

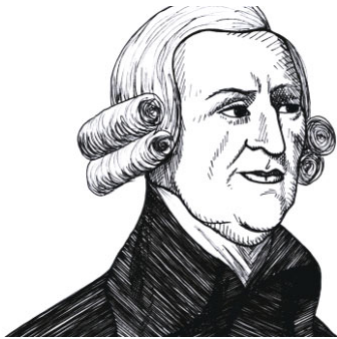
Cause and effect
and
Evidence-based policies

2024 Mackintosh
Lecture

David Card
UC Berkeley

Economics and policy advice

- from the beginning of the field, economists have proposed and argued over alternative public policies:



Adam Smith – argued against mercantilism
David Ricardo – argued for free trade
Francis Walker (founder of AEA) – argued
for restricting immigration

- the advice of the classical and neoclassical economists were largely based on models and assumptions
- but in the past 50 years, a gradual shift has occurred toward evidence-based policy recommendation

Evidence based policy advice

- evidence-based policy requires **evidence** on underlying causal questions. Example:
how does arrival of migrants affect native labor market opportunities?

today's agenda:

- scientific evidence
- randomized experiments and counterfactuals
- Social experiments
- natural experiments (non-experimental designs)
- applications to immigration and minimum wages
- are policy choices really affected by evidence?

Scientific evidence

- high school students are taught about the “scientific method” – propose and test hypotheses
- classic example: Boyle’s air bell experiment (pumping air from the jar causes the bird to die)

⇒single decisive measurement



- famous 20th C example: effect of gravity on light
- but in many areas, there is no single decisive measurement: too much natural variability in outcomes

Scientific evidence (2)

- if outcomes vary a lot, we could “treat” a larger group, and look at the *average outcome*

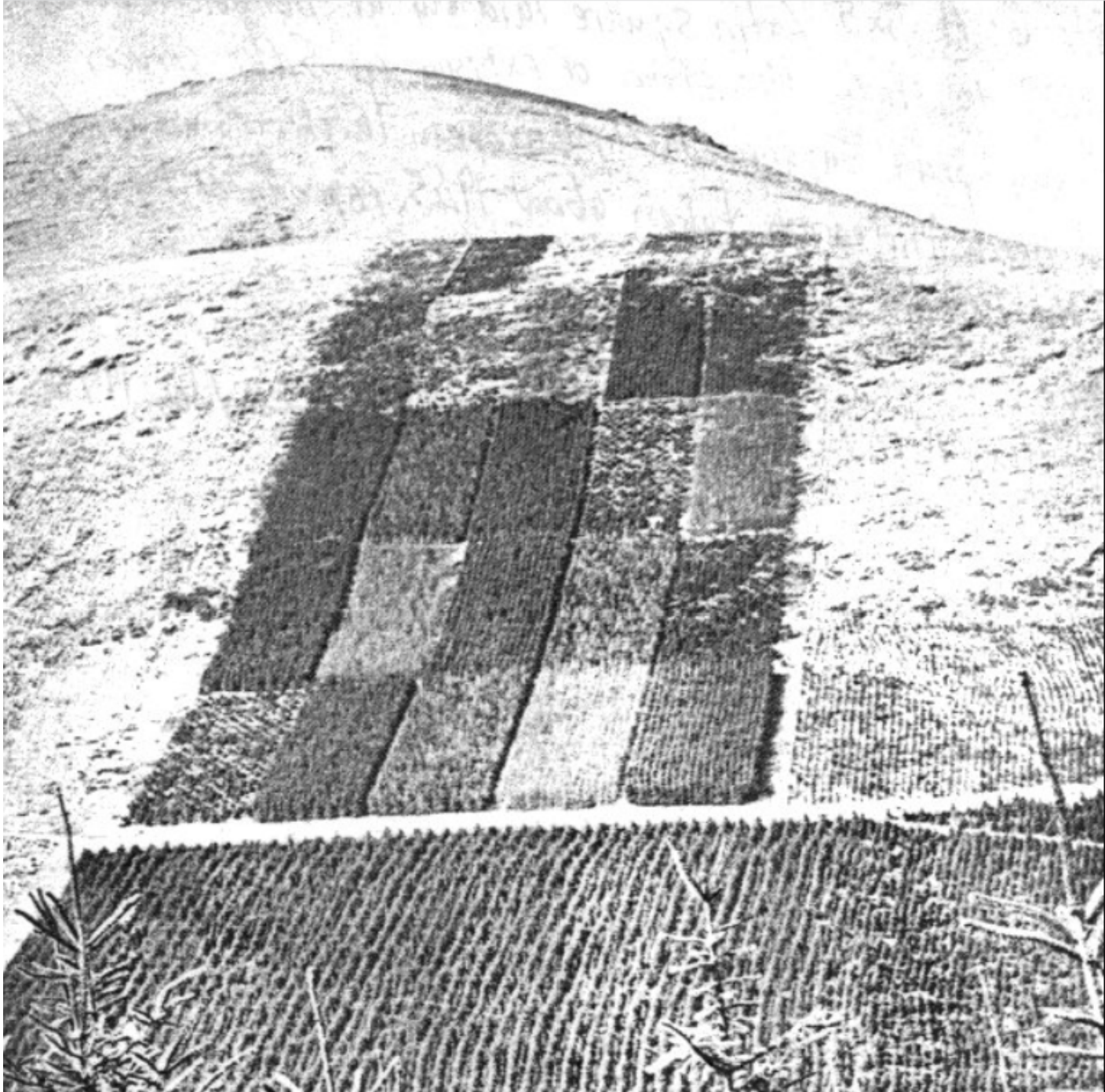
- for example – vaccinate one group of patients with Cow Pox

