Quasi Experimental Methods

Chris Conlon

September 17, 2024

NYU Stern

- What about an actual experiment that someone else ran?
- Something that isn't an experiment but sort of looks like one?
- We want as good as random variation.

Lotteries

Angrist, Pathak, Walters (2013)

Are charter schools effective at improving educational outcomes?

- Challenge: Simply being willing to sign up for a Charter school may mean you care more about education than the average student.
  This is the selection problem.
- Solution: Charters are over-subscribed, more students sign up that can be accomodated, and a lottery is held.
- Compare students who win the lottery to those that lose the lottery but otherwise look similar.

Angrist, Pathak, Walters (2013)

Table 2—Characteristics of Charter and Public Schools

	All charters (1)	Urban charters (2)	Nonurban charters (3)	Traditional public schools (4)
Panel A. Charter school characteristics				
Years open	10.1	8.6	12.4	—
Days per year	186	189	183	—
Average minutes per day	463	471	440	—
Saturday school	0.14	0.19	0.00	—
Average math instruction (min)	91	97	70	—
Average reading instruction (min)	84	91	52	—
No excuses	0.41	0.67	0.00	—
Panel B. Comparison with traditional public schools				
Proportion of teachers 32 and younger	0.56	0.70	0.33	0.20
Proportion of teachers 49 and older	0.17	0.08	0.31	0.42
Proportion of teachers licensed to teach assignment	0.64	0.62	0.67	0.98
Proportion of core classes taught by highly qualified teachers	0.94	0.92	0.98	0.98
Student/teacher ratio	12.0	12.6	11.1	15.2
Average per-pupil expenditure	$12,618	$13,668	$11,091	$13,047
Title I eligible	0.86	1.00	0.64	0.50
Observations (schools)	27	18	9	1,810

Notes: This table reports characteristics of Massachusetts charter and traditional schools. Charter school characteristics come from a survey of school administrators. Panel B compares charter schools to traditional public schools using data from state school profiles. Column 1 reports results from our statewide sample of charter schools with entry

Angrist, Pathak, Walters (2013)

TABLE 3—DESCRIPTIVE STATISTICS FOR STUDENTS

	Traditional public school students		Charter students		Charter lottery applicants	
	Urban (1)	Nonurban (2)	Urban (3)	Nonurban (4)	Urban (5)	Nonurban (6)
Panel A. Middle schools (5th–8th grade)						
Female	0.486	0.488	0.501	0.478	0.496	0.509
Black	0.183	0.027	0.381	0.035	0.479	0.022
Hispanic	0.319	0.038	0.246	0.039	0.233	0.025
Special education	0.191	0.165	0.167	0.158	0.176	0.185
Subsidized lunch	0.687	0.146	0.642	0.211	0.686	0.103
Limited English proficiency	0.160	0.017	0.082	0.022	0.085	0.008
Baseline math score	−0.427	0.210	−0.322	0.259	−0.356	0.305
Baseline ELA score	−0.466	0.232	−0.312	0.275	−0.375	0.391
Years in charter	0.00	0.00	2.09	1.97	1.59	1.25
Observations (students)	171,703	415,794	8,388	9,070	4,155	1,701
Observations (schools)	262	400	17	11	9	8
Panel B. High schools (10th grade)						
Female	0.499	0.494	0.557	0.545	0.548	0.538
Black	0.189	0.028	0.527	0.021	0.614	0.028
Hispanic	0.275	0.034	0.183	0.010	0.257	0.017
Special education	0.172	0.156	0.166	0.109	0.178	0.114
Subsidized lunch	0.612	0.126	0.608	0.146	0.717	0.123
Limited English proficiency	0.094	0.009	0.024	0.004	0.035	0.003
Baseline math score	−0.420	0.268	−0.371	0.321	−0.320	0.440
Baseline ELA score	−0.392	0.278	−0.318	0.412	−0.315	0.552
Years in charter	0.00	0.00	1.77	1.81	0.64	1.30
Observations (students)	132,774	357,733	2,676	909	3,029	351
Observations (schools)	104	316	6	2	4	2

Notes: This table reports descriptive statistics for the sample of public school students (columns 1 and 2), the sample of students in charter schools eligible for the study (columns 3 and 4), and our lottery sample of charter appli-

Angrist, Pathak, Walters (2013)

Table 4—Lottery Estimates of Charter Effects

Subject	All charter schools First stage (1)	All charter schools 2SLS (2)	Urban charter schools First stage (3)	Urban charter schools 2SLS (4)	Nonurban charter schools First stage (5)	Nonurban charter schools 2SLS (6)
Panel A. Middle school						
Math	1.02***	0.213***	1.03***	0.321***	1.01***	−0.123***
	(0.040)	(0.028)	(0.051)	(0.031)	(0.074)	(0.047)
N	16,543		11,941		4,602	
ELA	1.02***	0.075***	1.04***	0.146***	1.00***	−0.144***
	(0.040)	(0.025)	(0.051)	(0.028)	(0.074)	(0.039)
N	16,285		11,649		4,636	
Panel B. High school						
Math	0.565***	0.273***	0.508***	0.339***	1.13***	−0.020
	(0.085)	(0.071)	(0.090)	(0.077)	(0.197)	(0.071)
N	4,050		3,519		531	
ELA	0.565***	0.206***	0.508***	0.264***	1.14***	−0.046
	(0.086)	(0.060)	(0.090)	(0.067)	(0.196)	(0.059)
N	4,103		3,567		536	

Notes: This table reports 2SLS estimates of the effects of time spent in charter schools on test scores. The endogenous variable is years spent in charter schools and the instrument is a lottery offer dummy. Columns 1–2 show estimates for all schools, columns 3–4 show estimates for urban charter schools, and columns 5–6 show estimates for nonurban schools. The urban and nonurban estimates for a given subject come from a single regression with

How does it work?

