Empirical Methods

Prof. Richard Sweeney

ECON3391.01, Boston College

Outline

Intro

Additional Material

- A great undergraduate level treatment of the econometric concepts we'll cover is in "Mastering 'Metrics: The Path from Cause to Effect"
- For a reasonably accessible article discussing these methods in the context of environmental policy evaluation, see "Quasi-Experimental and Experimental Approaches to Environmental Economics" by Greenstone and Gayer

Policymakers and business leaders have lots of questions about energy markets

How much more will people drive when gas is \$2 vs \$4?

Does fracking affect property values?

How much money will people save if we invest in energy efficiency?

How much does pollution reduce worker / firm productivity?

Does deregulation lower electricity prices?

Policymakers and business leaders have lots of questions about energy markets

How much more will people drive when gas is $2 vs $4?

Does fracking affect property values?

How much money will people save if we invest in energy efficiency?

How much does pollution reduce worker / firm productivity?

Does deregulation lower electricity prices?

What can economists say about these?

Start with economic theory

- Sometimes we have an obvious prediction from economic theory
  - What will happen if we cap price? There will be excess demand (shortage)

Start with economic theory

- Sometimes we have an obvious prediction from economic theory
  - What will happen if we cap price? There will be excess demand (shortage)
- Other times theory is ambiguous
  - Is monopoly good for innovation (Schumpeter) or bad for innovation (Arrow)?

Start with economic theory

- Sometimes we have an obvious prediction from economic theory
  - What will happen if we cap price? There will be excess demand (shortage)
- Other times theory is ambiguous
  - Is monopoly good for innovation (Schumpeter) or bad for innovation (Arrow)?
- Even when the *sign* of the effect is theoretically clear, policymakers care about the magnitude of the effect, not just the direction.
- How can we quantify our answers to these questions?

Recap: How consumer prices are set

Regulated regions:

- Vertically integrated utility serves demand as it sees fit
- Reports **total costs** and total demand
- Regulator sets long run retail prices equal to the *average* cost plus a "fair" return

Recap: How consumer prices are set

Regulated regions:

- Vertically integrated utility serves demand as it sees fit
- Reports **total costs** and total demand
- Regulator sets long run retail prices equal to the *average* cost plus a "fair" return

Deregulated regions

- Wholesale auctions determine which plants get to operate (often every 15 minutes)
- All dispatched plants receive the *marginal* price
  - ie the marginal cost of the most expensive dispatched plants
- Retailers procure power from these auctions to keep the lights on
- Retail prices set to recover the average *auction* price, plus costs of transmission and distribution

Will deregulated prices be higher or lower?

Will deregulated prices be higher or lower?

- Short run: suspect markets will be better at reducing costs (-)
- But needs to be balanced against the fact that firms now may try to exert market power (price above marginal cost). (+)
- In long run, no capital bias, so the "right" plants will get built (-)
- Is the return guaranteed by the regulator lower that the average profit earned in deregulated regions?

The "engineering" / modeling approach

The electricity market game was a great example of how might answer this question with a model

- Create a detailed dataset with the capacity, variable costs and fixed costs of every plant.
- To simulate **regulation**
  - Set demand and the elasticity each period
  - Assume the cheapest marginal costs plants get dispatched each period
  - Compute the total operating costs over the year, add a return on capital, divide by quantity to get prices.

The "engineering" / modeling approach

The electricity market game was a great example of how might answer this question with a model

- Create a detailed dataset with the capacity, variable costs and fixed costs of every plant.
- To simulate **regulation**
  - Set demand and the elasticity each period
  - Assume the cheapest marginal costs plants get dispatched each period
  - Compute the total operating costs over the year, add a return on capital, divide by quantity to get prices.
- To simulate **deregulation**
  - Keep demand the same
  - Model the way in which power plants set *bids* each period, which may include attempts to exert market power.
  - Dispatch based on the lowest bids and pay each generator the marginal bid.

The engineering / modeling approach is **complicated** and **requires many assumptions**

- We know which plants are available *now* and their capacities, but we do not know their costs.
- We suspect the regulated utilities won't actually always use the cheapest plants or reduce costs.
- We do not know exactly how deregulated firms will bid. Will they be able to collude?
- Most importantly, we think that plant entry and exit will be a major determinant of long run prices.

An alternative approach would be to gather data and try to estimate the effect of interest.

The New York Times

The U.S. Economy Today | U.S. Hiring Surges | Mortgage Rates Fall | A Major Fed Slowdown | Economy Recor

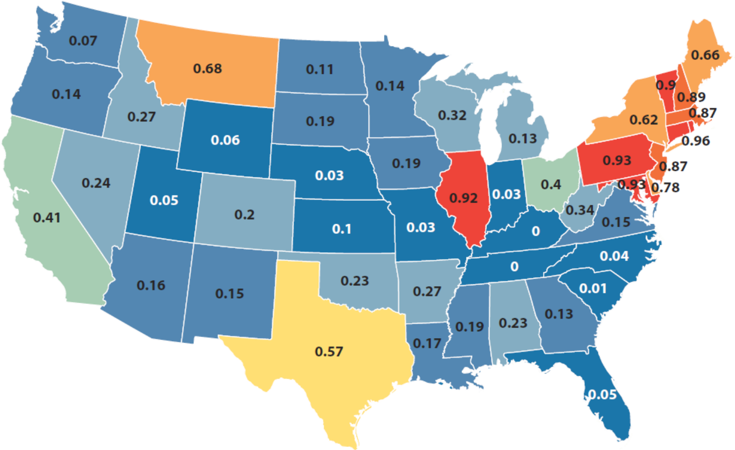
Why Are Energy Prices So High? Some Experts Blame Deregulation.

California and the 34 other states that have deregulated all or parts of their electricity system tend to have higher rates than the rest of the country.

Give this article

250

Empirical examination of deregulation



Wholesale deregulation share (Borenstein and Bushnell (2015))

- Electricity price data easily obtainable.
- Half of the country deregulated between 1995 and 2002.

Do deregulated regions have lower prices now?

```
. reg Price Deregulated if Year == 2019
```

Source	SS	df	MS
Model	91.2398127	1	91.2398127
Residual	728.974262	49	14.8770258
Total	820.214075	50	16.4042815

Number of obs	=	51
F(1, 49)	=	6.13
Prob > F	=	0.0168
R-squared	=	0.1112
Adj R-squared	=	0.0931
Root MSE	=	3.8571

Price	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Deregulated	2.837353	1.145722	2.48	0.017	.5349385	5.139767
_cons	10.31265	.6614829	15.59	0.000	8.983347	11.64195

Do deregulated regions have lower prices now?

```
. reg Price Deregulated if Year == 2019

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(1, 49)        =      6.13
       Model |  91.2398127         1  91.2398127   Prob > F        =    0.0168
    Residual |  728.974262        49  14.8770258   R-squared       =    0.1112
-------------+----------------------------------   Adj R-squared   =    0.0931
       Total |  820.214075        50  16.4042815   Root MSE        =    3.8571

------------------------------------------------------------------------------
       Price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 Deregulated |   2.837353   1.145722     2.48   0.017     .5349385    5.139767
       _cons |   10.31265   .6614829    15.59   0.000     8.983347    11.64195
------------------------------------------------------------------------------
```

What might be going on here?

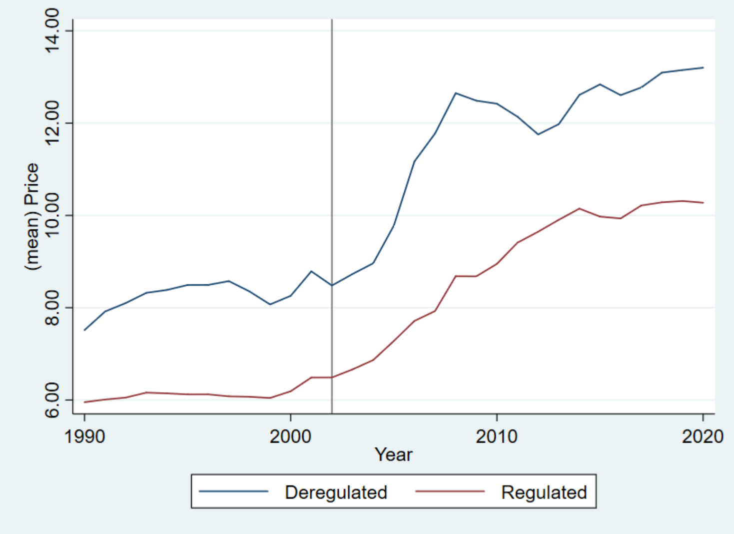
Did prices go down *after* regulation?

Did prices go down *after* regulation?

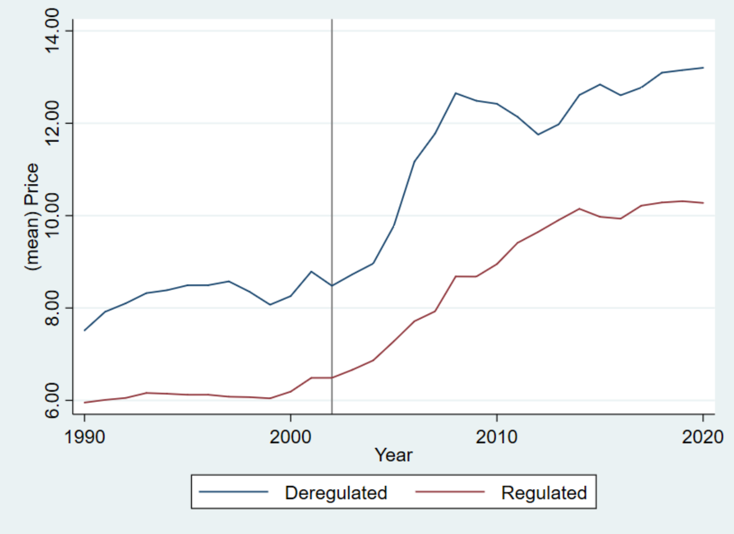
What might be going on here?

How did Borenstein and Bushnell estimate this?

How did Borenstein and Bushnell estimate this?



How did Borenstein and Bushnell estimate this?



What problems did they encounter?

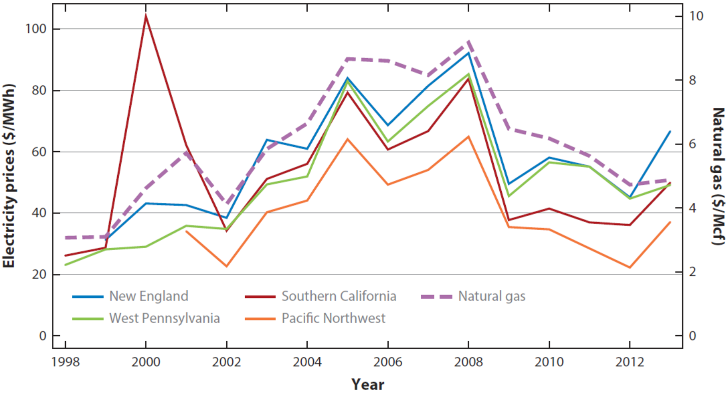
Early differences went in the wrong direction, but this was reversed over time

Table 1 Summary of retail price changes

Definition of restructured	Status	Average retail price (USD) 1997	2007	2012	Percent change 1997–2007	2007–2012	1998–2012
Power in the Public Interest definition	Not restructured	5.89	7.44	8.72	0.21	0.15	0.32
	Restructured	8.96	12.53	12.35	0.29	−0.01	0.27
At least 40% independent power producers in 2012	Not restructured	5.67	7.23	8.57	0.22	0.16	0.34
	Restructured	8.83	11.99	11.95	0.26	0.00	0.26

Source: Borenstein and Bushnell (2005)

Challenge: Nat gas prices increased after restructuring



- Although a small share of total output, natural gas is on the margin most hours
- Under average cost pricing, this has little effect; but under marginal cost pricing, this raises everyone's profits
- Ex ante, this increased exposure to natural gas prices was not appreciated

Regression review

Economist's approach to empirical policy analysis

Applied economists reduce policy questions to a relationship between two variables of interest.

