SCRIPPS INSTITUTION OF OCEANOGRAPHY

QUASI-EXPERIMENTAL METHODS FOR CLIMATE EPIDEMIOLOGY PART II

14 March 2025

Tarik Benmarhnia & Yiqun Ma

University of California, San Diego

Drexel Climate Change and Urban Health Research Center Workshop

GITHUB LINK FOR WORKSHOP MATERIALS

https://github.com/benmarhnia-lab/Drexel_CCUH_workshop/tree/main

- You will find:
  - Slides
  - Readings
  - Code and data

OUTLINE

- Natural experiments and quasi-experimental methods
- Difference-in-differences methods
- Interrupted Time Series
- Synthetic Control Methods

RANDOMIZATION AS A SOLUTION TO DEAL WITH CONFOUNDING

- Randomization has been proposed as a solution to deal with confounding
  - Complying with exchangeability between exposed and non-exposed to the policy/treatment
  - For both **measured & unmeasured** confounding

IDEAL RANDOMIZED EXPERIMENTS

- What makes an ideal randomized experiment:
  - No loss to follow-up,
  - Full adherence to the assigned treatment over the duration of the study,
  - A single version of treatment, and double-blind assignment
- Ideal randomized experiments are unrealistic **but useful to introduce some key concepts for causal inference**

THE IDEA BEHIND RANDOMIZATION

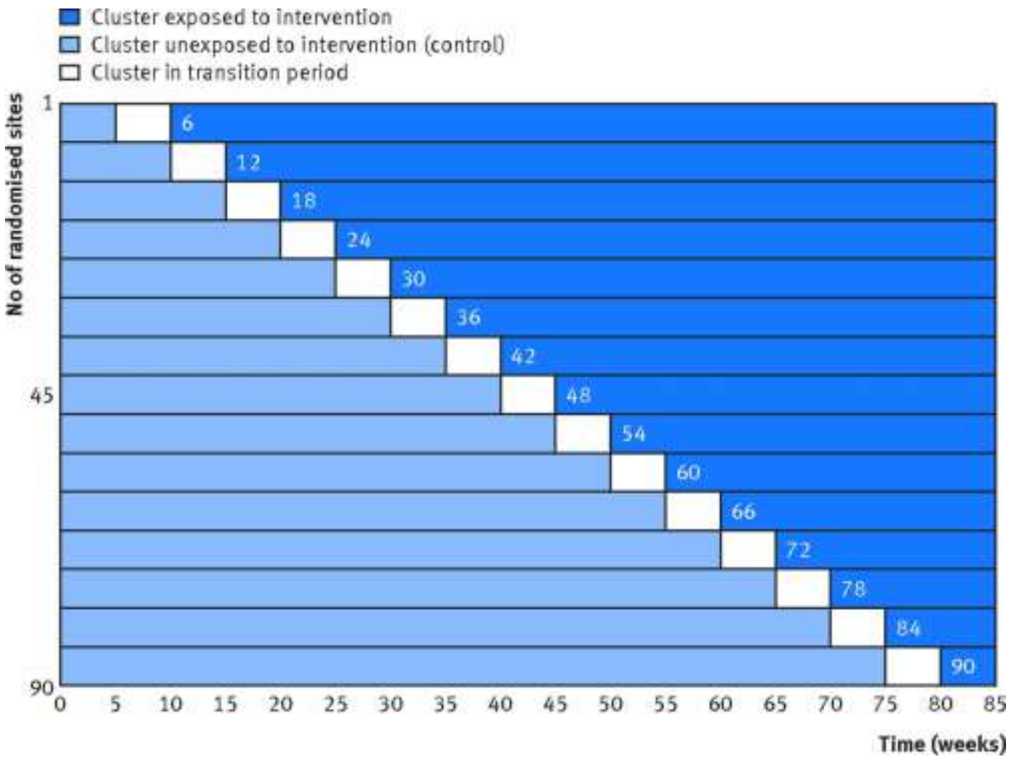
- How do RCT work …
- Randomization can be analyzed as an "**Instrumental Variable**"
- The overall aim is to deal with <u>both measured and unmeasured</u> confounding

$$\begin{tikzcd} & \textbf{C and U} \arrow[d] \arrow[dr, bend left] & \\ \textbf{IV} \arrow[r] & \textbf{X} \arrow[r] & \textbf{Y} \end{tikzcd}$$

- **The problem:** in many cases RCT are not feasible due costs, scale or ethical objections
- Or because it is simply too late ..

DIFFERENT TYPES OF RCTS

- Two-arm, parallel design
- Planned cross-over design
- Factorial design
- Cluster randomized trials
- **Stepped Wedge Designs**
  - **Based on the timing of the intervention**



Hemming, Karla, et al. "The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting." Bmj 350 (2015): h391.

Hulley et al. Designing Clinical Research. 2nd Edition. Lippincott Williams & Wilkins, 2001

USING NATURAL EXPERIMENTS

- By capitalizing on natural experiments, Quasi-experimental methods (QEM) can be used as alternatives to experimental methods to provide causal estimates from observational studies.
- The term quasi-experiment refers to:
  - "*experiments that have treatments, outcome measures, and experimental units, but do not use random assignment to create the comparisons from which treatment-caused change is inferred*" (Cook et al. 1979)
- Different QEM, different configurations, different assumptions to draw causal inference
- To <u>partially</u> deal with <u>both measured and unmeasured confounding</u>