Naive comparison (this has a selection problem):

$$\beta = \mathbb{E}[Score_i \mid \text{Attend}] - \mathbb{E}[Score_i \mid \text{Not Attend}]$$

What we want instead is:

$$\tau = \frac{\mathbb{E}[Score_i \mid Z_i = Win] - \mathbb{E}[Score_i \mid Z_i = Lose]}{\mathbb{P}[Attend_i \mid Z_i = Win] - \mathbb{P}[Attend_i \mid Z_i = Lose]} \approx \frac{\partial Score_i}{\partial Attend_i}$$

We learn about the average improvement in test scores for those people who attend if they win the lottery and don't attend if they lose (ie: those that apply).

Other Lotteries

- Angrist (1990): How does being a veteran affect long term earnings?
  - Problem: People who choose to serve in military are different from those who don't.
  - Solution: Vietnam War Draft Lottery numbers. Compare similar individuals with different lottery nubmers.
  - Answer: About 15% less over lifetime (not including service era).
- Do longer sentences affect recitivism?
  - Judges vary (in a predictable way) in how severe they setence for certain types of crimes
  - Compare individuals of similar age, background, and crime with strict vs lenient judges.

Supply and Demand (Graddy JEP)

$$\varepsilon_D = \frac{\mathbb{E}[\log Q_t \mid Z_t = \text{Rainy}] - \mathbb{E}[\log Q_t \mid Z_t = \text{Clear}]}{\mathbb{E}[\log P_t \mid Z_t = \text{Rainy}] - \mathbb{E}[\log P_t \mid Z_t = \text{Clear}]}$$

- What if we had something that randomly moved around supply but didn't affect demand.
- We could use this to trace out the demand curve or estimate the elasticity
- Weather offshore determines how many fishing boats go out, but not how much fish people want to eat (maybe?)

Supply and Demand (Graddy JEP)

Table 2

Ordinary Least Squares and Instrumental Variable Estimates of Demand Functions with Stormy Weather as an Instrument

	Ordinary least squares (dependent variable: log quantity)		*Instrumental variable*	
Variable	*(1)*	*(2)*	*(3)*	*(4)*
Log price	−0.54	−0.54	−1.08	−1.22
	(0.18)	(0.18)	(0.48)	(0.55)
Monday		0.03		−0.03
		(0.21)		(0.17)
Tuesday		−0.49		−0.53
		(0.20)		(0.18)
Wednesday		−0.54		0.58
		(0.21)		(0.20)
Thursday		0.09		0.12
		(0.20)		(0.18)
Weather on shore		−0.06		0.07
		(0.13)		(0.16)
Rain on shore		0.07		0.07
		(0.18)		(0.16)
R^2	0.08	0.23		
No. of Obs.	111	111	111	111

Source: The data used in these regressions are available by contacting the author.

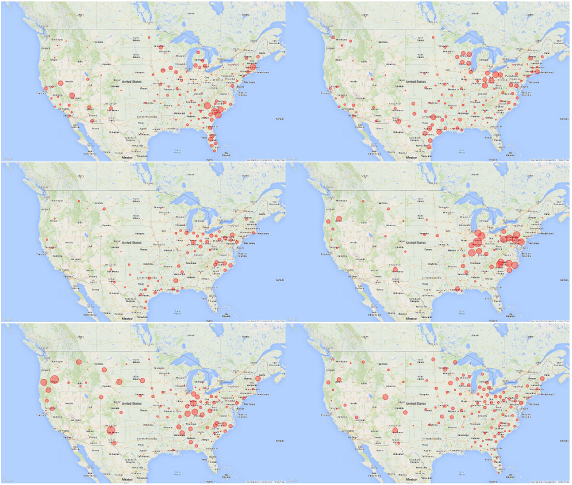
Sinkinson Starc (2021)

Back to advertising, do ads actually work?

- Problem: Ad spending is often chosen as a share of revenue (oops?)
- Problem: Ad spending is ideally chosen where it will be most effective.
- Idea: During elections, ads for cholesterol drugs are crowded out by political ads.
- Idea: They are crowded out even more in swing states
- Use this to move around the price/number of ads but not their effectiveness.