X	Y
gas prices	miles driven
proximity to wells	house prices
energy efficiency of capital	electricity consumption
exposure to pollution	worker productivity (wages)

All of these are **random variables**.

Inference about relationships

- Just as Y_i and X_i have randomness, the relationship between them is also not deterministic.
  - If we look at the data, we will likely find some very productive workers in both polluted and unpolluted areas.
- So we are often interested in the *average* relationship
- One way to summarize this relationship is with the *covariance*

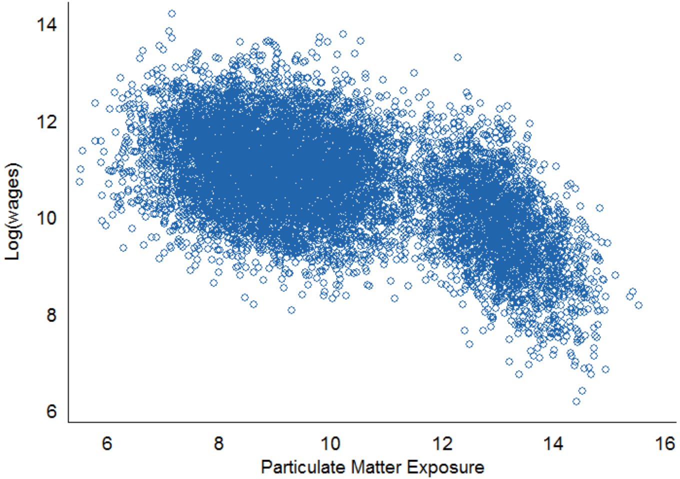
$$Cov(X_i, Y_i) = E[(X_i - E[X_i])(Y_i - E[Y_i])]$$

- So when X_i is above it's average:
  - we expect Y_i to be **above** its average $Cov(X_i, Y_i) > 0$
  - and we expect Y_i to be **below** its average $Cov(X_i, Y_i) < 0$

What if we want to know *how much* we expect Y_i to change if X_i changes?

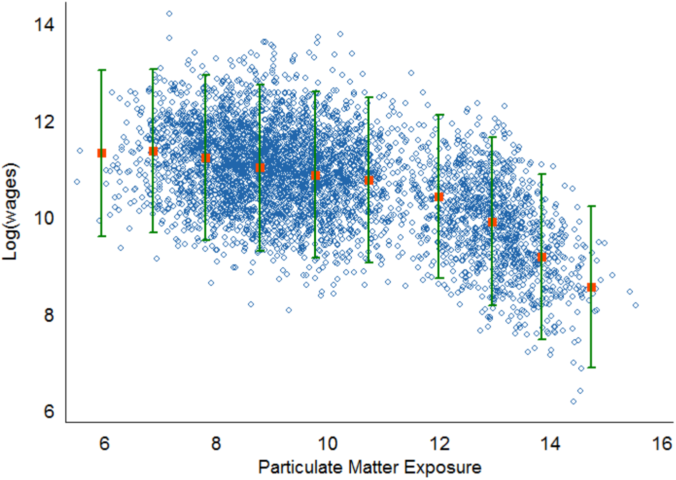
- $Cov(X_i, Y_i)$ only tells you about the sign of the relationship.
- To get magnitudes, we need to think about the *conditional expectation* $E[Y_i|X_i = x]$
  - For a review of conditional expectations, see the backup slides here.
- Can then compare the CE at different values of X to estimate continuous relationships

Good practice to start with a scatter plot of your data



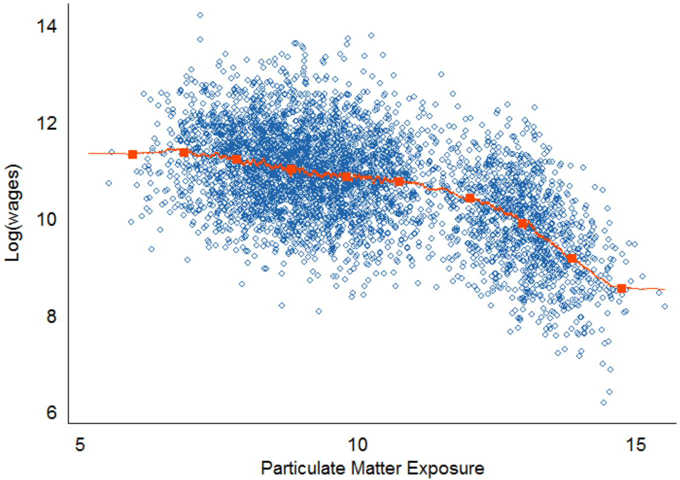
Can then take the conditional expectation of Y_i over different ranges of X_i

- This figure breaks the data up into 10 ranges and plots the sample means and 95% CIs

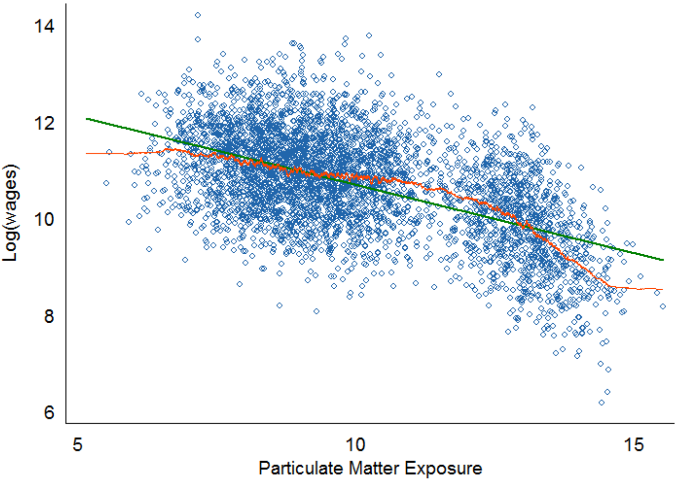


These conditional expectations trace out a conditional expectation function (CEF)

- We can now see how much our expectation of Y_i changes as we move over different values of X_i
- In this case, the relationship appears to be nonlinear



We typically approximate the CEF using linear regression



Linear regression gives a line that "best" fits the data

- Assume the CEF is linear

$$E[Y_i|X_i] = \beta_o + \beta_1 X_i$$

- Implies

$$Y_i = \beta_o + \beta_1 X_i + e_i$$

- This difference (e_i) between our expectation and what observe in the data is called the *residual*
- OLS picks the "best" line by minimizing the sum of $(e_i)^2$

$$Min_\beta \sum_{i=1}^{N} (y_i - E[Y_i|X_i = x_i])^2$$

- It turns out that the solution to this is

$$\beta_1 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} = \frac{Cov(X_i, Y_i)}{Var(X_i)}$$

Causality

Problem: correlation does not equal causation

$$E[Y_i|X_i] = \beta_o + \beta_1 X_i$$

- So regression provides a linear summary of the CEF
- We can interpret β_1 as saying: if we compare two individuals whose X differ by 1 unit, we expect their Y to differ by β_1 units **in the data.**
  - ie when we compare two workers whose pollution exposure differs by 1 unit, their wages differ by β_1 units on average.

Problem: correlation does not equal causation

$$E[Y_i|X_i] = \beta_o + \beta_1 X_i$$

- So regression provides a linear summary of the CEF
- We can interpret β_1 as saying: if we compare two individuals whose X differ by 1 unit, we expect their Y to differ by β_1 units **in the data.**
  - ie when we compare two workers whose pollution exposure differs by 1 unit, their wages differ by β_1 units on average.
- However, this does NOT necessarily mean changes X *cause* changes in Y.
- This is because most outcomes of interest have MANY determinants, and those could be driving the results.
  - For example, if some workers exposed to pollution are also less educated, want to make sure we control for that and just measure the impact of pollution

Other things equal

- A major difference between econometricians and statisticians is that economists spend **a lot** of time trying to determine which correlations are actually causal.

Why is this distinction important for policy analysis?

Other things equal

- A major difference between econometricians and statisticians is that economists spend **a lot** of time trying to determine which correlations are actually causal.

Why is this distinction important for policy analysis?

- Most policy and business questions have a causal tone: People want to know "If I do X, what will happen to Y?"
- Implicit in this question is the statement, "holding everything else constant". That's because policy typically only effects one factor.
  - ie a policy to clean up air quality won't change education, neighborhood safety, etc.
  - MA has high electricity prices and OK has low prices. MA can adopt OK's electricity regulation practices; it cannot adopt OK's whether and geography.
- Failure to account for this will lead to biased predictions and poor policy advice.

Omitted variable bias

- Example: imagine that wages have only two determinants: education and whether you live in a polluted area ($D_i = 1$)

$$\text{Wage}_i = \beta_o + \beta_1 D_i + \beta_2 \text{Educ}_i + e_i$$

- What if we try to estimate the impact of moving to a polluted area without conditioning on education? This is equivalent to:
  1. computing two conditional expectations:

$$\begin{aligned} E[\text{Wage}_i|D_i = 0] &= \beta_0 + \beta_1 \times 0 + \beta_2 \times E[\text{Educ}_i|D_i = 0] \\ E[\text{Wage}_i|D_i = 1] &= \beta_0 + \beta_1 \times 1 + \beta_2 \times E[\text{Educ}_i|D_i = 1] \end{aligned}$$

  2. And taking the difference

$$\begin{aligned} \hat{\beta}_1 &= E[\text{Wage}_i|D_i = 1] - E[\text{Wage}_i|D_i = 0] \\ &= \beta_1 + \beta_2(E[\text{Educ}_i|D_i = 1] - E[\text{Educ}_i|D_i = 0]) \end{aligned}$$

Omitted variable bias

True model: $\text{Wage}_i = \beta_o + \beta_1 D_i + \beta_2 \text{Educ}_i + e_i$

Estimate: $\hat{\beta}_1 = \beta_1 + \beta_2(E[\text{Educ}_i | D_i = 1] - E[\text{Educ}_i | D_i = 0])$

- Our estimate now includes the true effect of pollution, β_1, plus the effect of the difference in education across polluted and unpolluted areas.
- Would this cause us to over or underestimate the effect of pollution?
  - If β_2 and $(E[\text{Educ}_i | D_i = 1] - E[\text{Educ}_i | D_i = 0])$ are the same sign, we over-estimate (ie $\hat{\beta}_2 > \beta_2$)
  - If different signs, we under-estimate (ie $\hat{\beta}_2 < \beta_2$)

When is this *not* a problem

True model: $\text{Wage}_i = \beta_o + \beta_1 D_i + \beta_2 \text{Educ}_i + e_i$

Estimate: $\hat{\beta}_1 = \beta_1 + \beta_2 (E[\text{Educ}_i | D_i = 1] - E[\text{Educ}_i | D_i = 0])$

- If $\beta_2 \approx 0$. This is just another way of saying education really doesn't affect wages much, so it's ok if we omit it.
- $E[\text{Educ}_i | D_i = 1] - E[\text{Educ}_i | D_i = 0] = 0$. This is just saying the education is uncorrelated with pollution.
- To summarize, in order to bias our estimate of some included variable D, the omitted variable Z
  1. Has to effect the outcome ($\beta_z \neq 0$)
  2. Has to be correlated with variable we are interested in (D)

Of course, there is only a problem is omit important variables from our regression

We can extend the CEF to condition on many factors

- For example,

$$\text{Wage}_i = \beta_o + \beta_1 D_i + \beta_2 \text{Educ}_i + \beta_2 Industy_i + \beta_4 Age_i + ... + e_i$$

- Now we have many potential sources of omitted variable bias.
- We are still ok as long as all omitted factors are *uncorrelated* with pollution.
- The problem is that it is hard to condition on everything

Always useful to think through what the impact of OVB *might* be, even if you can't fix it

- Lets return to the topic of electricity deregulation.

$$\text{Price}_s = \beta_0 + \beta_1 * \text{Deregulated}_s + \beta_2 X_s + \epsilon_s$$

- What are some factors (X) that determine electricity prices? (independent of regulation)

Always useful to think through what the impact of OVB *might* be, even if you can't fix it

- Lets return to the topic of electricity deregulation.

$$\text{Price}_s = \beta_0 + \beta_1 * \text{Deregulated}_s + \beta_2 X_s + \epsilon_s$$

- What are some factors (X) that determine electricity prices? (independent of regulation)
- Do you think those factors generally cause electricity prices to be higher or lower? (sign of β_2)

Always useful to think through what the impact of OVB *might* be, even if you can't fix it

- Lets return to the topic of electricity deregulation.

$$\text{Price}_s = \beta_0 + \beta_1 * \text{Deregulated}_s + \beta_2 X_s + \epsilon_s$$

- What are some factors (X) that determine electricity prices? (independent of regulation)
- Do you think those factors generally cause electricity prices to be higher or lower? (sign of β_2)
- Do you think X is generally higher or lower in deregulated states?