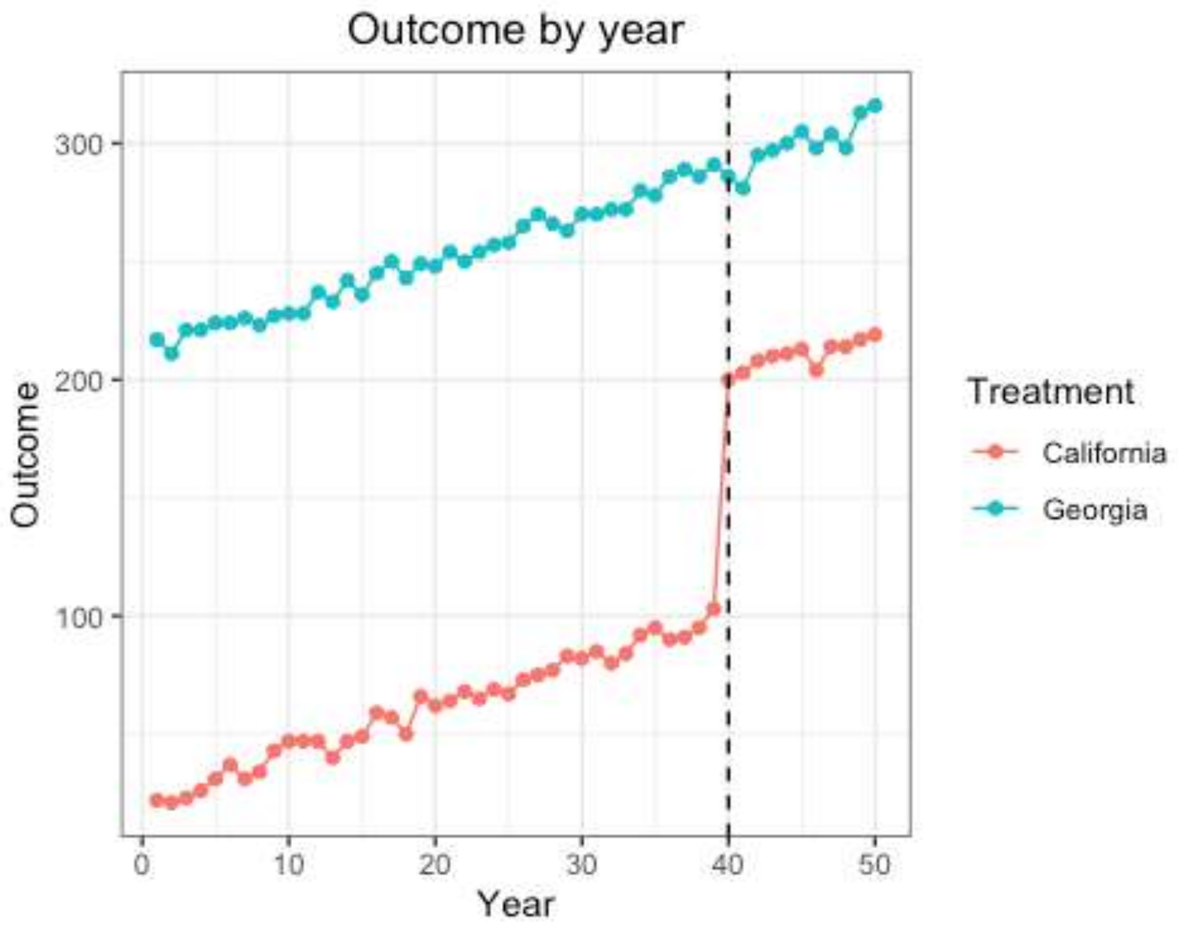
EXAMPLES OF NATURAL EXPERIMENTS
NOT ONLY FOR POLICY EVALUATION

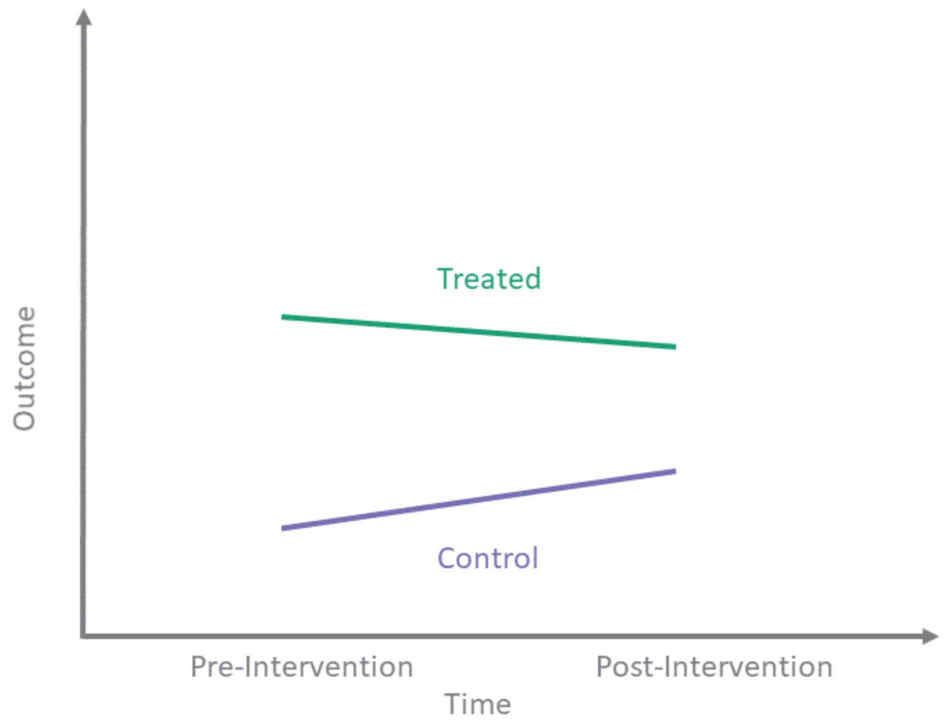
- Public policies:
  - Smoking ban
  - Legalization of marijuana
- Clinical Treatments, Vaccination
- Conditional Cash Transfers
- Natural Hazards
  - Earthquakes
  - Wildfires

2 main types of natural experiments
- **Timing of the intervention**
- Eligibility to a specific policy (e.g. age for vaccination) and other IV strategies

DIFFERENCE-IN-DIFFERENCES

The difference-in-differences idea





By Bret Zeldow and Laura Hatfield

DIFFERENCE-IN-DIFFERENCES
THE STANDARD APPROACH

- To estimate the effect of interest

$$(\mu_{11} - \mu_{10}) - (\mu_{01} - \mu_{00})$$

$i = 0$ is control group, $i = 1$ is treatment.

$t = 0$ is pre-period, $t = 1$ is post-period.

- Simple Regression modeling approach (with 2 groups)
  - $E[Y \mid a, t, did] = \beta_0 + \beta_1 a + \beta_2 t + \beta_3 did$
  - Where:
    - a represents the group with the policy
    - t represents the period after the policy implementation
    - DID is the interaction between a and t

HOW DOES IT WORK?

	T=0	T=1
A=0	β_0	$\beta_0 + \beta_1$
A=1	$B'_0 + \beta_2$	$B'_0 + \beta'_1 + \beta'_2 + \beta_3$

$$E[Y \mid a, t, did] = \beta_0 + \beta_1 t + \beta_2 a + \beta_3 [a*t]$$

0 0
1 0
0 1
1 1

You can also include more complex time trends

Using double or 2-way fixed effect

DID ASSUMPTIONS

- The key assumptions of the DiD analysis are:
  1. The trend in the control group represents a good approximation for the counterfactual trend of the treated group in the absence of the treatment.
  2. Common Shock Assumption
  3. No spillover

MANY APPLICATIONS

Removing user fees for facility-based delivery services: a difference-in-differences evaluation from ten sub-Saharan African countries

Britt McKinnon,[1]* Sam Harper,[1] Jay S Kaufman[1] and Yves Bergevin[2]

Gotta catch'em all! Pokémon GO and physical activity among young adults: difference in differences study

Katherine B Howe,[1,2] Christian Suharlim,[3] Peter Ueda,[4,5] Daniel Howe, Ichiro Kawachi,[2] Eric B Rimm[1,6,7]

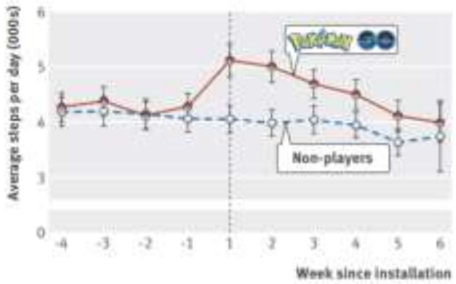


Fig 1 | Average number of daily steps and 95% confidence intervals by week before and after installation of Pokémon GO (median 8 July 2016)

JAMA Pediatrics | Original Investigation

Difference-in-Differences Analysis of the Association Between State Same-Sex Marriage Policies and Adolescent Suicide Attempts

Texting Bans and Fatal Accidents on Roadways: Do They Work? Or Do Drivers Just React to Announcements of Bans?†

By Rahi Abouk and Scott Adams*

Evaluating the effect of hierarchical medical system on health seeking behavior: A difference-in-differences analysis in China

Zhongliang Zhou [a], Yaxin Zhao [b,*], Chi Shen [a], Sha Lai [a], Rashed Nawaz [a], Jianmin Gao [a]

A NOTE ON THE 2-WAY FIXED EFFECTS (TWFE) TERMINOLOGY

- In this workshop (and in epidemiology applications more generally), we used the term TWFE to describe the previous setting:
  - One fixed effect for before/after the policy
  - One fixed effect for the intervention/control units
  - And then an interaction term (DID term) between these 2 fixed effects
  - This is an analytical tool, not a design: **using a TWFE does not mean this is a Difference-in-differences analysis**

Long-Term Exposure to Ambient $PM_{2.5}$ and Hospitalizations for Myocardial Infarction Among US Residents: A Difference-in-Differences Analysis

Yichen Wang, MS; Xinye Qiu, PhD; Yaguang Wei, PhD; Joel D. Schwartz, PhD

BACKGROUND: Air pollution has been recognized as an untraditional risk factor for myocardial infarction (MI). However, the MI risk attributable to long-term exposure to fine particulate matter ≤2.5 µm in aerodynamic diameter ($PM_{2.5}$) is unclear, especially in younger populations, and few studies have represented the general population or had power to examine comorbidities.