Sinkinson Starc (2021)

Figure 2: Political Ad Levels, January-June 2008



Notes: The above maps show a dot for each DMA in the USA. The diameter of each dot is proportional to the number of political ads aired in that market, in that month, for all races (Presidential, Senatorial, House, Gubernatiorial). The first row are January and February; second

Figure 3: National Pharmaceutical Ad Levels for Statins

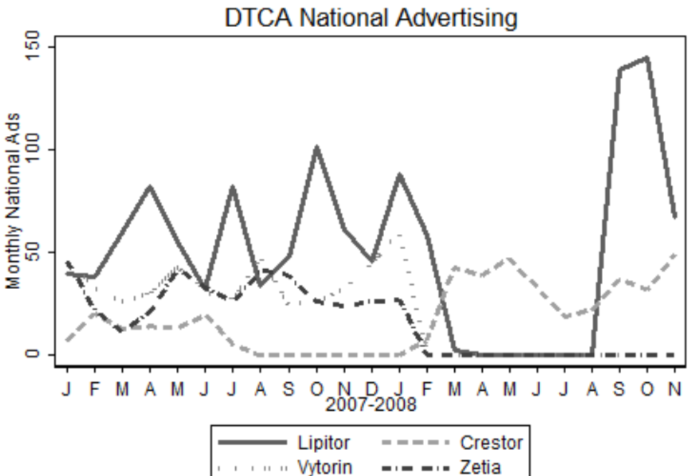


Figure 4: National Pharmaceutical Ad Levels for Statins

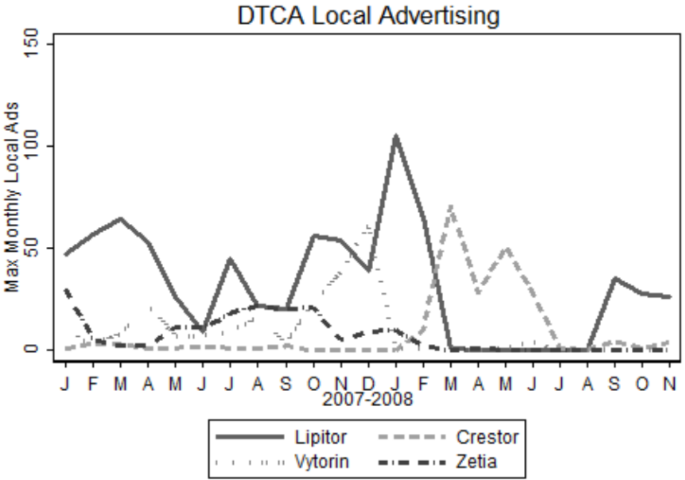
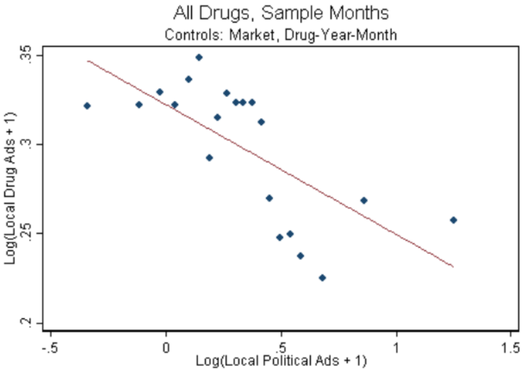
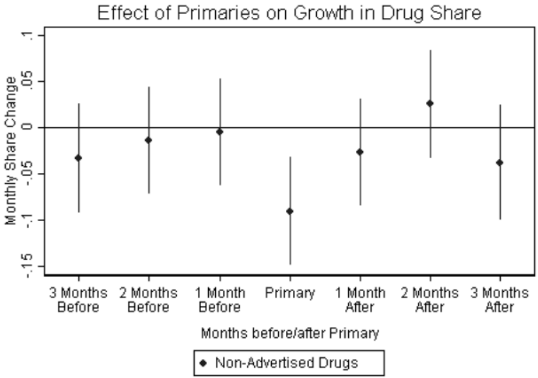


Figure 5: Political Ads Displace Local Drug Ads, Binned Scatter plot



Notes: The above plots bins of observations from July 2007 to November 2008 at the market-month level after residualizing by market and year-month fixed effects, and adding back the sample mean. Twenty bins are used. The fitted line is based on a regression of all underlying

Figure 6: Effect of Primary Timing on Non-Advertised Statins



Note: The above plots estimated coefficients for timing dummies relative to a market's primary month. The dependent variable is the (one-month) change in market share, defined as the percentage of the population taking a non-advertised statin.

Sinkinson Starc (2021)

Table 4: OLS Revenue Regressions for Advertised Drugs

	Dependent Variable: Log(Revenue per Insured)					
	This Month			Two Month Trailing Average		
Own Ads	0.0194***	0.0064***	0.0114***	0.0239***	0.0066***	0.0096***
	(0.0022)	(0.0020)	(0.0023)	(0.0021)	(0.0021)	(0.0023)
Rival Ads	0.0042*	0.0056**	0.0061**	0.0016	0.0036	0.0039
	(0.0025)	(0.0023)	(0.0026)	(0.0027)	(0.0026)	(0.0028)
Controls:						
Market FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Drug FEs	X	X	X	X	X	X
Drug FE*Time Trend		X	X		X	X
Drug FE*Time Trend^2			X			X
Drug-Year FEs	X			X		
N	11,551	11,551	11,551	11,550	11,550	11,550
R^2	0.842	0.847	0.848	0.843	0.847	0.848

Sinkinson Starc (2021)