Always useful to think through what the impact of OVB *might* be, even if you can't fix it

- Lets return to the topic of electricity deregulation.

$$\text{Price}_s = \beta_0 + \beta_1 * \text{Deregulated}_s + \beta_2 X_s + \epsilon_s$$

- What are some factors (X) that determine electricity prices? (independent of regulation)
- Do you think those factors generally cause electricity prices to be higher or lower? (sign of β_2)
- Do you think X is generally higher or lower in deregulated states?
- If correlation and effect β_2 have the same sign, $\hat{\beta}_1 > \beta_1$. If these have opposite signs, the opposite is true.

In class exercise: Do this for your blog post

- Pick *one* question from your blog post. State it as a question about the *causal* relationship between two variables (X and Y):
  - Do EVs cause local electricity prices to increase?
    - Silvia: What's the mechanism here?

In class exercise: Do this for your blog post

- Pick *one* question from your blog post. State it as a question about the *causal* relationship between two variables (X and Y):
  - Do EVs cause local electricity prices to increase?
    - Silvia: What's the mechanism here?
- Imagine we had perfect data on both EV adoption and electricity prices, but nothing else. What are some other factors that could confound a simple empirical comparison?

In class exercise: Do this for your blog post

- Pick *one* question from your blog post. State it as a question about the *causal* relationship between two variables (X and Y):
  - Do EVs cause local electricity prices to increase?
    - Silvia: What's the mechanism here?
- Imagine we had perfect data on both EV adoption and electricity prices, but nothing else. What are some other factors that could confound a simple empirical comparison?
  - Wealthier people buy more EVs. Increases in income in an area could be associated with wage increases and increased electricity demand outside of EVs

In class exercise: Do this for your blog post

- Pick *one* question from your blog post. State it as a question about the *causal* relationship between two variables (X and Y):
  - Do EVs cause local electricity prices to increase?
    - Silvia: What's the mechanism here?
- Imagine we had perfect data on both EV adoption and electricity prices, but nothing else. What are some other factors that could confound a simple empirical comparison?
  - Wealthier people buy more EVs. Increases in income in an area could be associated with wage increases and increased electricity demand outside of EVs

Spend about 5 minutes each on this, helping each other with these two questions. Then we will discuss a few.

In class discussion of some blog posts

- Can Developing Economies Have High Growth Without Using Coal? (Liam)
- How will the Alaskan Willow project effect energy prices? (many)
- What is the impact of EV subsidies on EV sales? (Justin)

Counterfactual Thinking

- all causal policy questions have to rely on some counterfactual comparison
    - What would MA electricity prices be if it re-regulated the electricity sector?(holding everything else unchanged)
- define "treatment" as an indicator (D) for whether or not some unit of observation was exposed to the intervention of interest
    - ie is state i deregulated
- We have a dataset where some units were (weren't) treated, and we observe their outcomes Y
- We want to estimate what Y would have been if had not (had) been treated, holding every thing else constant

Linear regression predicts counterfactual outcomes based on observable factors

- Assume Y is linear and additive in all factors. Assume that the impact of X on outcomes is the same for each group.

$$Y_i = \alpha D + \beta' X_i + e_i$$

- If we omit some important factor X^j, our estimate of the true "treatment effect" α is biased:

$$\hat{\alpha} = \alpha + \beta_j (E[X_i^j | D_i = 1] - E[X_i^j | D_i = 0])$$

- Unfortunately it is often impossible to condition on all important factors
  - classic example is "ability"
- Instead, economists often try to ensure / argue that the omitted factor is uncorrelated with treatment (so the second term is zero)

This observed difference contains both "treatment" and "selection" effects

- Add and subtract $E[Y_{0i}|D=1]$ to the above equation
- This is the expected outcome for treated individuals had they not received treatment (the *counterfactual*)

$$\delta = \underbrace{E[Y_{1i} - Y_{0i}|D=1]}_{Causal \quad Effect} + \underbrace{E[Y_{0i}|D=1] - E[Y_{0i}|D=0]}_{Selection}$$

- The first term is the average causal effect
- The second term is the "selection effect"
  - This captures the fact that the treated and untreated units might have different outcomes absent the treatment.
- So if the goal is to just recover the first term, we need to convince ourselves that the second term is equal to zero.

Experiments

Randomized control trials

- One way around this problem is to run an experiment
- Could take a sample from the population and **randomly** assign them to either a treated state $(D_i = 1)$ or not $(D_i = 0)$
  - By construction: $E[X_i^j | D = 1] = E[X_i^j | D = 0]$ for all X^j
- The counterfactual for each individual is just the average outcome in the other group.
- Comparing these averages now has a clear causal interpretation

Example: Covid vaccines

Search **Bloomberg** Sign In

Prognosis

Vaccine Pregnancy Trials Begin in Bid to Fill Data Void

By Suzi Ring and Naomi Kresge
February 18, 2021, 5:39 AM EST
Updated on February 18, 2021, 2:43 PM EST

- Pfizer-BioNTech set to start testing women in late pregnancy

- Why not just open up vaccine to pregnant women? Some will get them some won't. Can compare outcomes across the two groups.

© Nobel Media. Photo: A. Mahmoud

Abhijit Banerjee

Prize share: 1/3

© Nobel Media. Photo: A. Mahmoud

Esther Duflo

Prize share: 1/3

© Nobel Media. Photo: A. Mahmoud

Michael Kremer

Prize share: 1/3

The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2019 was awarded jointly to Abhijit Banerjee, Esther Duflo and Michael Kremer "for their experimental approach to alleviating global poverty."

Example Electricity RCT

Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use†

By Katrina Jessoe and David Rapson*

Imperfect information about product attributes inhibits efficiency in many choice settings, but can be overcome by providing simple, low-cost information. We use a randomized control trial to test the effect of high-frequency information about residential electricity usage on the price elasticity of demand. Informed households are three standard deviations more responsive to temporary price increases, an effect that is not attributable to price salience. Conservation extends beyond pricing events in the short and medium run, providing evidence of habit formation and implying that the intervention leads to greenhouse gas abatement. Survey evidence suggests that information facilitates learning. (*JEL* D12, D83, L11, L94, Q41, Q54)

How much electricity do you use?

For people living off campus, do you know how much you used last month? Yesterday?

How much electricity do you use?

For people living off campus, do you know how much you used last month? Yesterday?

Let's say you and your roommates got a big bill. How would you go about identifying what to change?

How much electricity do you use?

For people living off campus, do you know how much you used last month? Yesterday?

Let's say you and your roommates got a big bill. How would you go about identifying what to change?

- This is hard because you get one bill a month, and you need to figure out how much comes from each activity /appliance
- In your case many people involved

Motivation: *Why* is electricity demand so inelastic?

- We talked a lot about demand for electricity being very inelastic
- This paper asks if this is because people are not very informed
- Specifically: If people had better information about how much they use, would they be more responsive to price changes (ie more elastic)?

What was the experimental setup?

How were people recruited?

- To be eligible for participation in the pilot a customer needed to reside in a townhouse or single family home, have a broadband Internet connection, and sign and return an end-user agreement indemnifying UI against litigation risk.
- As an additional participation incentive, we offered households \$40 to complete two surveys
- To recruit households into the pilot, UI e-mailed 60,000 customers that had enrolled in paperless billing, indicating the likely presence of Internet in their home.
- Estimate that approximately 7,000 households opened the e-mails.
- Recruited 1,152 households (approximately 1 in 6) to participate in the project
- 437 selected into the final sample.

External validity

- What do people think about this sample? Of 60,000 potential customers, less than 2 percent sign up. What type of person signs up for an experiment like this?

External validity

- What do people think about this sample? Of 60,000 potential customers, less than 2 percent sign up. What type of person signs up for an experiment like this?
- Within this group, treatment will be randomly assigned. However nothing guarantees that the response we see for this group applies to some other population.
    - Within this sample population, estimates are internally valid, unbiased causal effects, because treatment is randomly assigned.
    - But the sample population itself might be unique, such that if an experiment were performed on another population, the estimate there would differ (even though both would internally valid)
- This is probably the most contentious aspect of RCTs

Sample Groups

Control Group: A total of 207 households were assigned to the control group. These received a mailing that notified them they were in the pilot, informed them of their group assignment, and contained an energy conservation pamphlet documenting "101 Ways to Conserve Electricity."

Price-Only Treatment Group: The 130 households in this group experienced pricing events that varied in the magnitude of the price increase and the timing of event notification. Two event days:

1. **Day Ahead (DA):** told the day before that price of electricity would be increased by \$0.50/kWh (250 percent increase)
2. **Thirty Minutes (TM):** sent notification thirty minutes before a \$1.25/ kWh increase

Price + IHD Group: The 100 households get the price treatments above, but also get real-time information on their energy use via in-home display (IHD)

Here are the pricing events

TABLE 1—TREATMENT EVENTS

Event date	Desc	Type	Start hour	High temp	Mean temp	Humidity
07/21/11	4 hr $0.50	DA	12	89	82	75
07/22/11	4 hr $1.25	TM	12	103	90	61
08/04/11	2 hr $0.50	DA	15	80	74	68
08/10/11	2 hr $1.25	TM	16	88	80	63
08/17/11	2 hr $1.25	TM	16	86	75	64
08/26/11	4 hr $0.50	DA	12	84	78	69

What does the display really tell users?

Figure A.1: In-Home Display (1)

Customers view in real time the quantity of power being consumed, the price of electricity, and their estimated monthly bill-to-date.

What do we learn relative to the price only group?

Since treatment groups randomly assigned, can see they are the same

TABLE 2—SUMMARY STATISTICS BY CONTROL AND TREATMENT GROUP

	Control		Price			Price + IHD		
	Mean	Obs.	Mean	Obs.	Difference	Mean	Obs.	Difference
Panel A. Initial group								
Off-peak usage (kWh/h)	1.230	207	1.282	130	0.052	1.225	100	−0.005
	(0.738)		(0.739)		(0.629)	(0.658)		(0.058)
Peak usage (kWh/h)	1.519	207	1.533	130	0.014	1.413	100	−0.106
	(1.197)		(1.036)		(0.109)	(0.984)		(0.772)
TOU Rate (1 = yes)	0.184	207	0.200	130	0.016	0.240	100	0.056
	(0.388)		(0.402)		(0.373)	(0.429)		(1.153)
Home ownership (1 = yes)	0.768	203	0.798	129	0.030	0.773	97	0.005
	(0.423)		(0.403)		(0.641)	(0.42)		(0.091)
Annual income ($1,000)	72.00	203	74.00	129	2.000	71.00	97	−0.001
	(29.00)		(29.00)		(0.690)	(31.00)		(0.181)
Home size (1,000 square feet)	1.529	189	1.880	119	0.351**	1.451	91	−0.078
	(1.10)		(1.83)		(2.100)	(1.14)		(0.550)
Age of home (years)	52.423	156	57.619	97	5.195	52.239	71	−0.184
	(30.29)		(31.34)		(1.309)	(26.94)		(0.044)

We know this isn't all the X's that matter, but if these are "balanced", we infer that randomization worked.
[Note this paper also has some attrition that they have to deal with.]

