SCRIPPS INSTITUTION OF OCEANOGRAPHY

ECO-HEALTH WORKSHOP: INTRO TO CLIMATE EPI METHODS PART II

04 November 2025

Tarik Benmarhnia & Yiqun Ma

University of California, San Diego

UC Center for Climate Health and Equity

UNIVERSITY OF CALIFORNIA Center for Climate, Health and Equity

OUTLINE

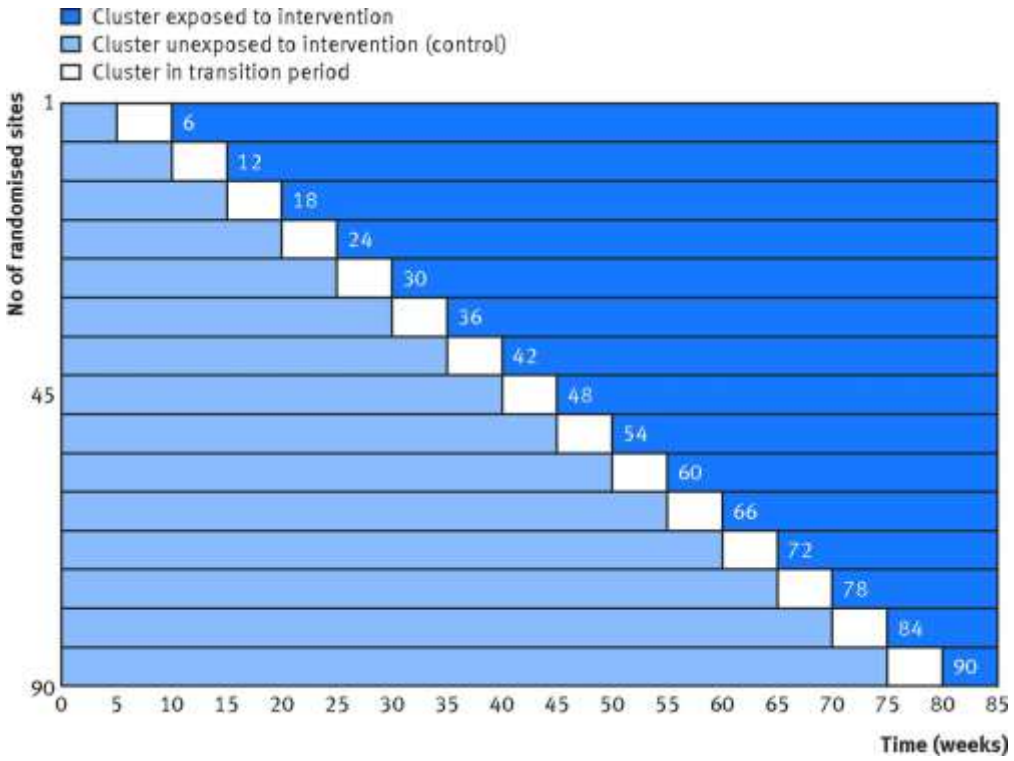
- Climate-sensitive exposures as natural experiments
  - Interrupted time series
  - Synthetic control analyses
- Effect modification including in high-dimensional settings
- Multi-stage approaches for drivers of spatial or temporal effect heterogeneity
- Mediation analyses to better understand the mechanisms in climate and health: heat and ozone as a case study

GITHUB LINK

- https://github.com/benmarhnia-lab/EcoHealth_Climate_and_Health_Research_Workshop

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

THE TARGET TRIAL FRAMEWORK

- Target trial emulation is the application of design principles from randomized trials to the analysis of observational data, thereby explicitly tying the analysis to the trial it is emulating.
- The purpose is to improve the quality of observational epidemiology through the application of trial design principles, even when, or perhaps especially when, a comparator trial is not yet available or feasible

Eur J Epidemiol (2017) 32:473–475
DOI 10.1007/s10654-017-0293-4

COMMENTARY

Target trial emulation: teaching epidemiology and beyond

Jeremy A. Labrecque[1] · Sonja A. Swanson[1]