Table 5: IV Revenue Regressions for Advertised Drugs

	Dependent Variable: Log(Revenue per Insured)					
	This Month			Two Month Trailing Average		
Own Ads	0.1559***	0.0808**	0.0923**	0.1252***	0.0764***	0.0734***
	(0.0251)	(0.0344)	(0.0373)	(0.0136)	(0.0258)	(0.0270)
Rival Ads	−0.1064***	−0.0492**	−0.0471**	−0.0966***	−0.0548***	−0.0425***
	(0.0179)	(0.0247)	(0.0261)	(0.0112)	(0.0212)	(0.0215)
Controls:						
Market FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Drug FEs	X	X	X	X	X	X
Drug FE*Time Trend		X	X		X	X
Drug FE*Time Trend^2			X			X
Drug-Year FEs	X			X		
N	11,551	11,551	11,551	11,550	11,550	11,550
R^2	0.755	0.819	0.822	0.788	0.824	0.832

A Famous Example: Card and Krueger (AER 1994)

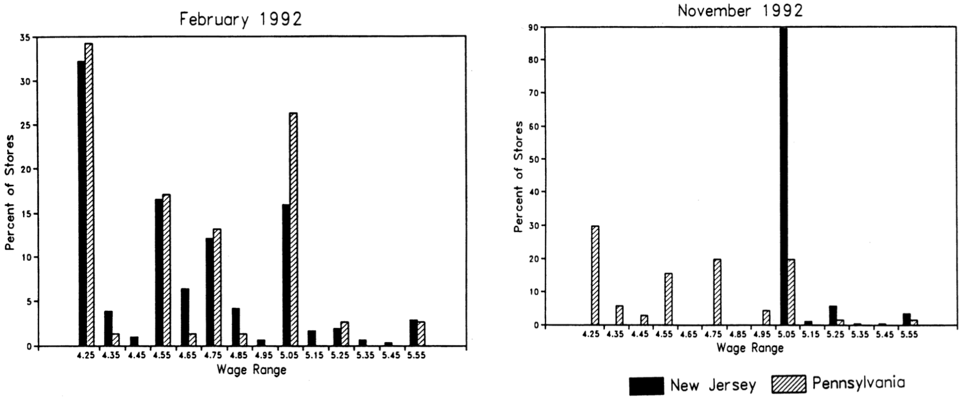
A Famous Example: Card and Krueger (AER 1994)

- On April 1, 1992 NJ raises its minimum wage from \$4.25 $\longrightarrow$ \$5.05 per hour.
- Question: Econ 101 predicts this will reduce demand for low wage workers
  - Focus on fast food restaurants (since they pay min wage)
  - Focus on starting wage (avoid tenure, high turnover)
- Survey 410 restaurants in NJ (treated group) and eastern PA (control group).
- Idea: Compare change in wages in *NJ* to *PA*: $\Delta_{DD} = \Delta_{NJ} - \Delta_{PA}$
  - Wave 1: February 15-March 4, 1992
  - Wave 2: November 5 - December 31, 1992

Balance Table: Covariates

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 (0.51)	23.3 (1.35)	−2.0
b. Percentage full-time employees	32.8 (1.3)	35.0 (2.7)	−0.7
c. Starting wage	4.61 (0.02)	4.63 (0.04)	−0.4
d. Wage = \$4.25 (percentage)	30.5 (2.5)	32.9 (5.3)	−0.4
e. Price of full meal	3.35 (0.04)	3.04 (0.07)	4.0
f. Hours open (weekday)	14.4 (0.2)	14.5 (0.3)	−0.3
g. Recruiting bonus	23.6 (2.3)	29.1 (5.1)	−1.0
3. *Means in Wave 2:*			
a. FTE employment	21.0 (0.52)	21.2 (0.94)	−0.2
b. Percentage full-time employees	35.9 (1.4)	30.4 (2.8)	1.8
c. Starting wage	5.08 (0.01)	4.62 (0.04)	10.8
d. Wage = \$4.25 (percentage)	0.0	25.3 (4.9)	—
e. Wage = \$5.05 (percentage)	85.2 (2.0)	1.3 (1.3)	36.1
f. Price of full meal	3.41 (0.04)	3.03 (0.07)	5.0
g. Hours open (weekday)	14.4 (0.2)	14.7 (0.3)	−0.8
h. Recruiting bonus	20.3 (2.3)	23.4 (4.9)	−0.6

Distribution of Wages



Differences in Wages : 2 x 2 Table

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

	Stores by state			Stores in New Jersey[a]			Differences within NJ[b]	
Variable	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.

Outcome Equation

- Differences lack any covariates (different fast food chains).
- Also $\Delta_{PA} < 0$ and $\Delta_{NJ} > 0$ (!)
- Recall i denotes stores, $t \in 1, 2$. Run the following regression:

$$Y_{it} = \beta X_{it} + \alpha \cdot [i \in \text{NJ}] + \gamma \cdot \text{After}_t + \delta \cdot NJ_i \times After_t + u_i$$

$$Y_{it} = \beta X_{it} + \alpha \cdot [\text{wage gap}_i] + \gamma \cdot \text{After}_t + \delta \cdot \text{wage gap}_i \times After_t + u_i$$

- α is mean difference between NJ and PA
- γ is mean difference between period 1 and 2
- δ is the parameter of interest, the difference in difference
- $\text{wage gap}_i = [\text{min wage}_{i,2} - w_{i1}]_+ = \max\{0, \text{min wage}_{i,2} - w_{i1}\}$.