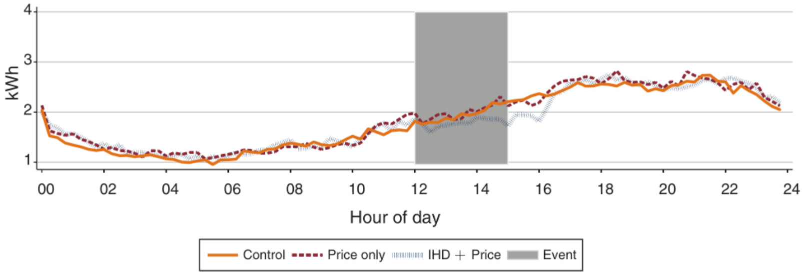


FIGURE 1. JULY 21, 2011: 4HR $0.50 INCREASE, DAY-AHEAD NOTICE

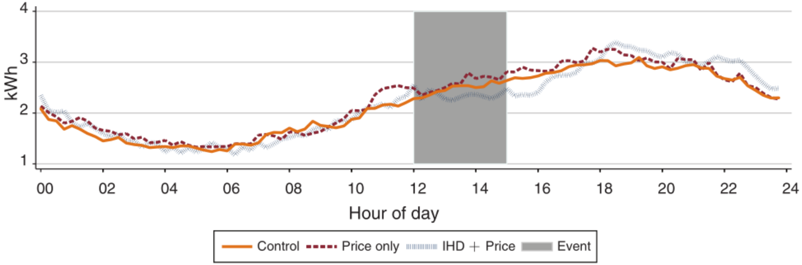


FIGURE 2. JULY 22, 2011: 4HR $1.25 INCREASE, 30-MIN NOTICE

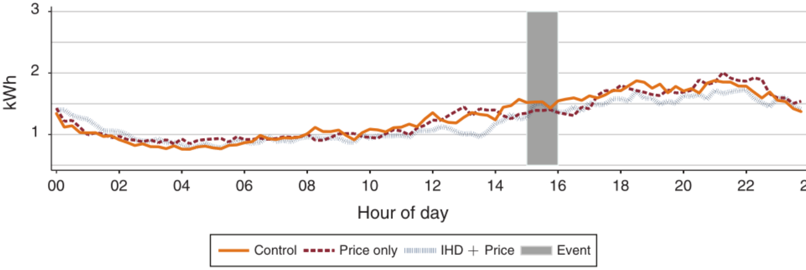


FIGURE 3. AUGUST 4, 2011: 2HR $0.50 INCREASE, DAY-AHEAD NOTICE

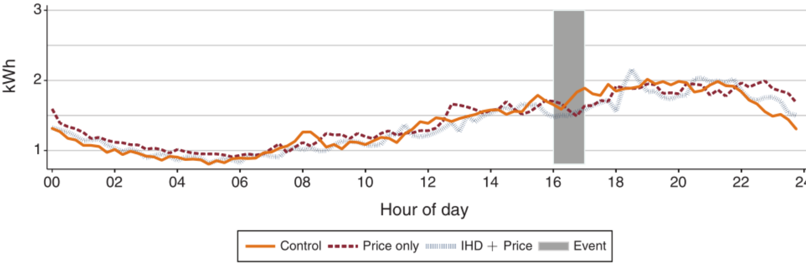


FIGURE 4. AUGUST 10, 2011: 2HR $1.25 INCREASE, 30-MIN NOTICE

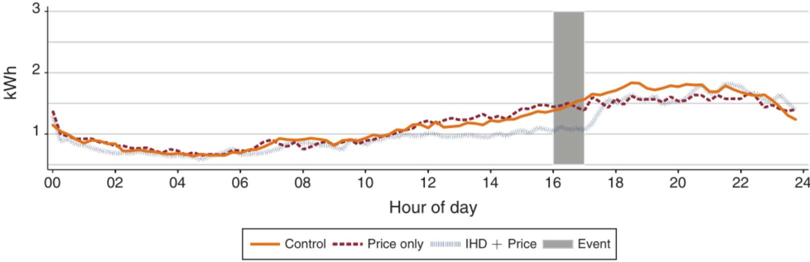


FIGURE 5. AUGUST 17, 2011: 2HR $1.25 INCREASE, 30-MIN NOTICE

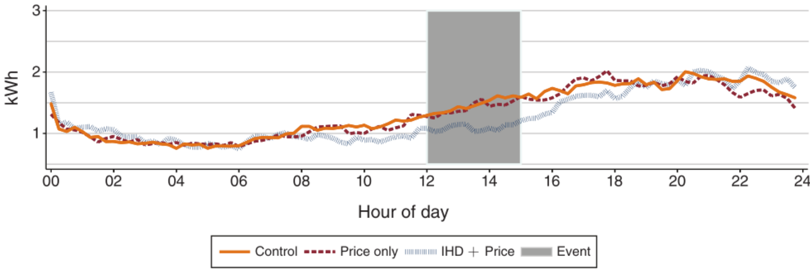


FIGURE 6. AUGUST 26, 2011: 4HR $0.50 INCREASE, DAY-AHEAD NOTICE

With an RCT, can just compare means

TABLE 4—MEAN KWH DIFFERENCES (WRT CONTROL) BY TREATMENT GROUP

		Mean kWh during events			Difference in mean kWh wrt control	
Event type	Variable	Control	Price	Price + IHD	Price	Price + IHD
Sample: Unbalanced panel						
DA	Mean	1.65	1.59	1.35	−0.06	−0.30*
	SD	(1.51)	(1.25)	(1.22)		
	Obs	207	130	100		
TM	Mean	2.07	1.99	1.79	−0.07	−0.28
	SD	(1.77)	(1.54)	(1.42)		
	Obs	186	128	87		

- Price-only group usage declines by 0-7 percent, relative to control.
- IHD group usage declines by 8 to 22 percent.

Having established a difference, then explore the mechanism

- Doesn't seem to be "salience" (ie people just more aware an event is happening)
- People seem to plan more for day ahead event (usage changes in hours before / after)
- Some evidence of habit formation. A result they emphasize is that usage declines in *non-event* hours more among this group as well.

What are some limitations of this experiment?

What are some limitations of this experiment?

- 6 days among 435 selected people in CT
- Would this effect be the same elsewhere?
- Can we really see any trends in one month?
- Would adaptation grow if we did this longer (learning) or fade (fatigue)?

Could we have answered this question without an experiment like this?

- Would need information on how aware people are about electricity consumption (this would likely be selected)
- Would need information on aware vs un-aware people at the same time (utility might have this)
- Seems unlikely this would be credible

Can we simply run RCT's to answer all questions of interest in this class?

How useful is and RCT in the setting you wrote your blog post about?

Natural Experiments

Unfortunately, RCTs are typically not feasible

- Although RCTs are increasingly being used in economics, in many settings they are either unethical or impractical
- Absent a controlled experiment, we are left with what we can recover from data
- We know from above that most of the variation we see in explanatory variables of interest will likely be correlated with other confounding factors
  - Education is correlated with pollution & productivity
- So for many important questions we're going to be forced to use "observational" data. But some sources of variation in D may be more useful than others.

2021 Nobel Prize awarded for showing " what conclusions about cause and effect can be drawn from natural experiments."

© Nobel Prize Outreach. Photo: Paul Kennedy

David Card

Prize share: 1/2

© Nobel Prize Outreach. Photo: Risdon Photography

Joshua D. Angrist

Prize share: 1/4

© Nobel Prize Outreach. Photo: Paul Kennedy

Guido W. Imbens

Prize share: 1/4

2021 Nobel Prize awarded for showing " what conclusions about cause and effect can be drawn from natural experiments."

© Nobel Prize Outreach. Photo: Paul Kennedy

David Card

Prize share: 1/2

© Nobel Prize Outreach. Photo: Risdon Photography

Joshua D. Angrist

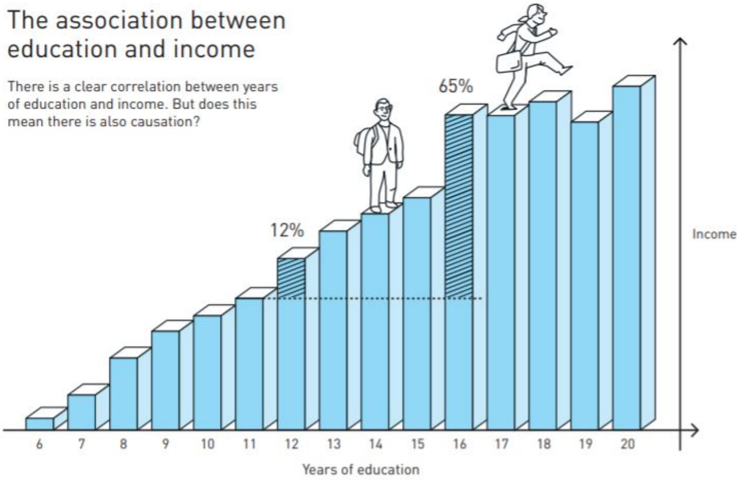
Prize share: 1/4

© Nobel Prize Outreach. Photo: Paul Kennedy

Guido W. Imbens

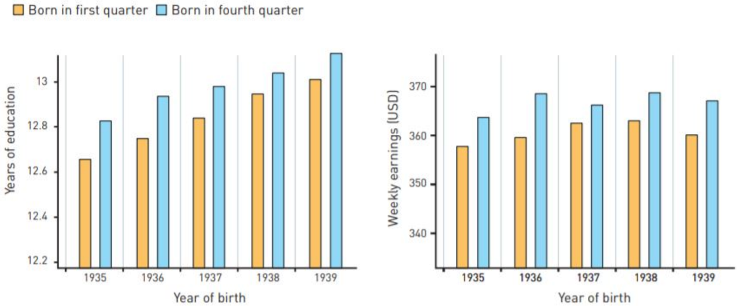
Prize share: 1/4

What is a natural experiment?


The association between education and income
There is a clear correlation between years of education and income. But does this mean there is also causation?
12%
65%
Income
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Years of education

People born late in the year have more years of education and higher incomes

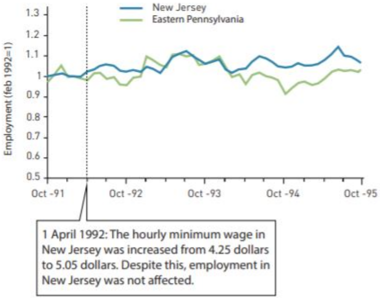
Additional years of education have a positive effect on income. The figure uses data from Angrist and Krueger (1991).

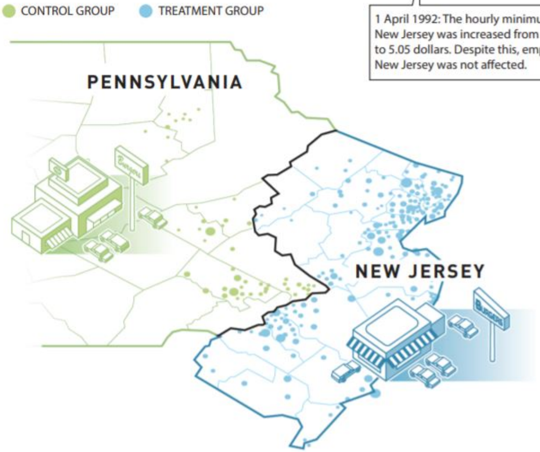


The effect of increasing the minimum wage

Card and Krueger used a natural experiment to study how increasing the minimum wage affects employment.

The researchers identified a treatment group (restaurants in New Jersey) and a control group (restaurants in eastern Pennsylvania) to measure the effect of increasing the minimum wage.