METHODS AND RESULTS: We applied the difference-in-differences approach to estimate the relationship between annual $PM_{2.5}$ exposure and hospitalizations for MI among US residents and further identified potential susceptible subpopulations. All hospital admissions for MI in 10 US states over the period 2002 to 2016 were obtained from the Healthcare Cost and Utilization Project State Inpatient Database. In total, 1 914 684 MI hospital admissions from 8106 zip codes were included in this study. We observed a 1.35% (95% CI, 1.11–1.59) increase in MI hospitalization rate for 1-µg/m³ increase in annual $PM_{2.5}$ exposure. The estimate was robust to adjustment for surface pressure, relative humidity, and copollutants. In the population exposed to ≤12 µg/m³, there was a larger increment of 2.17% (95% CI, 1.79–2.56) in hospitalization rate associated with 1-µg/m³ increase in $PM_{2.5}$. Young people (0–34 years of age) and elderly people (≥75 years of age) were the 2 most susceptible age groups. Residents living in more densely populated or poorer areas and individuals with comorbidities were observed to be at a greater risk.

CONCLUSIONS: This study indicates long-term residential exposure to $PM_{2.5}$ could increase risk of MI among the general US population, people with comorbidities, and poorer individuals. The association persists below current standards.

Statistical Analysis

We used the DID approach to estimate the relationship between long-term exposure to ambient $PM_{2.5}$ and the incidence of hospital admissions for MI. The analysis was limited to the zip code year combinations with a population of >100 to reduce the noise from low-population areas and increase the analytical power.[26] Air pollution data were almost complete (<1% missingness), whereas some covariates had small fractions of missingness (<5%) (eg, median household income, percent of residents having high school education or less, percent of residents >65 years of age living in poverty). We assumed any missingness of covariates to be at random and excluded those observations from further study. First, we calculated the annual aggregated counts of hospital admissions for MI for each zip code, year, and age group (0–34, 35–54, 55–64, 65–74, ≥75 years). We fitted a quasi-Poisson regression to account for the overdispersion of MI hospitalization counts. The equation is given below:

$$\ln(E(Y_{z,t,k})) = \beta_0 + \beta_1 PM_{2.5\,z,t} + \beta_{2,z} Z_z + \beta_{3,t} U_t + \beta_{4,k} I_k + \beta_{5,z,t} W_{z,t} + \ln(P_{z,t,k})$$

where $Y_{z,t,k}$ represents the aggregated count of hospitalizations for MI in zip code z, year t and age group k; $PM_{2.5\,z,t}$ represents the mean ambient $PM_{2.5}$ concentration for the same stratum of zip code z and year t; Z_z is a dummy variable for zip code, which captures all time-invariant or slowly changing variables that vary across zip code areas, measured or unmeasured; U_t is a dummy variable for year t, which represents the time-varying variables whose temporal variation is similar across zip code areas; I_k is a dummy variable for age groups; $W_{z,t}$ represents the variables that may vary differently over time across zip code areas; $\ln(P_{z,t,k})$ is an offset term that represents the natural log of the population in zip code z, year t, and age group k.

DID COUPLED WITH PROPENSITY SCORE METHODS

- When multiple control groups are available, it is possible to use available information on time-varying and time-fixed confounders
- We can use propensity score matching and IPTW for example
- By doing so, we aim at identifying similar observations at each time point in the control groups

Using propensity scores in difference-in-differences models to estimate the effects of a policy change

Elizabeth A. Stuart · Haiden A. Huskamp · Kenneth Duckworth · Jeffrey Simmons · Zirui Song · Michael E. Chernew · Colleen L. Barry

LETTER

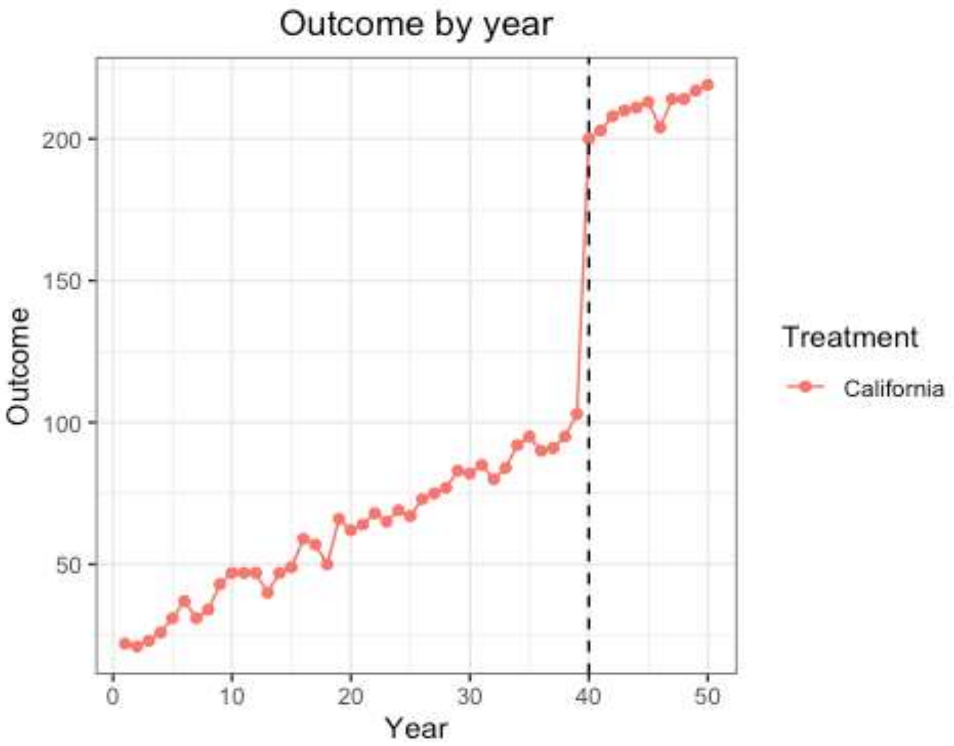
Quantifying the impact of changing the threshold of New York City heat emergency plan in reducing heat-related illnesses

Tarik Benmarhnia[1,3], Lara Schwarz[1], Amruta Nori-Sarma[2] and Michelle L Bell[2]

Evaluating the potential public health impacts of the Toronto cold weather program☆

Tarik Benmarhnia[a,*], Xu Zhao[b], John Wang[b], Melissa Macdonald[c], Hong Chen[b,d,e]

INTERRUPTED TIME SERIES: DID WITHOUT CONTROL GROUPS



Estimation

The traditional approach:

The following model can be used to estimate the effect of the policy on the outcome y

$$y = \alpha + \beta_1(\text{year}) + \beta_2(\text{post}) + \beta_3(\text{xi}) + \beta_4(\text{xt}) + \beta_5(\text{xit}) + \beta_6(\text{year} \times \text{post}) + \epsilon$$

Can be also done through a 2-stage approach:

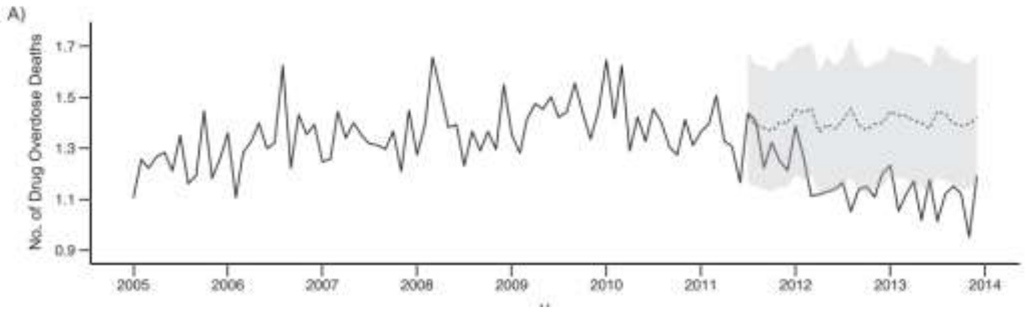
1. Building and optimizing a predictive model for Y in the pre-treatment period
2. Predicting *Y* in the post-treatment period using the model developed in stage 1 and compare with observed outcomes

Various approaches can be used for stage 1 (ARIMA, random forest, and other ML algorithms)

ITS EXAMPLES

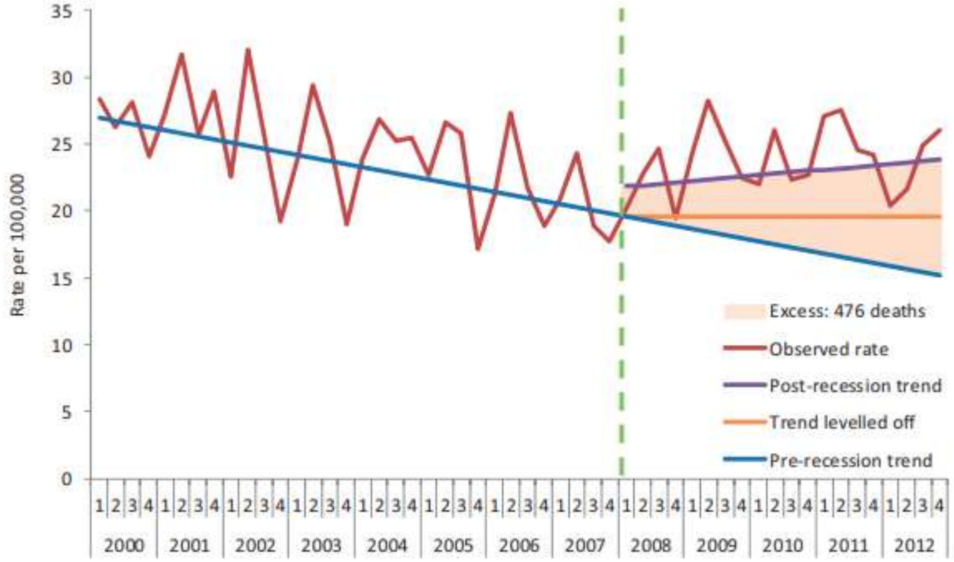
Florida's Opioid Crackdown and Mortality From Drug Overdose, Motor Vehicle Crashes, and Suicide: A Bayesian Interrupted Time-Series Analysis

Kenneth A. Feder, Ramin Mojtabai, Elizabeth A. Stuart*, Rashelle Musci, and Elizabeth J. Letourneau



Impact of the economic recession and subsequent austerity on suicide and self-harm in Ireland: An interrupted time series analysis

Paul Corcoran,[1,2*] Eve Griffin,[1] Ella Arensman,[1,2] Anthony P Fitzgerald,[2] and Ivan J Perry[2]



CONTROLLED ITS OR DID?

- When using one (or more) control group.s, there is no fundamental distinction between Controlled ITS and DID models
- They yield the same counterfactuals and identify the same treatment effects.
- The terminology simply varies across disciplines

Fry, C. E., & Hatfield, L. A. Birds of a feather flock together: Comparing controlled pre–post designs. *Health Services Research*. https://onlinelibrary.wiley.com/doi/epdf/10.1111/1475-6773.13697

ALTERNATIVE OPTIONS FOR SELECTING CONTROL GROUPS

- Besides geographical units that did not receive the treatment/policy, it is possible to consider different types of control groups
- It is possible to use different outcomes or population subgroups to contrast the change in the outcome over time
  - For example, if a given policy only targets individuals above 65 years, it will be possible to use the 64 years of age and below subgroup as a reference group

A Section 508–conformant HTML version of this article
is available at http://dx.doi.org/10.1289/EHP203.

A Difference-in-Differences Approach to Assess the Effect of a Heat Action Plan on Heat-Related Mortality, and Differences in Effectiveness According to Sex, Age, and Socioeconomic Status (Montreal, Quebec)

Tarik Benmarhnia,[1] Zinzi Bailey,[1] David Kaiser,[2] Nathalie Auger,[3] Nicholas King,[4,5] and Jay S. Kaufman[1,5]

- The Montreal heat warning system
  - Implemented in 2004
  - ‘active watch’ alert level, when daily max temperatures **<u>exceed 30°C</u>**
  - Focus on vulnerable populations
    - Age
    - SES
    - Gender as a placebo

THE ANALYTICAL APPROACH:

- Assigning days (the unit of analysis) to two groups: an "intervention" group that meets an eligibility criterion and a "non-intervention" group that does not.
- the "intervention" group: if 'active watch' alert level is present (heat wave days)
- The non-intervention group: non heat wave days
- The **counterfactual quantity** being estimated is:
  - The difference in the daily number of deaths between intervention (heat wave) and non-intervention (non-heat wave) days in the post-2004 period, had the heat warning system not been implemented

ANALYSES

- Quasi-Poisson Model to estimate a number of 'prevented' daily deaths during heat waves after the policy implementation.
- $\log\big(E(Y_{ct})\big) = \beta_0 + \beta_1 E_{ct} + \beta_2 I_t + \boldsymbol{\beta_3} E_{ct} I_t + f(confounders_{ct}) + offset_{ct}.$
  - E_{ct} to be an indicator variable taking the value of 1 if day t in community c (here Montreal) is an eligible day (i.e., exceeds the community's threshold for activating its HAP)
  - I_t to be an indicator variable taking the value of 1 if day t is post-HAP implementation and the value 0 otherwise
  - β_3 represents our coefficient of interest (DID estimate), capturing whether the HAP affected daily mortality after its implementation
- Cumulative heat effect (lag 0-5) and harvesting effect
- **Sensitivity analysis: Defining the policy implementation at arbitrary ('fake') policy implementation periods (2000 and 2002)**
- **Equity in the causal effect**
  - **To assess heterogeneity in the policy causal effect, we calculated differences-in-differences-in-differences (DIDID) estimates**

DID ASSUMPTIONS

- No time trends in daily mortality among non-eligible days
- Among non-eligible days (less than 30°C), daily mortality did not change before and after 2004
- Short interval of time (4 years before and 4 years after the initiation of the HAP intervention), to limit confounding due to population adaptation and urban changes

RESULTS

- Main effect: 2.52 deaths per day (95% CI: –0.34, 5.38)
- Represents ~50% of deaths attributable to HWs (using the same definition)

Table 2. Estimated effect of the heat action plan program on equity.

Potential modifiers of the program benefits	Heterogeneity in the program effect[a] estimate	95% CI	*p*-Value[b]
Sex (men vs. women)	1.38	(–1.60, 4.36)	0.36
Age (≥ 65 vs. < 65 years)	2.44	(0.27, 4.59)	0.03
Neighborhood SES (lowest SES tercile vs. highest SES tercile)	2.48	(0.69, 4.27)	< 0.01

[a]From DIDID (differences-in-differences-in-differences) estimates (Poisson model adjusted for temporal trends); 95% CIs were obtained by bootstrapping (1,000 samples).
[b]*p*-Values are obtained from a Wald test on the interaction term (i.e., DID estimate considering as health outcome the daily difference between two groups).

SENSITIVITY ANALYSES

Table 3. Sensitivity analyses for the estimated effects of the heat action plan program.

