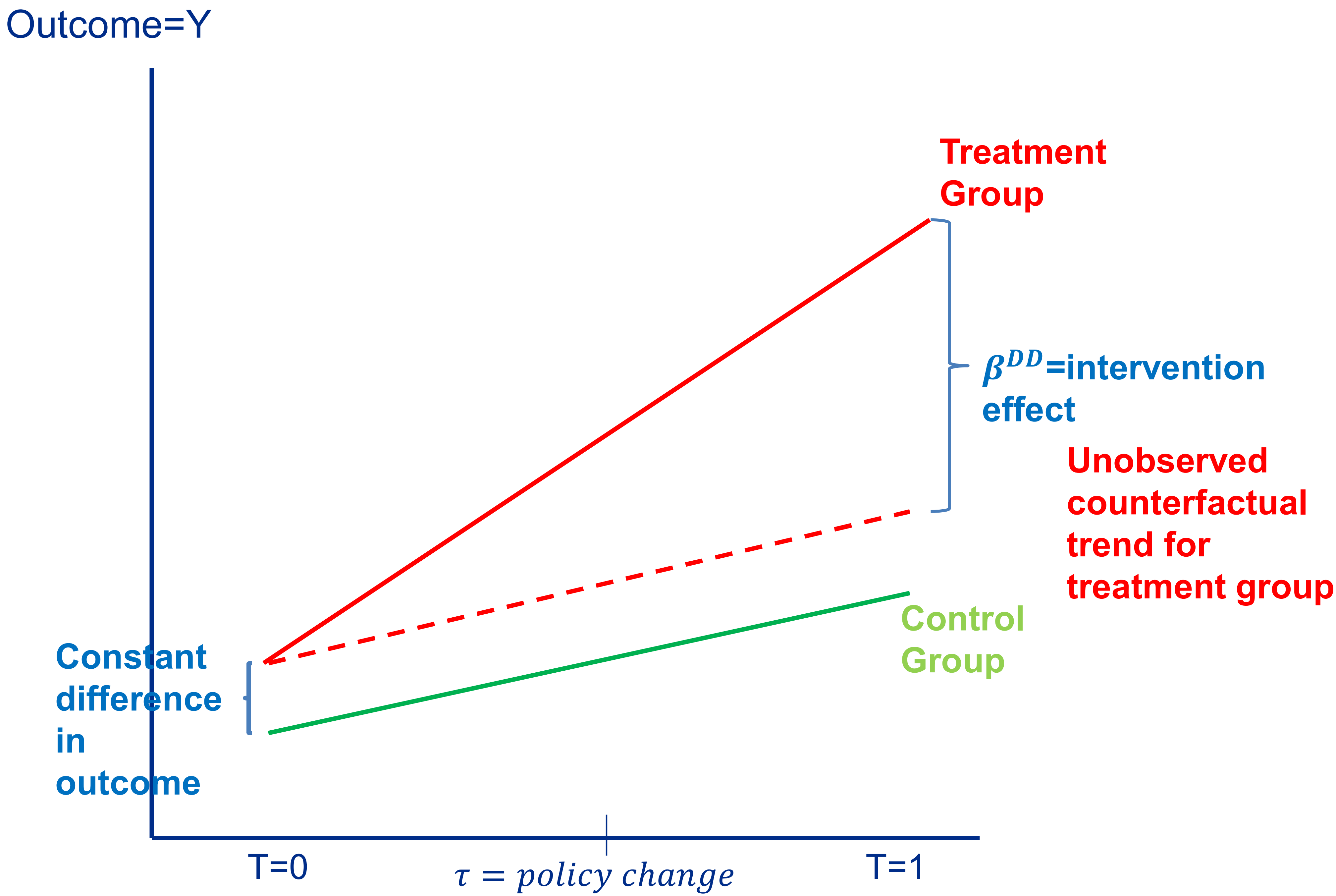
Difference-in-Differences

- To solve for this problem, suppose we have two groups:
  - People who are affected by the policy change (treated)
  - People who are not affected by the policy change (control)
- We can use the control groups to pick up the time changes:

$$\overline{Y_1{}^c} - \overline{Y_0{}^c}$$

- Then we can estimate our policy effect as a

$$\widehat{\beta^{DD}} = (\overline{Y_1{}^t} - \overline{Y_0{}^t}) - (\overline{Y_1{}^c} - \overline{Y_0{}^c})$$