  (How much do you need to raise $t = 1$ wages to achieve minimum wage in $t = 2$?)

Differences in Wages

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).

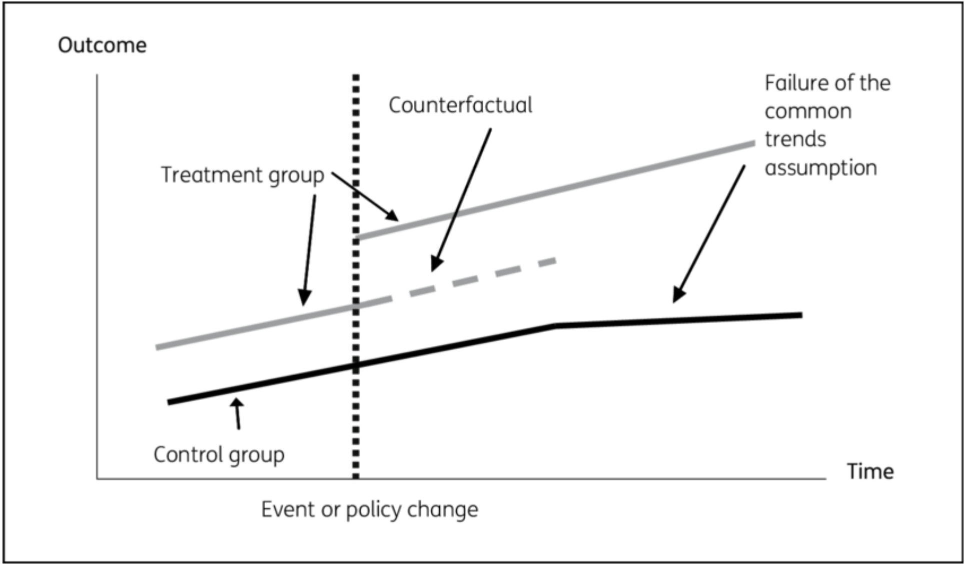
[a]Proportional increase in starting wage necessary to raise starting wage to new minimum rate. For stores in Pennsylvania the wage gap is 0.

[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.