- or plant an entire field with a new seed variety



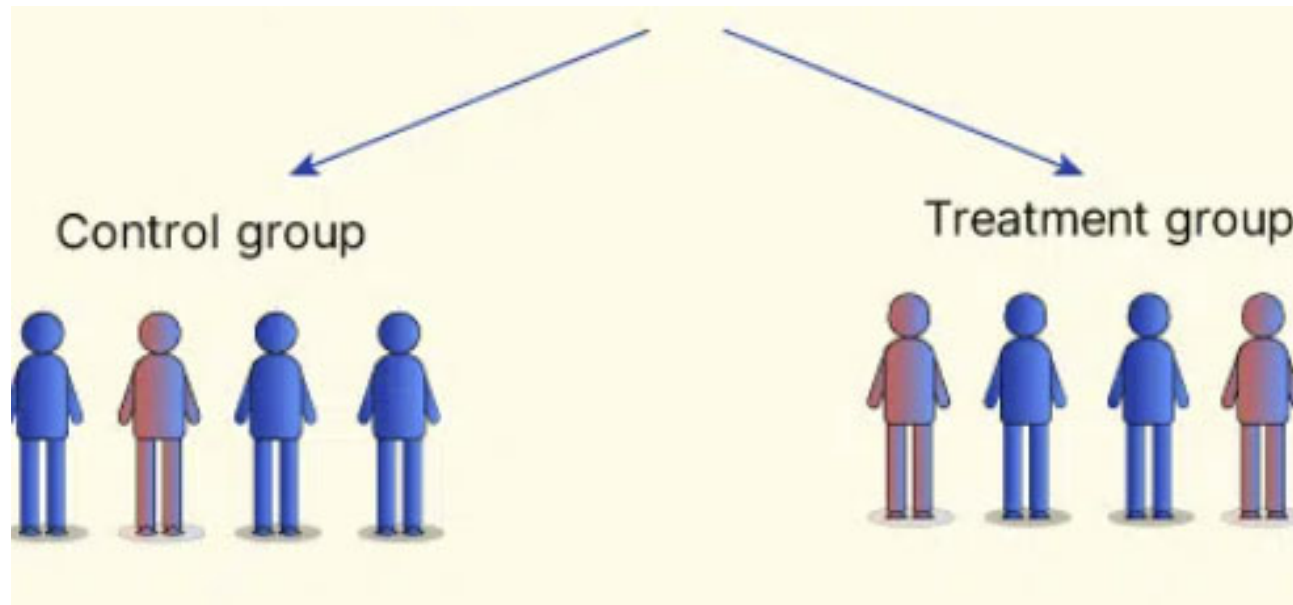
- but how will we know if the treatment works?
- R.A. Fisher (1925): divide a field into “well distributed” blocks and randomly assign different treatments

Rothamsted Experimental Station – Fisher's 5 x 5 design



Randomized controlled trial (RCT)

- divide population into 2 groups:



- ideally, subjects don't know status (placebo pills) and people measuring outcomes don't know status ("double blind" experiment)
- compare mean outcomes: treatments – controls

Why does random assignment work?