Outcome=Y
Treatment Group
β^{DD}=intervention effect
Unobserved counterfactual trend for treatment group
Control Group
Constant difference in outcome
T=0
$\tau = policy\ change$
T=1

Differences-in-Differences

- DID is usually implemented as an interaction term between time and treatment group dummy variables in a regression model:
- $Y = \beta_0 + \beta_1 Post + \beta_2 Treatment + \beta^{DD} Post * Treatment + \beta_4 X + e$
- Post=1 for the second period (after the intervention)
- Treatment=1 for the treatment group
- X = control variables

Differences-in-Differences

- $Y = \beta_0 + \beta_1 Post + \beta_2 Treatment + \beta^{DD} Post * Treatment + e$

E[Y]	T=0	T=1	Difference
Treatment Group	$\beta_0 + \beta_2$	$\beta_0 + \beta_1 + \beta_2 + \beta^{DD}$	$\beta_1 + \beta^{DD}$
Control Group	β_0	$\beta_0 + \beta_1$	β_1
	Difference in Differences=		β^{DD}

$$\widehat{\beta^{DD}} = \left(\overline{Y_1{}^t} - \overline{Y_0{}^t}\right) - \left(\overline{Y_1{}^c} - \overline{Y_0{}^c}\right)$$

DID Example

- Suppose you are interested in the effect of minimum wages on employment (a classic and controversial question in labor economics).
- In a competitive labor market, increases in the minimum wage would move us up a downward-sloping labor demand curve.
  – Employment would fall.

DID Example: Card & Krueger (1994)

- Card & Krueger (1994) analyze the effect of a minimum wage increase in New Jersey using a differences-in-differences methodology.
- In April 1992 NJ increased the state minimum wage from \$4.25 to \$5.05. Pennsylvania's minimum wage stayed at \$4.25.

- They surveyed about 410 fast food stores (Burger King, Roy Rogers, Wendy's and KFC) both in NJ and in PA both before (Feb) and after (Nov) the minimum wage increase in NJ.
- The survey included questions on employment, starting wages, prices, and other store characteristics (company-owned)

Question:

- Why fast food restaurants?
  - Fast food restaurants are the leading employer of low-wage (min-wage) workers.
  - Job requirements in fast food restaurants are homogenous: easier to obtain reliable measures of employment, wages and product prices.
  - No tips

TABLE 2—MEANS OF KEY VARIABLES

Variable	Stores in: NJ	PA	t^{a}
1. *Distribution of Store Types (percentages):*			
a. Burger King	41.1	44.3	−0.5
b. KFC	20.5	15.2	1.2
c. Roy Rogers	24.8	21.5	0.6
d. Wendy's	13.6	19.0	−1.1
e. Company-owned	34.1	35.4	−0.2
2. *Means in Wave 1:*			
a. FTE employment	20.4	23.3	−2.0
	(0.51)	(1.35)	
b. Percentage full-time employees	32.8	35.0	−0.7
	(1.3)	(2.7)	
c. Starting wage	4.61	4.63	−0.4
	(0.02)	(0.04)	
d. Wage = $4.25 (percentage)	30.5	32.9	−0.4
	(2.5)	(5.3)	
e. Price of full meal	3.35	3.04	4.0
	(0.04)	(0.07)	
f. Hours open (weekday)	14.4	14.5	−0.3
	(0.2)	(0.3)	
g. Recruiting bonus	23.6	29.1	−1.0
	(2.3)	(5.1)	

DID Example: Card & Krueger (1994)

- They write the following specification:

$$Y_{ist} = \alpha + NJ_s + \lambda d_t + \delta(NJ_s * d_t) + e_{ist}$$

- Y_{ist} is the observed employment at restaurant i, in state s, at time t
- NJ_s is a dummy=1 if the obs is from NJ (high min wage state)
- d_t is a dummy=1 if the obs is from Nov (post policy change)
- Diff-in-diff estimate:
- (Y_NJ post – Y_NJ pre) – (Y_PA post – Y_PA pre) = δ

DID Strategy

- The differences-in-differences strategy amounts to comparing the change in employment in NJ to the change in employment in PA.
- The population differences-in-differences are:
- $E[Y_{ist} \,|s = NJ, t = Nov] - E[Y_{ist} \,|s = NJ, t = Feb]$
$-E[Y_{ist} \,|s = PA, t = Nov] - E[Y_{ist} \,|s = PA, t = Feb] = \delta$

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. Change in mean FTE employment	−2.16 (1.25)	0.59 (0.54)	2.76 (1.36)

- Surprisingly, employment rose in NJ relative to PA after the minimum wage change