Practice of Epidemiology

Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available

Miguel A. Hernán* and James M. Robins

NATURAL EXPERIMENTS FOR CLIMATE EPIDEMIOLOGISTS

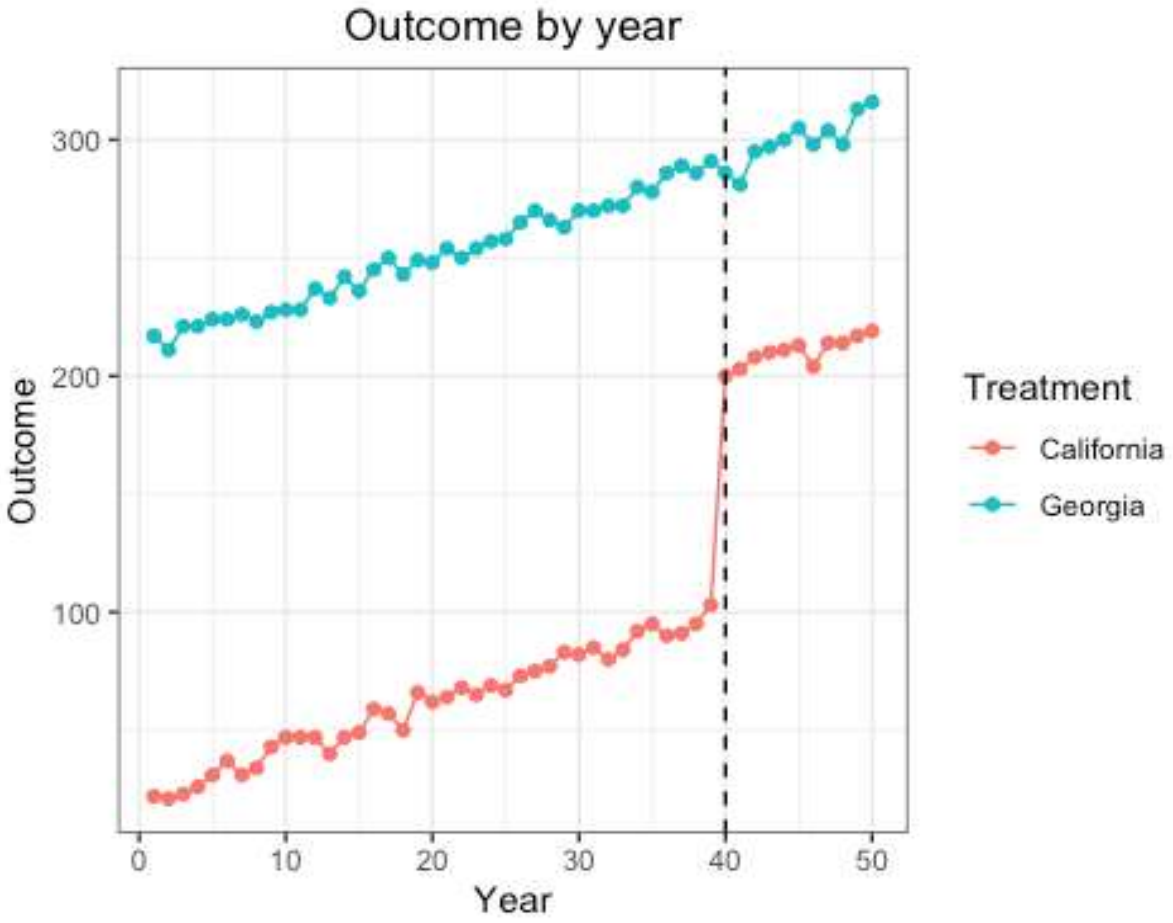
CONSIDERING CLIMATE HAZARDS AS NATURAL EXPERIMENTS

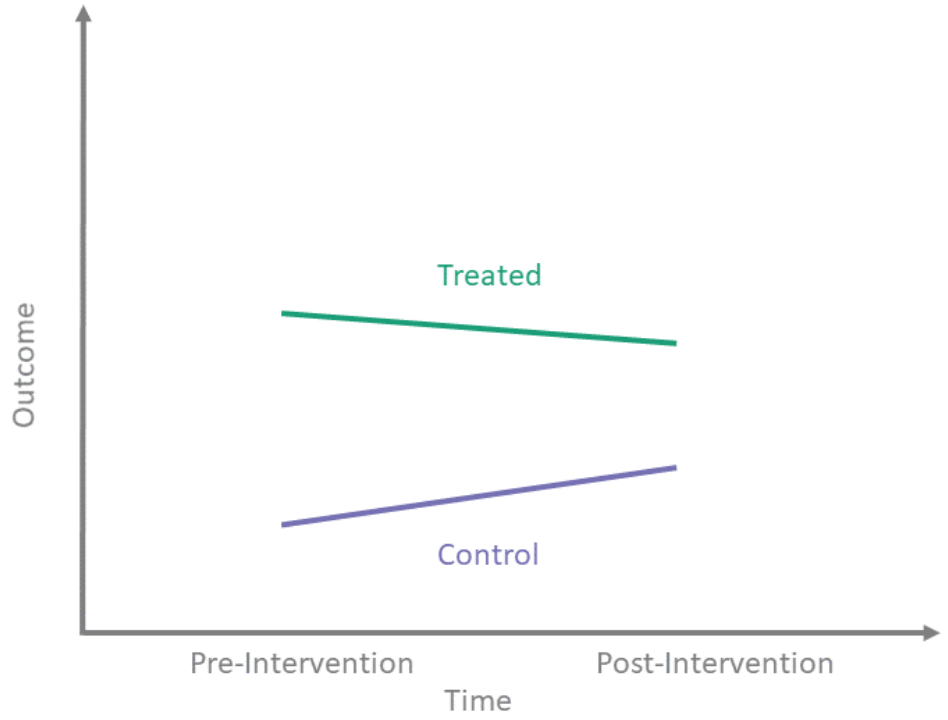
Using the Generalized Synthetic Control Method to Estimate the Impact of Extreme Weather Events on Population Health

Paige Sheridan,[a,b] Sara McElroy,[a,b] Joan Casey,[c] and Tarik Benmarhnia[a,d]

- If we can argue that the timing of a given extreme weather event is independent with the trend for the outcome of interest, we can capitalize on this as a natural experiment
- Multiple approaches exist to capitalize on such natural experiments
  - Difference-in-differences methods
  - Synthetic control methods
  - Interrupted time series analyses

The difference-in-differences idea



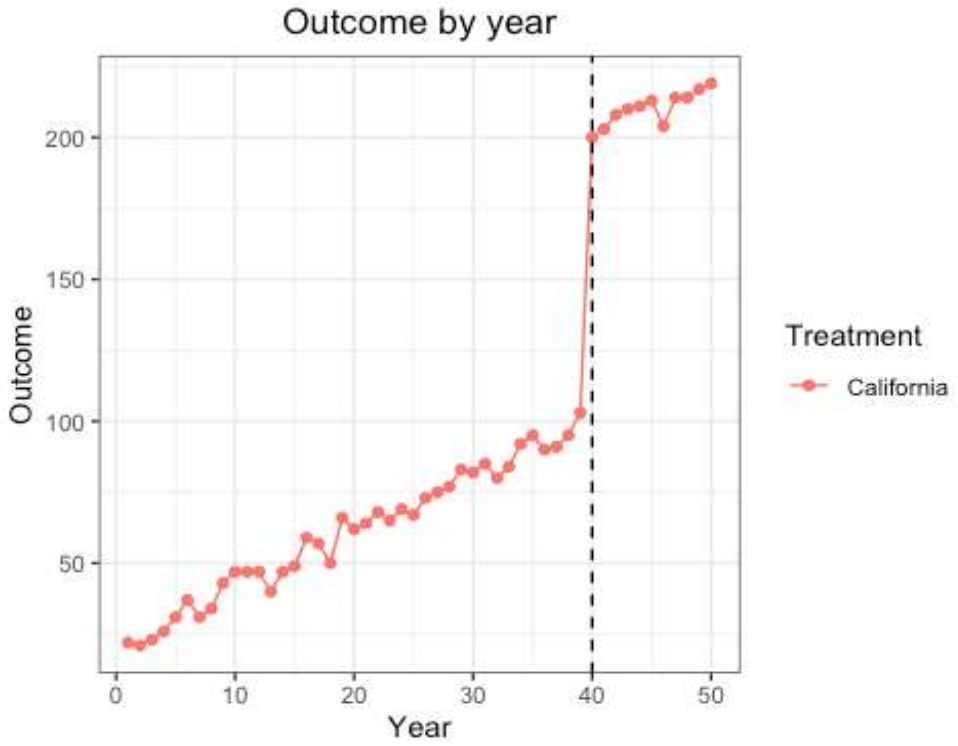


By Bret Zeldow and Laura Hatfield

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

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)