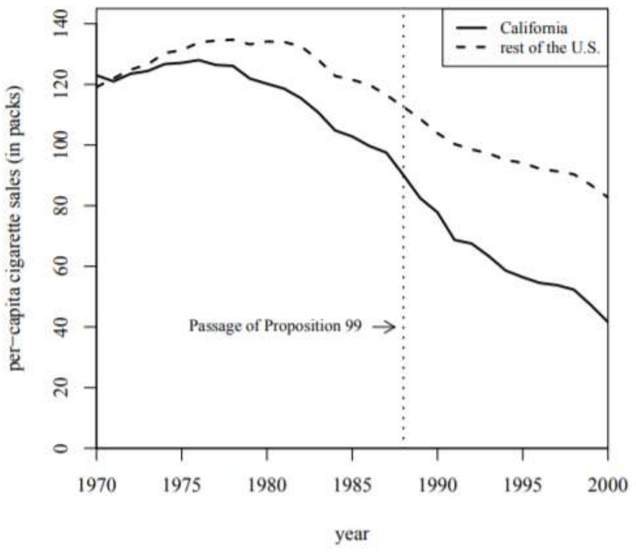
Sensitivity analyses	DID estimate	95% CI	*p*-Value[a]
Arbitrary programs			
Program implemented in 2000[b]	0.94	(–2.08, 3.96)	0.54
Program implemented in 2002[c]	0.42	(–3.62, 2.77)	0.80
Other hot days definitions			
When maximum temperature is above 28°C	0.58	(–1.77, 2.93)	0.63
When maximum temperature is above 32°C	2.79	(–2.65, 8.23)	0.32
Cumulative heat[d]	4.87	(0.67, 8.20)	0.03
Accounting for displacement ratio[e]	1.87	(0.29, 3.47)	0.02
Restriction to non-eligible days above 25°C	2.23	(–0.80, 5.27)	0.15

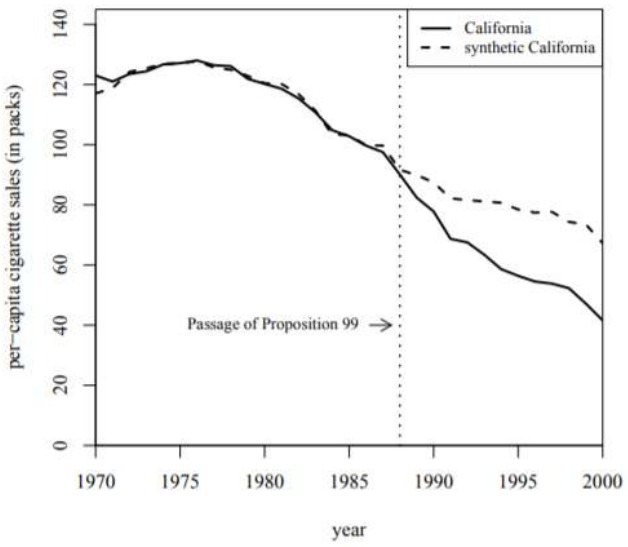
[a] *p*-Values are obtained from a Wald test on the interaction term (i.e., DID estimate).
[b] Using mortality and temperature data for periods 1996–1999 vs. 2000–2003.
[c] Using mortality and temperature data for periods 1998–2001 vs. 2002–2005.
[d] Considering a cumulative heat effect up to 5 consecutive hot days (lag 0–5).
[e] The displacement ratio (Saha et al. 2014) was 0.65.

SYNTHETIC CONTROL METHODS

THE INTUITION

- When using a DID, it is '*sometimes*' difficult to establish whether the parallel trends assumption is met and whether the control group is a sufficiently accurate representation of what would have happened in the treated area without the intervention
- Synthetic control methodology (SCM) allows the construction of a counterfactual by selecting a **weighted average** of the outcome variable from a group of units similar to the treated unit
- The intuition behind this method using the original paper by Abadie et al. 2010





HOW DOES IT WORK?

- The synthetic control is based on the vector of weights **W** that minimizes the imbalance between the treated unit and a weighted average of the controls across a set of variables **X** (e.g. pre-intervention outcomes and/or covariates), $(\boldsymbol{X_1} - \boldsymbol{X_0 W})'\boldsymbol{V}(\boldsymbol{X_1} - \boldsymbol{X_0 W})$
  - X_1 and X_0 contain the pre-treatment outcomes and covariates for the treated unit and control units respectively, and $\boldsymbol{V}$ captures the relative importance of these variables as predictors of the outcome of interest.
- In this setting, we assume the weights W to be positive and summing to 1 to avoid extrapolations issues (Abadie et al 2010). More recent approaches relaxed this assumption
- The treatment effect for the treated unit ($i = 1$), τ_{1t}, can then be estimated by $(Y_{1t} - \hat{Y}_{1t}^{0})$ for each post-intervention period separately, and these can be averaged over time to obtain an ATT over the post-intervention period
- This is a non-parametric approach, but statistical inference can be obtained via permutation tests
- What happens when we have multiple treated units?
  - Discussed later …

STEPS IN CONDUCTING A SYNTHETIC CONTROL STUDY

1. Ensure the theory behind the intervention is well understood. Develop or present a conceptual model to make the theory transparent.
   - To ensure areas that have also been exposed to a similar intervention are excluded from the pool of potential controls
2. Identify potential control units that are plausibly eligible
3. Develop the synthetic control.
   - An optimization procedure using the outcome variables from the potential control areas to select the best weighting of units from the donor pool to create a synthetic control
4. Run outcome analysis and present results
5. Run robustness checks (discussed later)

KEY ASSUMPTIONS

Synthetic control methodology as a tool for evaluating population-level health interventions

Janet Bouttell,[1] Peter Craig,[2] James Lewsey,[1] Mark Robinson,[3] Frank Popham[2]

Table 2 Key assumptions of synthetic control methodology

Assumption	Assessment
1. Treated units and potential control units in the donor pool are similar.	Similar levels in variables known to influence outcome variable (see box 1 for objective and subjective elements of this assessment).
2. There is no contamination – spillover of effects of intervention into potential control units.	Based on background knowledge of researchers.
3. No external shocks in potential control units.	Based on background knowledge of researchers informed by review of trends in outcome variable.

EXAMPLES

SEX WORK REGULATION AND SEXUALLY TRANSMITTED INFECTIONS IN TIJUANA, MEXICO

TROY QUAST[a,*] and FIDEL GONZALEZ[b]
[a]Health Policy and Management, College of Public Health, University of South Florida, Tampa, FL, USA
[b]Economics and International Business, Sam Houston State University, Huntsville, TX, USA

A New Tool for Case Studies in Epidemiology—the Synthetic Control Method

David H. Rehkopf,[a] and Sanjay Basu[b]

Original Contribution

Health Behaviors, Mental Health, and Health Care Utilization Among Single Mothers After Welfare Reforms in the 1990s

Sanjay Basu*, David H. Rehkopf, Arjumand Siddiqi, M. Maria Glymour, and Ichiro Kawachi

Do medical marijuana laws reduce addictions and deaths related to pain killers?☆

David Powell[a,*], Rosalie Liccardo Pacula[a,b], Mireille Jacobson[b,c]
[a] RAND, Santa Monica, United States
[b] NBER, Cambridge, MA, United States
[c] University of California, Irvine, United States

Effects of changes in permit-to-purchase handgun laws in Connecticut and Missouri on suicide rates