- if we **randomly** split the population then the average for the C's is a good estimate of what the outcome would be for the T's *if they had not been treated*

- the C's provide an estimate of the “counterfactual” for the T's if they had not been treated (the “mirror universe” in Star Trek)



- a lot of policy debates continue because we don't have a good way to estimate the counterfactual (e.g., questions in macro policy)

Social Experiments

- RCT's became widely used in medicine in the 1950s.
famous example: 1954 Salk vaccine polio trial – 1.8 million children enrolled (Placebo injections for C's)

Why so big a sample?

- low incidence
(50/100,000)
- risk varied
widely across
places



Social Experiments

- in mid-1960s, much enthusiasm in the US for re-imagining social programs like welfare, health insurance (the “Great Society”)



- “Negative Income Tax” experiment proposed to HHS by MIT grad student Heather Ross in 1966
- Eventually 4 different NIT experiments and the RAND Health Insurance experiment implemented in US. A parallel experiment was implemented in Manitoba

Social Experiments (2)

- by the late 1970s, enthusiasm (and funding) for “grand visionary” social reforms faded
- and people realized there were problems in the NIT experiments
- next generation of Social Experiments were more modest in scope and used very simple designs
- these experiments set the stage for widespread use of RCT’s in developing countries – JPAL initiative etc.



Non-experimental “research designs”

- RCT’s can credibly estimate the counterfactual
- but there are limitations
 - (i) can’t be sure results will ‘scale up’
(or that results apply outside the trial setting)
 - (ii) many policies can’t be randomized easily
- in many ‘real life’ situations, one group of people (or schools, firms...) is exposed to a new treatment and other groups are not. For example:
 - rise in the minimum wage in one state
 - one state opts out of Daylight Savings Time

Natural experiments

- situations where one group is exposed to a new treatment (and other groups are not) are sometimes called “natural experiments”
- e.g. we want to know the effect of a jobs program for long term unemployed
- there is a treatment group and a comparison group
- BUT: *does the comparison group provide a valid counterfactual?*
- the groups may have other differences that need to be taken into account in some way (Random assignment to T and C groups solves this)

Difference in differences

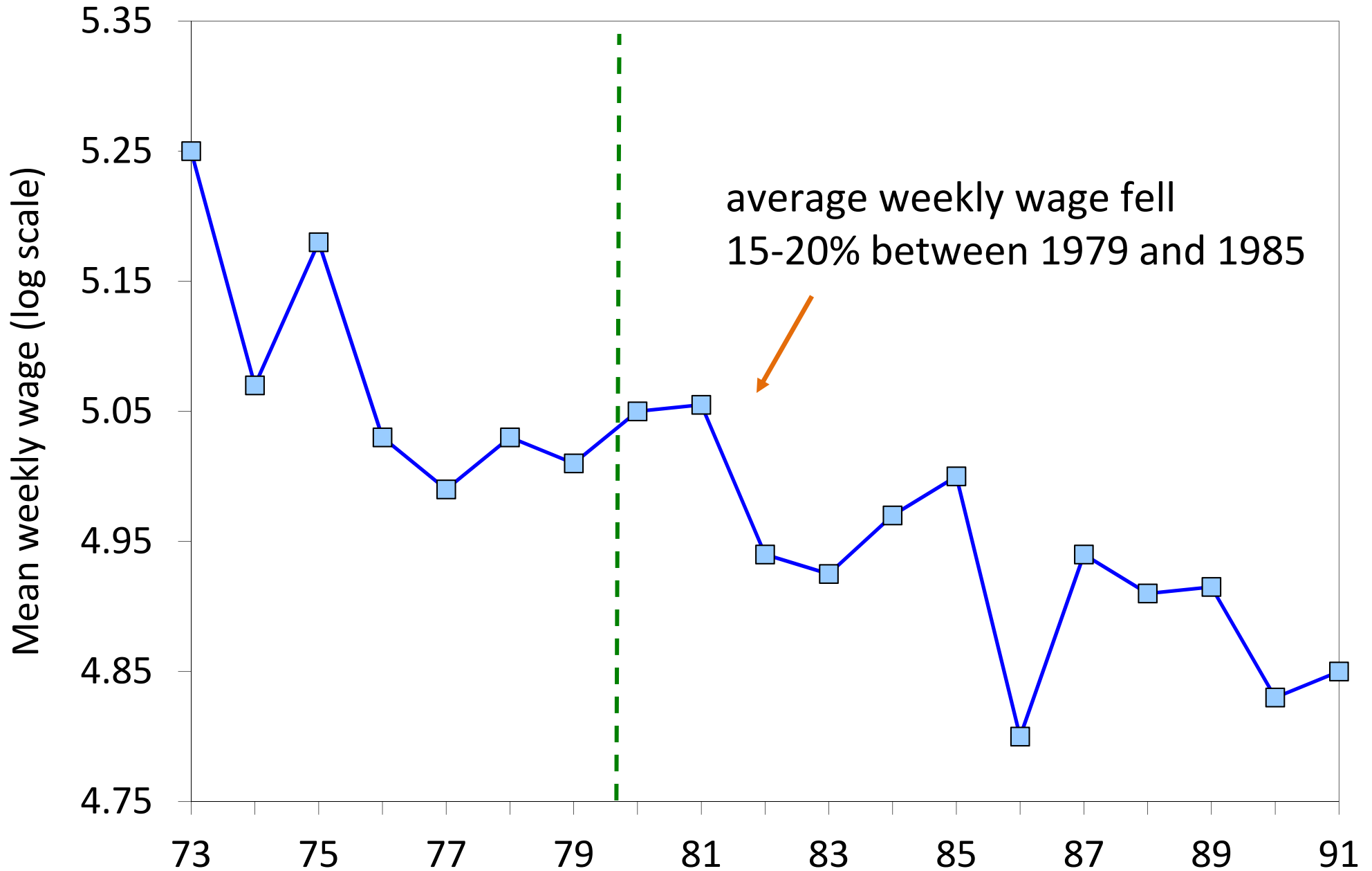
- treated and comparison groups usually have other differences that lead to a difference in outcomes even with no treatment effect!
- if these other difference remain constant, we would expect the difference in outcomes to stay constant if there was no treatment
- so: look at how the difference changes after the treatment – a “difference in differences”
- DST example: compare energy use in summer months in AZ vs other states before and after 1968 (when AZ opted out)