TWO-STAGE TIME SERIES ANALYSIS COUPLED WITH MACHINE LEARNING: EVALUATING THE HEALTH EFFECTS OF THE 2018 WILDFIRE SMOKE EVENT IN SAN FRANCISCO COUNTY AS A CASE STUDY

- We compared multiple algorithms in a 2-stage ITS approach
  - Autoregressive Integrated Moving Average (ARIMA)
  - NNETAR (Neural Network)
  - Prophet-XGBoost

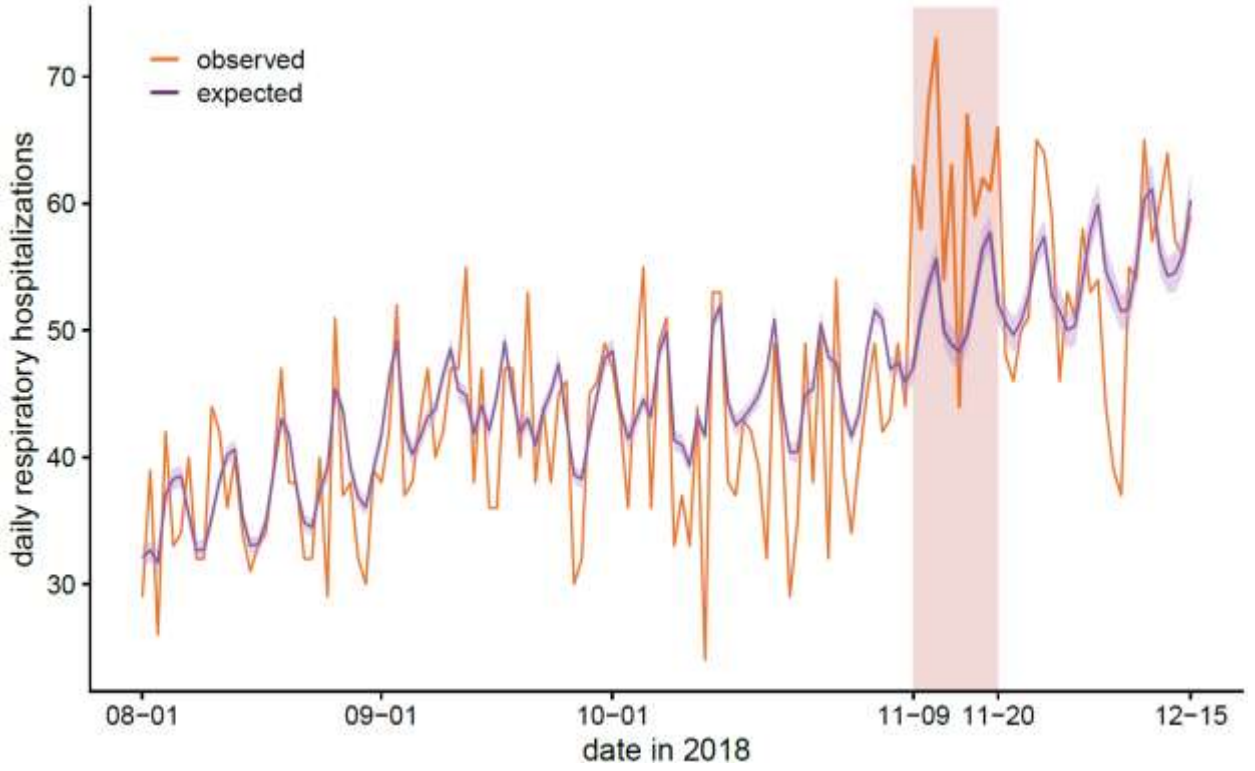


Table 1. Performance metrics of the ARIMA, NNETAR, and Prophet-XGBoost model

	Training (2009-01-01 to 2016-11-07)			Testing (2016-11-08 to 2018-11-07)		
	ARIMA	NNETAR	Prophet-XGBoost	ARIMA	NNETAR	Prophet-XGBoost
R^2	0.71	0.83	0.83	0.65	0.58	0.71
MAE	7.08	5.52	5.33	8.67	9.36	8.11
RMSE	9.19	7.12	7.02	11.84	13.00	10.85
MAPE	0.14	0.11	0.11	0.16	0.16	0.15
SMAPE	0.13	0.11	0.10	0.16	0.17	0.15

R^2: coefficient of determination
MAE: mean absolute error
RMSE: root mean square error
MAPE: mean absolute percentage error
SMAPE: symmetric mean absolute percentage error

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]

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

Research

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.