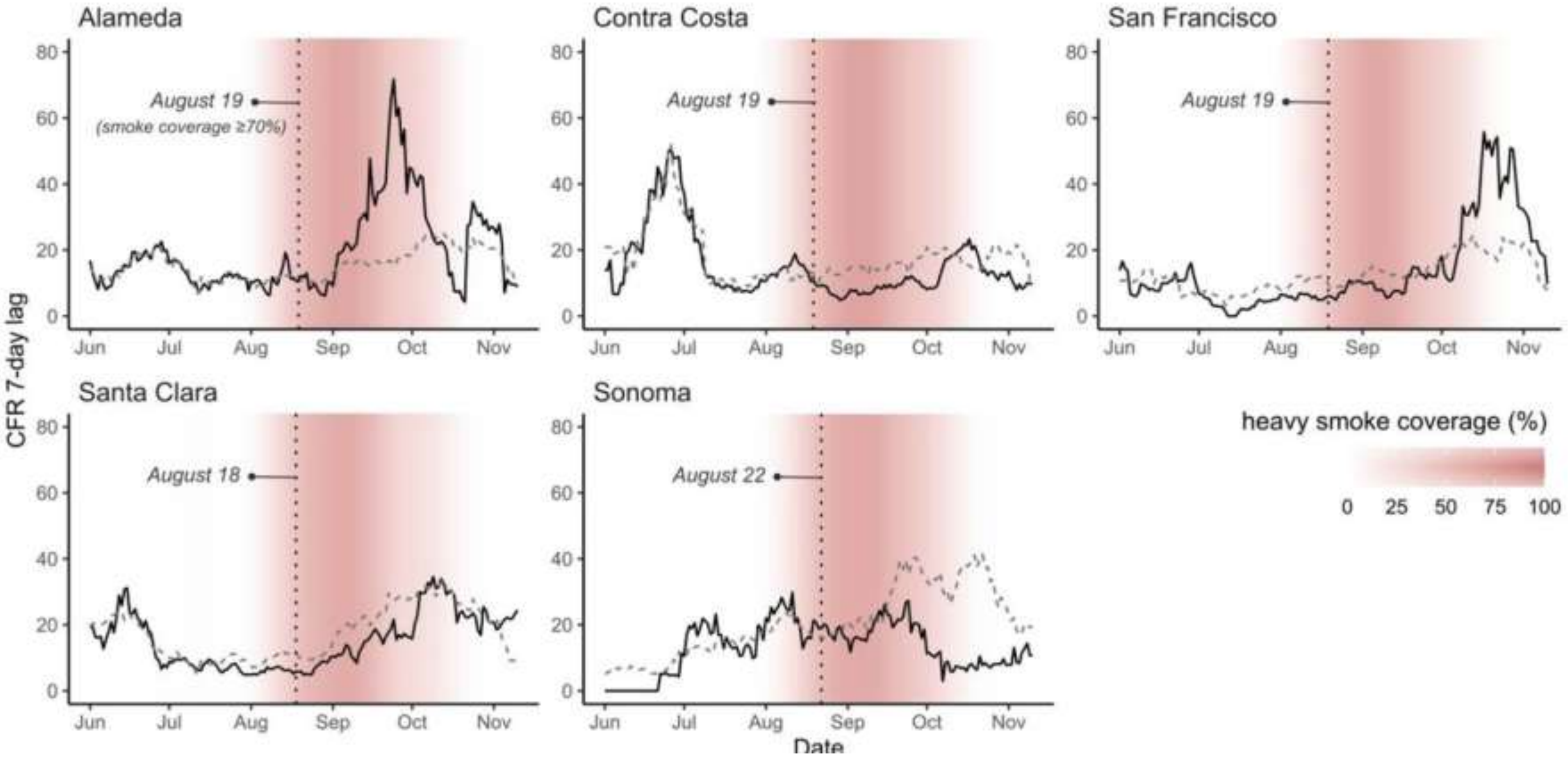
CrossMark

Cassandra K. Crifasi *, John Speed Meyers, Jon S. Vernick, Daniel W. Webster
Johns Hopkins Center for Gun Policy and Research, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

LETTER

Smoke and COVID-19 case fatality ratios during California wildfires

Lara Schwarz[1,2,*], Anna Dimitrova[3], Rosana Aguilera[3], Rupa Basu[4], Alexander Gershunov[3] and Tarik Benmarhnia[2,3]



OVERVIEW:
COMPARING DIFFERENT APPROACHES TO EVALUATE NATURAL EXPERIMENTS BASED ON THE TIMING OF THE INTERVENTION

- Uncontrolled before and after designs
  - Single or multiple measurements
- Interrupted time series design

No Control Group

- Controlled Before and after with single measurements
- Difference-in-Differences

With one Control Group

- Augmented DID approaches
  - With propensity score methods
- Synthetic Control Methods

With >1 Control Groups

MULTIPLE TREATED GROUPS

- When dealing with multiple treated units, there are two distinct settings:
  - All treated units received the intervention of interest at the same time
    - Discussed next week
  - Treated units received the intervention at different times
    - Discussed next week
- When multiple units receive the intervention at the same time:
  - A simple approach is to estimate a separate ATT for each treated unit and then conduct a meta-analysis to get a pooled estimate (and also information about heterogeneity across units)
  - Or aggregate the treated units and form a synthetic control for the aggregated treated unit (Acemoglu et al. 2013; Dube and Zipperer, 2015)

EXTENSIONS OF TRADITIONAL SC: GENERALIZED SYNTHETIC CONTROL (GSC)

Xu et al. 2017 proposed generalized synthetic control (**GSC**): estimates the average treatment effect on the treated using time-series cross sectional data

Improves efficiency and interpretability from SC, and can be used with multiple treated units and time varying confounders

GSC overcomes limitations in SC:

- Only unbiased when weights yield exact balance on lagged outcomes (and unidirectional weights)
- Only handles one treated unit at a time
- No formal measures of variance in traditional SC

HOW DOES GSC WORK?

- Generalized synthetic control (**GSC**) methods estimate the average treatment effect on the treated (ATT) using time-series cross sectional data
- Well suited for:
  - Time varying confounding: temperature, other weather events
  - Widespread exposure - multiple exposed units
- **The intuition behind this approach**

$$Y_{it} = \delta_{it} D_{it} + x'_{it}\beta + \lambda'_i f_t + e_{it}$$

- D_{it} treatment indicator
- x_{it} observed covariates
- f_t latent factors
- λ_i factor loadings

$$Y^0_{it} = x'_{it}\beta + \lambda'_i f_t + e_{it}$$

$$Y^1_{it} = \delta_{it} + x'_{it}\beta + \lambda'_i f_t + e_{it}$$

$$\widehat{\Delta}_{it} = Y^1_{it} - \hat{Y}^0_{it}$$

CASE STUDY: 2007 SOUTHERN CALIFORNIA WILDFIRES ON RESPIRATORY HOSPITALIZATIONS [SHERIDAN ET AL.]

As the climate changes, wildfires are expected to increase in frequency, intensity and duration (especially in California)

In this case study we use:

- Satellite based smoke plume data and burn area data to classify wildfire exposure
- OSHPD respiratory hospitalization data by zip code for outcome