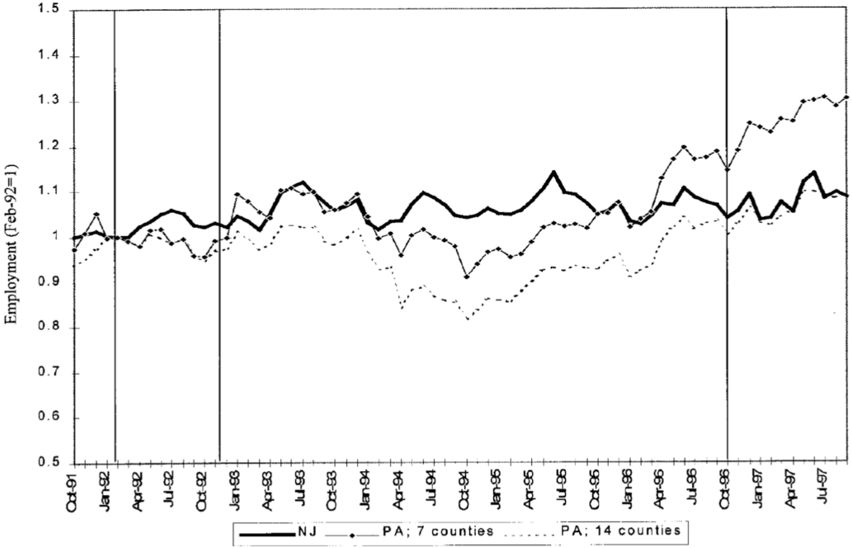
Parallel Trends



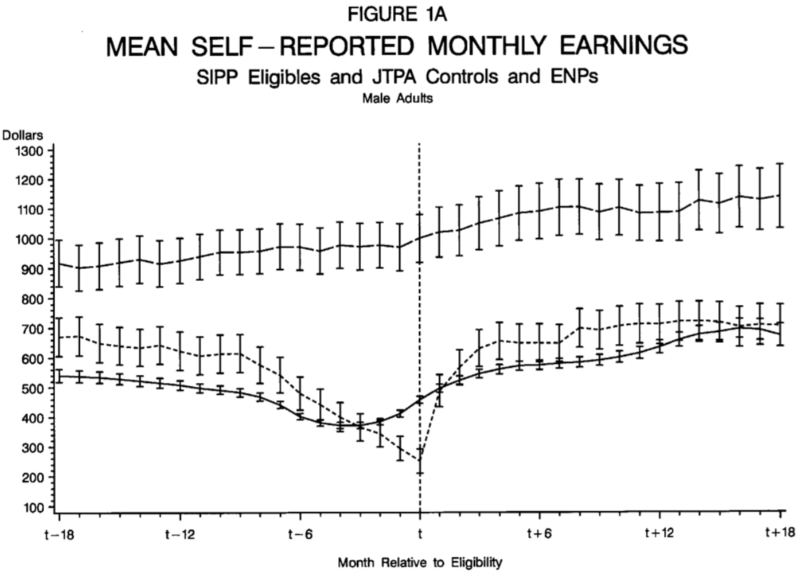
Difference in Differences: Limitations

1. Functional form restrictions
   - Parallel trends assumes that absent treatment that we add $\gamma_2 - \gamma_1$ to each unit
   - Because this is additive it is not invariant to transformations $f(Y_{it})$ (ie: taking logs)
2. Parallel Trend Assumption is not testable
   - Best we can hope is that it looks similar in the pre-period
3. Compositional Effects: the treatment may affect who is in each group
   - Restaurants could close in NJ and open nearby in PA to avoid minimum wage.
   - A good job training program may lead to migration, etc.
   - One approach: redefine the population so that it doesn't endogenously respond to treatment
     - Recover something, but probably not ATT anymore...

Checking Pre-Trend: Card Krueger (2000)



The “Ashenfelter Dip” (Heckman and Smith 2000)



Motivation: Recap

Difference in Difference approaches have some drawbacks:

- We need to really believe parallel trends
  - Is ΔPA really a good counterfactual for ΔNJ?
  - Obvious question: why not pick ΔDE or ΔNY?
  - Measured effect shouldn't change if our assumption is valid (but it probably will!)

Example: Abadie, Diamond, Hainmueller (2010)

The Question

In 1988 California passes anti-smoking Prop 99

- increased excise tax by 25 cents per pack,
- earmarked the tax revenues to health and anti-smoking education budgets and funded anti-smoking ads
- led to indoor smoking bans in restaurants and bars city by city

What was the effect on per capita cigarette sales?

- Already a bunch of pre-existing trends.
- What is a good control for California?

Use state-level data from 1970-2000.

The Idea

Use a convex combination of other states to construct a synthetic counterfactual California.

- We observe Y_{it}, X_{it}, T_{it}.
- Assume only $i = 1$ and $t > T_0$ are treated.
- Construct a donor pool of potential controls subscripted by j.
- Choose some weights w_j for each entity (state) in donor pool. How?
  - Same $\mathbf{X_1} = \sum_j w_j \mathbf{X_j}$ as treated observations (like matching).
  - Same $\left(Y_{1,1}, \ldots, Y_{1,T_0}\right) = \sum_j w_j \cdot \left(Y_{j,1}, \ldots, Y_{j,T_0}\right)$ (like parallel trends).
  - Weights sum to one $\sum_j w_j = 1$ and maybe are non-negative (or not!)
- Idea is to match all of the X's and all of the Y_{it}'s in the pre-period

Covariate Balance

Table 1. Cigarette sales predictor means

Variables	California Real	California Synthetic	Average of 38 control states
Ln(GDP per capita)	10.08	9.86	9.86
Percent aged 15–24	17.40	17.40	17.29
Retail price	89.42	89.41	87.27
Beer consumption per capita	24.28	24.20	23.75
Cigarette sales per capita 1988	90.10	91.62	114.20
Cigarette sales per capita 1980	120.20	120.43	136.58
Cigarette sales per capita 1975	127.10	126.99	132.81

NOTE: All variables except lagged cigarette sales are averaged for the 1980–1988 period (beer consumption is averaged 1984–1988). GDP per capita is measured in 1997 dollars, retail prices are measured in cents, beer consumption is measured in gallons, and cigarette sales are measured in packs.

Donor Weights

Table 2. State weights in the synthetic California

State	Weight	State	Weight
Alabama	0	Montana	0.199
Alaska	–	Nebraska	0
Arizona	–	Nevada	0.234
Arkansas	0	New Hampshire	0
Colorado	0.164	New Jersey	–
Connecticut	0.069	New Mexico	0
Delaware	0	New York	–
District of Columbia	–	North Carolina	0
Florida	–	North Dakota	0
Georgia	0	Ohio	0
Hawaii	–	Oklahoma	0
Idaho	0	Oregon	–
Illinois	0	Pennsylvania	0
Indiana	0	Rhode Island	0
Iowa	0	South Carolina	0
Kansas	0	South Dakota	0
Kentucky	0	Tennessee	0
Louisiana	0	Texas	0
Maine	0	Utah	0.334
Maryland	–	Vermont	0
Massachusetts	–	Virginia	0
Michigan	–	Washington	–
Minnesota	0	West Virginia	0
Mississippi	0	Wisconsin	0
Missouri	0	Wyoming	0