Example 1: the Mariel Boatlift

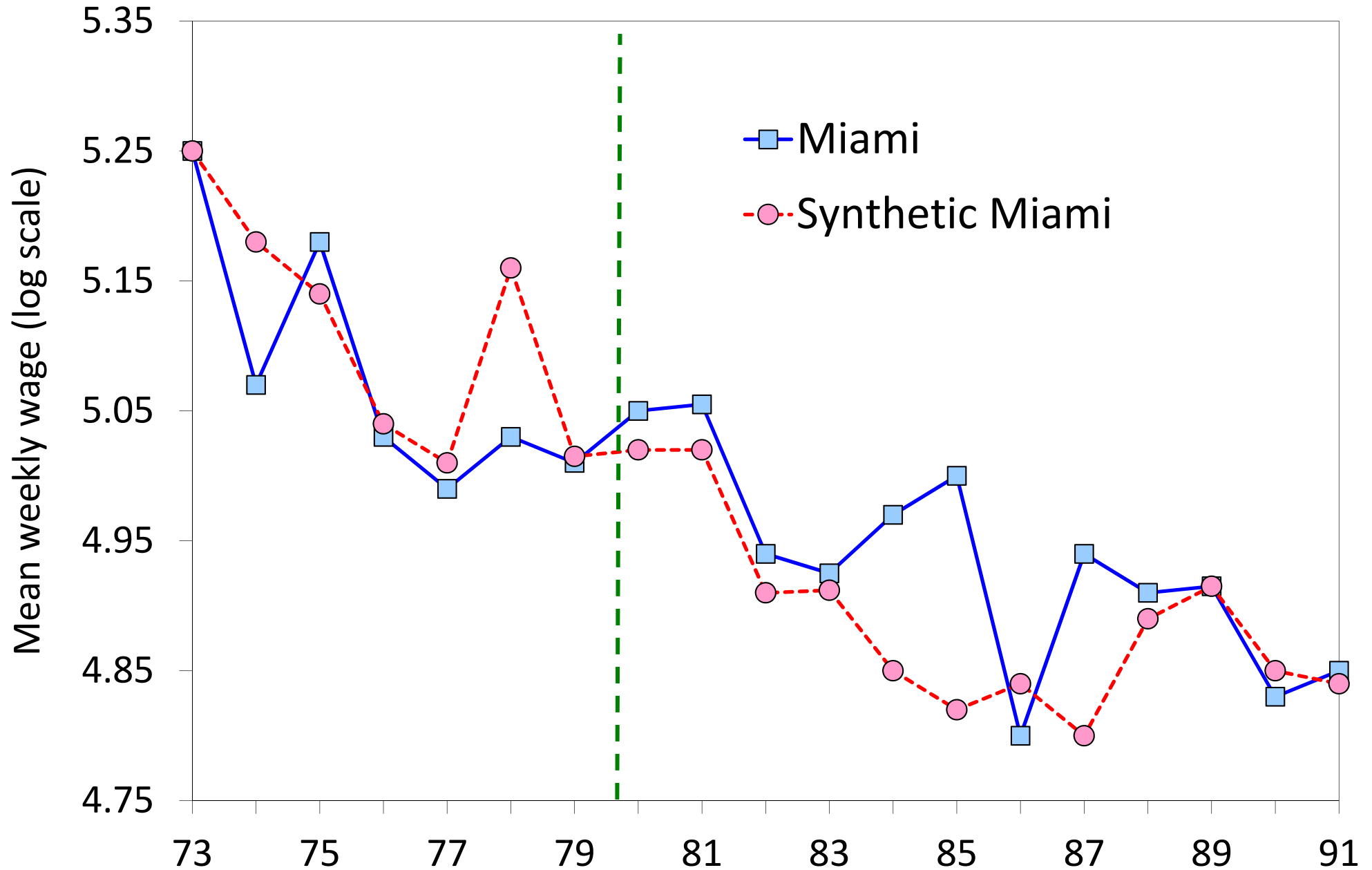
- in March 1980, a demonstration in Havana ended when Fidel Castro agreed to allow people to leave Cuba from the port of Mariel (Havana). Relatives in Miami set up a flotilla of boats and brought 125,000 people to the US, including some prisoners
- result was a *large* (+7%) increase in the Miami workforce designed a comparison group of 4 cities that had trended similar to Miami in previous decade: Tampa, Atlanta, Houston, LA