EPIDEMIOLOGY

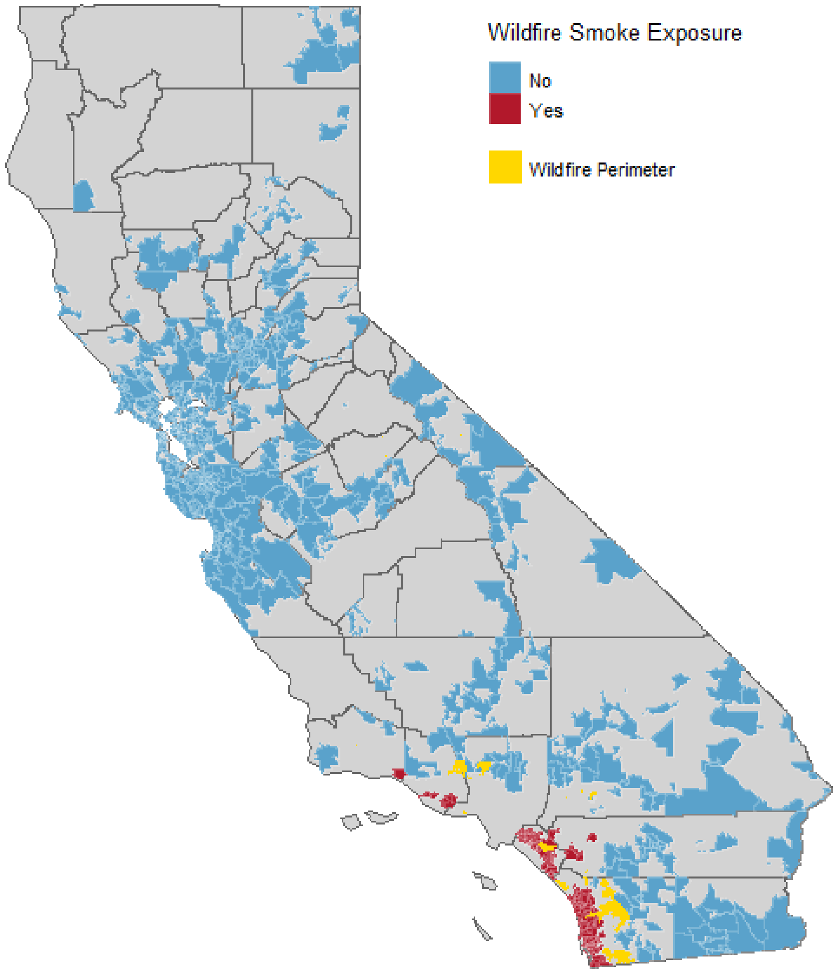
Sheridan P, McElroy S, Casey J, Benmarhnia T. Using generalized synthetic control to estimate the impact of extreme weather events on population health. Epidemiology [In Press]

RESULTS

Out of 1779 zip code tabulation areas in California

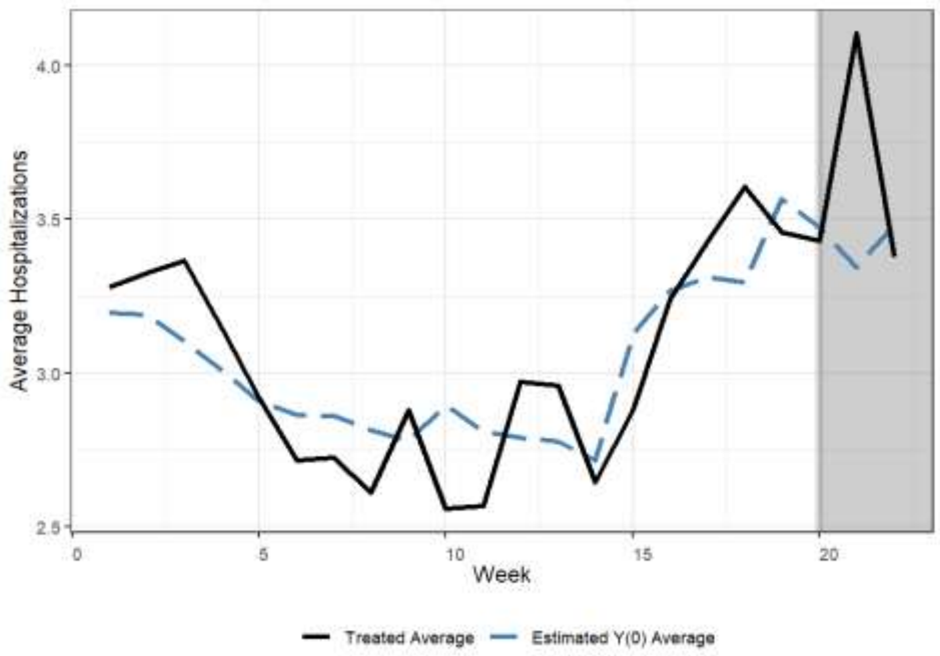
685 were included:

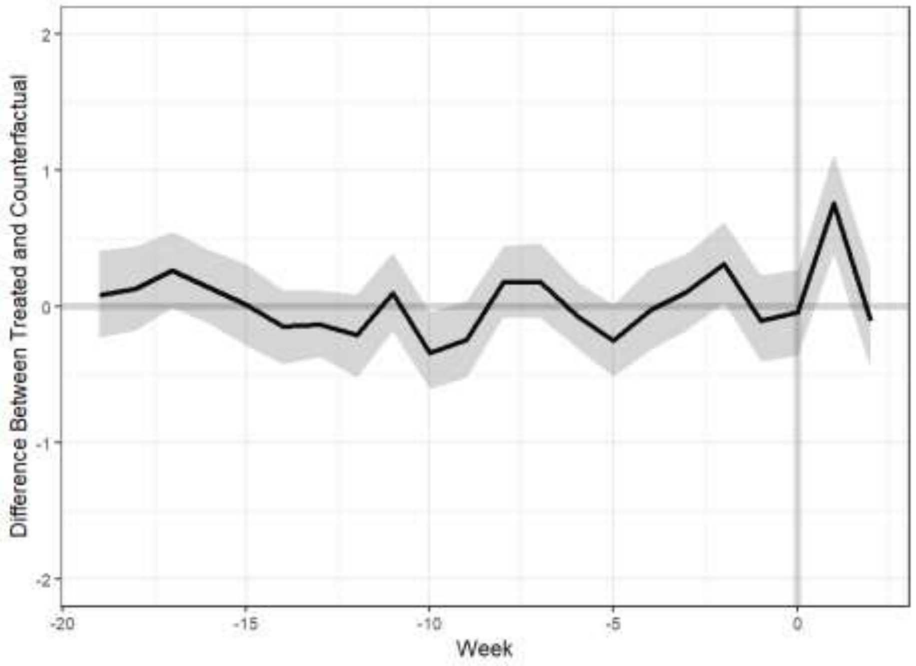
- 130 exposed
- 555 unexposed



Map of California zip codes exposed to wildfire smoke week of October 20th 2007.

RESULTS





We found an **18% (95%CI: 10%-29%)** average increase in respiratory hospitalizations as a result of the wildfire storm