Trend Check and Treatment Effects

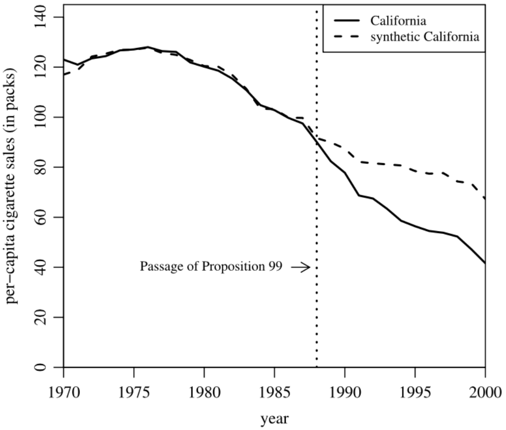
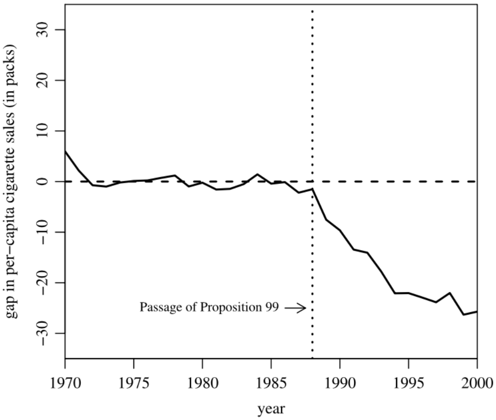


Figure 2. Trends in per-capita cigarette sales: California vs. synthetic California.



But still some issues

- How sensitive are weights estimates to different covariates?
  - "state-level measures of unemployment, income inequality, poverty, welfare transfers, crime rates, drug related arrest rates, cigarette taxes, population density, and numerous variables to capture the demographic, racial, and social structure of states".
- Can we run a placebo check? Do we detect effects where we know there is a null effect?
  - Put California in the donor pool.
  - Pick a state from the donor pool at pretend that receives the treatment after T_0
  - Choose w_j following the synthetic control procedure.
  - Compute the treatment effects in the same way.
  - Repeat for all states in donor pool.
  - Compare mean-square prediction error (MSPE) for $(Y_{1,1}, \ldots, Y_{1,T_0})$
  - This doubles as inference.

Regression Discontinuity Design

- Another popular research design is the Regression Discontinuity Design.
- In some sense this is a special case of IV regression. (RDD estimates a LATE).
- Most of this is taken from the JEL Paper by Lee and Lemieux (2010).

RDD: Basics

- We have a running or forcing variable x such that

$$\lim_{x \longrightarrow c^+} P(T_i|X_i = x) \neq \lim_{x \longrightarrow c^-} P(T_i|X_i = x)$$

- The idea is that there is a discontinuous jump in the probability of being treated.
- For now we focus on the sharp discontinuity:
  $P(T_i|X_i \geq c) = 1$ and $P(T_i|X_i < c) = 0$
- There is no single x for which we observe treatment and control.

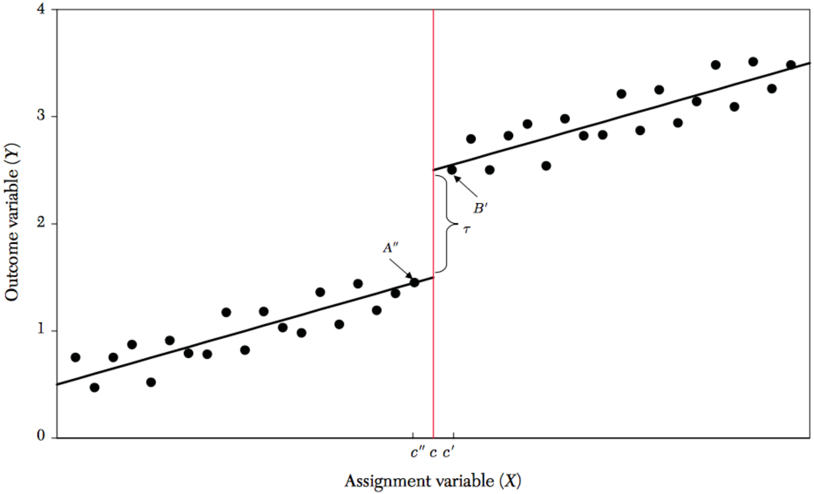
RDD: Basics

- Example: a social program is available to people who earned less than $25,000.
  - If we could compare people earning $24,999 to people earning $25,001 we would have as-if random assignment. (MAYBE)
  - But we might not have that many people...
- We are going to label the treatment effect τ_i.
  Note: my lack of precision here!
- The most important assumption is that of no manipulability $\tau_i \perp T_i$ in some neighborhood of c.
  - If agents can choose x_i we are in trouble: underreporting income, avoiding "possession with intent to distribute" for drugs, etc.

RDD: Continuity

- The central idea in RDD is that of continuity
- We need that $E[Y(1)|X]$ and $E[Y(0)|X]$ both be continuous at $X = c$.
  - We expect that $Y_i = f(x_i)$ to be a smooth, continuous function of x_i
  - The only departure from that is the treatment $\tau_i \cdot I(x_i \geq c)$.
- We want to be as agnostic as possible about functional form
  - Don't want to restrict ourselves to $f(x_i) = \beta_0 + \beta_1 x_i$.
  - The central idea: we know $f(x_i)$ absent the treatment!

RDD: In Pictures



RDD: Sharp RD Case

RDD uses a set of assumptions distinct from our LATE/IV assumptions. Instead it depends on continuity.