Effect of the Mariel Boatlift on Real Log Weekly Wages of Low-Skilled Natives



Effect of the Mariel Boatlift on Real Log Weekly Wages of Low-Skilled Natives

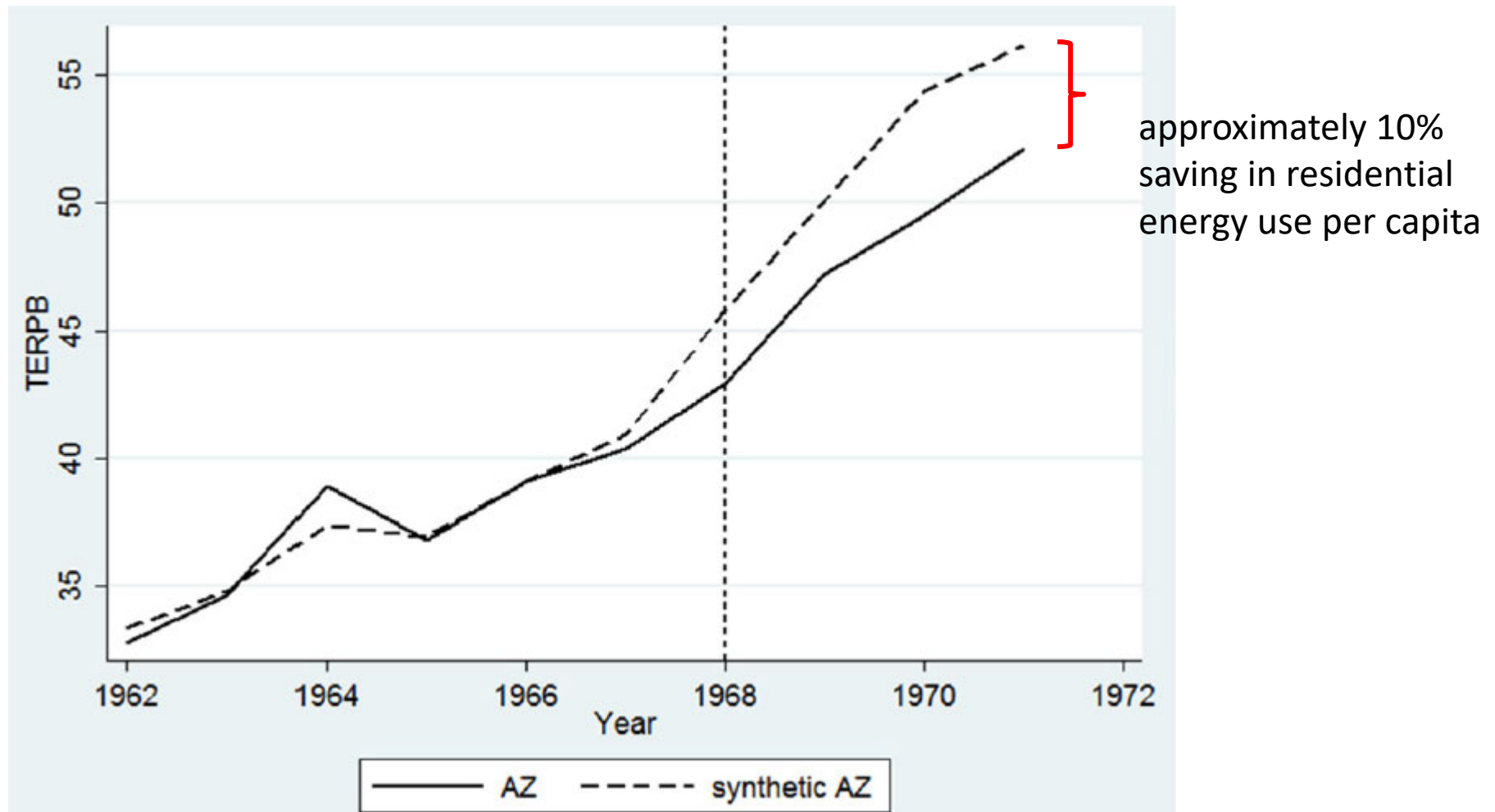


Later studies

- followed up with other studies of mass migrations:
 - i) repatriation of French nationals at end of Algerian War
 - ii) repatriation of Portuguese nationals at end of Angolan War
 - iii) migration of Russian Jews to Israel after collapse of USSR