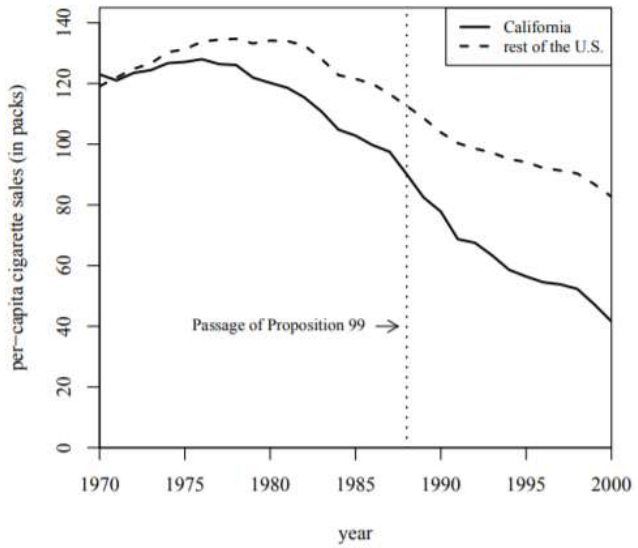
Sensitivity analyses	DID estimate	95% CI	*p*-Value[a]
Arbitrary programs			
Program implemented in 2000[b]	0.94	(–2.08, 3.96)	0.54
Program implemented in 2002[c]	0.42	(–3.62, 2.77)	0.80
Other hot days definitions			
When maximum temperature is above 28°C	0.58	(–1.77, 2.93)	0.63
When maximum temperature 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

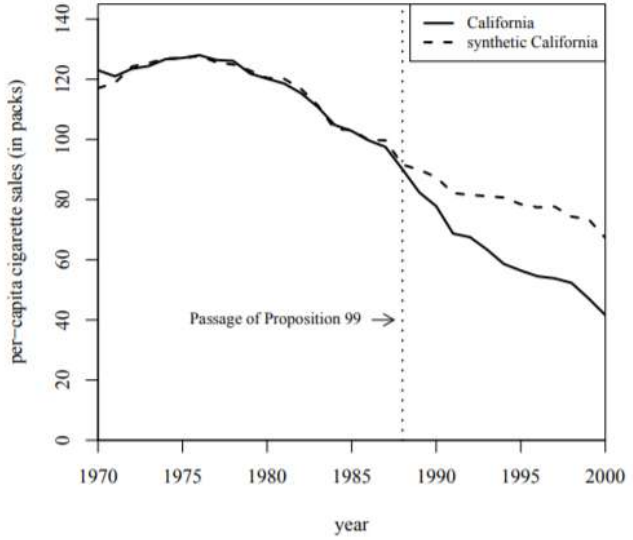
[a]*p*-Values are obtained from a Wald test on the interaction term (i.e., DID estimate).
[b]Using mortality and temperature data for periods 1996–1999 vs. 2000–2003.
[c]Using mortality and temperature data for periods 1998–2001 vs. 2002–2005.
[d]Considering a cumulative heat effect up to 5 consecutive 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), $(\boldsymbol{X_1 - X_0 W})' \boldsymbol{V} (\boldsymbol{X_1 - X_0 W})$
  - X_1 and X_0 contain the pre-treatment outcomes and covariates for the treated unit and control units respectively, and ***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

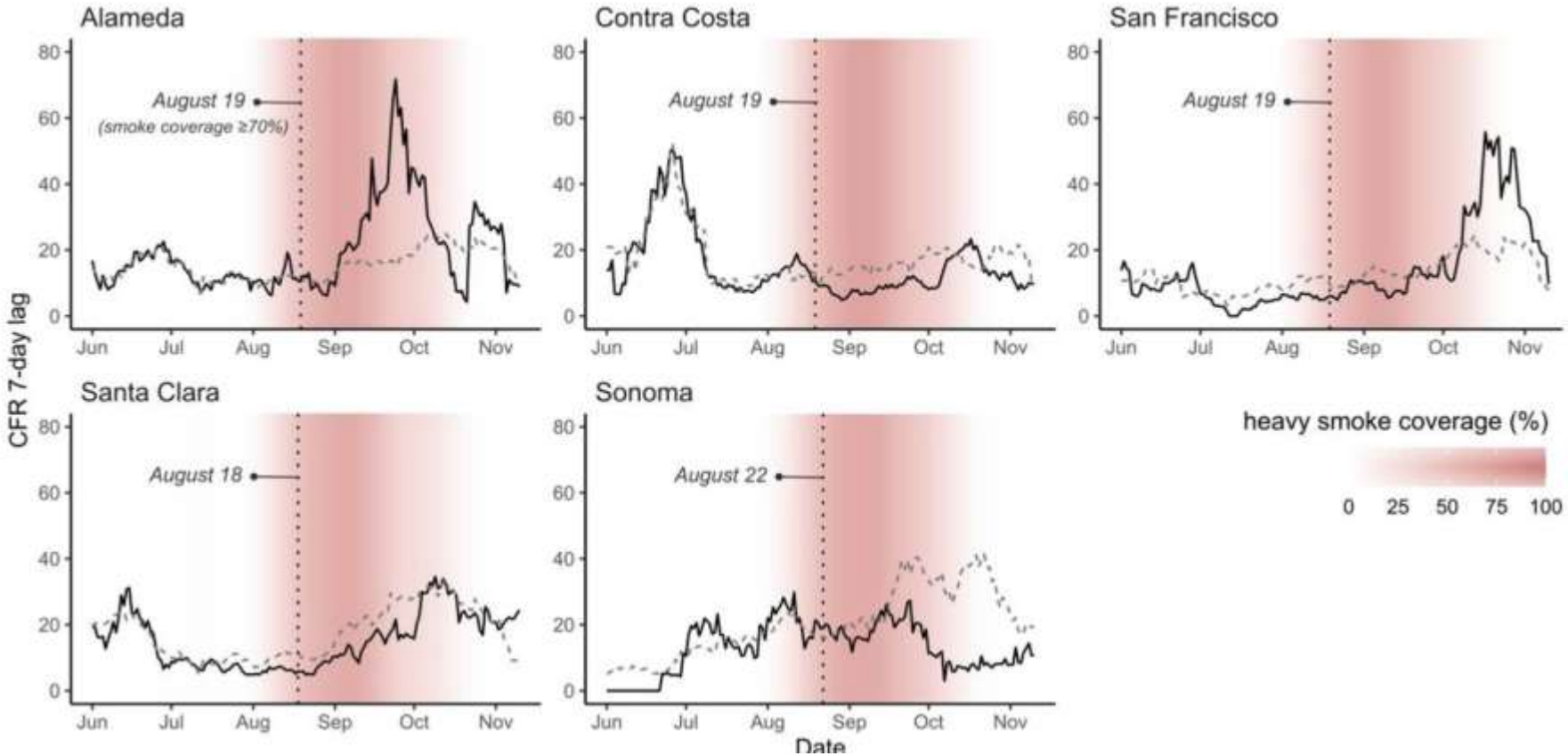
Cassandra K. Crifasi *, John Speed Meyers, Jon S. Vernick, Daniel W. Webster