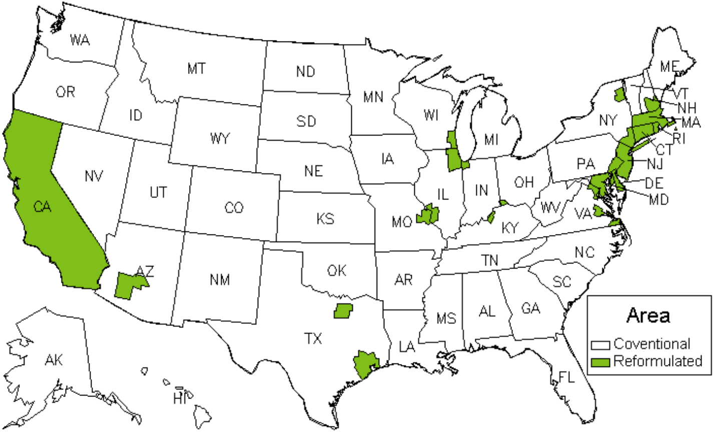
Difference in differences

Difference-in-Differences

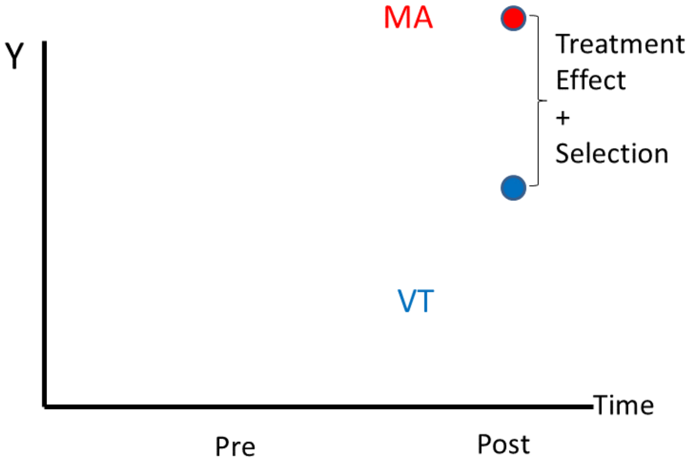
- Cross-sectional comparisons (where we compare treated to untreated units at a given point in time) subject to selection bias.
- Often we have another comparison available: the treated units before the treatment was implemented (or after it was removed)
- This is promising, but potentially confounded by time trends.
- Difference-in-Differences attempts to isolate the causal effect of interest by combining both comparisons

Example: Reformulated Gasoline

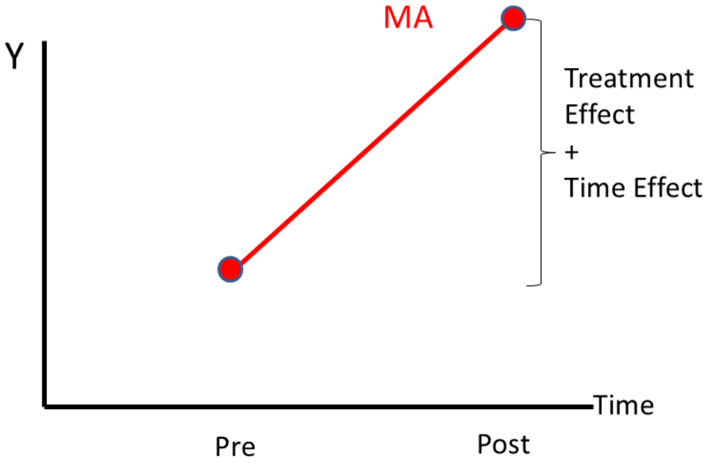
- In 1995 Massachusetts adopted a new clean gasoline (RFG)
  - Vermont did not
- Did RFG increase worker productivity (wages) in MA?



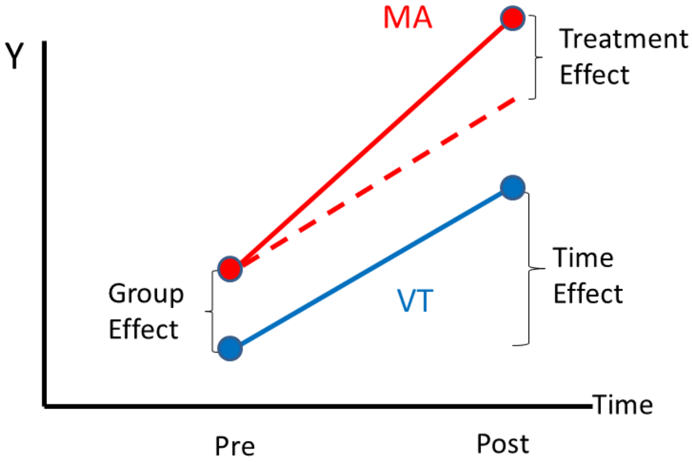
Comparing MA and VT *after* RFG adopted is likely biased by selection



Can compare MA wages pre and post RFG, but many things could have changed across time.



Diff-in-Diff creates a *counterfactual* prediction for MA using VT's pre-post trend.



Difference in differences formally

$$Y_{it} = a + b\text{Treat}_i + c\text{Post}_t + d\text{Treat}_i \times \text{Post}_t + \epsilon_{it}$$

- $\text{Treat}_i = 1$ if you are *ever* in the treated group (ie MA in our example).
- $\text{Post}_t = 1$ if t is in the post treatment period for *either* group.
- $\text{Treat}_i \times \text{Post}_t = 1$ if both are true, ie for the treatment group when the treatment is implemented. (ie MA after RFG)

Difference in differences formally

$$Y_{it} = a + b\text{Treat}_i + c\text{Post}_t + d\text{Treat}_i \times \text{Post}_t + \epsilon_{it}$$

If we compute expectations in the pre period, we get

$E[Y|\text{MA,Pre}] = a + b$

$E[Y|\text{VT,Pre}] = a$

Differencing these gives us the pure selection effect: b

Difference in differences formally

$$Y_{it} = a + b\text{Treat}_i + c\text{Post}_t + d\text{Treat}_i \times \text{Post}_t + \epsilon_{it}$$

If we compute expectations in the post period, we get

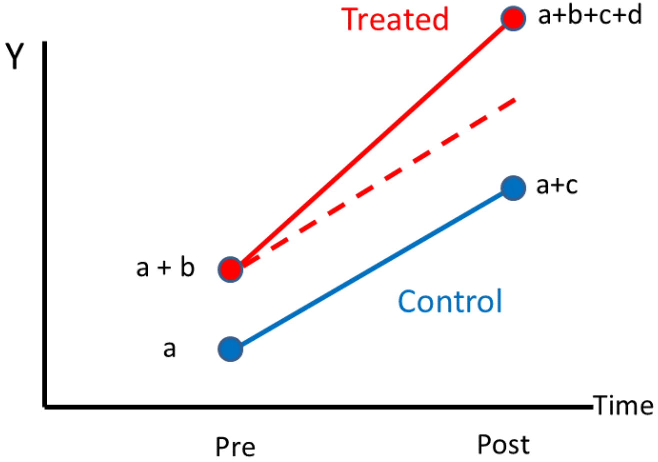
$E[Y|\text{MA,Post}] = a + b + c + d$

$E[Y|\text{VT,Post}] = a + c$

Differencing these removes the post period effect c, and gives $b + d$.

Since we just found b, can difference that to get causal treatment effect d.

Mapping regression to the picture



Diff-in-Diff Assumptions

- The key assumption for diff-in-diff to recover an unbiased estimate of the causal effect is parallel trends
- This is really sort of two assumptions:
  1. Absent the intervention, the pre vs post difference across the two groups would have been identical (ie VT and MA wages both would have grown by c)
  2. Implicitly, this says that selection across the two groups is time invariant. All the things that make MA and VT wages different are the same in each period (b)

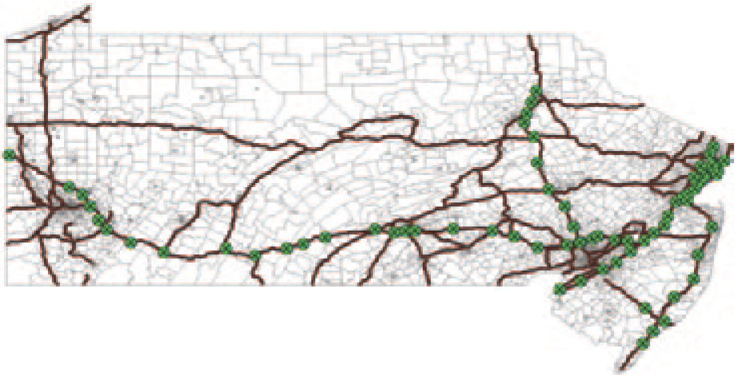
When are DiD estimates truly causal?

- Difference in differences controls for time invariant observation characteristics, and time varying factors that affect all observations similarly
- Before simply accepting these answers, we might question what is generating variation over time?
  - Why are plants opening/ closing in some cities not others?
  - Why do cities decide to suddenly enact new policies?
- Many changes are intentional, but sometimes variables we care about shift for exogenous reasons.
- Goal is to look for such natural experiments

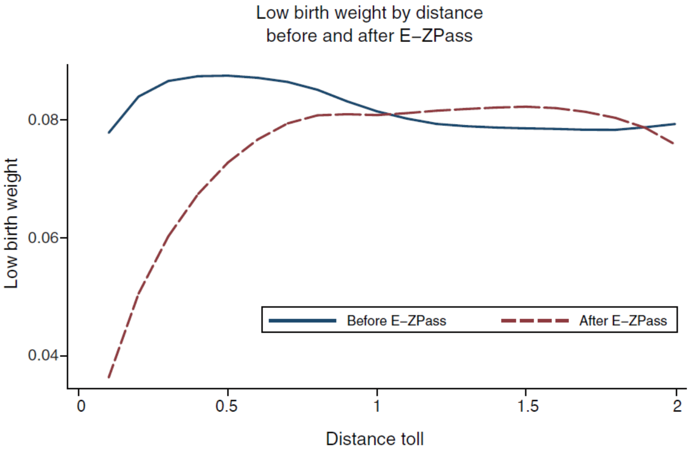
Traffic Congestion and Infant Health: Evidence from E-ZPass

- How does air pollution affect infant health?
- Lots of reasons to suspect the effect is large..
- But it's also clear that infant pollution exposure is not randomly assigned.
- Currie and Walker (2008) take advantage of the introduction of E-Zpass in New Jersey and Pennsylvania
- E-Zpass significantly reduced the time cars had to spend idling near tolls
- They compare infant health across households near and far to tolls before and after the program

Locations of Toll Plazas and Major Roadways in New Jersey and Pennsylvania



Infants born near tolls were much less likely to have low birth weight after E-ZPass



Types of economic datasets

cross-section: you see many units (individuals) at the same time

time-series: you see one unit many time periods

repeated cross-section: you see many periods and units, but not the same units over time

panel data: you see the same units many periods

Diff-in-diff logic extends natural to panel data

- Can have some units treated at different times (ie states adopt laws at different times), some never treated (controls)
- Picture is more complicated but the logic is the same.
- Create vector of dummy variables $\mu_{\mathbf{t}}$ for each time period t.
- Create another vector of dummy variables $\eta_{\mathbf{i}}$ for unit i.
- Create a treatment indicated $D_{it} = 1$ if unit i is treated in period t.

$$Y_{it} = \alpha_0 + \beta D_{it} + \mu_{\mathbf{t}} + \eta_{\mathbf{i}} + \epsilon_{it}$$

This is called fixed effects regression, and Stata/R will make the dummies for you

Example: 1600 "treatments"

Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings†

By Janet Currie, Lucas Davis, Michael Greenstone, and Reed Walker*

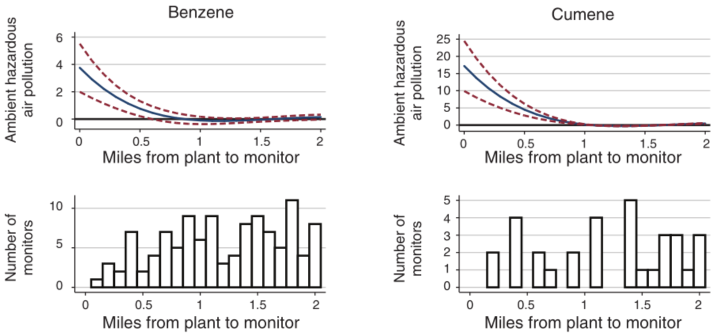
Regulatory oversight of toxic emissions from industrial plants and understanding about these emissions' impacts are in their infancy. Applying a research design based on the openings and closings of 1,600 industrial plants to rich data on housing markets and infant health, we find that: toxic air emissions affect air quality only within 1 mile of the plant; plant openings lead to 11 percent declines in housing values within 0.5 mile or a loss of about $4.25 million for these households; and a plant's operation is associated with a roughly 3 percent increase in the probability of low birthweight within 1 mile. (*JEL* I12, L60, Q52, Q53, Q58, R23, R31)