 - iv) migration of ethnic Germans to Germany after collapse of USSR
- most show very small effect on natives (some negative effect in the case of the German migration)
- other research designs based on comparisons across US cities also show little effect on natives

Aside – what about daylight savings time?

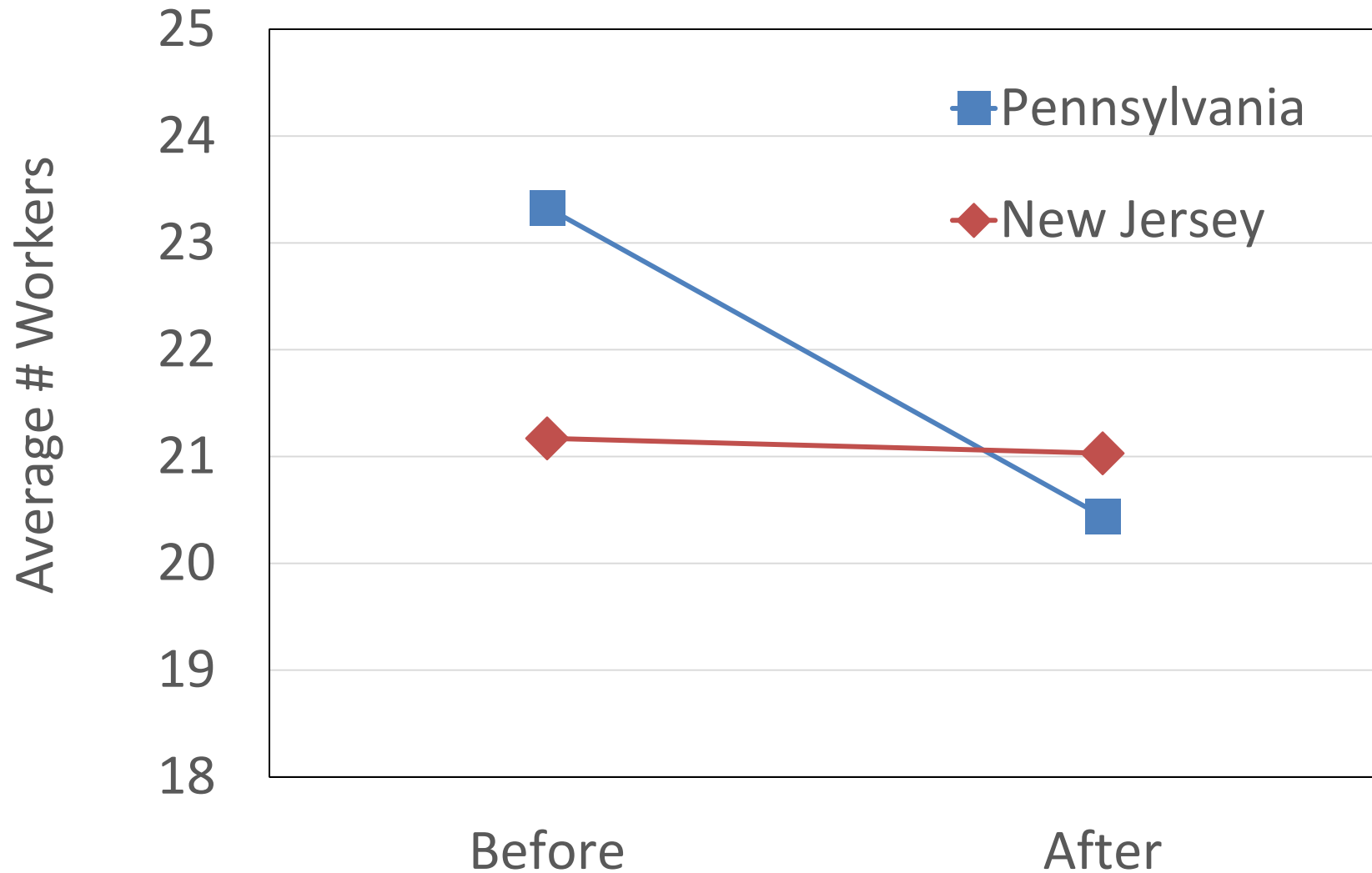
- July 2023 paper by Chenghao Li
- same design as Boatlift: compare AZ and 'synthetic AZ' (an average of other states) before and after 1968