DID Example

- We can also estimate δ

$$Y_{ist} = \alpha + NJ_s + \lambda d_t + \delta(NJ_s * d_t) + e_{ist}$$

By doing first differences:

$$Y_{ist=1} = \alpha + NJ_s + \lambda + \delta(NJ_s) + e_{ist} \text{ (Nov)}$$

$$Y_{ist=0} = \alpha + NJ_s + e_{ist} \text{ (Feb)}$$

$$\Delta Y_{is} = \lambda + \delta(NJ_s) + \Delta e_{ist}$$

sum dfte_NJ dfte_PA DID

Variable	Obs	Mean	Std. Dev.	Min	Max
dfte_NJ	392	.6993666	0	.6993666	.6993666
dfte_PA	392	-2.743421	0	-2.743421	-2.743421
DID	392	3.442787	0	3.442787	3.442787

. reg cfte state

Source	SS	df	MS
Model	726.164819	1	726.164819
Residual	41348.6866	390	106.022273
Total	42074.8514	391	107.608316

Number of obs	=	392
F(1, 390)	=	6.85
Prob > F	=	0.0092
R-squared	=	0.0173
Adj R-squared	=	0.0147
Root MSE	=	10.297

cfte	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
state	3.442788	1.315501	2.62	0.009	.8564268	6.029149
_cons	-2.743421	1.181114	-2.32	0.021	-5.065568	-.4212741

DID Example

Advantages of regression:

- It is easy to calculate standard errors.
- We can control for other variables which may reduce the residual variance (lead to smaller standard errors).
- It is easy to include multiple periods.
- We can study treatments with different treatment intensity. (e.g. varying increases in the minimum wage for different firms)

```
.   esttab reg1 reg2, se title("Replication of Card-Krueger Paper")
```

Replication of Card-Krueger Paper

	(1) cfte	(2) cfte
state	3.443**	3.429**
	(1.316)	(1.319)
co_owned		1.065
		(1.229)
1.chain		0
		(.)
2.chain		0.129
		(1.453)
3.chain		-2.668
		(1.473)
4.chain		-1.585
		(1.628)
_cons	-2.743*	-2.260
	(1.181)	(1.335)
N	392	392

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

TABLE 4—REDUCED-FORM MODELS FOR CHANGE IN EMPLOYMENT

Independent variable	Model (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

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).

Intensity of the treatment

$$
\begin{aligned}
\mathrm{GAP}_i &= 0 \quad \text{for stores in Pennsylvania} \\
&= 0 \quad \text{for stores in New Jersey with } W_{1i} \geq \$5.05 \\
&= (5.05 - W_{1i}) / W_{1i} \quad \text{for other stores in New Jersey.}
\end{aligned}
$$

- The value of GAP, is a strong predictor of the actual proportional wage change between waves 1 and 2 or the increase in wages at store i necessary to meet the new minimum wage rate.
- The mean value of GAP among New Jersey stores is 0.10. Thus the estimate in column (2) implies a 1.39 increase in FTE employment in New Jersey relative to Pennsylvania.

$$\overline{\Delta Y_{is}} = 13.94(\overline{GAP})$$

	(1)	(2)
	cfte	cfte
gap	15.64*	14.22*
	(6.788)	(6.924)
co_owned		1.112
		(1.234)
1.chain		0
		(.)
2.chain		0.279
		(1.456)
3.chain		-2.319
		(1.480)
4.chain		-1.145
		(1.656)
_cons	-1.233	-0.838
	(0.757)	(1.045)
N	392	392