Empirical Model

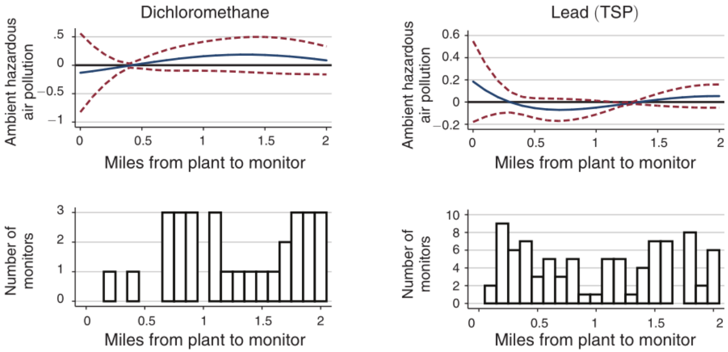
$$
(3) \quad \begin{aligned} Y_{jdt} = {} & \beta_0 + \beta_1 1[\textit{Plant Operating}]_{jt} + \beta_2 1[\textit{Near}]_{jd} \\ & + \beta_3 \left(1[\textit{Plant Operating}]_{jt} \times 1[\textit{Near}]_{jd}\right) + \eta_{jd} + \tau_t \\ & + \beta_4 (X1990_{jd} \times T_t) + \varepsilon_{jdt}, \end{aligned}
$$

where Y_{jdt} denotes the natural log of average housing values near plant site j, within distance group d, in year t. For each plant j, there are two observations per year. In each plant-year, one observation consists of average housing prices "near" a plant (i.e., within 0.5, 0.5 to 1.0, or 1 mile of the plant). The second observation per plant-year consists of average house prices for houses within 1–2 miles of the plant; this second group provides a counterfactual for housing prices near the plant. The availability of these two groups allows for a difference-in-differences-style estimator.

How far is far enough to be the control?



How far is far enough to be the control?



Average effect over 84 pollutants

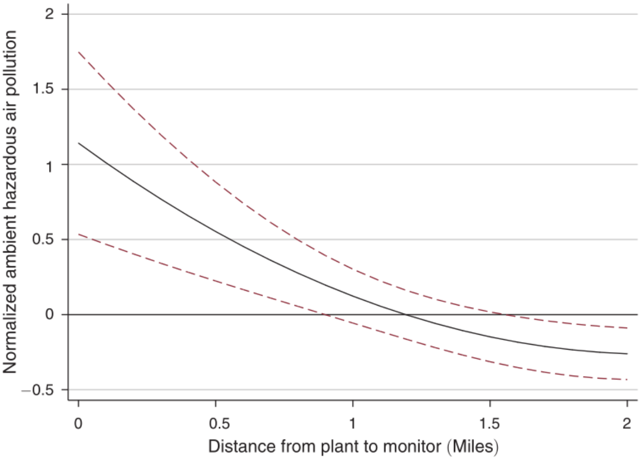


FIGURE 2. THE EFFECT OF TOXIC PLANTS ON AMBIENT HAZARDOUS AIR POLLUTION, ALL POLLUTANTS

Note that most previous studies used county level data, which made it impossible to detect these effects.

House price event study

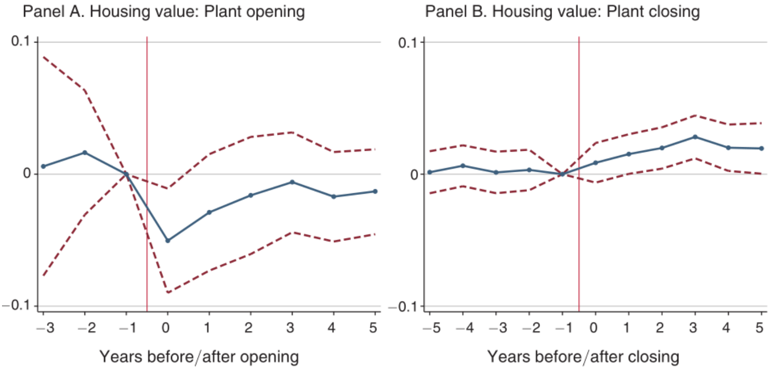


FIGURE 3. EVENT STUDY: THE EFFECT OF TOXIC PLANT OPENINGS AND CLOSINGS ON LOCAL HOUSING VALUES

	0–0.5 Miles		0.5–1 Miles		0–1 Miles	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Estimated effect of plant operation						
1(Plant Operating) × Near	−0.030***	−0.022***	−0.010**	−0.012***	−0.015***	−0.014***
	(0.007)	(0.006)	(0.005)	(0.004)	(0.005)	(0.004)
Observations (plant-distance-year cells)	34,736	34,736	34,736	34,736	34,736	34,736
Plant × distance-bin FE	X	X	X	X	X	X
State × year FE	X		X		X	
Plant × year FE		X		X		X
Panel B. First difference: Estimated effect of plant operation						
1(Plant Operating) × Near	−0.020**	−0.014**	−0.008*	−0.003	−0.010**	−0.005
	(0.010)	(0.007)	(0.004)	(0.004)	(0.005)	(0.004)
Observations	1,114,248	1,114,248	1,305,780	1,305,780	1,375,751	1,375,751
Panel C. First difference: Estimated effect of plant openings and closings						
1(Plant Opening) × Near	−0.096***	−0.107***	−0.007	−0.008	−0.020	−0.022
	(0.036)	(0.034)	(0.023)	(0.020)	(0.022)	(0.019)
1(Plant Closing) × Near	0.017	0.010	0.008	0.003	0.010*	0.005
	(0.011)	(0.009)	(0.005)	(0.004)	(0.006)	(0.005)
H_0: Opening = −Closing (p-value)	0.051	0.013	0.968	0.827	0.688	0.438
Observations	1,114,248	1,114,248	1,305,780	1,305,780	1,375,751	1,375,751
State × year fixed FE	X		X		X	
County × year FE		X		X		X

Birthweight event study

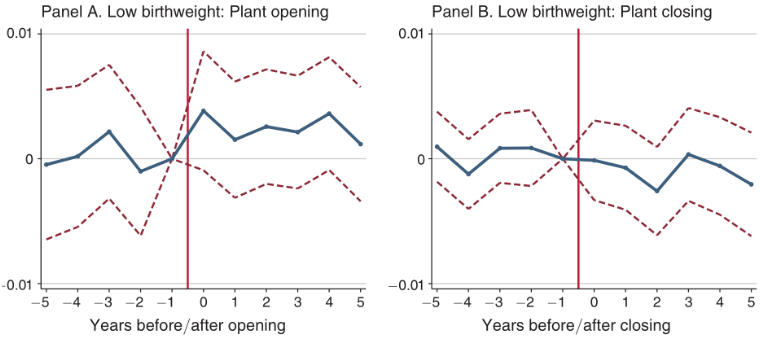


FIGURE 4. EVENT STUDY: THE EFFECT OF TOXIC PLANT OPENINGS AND CLOSINGS ON THE INCIDENCE OF LOW BIRTHWEIGHT

	0–0.5 Miles		0.5–1 Miles		0–1 Miles	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Estimated effect of plant operation						
1(Plant Operating)	0.0010	0.0012	0.0014**	0.0015**	0.0013**	0.0014**
× Near	(0.0010)	(0.0012)	(0.0006)	(0.0006)	(0.0006)	(0.0007)
Observations	88,958	88,958	88,958	88,958	88,958	88,958
Plant count	3,438	3,438	3,438	3,438	3,438	3,438
Panel B. Estimated effect of plant openings and closings						
1(Plant Opened)	0.0025	0.0022	0.0024***	0.0027***	0.0024**	0.0024***
× Near	(0.0019)	(0.0018)	(0.0009)	(0.0010)	(0.0009)	(0.0008)
1(Plant Closed)	−0.0002	−0.0007	−0.0009	−0.0009	−0.0007	−0.0009
× Near	(0.0016)	(0.0016)	(0.0009)	(0.0010)	(0.0009)	(0.0009)
H_0: Opening = −Closing (p-value)	0.44	0.56	0.32	0.28	0.22	0.24
Observations	88,958	88,958	88,958	88,958	88,958	88,958
Plant count	3,438	3,438	3,438	3,438	3,438	3,438
Plant × Distance-bin FE	X	X	X	X	X	X
State × Year FE	X		X		X	
Plant × Year FE		X		X		X

Davis and Wolfram (2012)

"Deregulation, Consolidation, and Efficiency: Evidence from US Nuclear Power"

What is the research question?

- What motivates this question?

Taking this question to the data

- What data do they have?

Taking this question to the data

- What data do they have?
    - 40 years of data on nuclear plants
- What are the main outcome and explanatory variables?

Taking this question to the data

- What data do they have?
  - 40 years of data on nuclear plants
- What are the main outcome and explanatory variables?
  - Efficiency measure: "capacity factor"
  - Treatment: some plants privatized in the 90's

Taking this question to the data

- What data do they have?
  - 40 years of data on nuclear plants
- What are the main outcome and explanatory variables?
  - Efficiency measure: "capacity factor"
  - Treatment: some plants privatized in the 90's

Empirical Model:

$$Output_{it} = \beta_0 + \beta_1[Divested]_{it} + \delta_i + \omega_t + \epsilon_{it}$$

- Plant fixed effects δ_i capture all time invariant difference across plants.
- Time fixed effects ω_t capture all time varying factors share by all plants.
- Divestiture "treatment" dummy
  - Hypothesis $\beta_1 > 0$: Private ownership leads to lover costs/ more efficient operation.

Are deregulated plants different?

TABLE 1—COMPARING DIVESTED WITH NONDIVESTED NUCLEAR REACTORS

	Reactors Divested 1999–2007 ($n = 48$) (1)	All other reactors ($n = 55$) (2)	p-value (1) versus (2) (3)
Mean Reactor Characteristics			
Design capacity (in MWe)	921.9	959.7	0.38
Reactor age as of December 1998	18.8	18.4	0.74
Number of reactors operated by the same reactor operator as of December 1998	3.8	4.0	0.67
Reactor Type, share that are:			
Pressurized water reactor	0.54	0.78	0.01
Boiling water reactor	0.46	0.22	0.01
Reactor Manufacturer, share made by:			
Westinghouse	0.42	0.51	0.35
General Electric	0.46	0.22	0.01
Combustion Engineering	0.08	0.18	0.15
Babcock and Wilcox	0.04	0.09	0.33
Reactor Location, share in:			
Northeast census region	0.50	0	< 0.01
Midwest census region	0.38	0.18	0.03
South census region	0.13	0.67	< 0.01
West census region	0	0.15	0.01

Notes: The sample includes all 103 nuclear power reactors operating in the United States as of January 1, 2000. Column 3 reports p-values from tests that the means are equal in the two

Graphical evidence

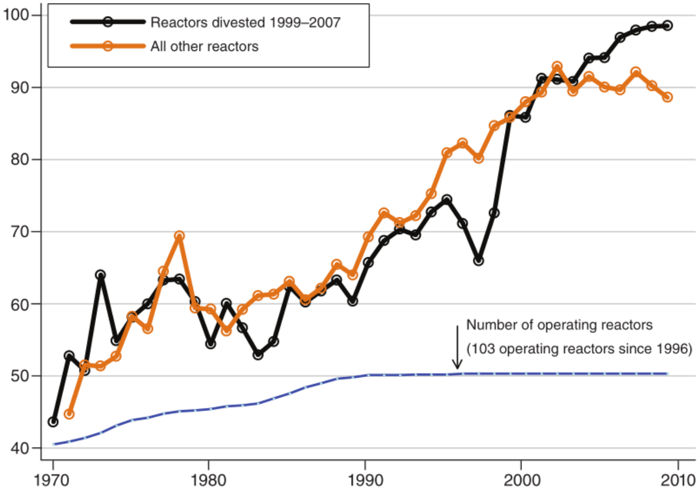


FIGURE 1. NET GENERATION SCALED BY REACTOR DESIGN CAPACITY

Estimates

TABLE 2—THE EFFECT OF DIVESTITURE ON NUCLEAR OPERATING PERFORMANCE

	(1)	(2)	(3)	(4)	(5)
$1[Divested]_{it}$	6.3***	10.2***	10.0***	10.1***	9.5***
	(1.2)	(2.1)	(2.1)	(2.1)	(2.0)
Month-of-sample fixed effects (480 months)	Yes	Yes	Yes	Yes	Yes
Reactor fixed effects (103 reactors)	No	Yes	Yes	Yes	Yes
Reactor age (cubic)	No	No	Yes	Yes	Yes
Observations weighted by reactor capacity	No	No	No	Yes	No
Dataset collapsed to plant level	No	No	No	No	Yes
Number of cross sectional units	103	103	103	103	65
Observations	36,667	36,667	36,667	36,667	23,796
R^2	0.18	0.22	0.22	0.22	0.26

Notes: This table reports coefficient estimates and standard errors corresponding to an indicator variable for reactors that have been divested from five separate regressions. In all regressions the dependent variable is net generation as a percent of design capacity. The sample includes monthly observations 1970–2009 for all 103 nuclear power reactors operating in the United States as of January 1, 2000. Standard errors are clustered at the plant level.