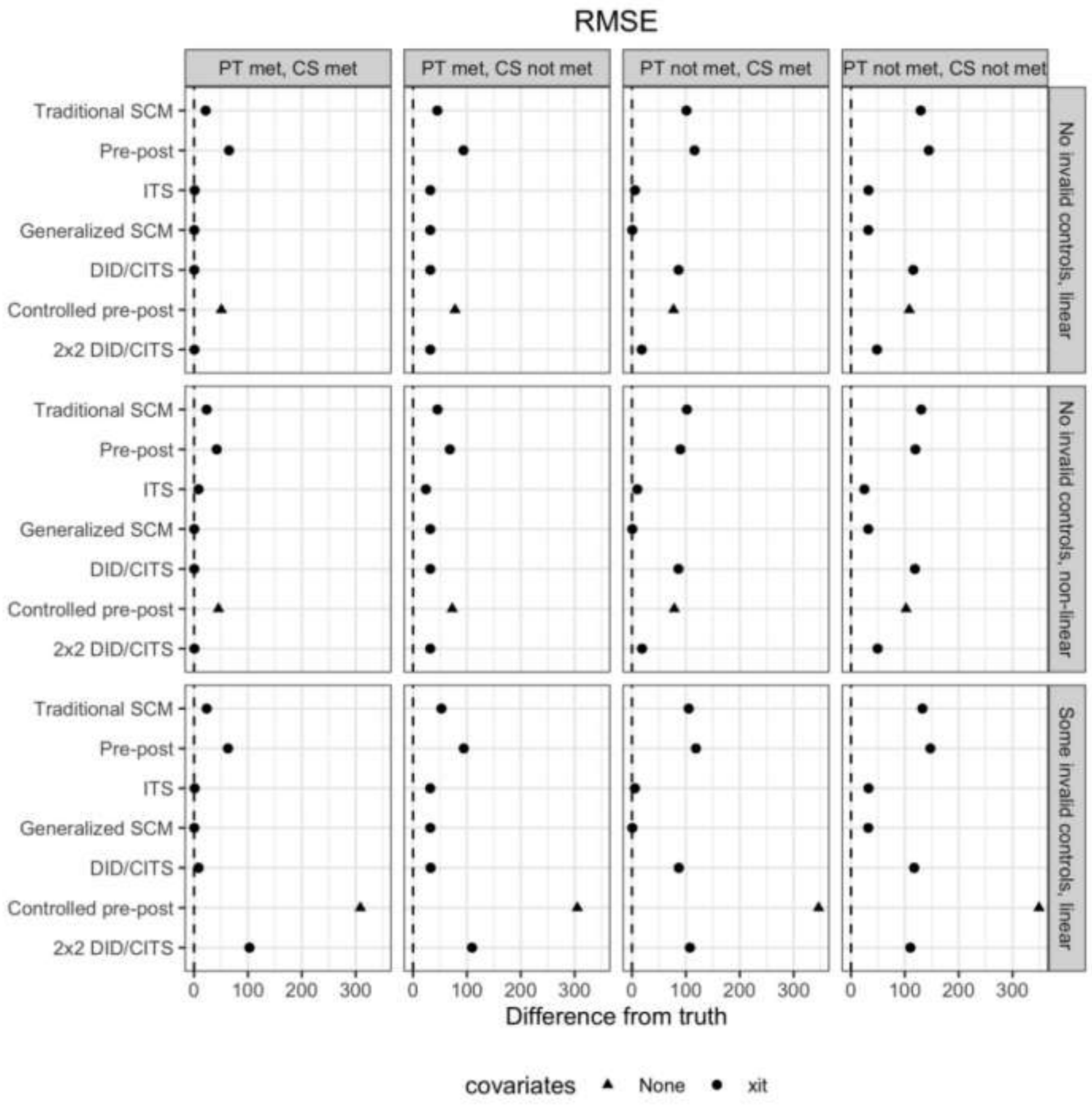
A COMPARISON OF THE DIFFERENT APPROACHES DISCUSSED TODAY

- We compared several quasi-experimental methods that use data before and after an intervention and contrast their performance within a simulation framework
  - Root mean squared error as our metric of interest
- We conducted a comprehensive simulation to assess:
  - The parallel trend assumption
  - The common shock assumption
  - Different sets of control groups
  - Different types of time trends
  - Time-varying confounding

A comparison of quasi-experimental methods with data before and after an intervention: an introduction for epidemiologists and a simulation study

Roch A Nianogo,[1,2]* Tarik Benmarhnia[3] and Stephen O'Neill[4]

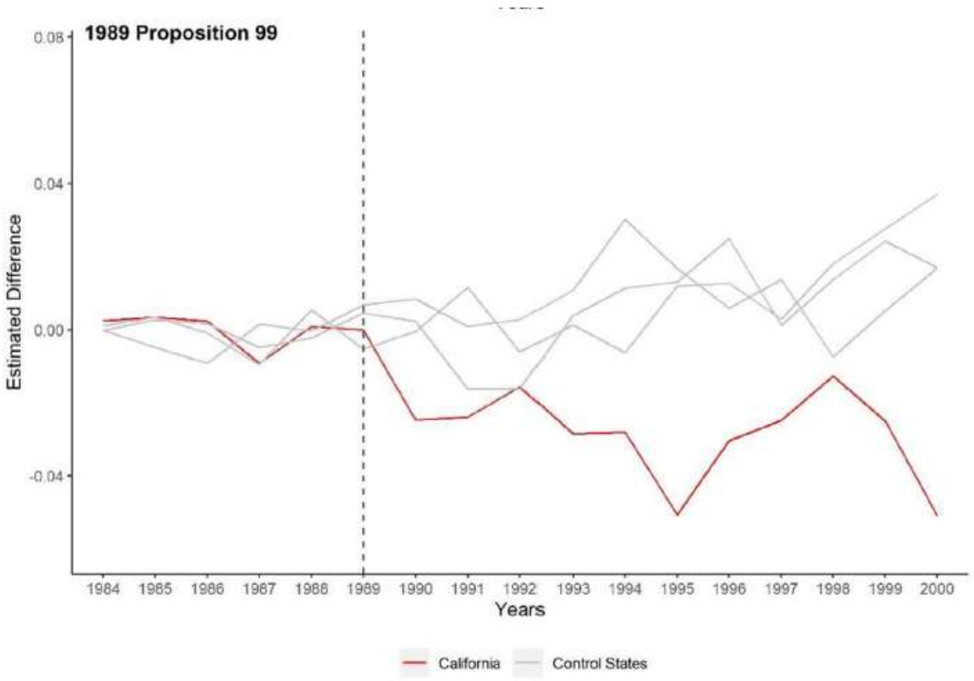
International Journal of Epidemiology, dyad032.



RMSE
PT met, CS met
PT met, CS not met
PT not met, CS met
PT not met, CS not met
Traditional SCM
Pre-post
ITS
Generalized SCM
DID/CITS
Controlled pre-post
2x2 DID/CITS
No invalid controls, linear
No invalid controls, non-linear
Some invalid controls, linear
0
100
200
300
Difference from truth
covariates
None
xit

THE IMPORTANCE OF FALSIFICATION TESTS

- <u>Permutation tests</u>
- Many assumptions cannot be checked empirically
  - It is therefore important to design a set of falsification/placebo tests to improve the inference of interest
- Negative control approaches
  - Lipsitch M et al. (2012): Negative controls: a tool for detecting confounding and bias in observational studies.



Sheridan et al. "Evaluating the impact of the California 1995 smoke-free workplace law on population smoking prevalence using a synthetic control method." Preventive medicine reports 19 (2020): 101164.

REFERENCES

- Nianogo, R. A., Benmarhnia, T., & O'Neill, S. (2023). A comparison of quasi-experimental methods with data before and after an intervention: an introduction for epidemiologists and a simulation study. International Journal of Epidemiology, dyad032.
- Craig, P., Katikireddi, S. V., Leyland, A., & Popham, F. (2017). Natural experiments: an overview of methods, approaches, and contributions to public health intervention research. Annual review of public health, 38, 39-56.
- De Chaisemartin, C., & d'Haultfoeuille, X. (2022). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey (No. w29734). National Bureau of Economic Research.
- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American statistical Association, 105(490), 493-505.
- Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. Political Analysis, 25(1), 57-76.
- Acemoglu, D., Johnson, S., Kermani, A., Kwak, J. and Mitton, T., 2016. The value of connections in turbulent times: Evidence from the United States. Journal of Financial Economics, 121(2), pp.368-391.
- Dube A, Zipperer B. Pooled synthetic control estimates for recurring treatment: an application to minimum wage studies, University of Massachusetts. Amherst Working Paper.2013
- Kreif, N., Grieve, R., Hangartner, D., Turner, A.J., Nikolova, S. and Sutton, M., 2016. Examination of the synthetic control method for evaluating health policies with multiple treated units. Health economics, 25(12), pp.1514-1528.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics.
- Callaway, B., & Sant'Anna, P. H. (2020). Difference-in-differences with multiple time periods. Journal of Econometrics.
- Roth, J., & Sant'Anna, P. H. (2020). When Is Parallel Trends Sensitive to Functional Form?. arXiv preprint arXiv:2010.04814.
- Lipsitch, M., Tchetgen, E. T., & Cohen, T. (2010). Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology (Cambridge, Mass.), 21(3), 383.

THANK YOU

YIM022@UCSD.EDU
TBENMARHNIA@UCSD.EDU