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]



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

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

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

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

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

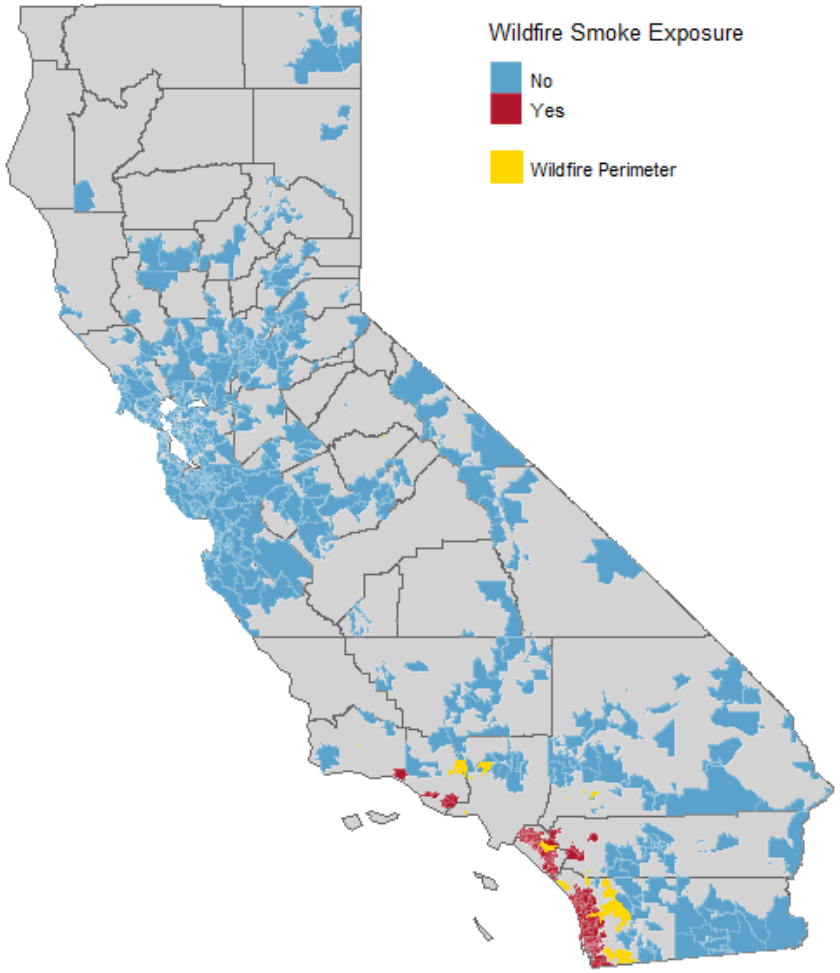
Sheridan P, McElroy S, Casey J, Benmarhnia T. Using generalized synthetic control to estimate the impact of extreme weather events on population health. Epidemiology [In Press]

RESULTS

Out of 1779 zip code tabulation areas in California

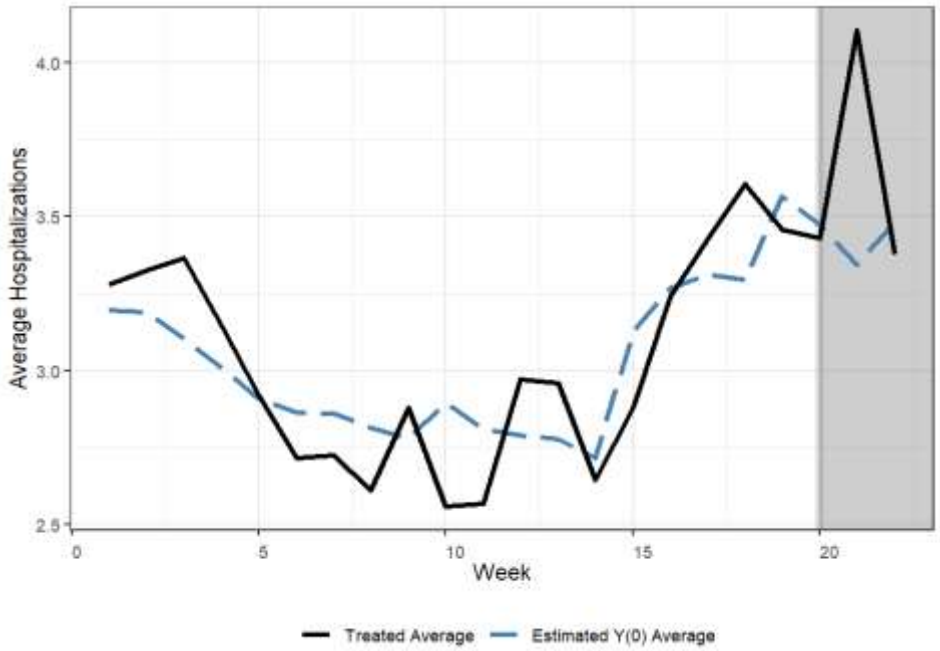
685 were included:

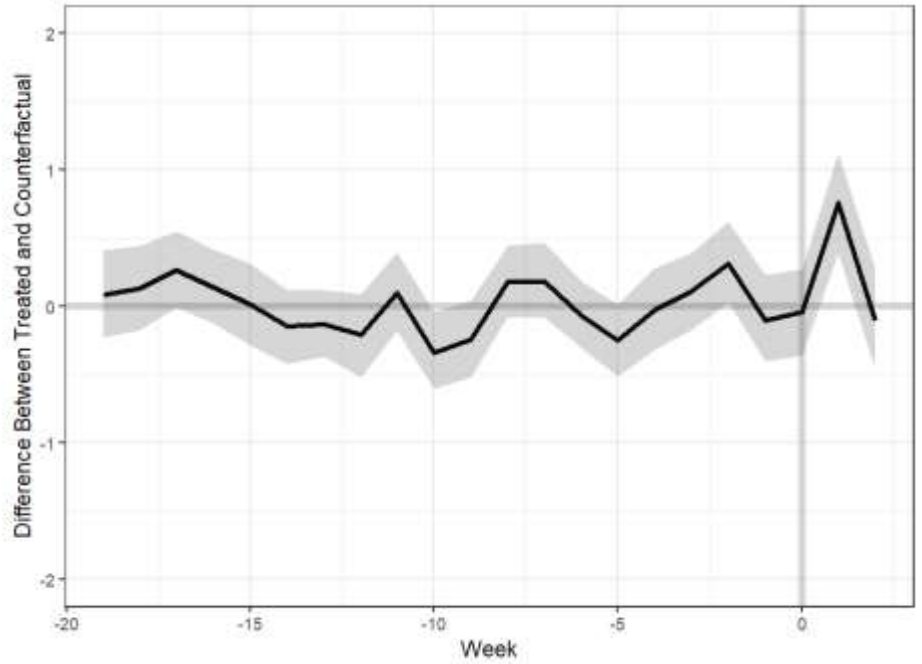
- 130 exposed
- 555 unexposed



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

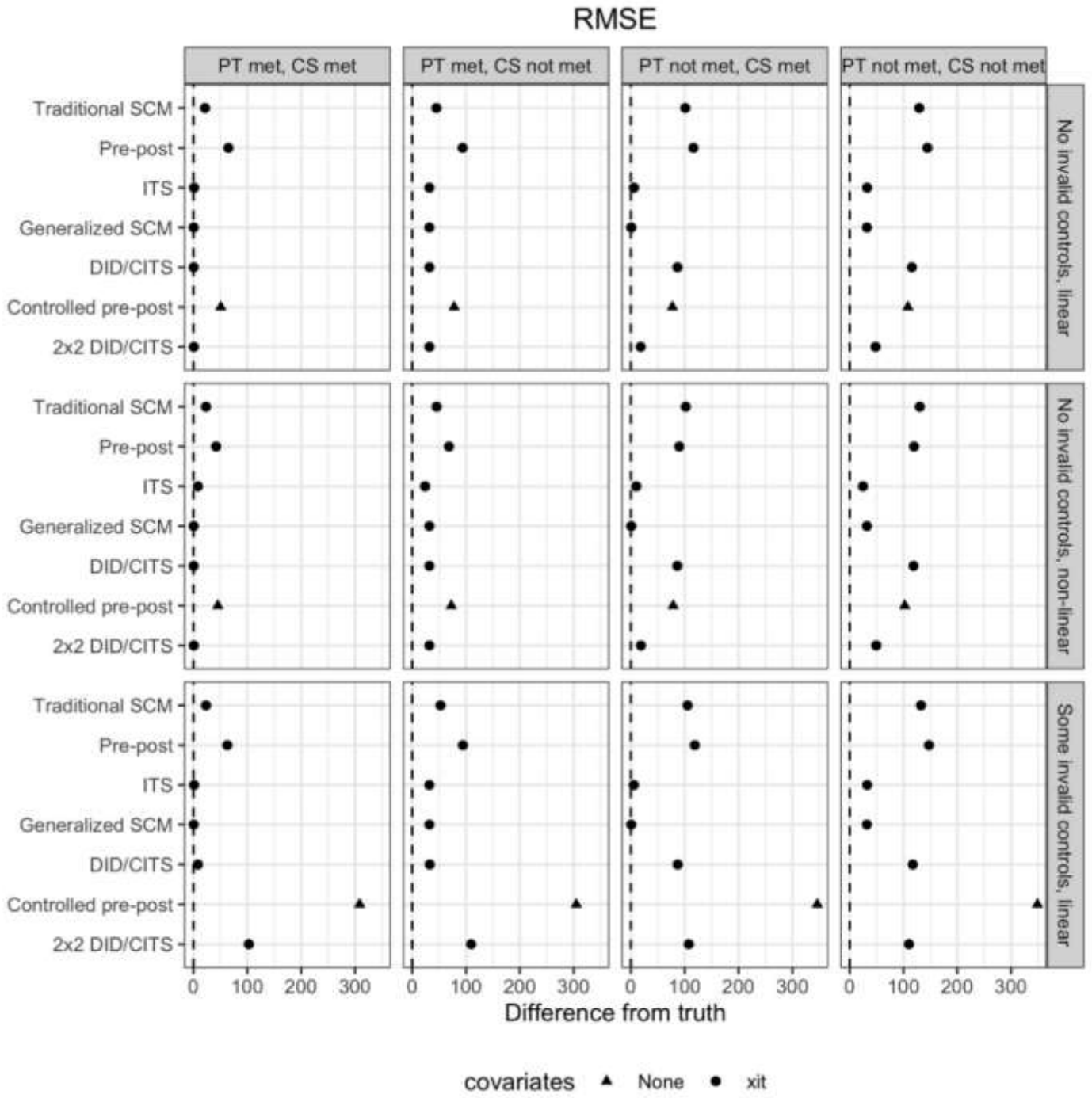
A COMPARISON OF THE DIFFERENT APPROACHES DISCUSSED TODAY

- 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

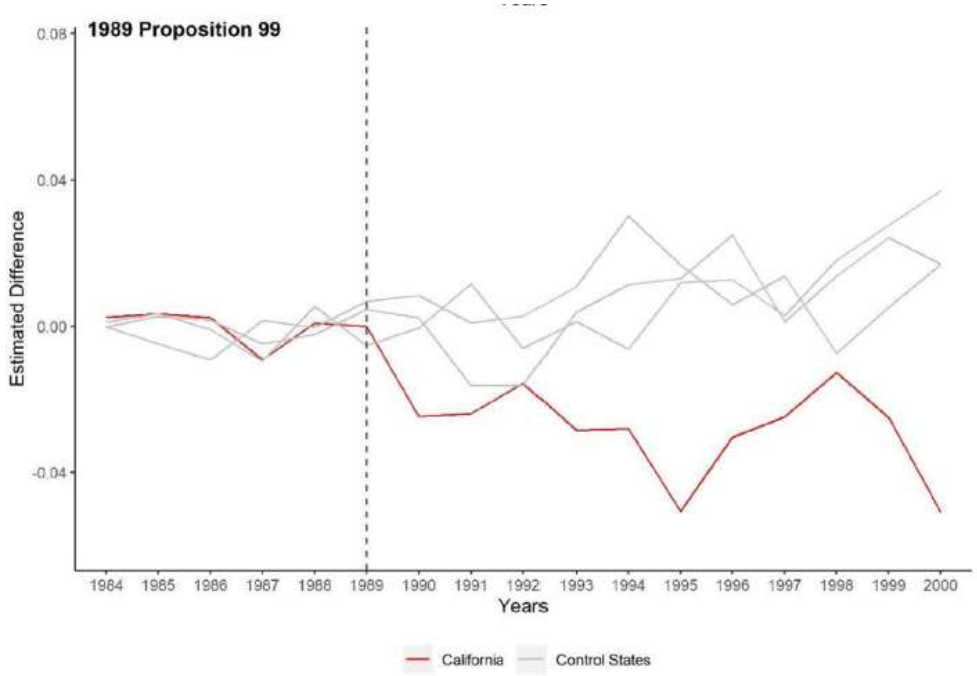
Roch A Nianogo ,[1,2]* Tarik Benmarhnia[3] and Stephen O'Neill[4]

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
No invalid controls, linear
No invalid controls, non-linear
Some invalid controls, linear
Traditional SCM
Pre-post
ITS
Generalized SCM
DID/CITS
Controlled pre-post
2x2 DID/CITS
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.

BREAK

A QUICK OVERVIEW ON EFFECT MEASURE MODIFICATION

EFFECT MODIFICATION

- We say that M is a modifier of the effect of A on Y when the average causal effect of A on Y varies across levels of M.
- Since the average causal effect can be measured using different effect measures (e.g., risk difference, risk ratio), the presence of effect modification depends on the effect measure being used
  - This is why we talk about effect measure modification

Additive effect modification:

$$\mathrm{E}[Y^{a=1} - Y^{a=0} | M = 1] \neq \mathrm{E}[Y^{a=1} - Y^{a=0} | M = 0]$$

Multiplicative effect modification:

$$\frac{\mathrm{E}[Y^{a=1}|M=1]}{\mathrm{E}[Y^{a=0}|M=1]} \neq \frac{\mathrm{E}[Y^{a=1}|M=0]}{\mathrm{E}[Y^{a=0}|M=0]}$$

THE DIFFERENCE BETWEEN CONFOUNDING AND EFFECT MEASURE MODIFICATION

- We want to condition on or control/adjust for confounding
  - This is a bias
  - Using different techniques
- EMM is not a bias so we do not need to control/adjust for it
- A variable can be both a confounder and an EMM
- Instead, we assess EMM to better understand the mechanisms underlying a specific causal association

Example with RCTs

THE DIFFERENCE BETWEEN INTERACTION AND EFFECT MODIFICATION

- The concept of effect (measure) modification refers to the causal effect of A, not to the causal effect of E
  - Only A is considered to be a variable on which we could hypothetically intervene.
- The concept of interaction refers to the joint causal effect of two treatments A and E
  - Interaction involves the counterfactual outcomes $Y^{a,e}$ under a joint intervention
  - Identifying interaction requires exchangeability, positivity, and consistency for both treatments.