Robustness

TABLE 5—CONSIDERING POSSIBLE CONCERNS ABOUT SELECTION BIAS

	Excluding Michigan (1)	Excluding California (2)	Excluding Iowa and Wisconsin (3)	Divest date 1/2001 for all reactors (4)	Excluding the north-east census region (5)	Propensity score weighting (6)
$1[Divested]_{it}$	9.5***	10.3***	10.1***	7.7***	9.9***	10.9***
	(2.1)	(2.1)	(2.1)	(2.5)	(2.6)	(3.1)
Month-of-sample fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Reactor fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Reactor age (cubic)	Yes	Yes	Yes	Yes	Yes	Yes
Number of reactors	99	99	99	103	79	71
Observations	35,459	35,155	34,905	36,667	27,825	25,484
R^2	0.23	0.22	0.23	0.22	0.21	0.22

Notes: This table reports coefficient estimates and standard errors corresponding to an indicator variable for reactors that have been divested from six separate regressions. In all regressions, the dependent variable is net generation as a percent of design capacity. The sample includes monthly observations for the period 1970–2009 for all nuclear power reactors operating in the United States as of January 1, 2000, excluding reactors or reactor-month observations as indicated in the column headings. Standard errors are clustered at the plant level.

Mechanism

TABLE 7—UNDERSTANDING THE MECHANISMS BEHIND POST-DIVESTITURE GAINS

	(1)	(2)	(3)
A. Maximum generating capacity			
Maximum generation over last 12 operating months	2.4***	1.5	1.6
[Sample mean: 100.4]	(0.9)	(1.5)	(1.4)
Maximum licensed thermal capacity (MWt)	1.8	2.0*	1.9*
[Sample mean: 102.0]	(1.1)	(1.1)	(1.1)
B1. Operating days			
$1[\mathit{Operating}]_{it} \times 100$	3.9***	3.5*	3.8**
[Sample mean: 91.0]	(0.7)	(2.0)	(1.9)
B2. Length versus number of outages			
Number of outages per year	–0.17	–0.13	–0.13
[Sample mean: 1.7]	(0.11)	(0.16)	(0.16)
Mean outage length in days	–6.4***	–6.2	–6.9
[Sample mean: 19.1]	(1.3)	(5.4)	(5.3)
C. Capacity factor when operating			
Capacity factor in percent excluding zeros	–0.3	0.5	0.4
[Sample mean: 97.7]	(0.3)	(0.3)	(0.3)
Time effects (4,017 days/11 years)	Yes	Yes	Yes
Reactor fixed effects (103 reactors)	No	Yes	Yes
Reactor age (cubic)	No	No	Yes

Why do we care?

Why do we care?

- Increased productivity worth about $2.5 billion annually
  - this required no investment
- Nuclear does not emit CO2
- Any concerns about this?

American Economic Journal: Economic Policy 2014, 6(3): 178–206
http://dx.doi.org/10.1257/pol.6.3.178

Corporate Incentives and Nuclear Safety†

By Catherine Hausman*

Following electricity market restructuring, approximately half of all commercial US nuclear power reactors were sold by price-regulated public utilities to independent power producers. At the time of the sales, some policymakers raised concerns that these corporations would ignore safety. Others claimed that the sales would bring improved reactor management, with positive effects on safety. Using data on various safety measures and a difference-in-differences estimation strategy, I find that safety improved following ownership transfers and the removal of price regulations. Generation increased, and this does not appear to have come at the cost of public safety. (*JEL* D24, L51, L94, L98, Q42, Q48)

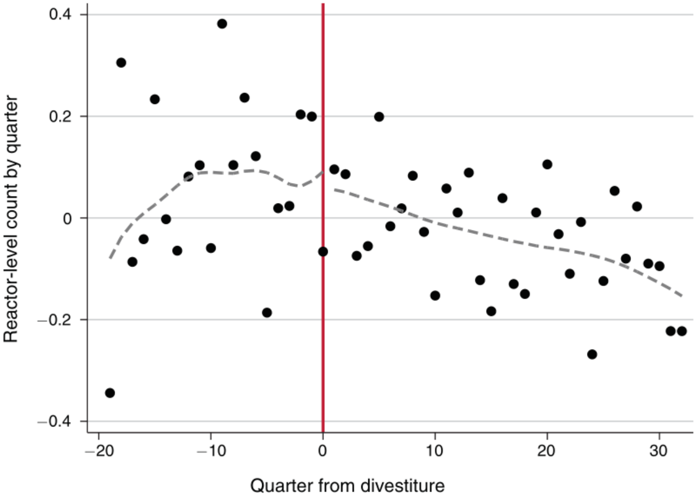


FIGURE 1. EFFECT OF DIVESTITURE ON UNSAFE EVENTS, QUARTERLY EVENT STUDY

Notes: This figure plots unsafe events at divested units relative to nondivested units. Time is normalized relative to divestiture. Unsafe events include initiating events, fires, and escalated enforcement. The median divestiture is in 2001. Quarter-of-sample effects have been removed.

Summary on difference-in-differences

- Requirements: need at least two periods of data for at least two groups of observations, one of which gets treated in the middle.
- Can easily accommodate many years, treatments at different times etc
- Key assumption is no parallel trends. With a long panel, easy to just check this by seeing if the two groups move similarly *before* the treatment.

Other Methods

Instrumental variables

- Remember, the problem stems from the fact that there is an omitted variable that is
  - correlated with the independent variable of interest
  - **and** correlated with the outcome of interest.
- Imagine that we knew something (Z) was correlated with X but not correlated with Y.
- Could just use that variation in to learn about the effect of interest.
  - 1: Regress X_i on Z_i
  - 2: Regress Y_i on predicted value of $\hat{X}$.
  - Since the predicted X is only a function of Z, and Z does not effect Y, we are now all set.
- Key assumptions:
  - $Cov(X, Z) \neq 0$
  - $Cov(e, Z) = 0$

An IV example

- Irrigation dams are one of the largest and most common public infrastructure projects in the developing world.
- A natural question is whether they benefit the surrounding economy
- Evaluation is complicated by the fact that dams are not randomly located
  - what are some stories here?
- Dufflo and Pande (2007) note that only certain terrain can accommodate dams
- They therefore use land gradient as an instrument for dam placement in India
- Assumption: gradient affects the probability of a dam being located, but does not otherwise influence economic activity

Does Air Quality Matter? Evidence from the Housing Market (Chay & Greenstone 2005)

- Want to infer benefits of the Clean Air Act from changes in housing prices due to changes in pollution in surrounding areas (hedonic approach)
- 30 years of cross-sectional studies find only very weak associations
  - why might that be?
- Clean Air Act imposed strict pollution reduction requirements for non-attainment counties
- Chay & Greenstone use attainment status as an instrument for pollution changes
- Find significant negative effect on housing prices

Airports, Air Pollution, and Contemporaneous Health (Schlenker and Walker 2015)

- What are the short term impacts of pollution on human health?
  - Surprisingly little evidence after all these years
- Look at some of the largest polluting facilities in the US, airports
- Major source of variation comes from idling/ taxiing on the runway
- S&W's intuition: much of this variation is due to congestion/ delay at *other airports*
- Use delay conditions at connected airports as an instrument for time spent idling/ taxiing
- Find that a 1 s.d. increase in daily pollution leads to a $540 thousand increase in hospitalization costs for respiratory and heart related admissions for people living near airports in California.

Regression discontinuity

- In many settings, there is a sharp cutoff point around which very similar entities are assigned differing treatments
  - Many state schools offer guaranteed admission / aid if students score above a certain level on the SAT
- RD intuition is that people very close to the cutoff on either side are very similar
  - a student that gets a 999 on the SAT is basically identical to one that gets a 1000
- $\underbrace{E[Y_{0i}|D=1, R<r] = E[Y_{0i}|D=0, R<r]}$
  - R is the absolute distance to the cutoff

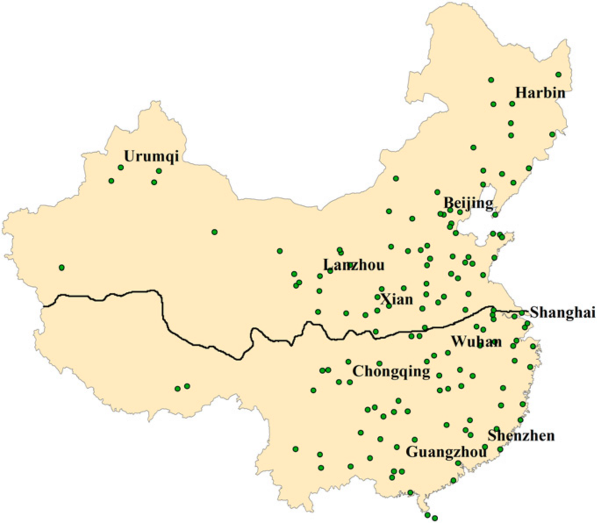
Does hazardous waste matter? Evidence from the Superfund program (Greenstone & Gallagher 2008)

- How much is WTP to avoid hazardous waste?
  - what is ovb here?
- Looks at the Superfund program (CERCLA)
- Funded to clean up 400 sites
  - Sites chosen by an internal scorecard
- G&G got access to these scores, and looked at areas that just missed the cut
- Find small, insignificant effects

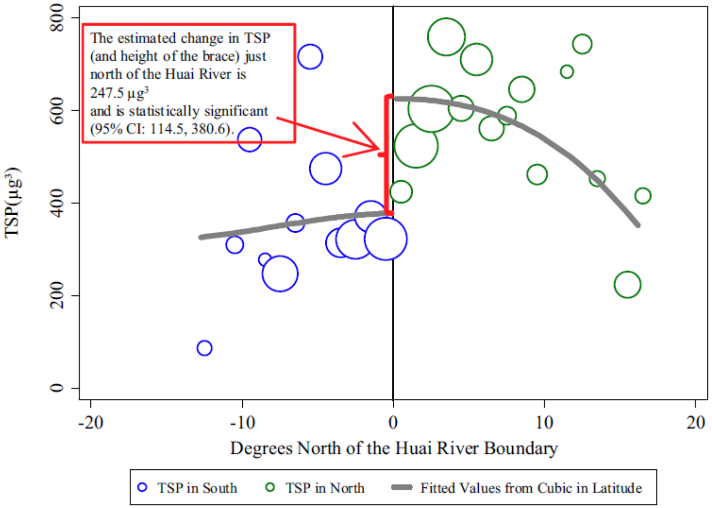
What are the long term impacts of air pollution on life expectancy? (Chen et. al 2013)