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

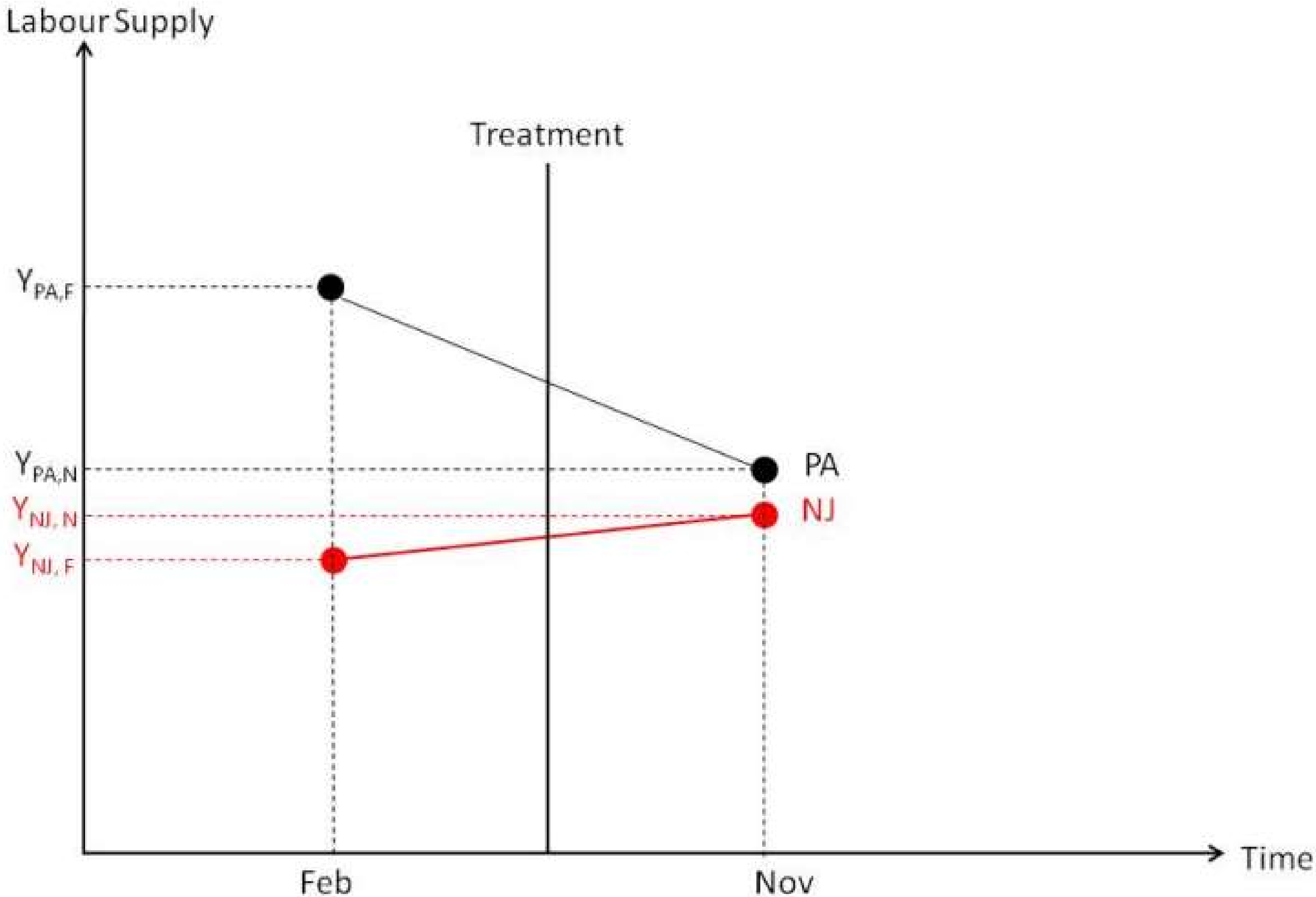
TABLE 4—REDUCED-FORM MODELS FOR CHANGE IN EMPLOYMENT

Independent variable	Model (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