  - When treatment E is randomly assigned, then the concepts of interaction and effect modification coincide

WHY DO WE WANT TO ASSESS EFFECT MODIFICATION ?

1. The identification of effect modification may help understand the biological, social, or other mechanisms
2. Understanding Disparities: evaluating the presence of effect modification is helpful to identify the groups of subjects that would benefit the most from an intervention
   - Additive, but not multiplicative, effect modification is the appropriate scale to identify the groups that will benefit the most from intervention (see next slide)
3. If the average causal effect differ between populations with different prevalence of M, it is important for generalizability/transportability:
   - in the presence of an EMM, the average causal effect in this population may not be transportable to other populations with a different distribution of effect modifiers.

METHODS FOR ASSESSING EFFECT MODIFICATION

- Stratified analyses
  - You need to conduct a heterogeneity test
  - Wald test, Cochran Q test…
- Introducing an interaction term in statistical models
  - For additive models (linear models), additive interactions is estimated
  - Not for multiplicative models
- Novel methods for high-dimensional heterogeneous effects

Cochran Q test

$$Cochran's\ Q = \left[\frac{(\beta_1-\beta_P)^2}{VAR(\beta_1)} + \frac{(\beta_2-\beta_P)^2}{VAR(\beta_2)}\right]$$

Where $\beta_1 = \ln(RRstrata_1)$; $\beta_2 = \ln(RRstrata_2)$; VAR is the variance. For the Cochran Q estimation it is necessary to conduct a χ^2 test statistic (with degrees of freedom equal to the number of strata minus 1)

NOVEL APPROACHES TO ADDRESS HIGH-DIMENSIONAL EFFECT MODIFICATION

- Traditionally estimated by a priori specification of covariates/effect modifiers
  - stratification/subgroup analysis (coupled with heterogeneity tests) or including interaction terms
- But in practice, there may be multiple effect modifiers of interest (as well as their combination)
- There is a need for automated, data-driven method to identify and quantify interpretable heterogeneity across multiple effect modifiers and their combination
- A few approaches have been recently proposed
  - Generalized Random Forests (GRF)
  - Bayesian Additive Regression Trees (BART)
  - Bayesian Causal Forests (BCF)

An overview of modern machine learning methods for effect measure modification analyses in high-dimensional settings