- People just to the left of c are a valid control for those just to the right of c.
- This is not a testable assumption $\longrightarrow$ draw pictures!
- We could run the regression where $T_i = \mathbf{1}[X_i > c]$.

$$Y_i = \beta_0 + \tau_i \cdot T_i + X_i\beta + \epsilon_i$$

- This puts a lot of restrictions (linearity) on the relationship between Y and X.
- Also (without additional assumptions) we only learn about τ_i at the point $X = c$.

Application: Lee (2008)

Looked at incumbency advantage in the US House of Representatives

- Running variable was vote share in previous election
  - Problem of naive approach: good candidates get lots of votes!
  - Compare outcomes of districts with barely D to barely R.
- First we plot bin-scatter plots and quartic (from each side) polynomials.
- Discussion about how to choose bin-scatter bandwidth (CV).

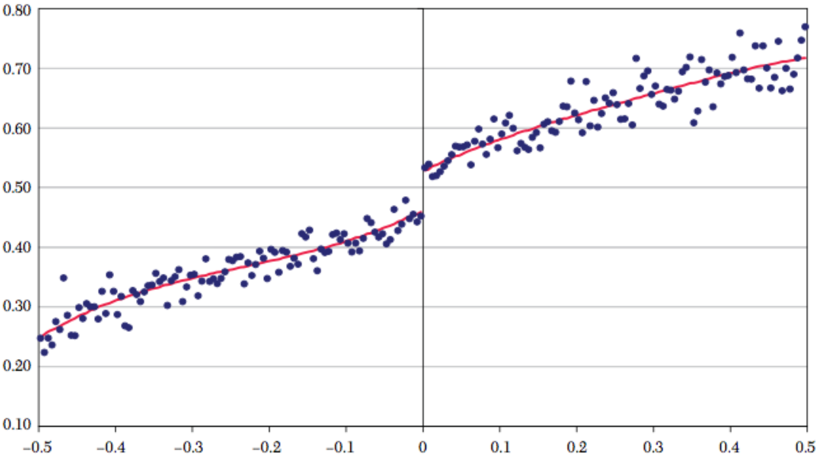


Figure 8. Share of Vote in Next Election, Bandwidth of 0.005 (200 bins)

Lee (2008)

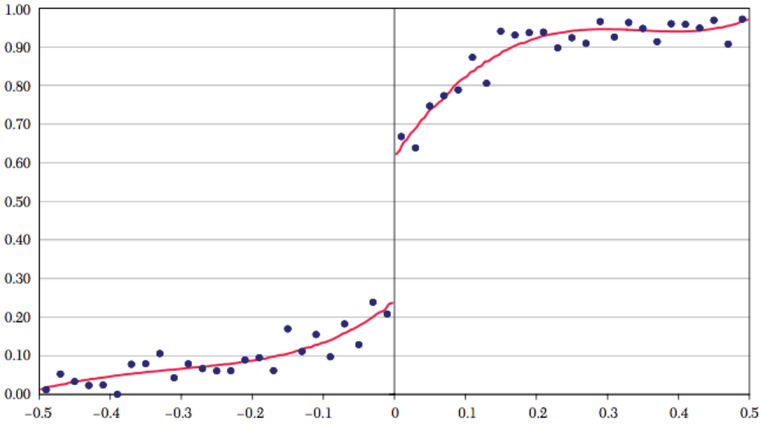


Figure 9. Winning the Next Election, Bandwidth of 0.02 (50 bins)

Lee (2008)

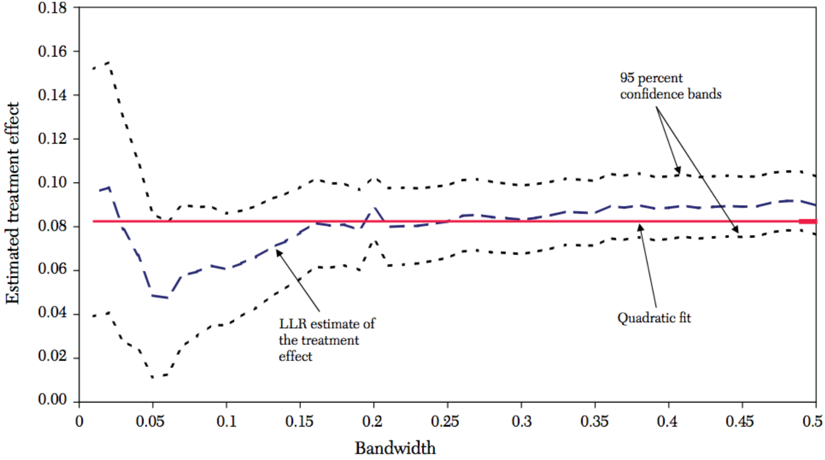


Figure 18. Local Linear Regression with Varying Bandwidth: Share of Vote at Next Election

Other Examples

Luca on Yelp

- Have data on restaurant revenues and yelp ratings.
- Yelp produces a yelp score (weighted average rating) to two decimals ie: 4.32.
- Score gets rounded to nearest half star
- Compare 4.24 to 4.26 to see the impact of an extra half star.
- Now there are multiple discontinuities: Pool them? Estimate multiple effects?

Final Thoughts

Always be asking

- What am I trying to measure?
- What is my endogeneity problem?
- What plausibly random variation can I use?

We don't want to look for our keys where the light is on, but you could do a lot worse than following one the scripts above !