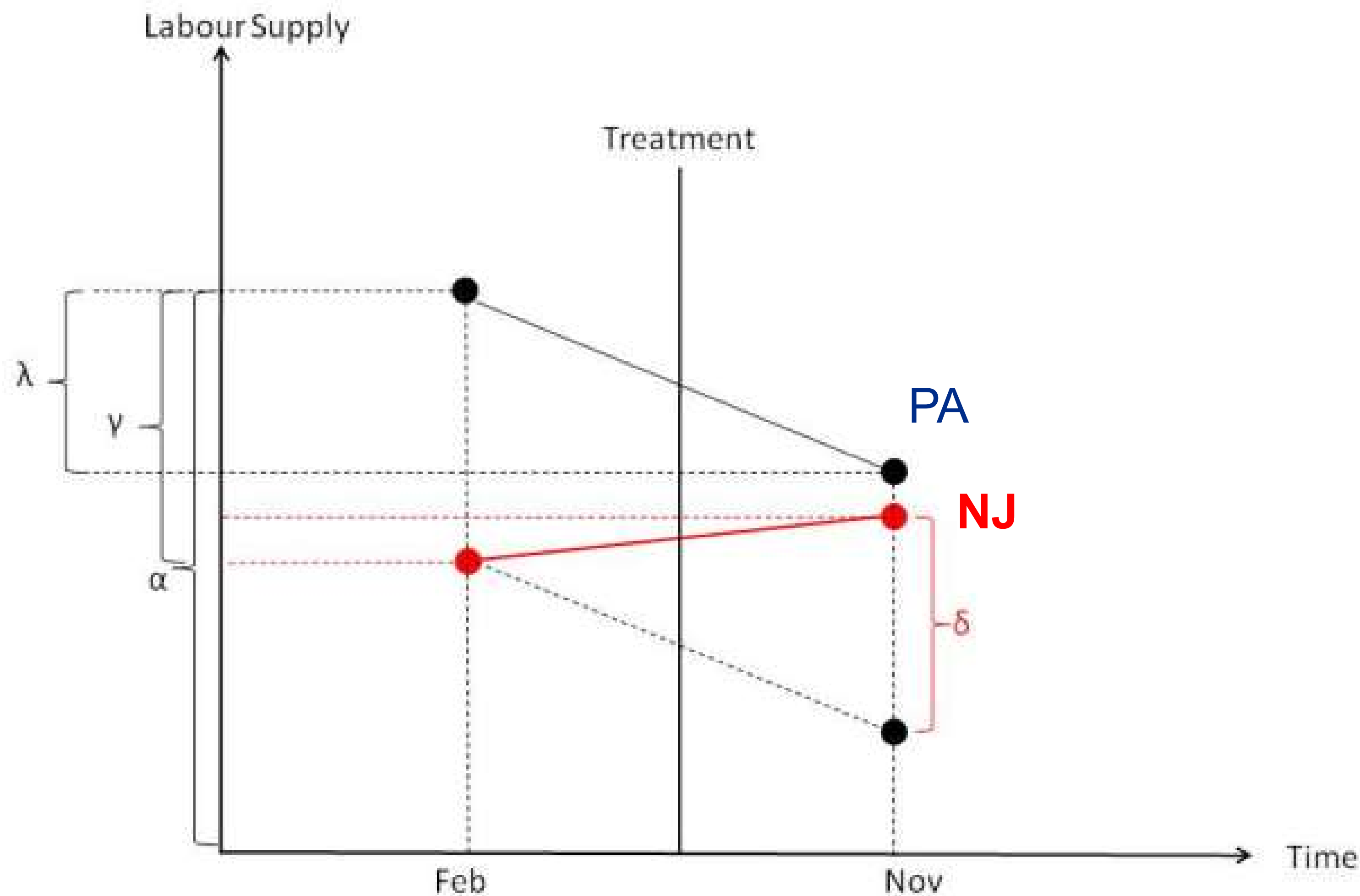
Notes: Standard errors are given in parentheses. The sample consists of 357 stores with available data on employment and starting wages in waves 1 and 2. The dependent variable in all models is change in FTE employment. The mean and standard deviation of the dependent variable are -0.237 and 8.825, respectively.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

DID Example: Card & Krueger (1994)

- How convincing is this evidence against the labor demand story?
  - The key identifying assumption in DD is that employment trends would be the same in both states in the absence of the treatment.


Labour Supply
Treatment
$Y_{PA,F}$
$Y_{PA,N}$
$Y_{NJ,N}$
$Y_{NJ,F}$
PA
NJ
Feb
Nov
Time

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta(NJ_s * d_t) + \varepsilon_{ist}$$



Card & Krueger (1994): Credible results?

- The key assumption for any DD strategy is that the outcome in treatment and control group would follow the same time trend in the absence of the treatment.
- Common trend assumption is difficult to verify but one often uses pre-treatment data to show that the trends are the same.
- Even if pre-trends are the same one still has to worry about other policies changing at the same time.
- Card and Krueger (2000) obtained administrative payroll data for restaurants in New Jersey and Pennsylvania for a number of years.