- From 1950-1980, China supplied free winter heating to homes north of the Huai River
- Heating was provided via extremely dirty coal boilers
- Authors look at change in life expectancy around this arbitrary border decades later

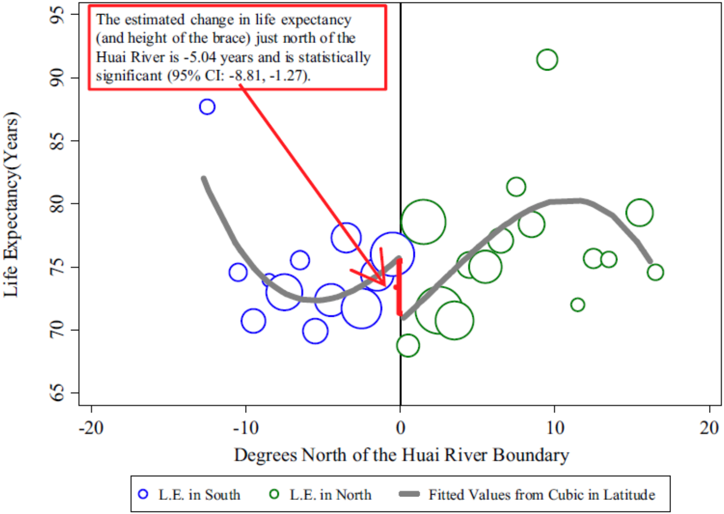
What are the long term impacts of air pollution on life expectancy? (Chen et. al 2013)



Pollution increased significantly north of the river



And life expectancy decreases significantly north of the river



Summary

Summary - not all empirical approaches are equal

"We can make the agreement or disagreement between theory and facts depending on two things: the facts we chose to consider, as well as our theory about them"

- Trygve Haavelmo (1944)

Summary - our approach

- Translate policy question into an X and aY
- Think about the ideal experiment you'd like to run if you could
  - what would the population be? what would you vary?
  - if you can run this experiment, stop here and go do it...
- If you can't run that experiment, what can you learn from data?
- What data is available?
  - do you have a panel or a cross section?
  - what generates variation in X in your data?
- If you control for other variables, can you convince yourself that X is essentially randomly assigned?
  - If not, are there any natural experiments you can take advantage of?
  - What else changes X, is it correlated with Y?
  - What are the rules for how X is assigned? Is there an RD?

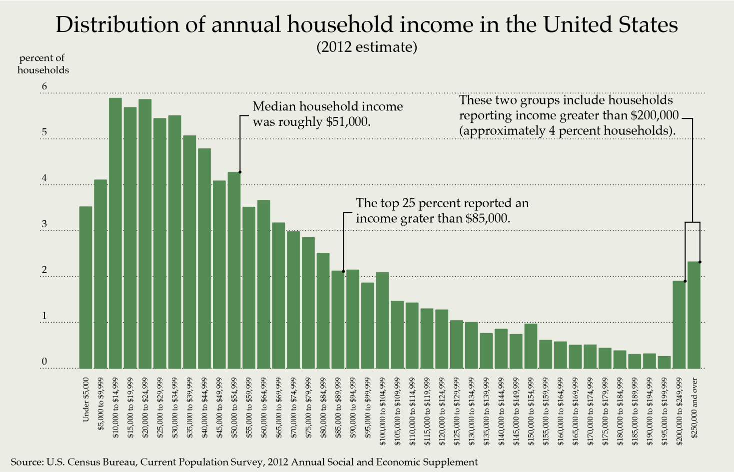
Backup

BACKUP

Inference: Learning from Data

- Say we want to estimate the impact of pollution on labor productivity.
- Imagine we have a sample of wages from N workers.
- Wages have many determinants, and our data will therefore contain sampling variation.
- What can we say about the true average level of productivity in the population from this sample?

Distribution of wages in the US is not normally distributed



Source: VOX

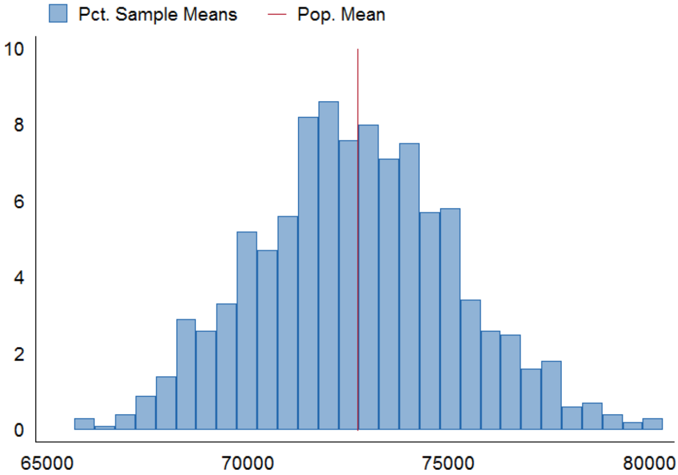
Nevertheless, the **sampling distribution** of $E[Y_i]$ will be normal

Sampling

- Let $n = 1000$
- Denote the mean from a given sample $\bar{Y}$
  - We refer to this sample mean as a sample *statistic*
- Under random sampling, the expectation of this sample statistic is an *unbiased estimator* of the population parameter we are after
  - $E[\bar{Y}] = E[Y_i]$

Central Limit Theorem

- If we draw 10,000 samples, each with sample size $n = 1000$, and plot the distribution



Sampling variance

- Notice that even though we drew these samples randomly from the same distribution, the sample means are not always equal to the true population mean.
- This sampling variance is a function of the population variance and the sample size

$$V(\bar{Y}) = \frac{\sigma_Y^2}{n}$$

- In practice, we typically work with the *standard error*

$$SE(\bar{Y}) = \frac{\sigma_Y}{\sqrt{n}}$$

- Of course, we don't know the population standard deviation, so we estimate it with the sample standard error:

$$\hat{SE}(\bar{Y}) = \frac{S(Y_i)}{\sqrt{n}}$$

Hypothesis testing

- Why do we care about this?
- Suppose we believe the data we have comes from a population with mean μ
  - This is our *hypothesis*
- Then we can construct something called a *t-statistic*

$$t(\mu) = \frac{\bar{Y} - \mu}{\hat{SE}(\bar{Y})}$$

  which has a distribution very close to standard normal (ie $N(0,1)$)
- This means that we can say how likely it is that we would get a sample at least as unusual as $\bar{Y}$ if the hypothesis were true.
- In economics, we often want to test if $\mu = 0$
  - $t > 1.96$ means that the probability we'd get that result by chance is less than 5%

Comparing groups

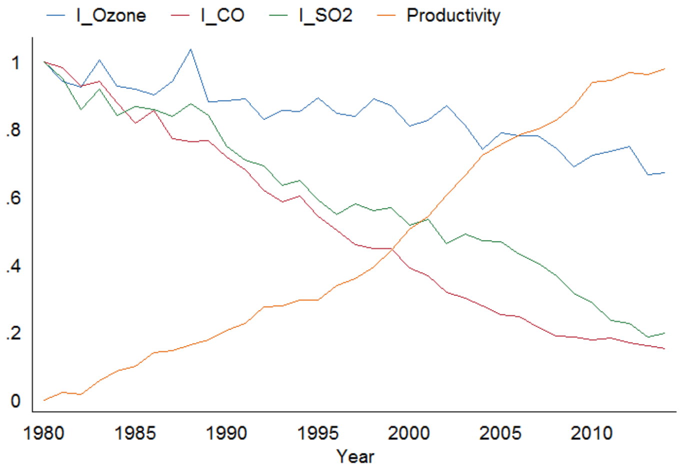
- Often we are interested in comparing whether two groups are the same
  - Are workers in Austin and Houston equally productive?
- Our hypothesis is $\mu = \bar{\mu}_A - \bar{\mu}_H = 0$
- We need a slightly different (typically pooled) standard error:

$$\hat{SE}(\bar{Y_A} - \bar{Y_H}) = S_P(Y_i)\sqrt{\frac{1}{n_A} + \frac{1}{n_H}}$$

- *t-statistic*

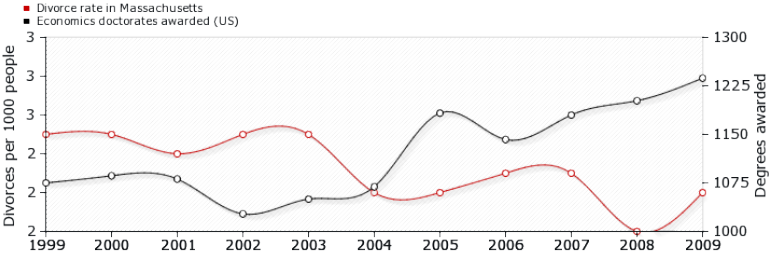
$$t(\mu) = \frac{\bar{Y_A} - \bar{Y_H} - \mu}{\hat{SE}(\bar{Y_A} - \bar{Y_H})}$$

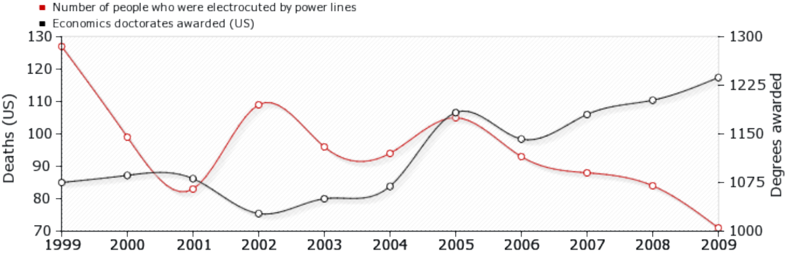
Pollution has declined considerably in the US, while productivity has increased



So can we use this graph as evidence that pollution harms productivity?

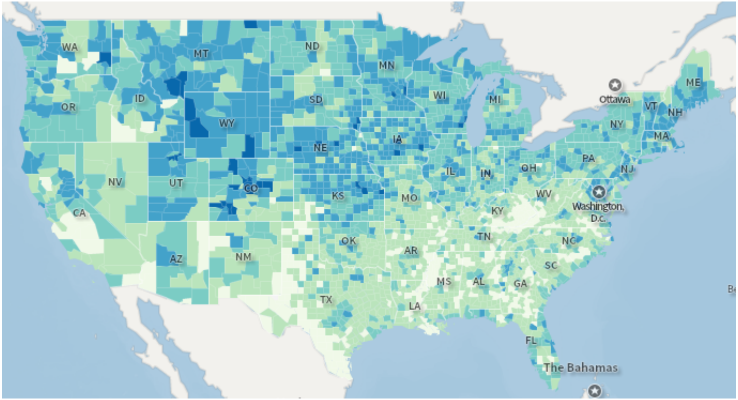
Be weary of comparing two trending variables





Source: Spurious correllations

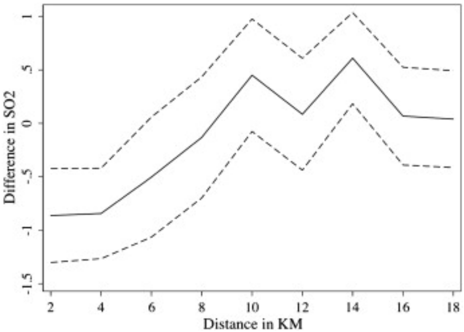
Pollution also varies considerably across the US at any given point in time



Could combine with productivity data from each Census tract.
Would that give us a causal estimate?

The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City (Hanna and Oliva 2015)

- Does pollution reduce labor supply?
- If anything, omitted variable bias might go the other way here
- Hanna and Oliva (2015) look at what happens to hours worked after one of the dirtiest refineries in the world closed in Mexico City in 1991.



Workers near the refinery saw a greater increase in hours after its closure