Example 2: New Jersey minimum wage

- in early 1992, the New Jersey legislature decided to raise the state minimum wage from \$4.25 to \$5.05 per hour
- Alan Krueger and I decided to try to conduct a “prospective” DD analysis, using a survey of fast food restaurants in NJ and eastern Pennsylvania (PA):
 - Wave 1, conducted February 1992
 - Wave 2, conducted November 1992
- We had only a single observation prior to treatment. But we had an alternative way to double check our results: some restaurants in NJ were already paying \$5.00 or more per hour – so we had 2 comparison groups.

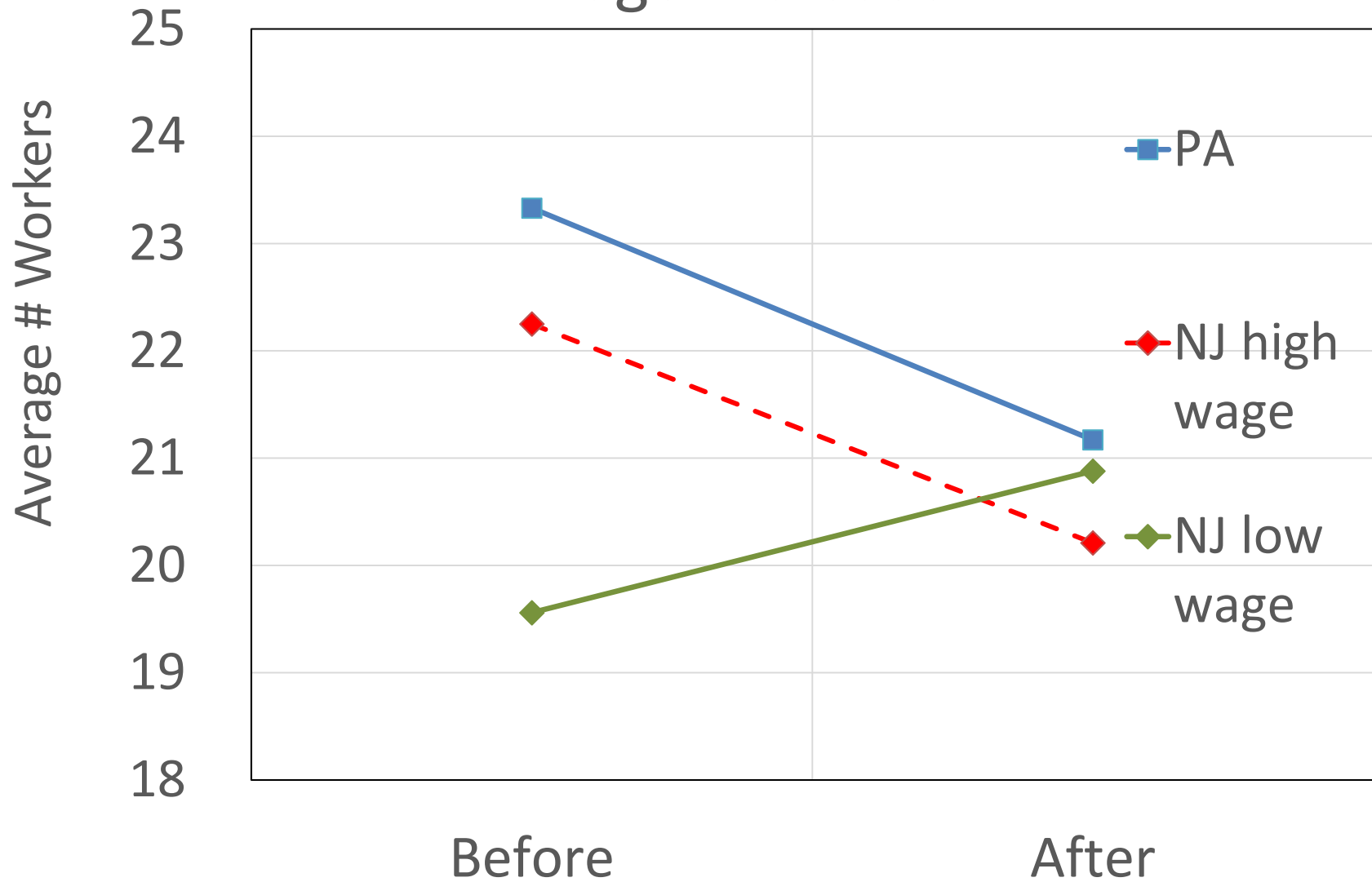
Effect of New Jersey Minimum Wage Increase