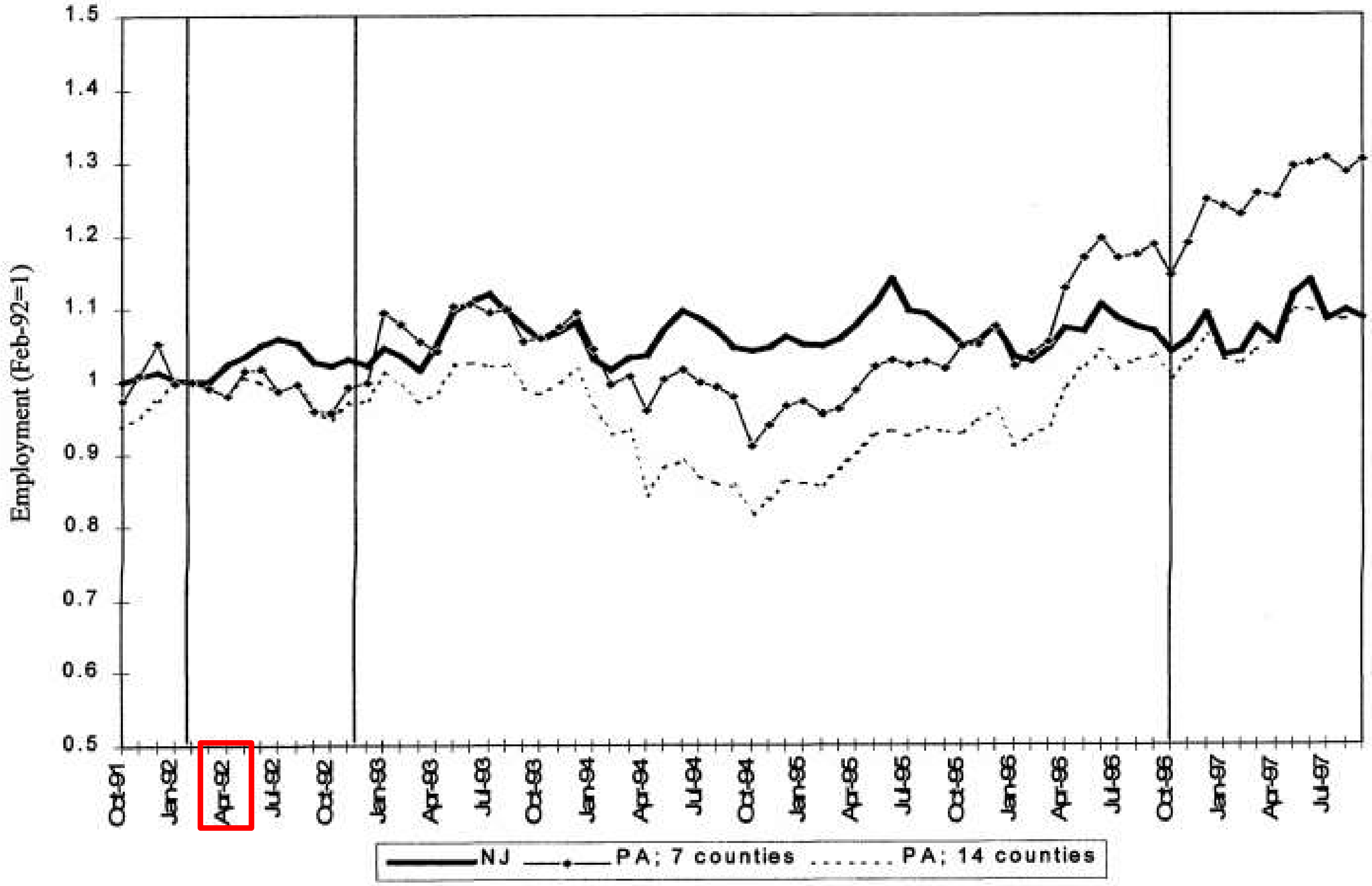


Figure 5.2.2: Employment in New Jersey and Pennsylvania fast-food restaurants, October 1991 to September 1997 (from Card and Krueger 2000). Vertical lines indicate dates of the original Card and Krueger (1994) survey and the October 1996 federal minimum-wage increase.

- The data reveals fairly substantial year-to-year employment variation in other periods.
- Employment trends in the two areas did not significantly differ, particularly for the 14 counties sample.

How can we explain these results?

An alternative to the conventional competitive model:

- NJ Stores were supply constrained (monopsony). Stores were in need of more workers, but no worker wanted to work at the old minimum wage.
- Price effects. Stores in NJ simply pass-through the minimum wage with a price increase. They generated more revenue, thus increase the number of employees.

Card & Krueger (1994): Other Critiques

- Min wage increase decided in early 1990 and implemented in April 92. The “Before” survey is conducted in Feb 1992.
- How could this announcement invalidate the identification strategy?
  - It might be the case that some restaurants start implementing the min wage policy earlier.
- Also, Nov might be too early to capture the treatment effect on employment (labor market frictions)
- **Note:** We don’t know which biases actually happened, and how big the effects were.
  - A good paper will check & rule out as many possible biases given the data.
  - Card and Krueger (1994,2000) present a wide variety of alternative specifications to probe the robustness of their initial conclusions.