New Jersey minimum wage (continued)

- what about the 2 groups of stores in NJ?
- low wage stores were paying the old min (\$4.25). They had to raise wages by 19%
- high wage stores were paying \$5.00 or more. They had to raise wages only a few cents at most

Effect of NJ Minimum Wage Increase



New Jersey minimum wage (continued)

- the two comparison groups looked pretty similar:
 - in PA: employment fell by 2.2 workers/store
 - in high-wage NJ Stores: emp. fell by 2.0 workers/store
- the two designs (NJ-PA, and low-v. high-wage stores) give similar DD's
- later studies have combined many state-specific minimum wage increases, and found generally very small effects of higher minimum wages on employment. (Some researchers manage to find effects)
- other studies look at effects on prices (definitely +) and store closures (very small)

Evidence-based policy and immigration

- despite research findings, immigration policies in US and many other countries have tightened since 2000
- does “evidence” not matter?
- *Card-Dustmann-Preston* study (using Euro. Social Survey)
Idea: what if people care about economic effects of immigration and also about the effects on the “composition” of the population (race, ethnicity, religion, language...)

Recent poll in US: 29% of respondents agree that “immigrants are invading our country and replacing our cultural and ethnic background”

Evidence-based immigration policy (2)

- we asked questions on economic concerns:
do you agree that immigrants lower wages?
- and compositional concerns:
do you agree it is better if everyone shares the same language? same religion? ...
- we then related these to views on immigration policy:

80% of the differences in views about allowing more immigrants are driven by *compositional concerns*.
- so maybe evidence on the economic impacts of immigration matters a little, but it does not address the real concerns that underlie reactions to immigration

Evidence-based policy and min. wages

- In the US: many states have min. wages > Fed. floor
- BUT: Federal min. is stuck. And the opposition still uses concerns over job losses as the lead argument.

- Europe: UK introduced its first minimum wage in 1998. Germany introduced its first in 2015. (In both cases, studies found little or no effect on employment but positive effects on wages). Last year the EU introduced a target for member countries of a min. wage set at 60% of median wage

- were these policy changes driven by evidence?

Summary

- we can use experiments and natural experiments (and other research designs) to answer difficult causal questions about how policies will affect various outcomes
- methods like DD are widely used by academic researchers, policy analysts and business to try to learn from past policy changes, especially when RCT's ("A/B testing") is not feasible
- whether the evidence from RCT's and other non experimental methods is really affecting how policies are made is less certain