Michael Cheung *, Anna Dimitrova, Tarik Benmarhnia

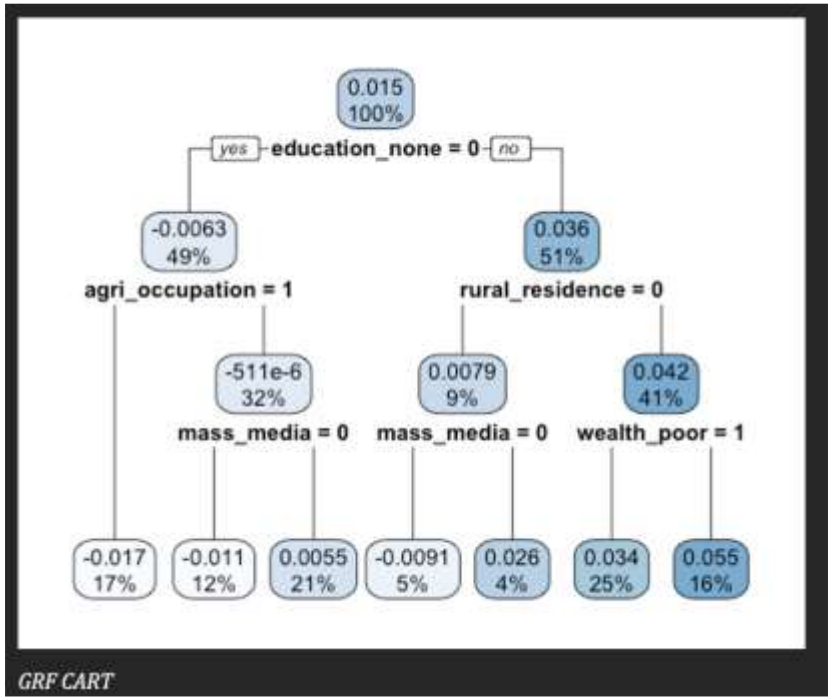
https://doi.org/10.1016/j.ssmph.2025.101764

- We summarize common machine learning methods for heterogeneous effect estimation
- We provide a practical guide to use these methods
- We demonstrate their application in R with a case study
  - Data and codes available

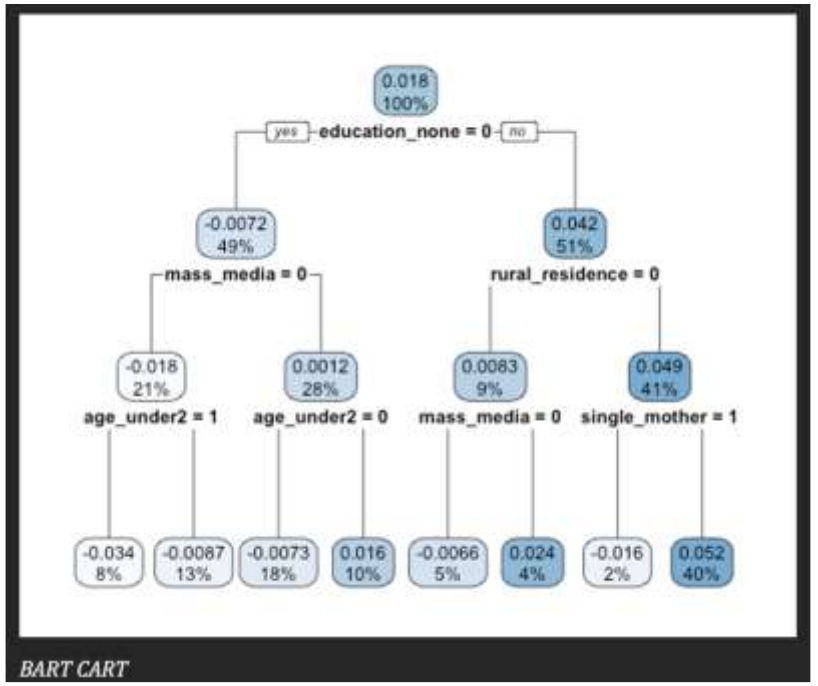
A CASE STUDY ON THE EFFECT OF DROUGHTS ON CHILD STUNTING IN SUB-SAHARAN AFRICA

- Using Demographic and Health Survey (DHS) data across 22 countries
- Children under 3 years of age
- Exposure: drought (SPEI <1 over 8 weeks)
- Outcome: stunted child growth
- Possible effect modifiers:
  - Mother's education
  - Child's age and sex
  - Birth size
  - Breastfeeding history
  - Rural/Urban
  - Agricultural occupation
  - Wealth index ..

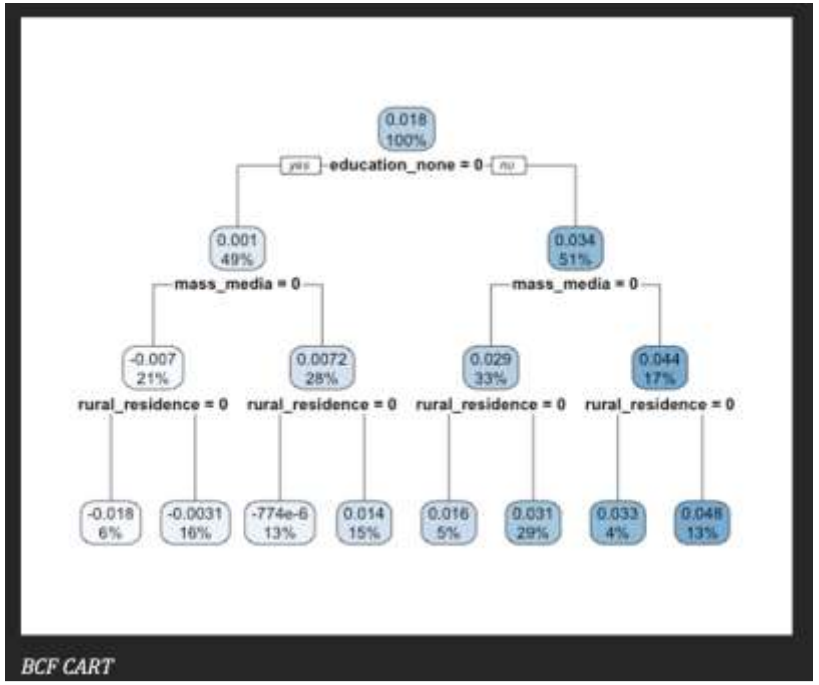
CLASSIFICATION TREES



Generalized Random Forests (GRF)

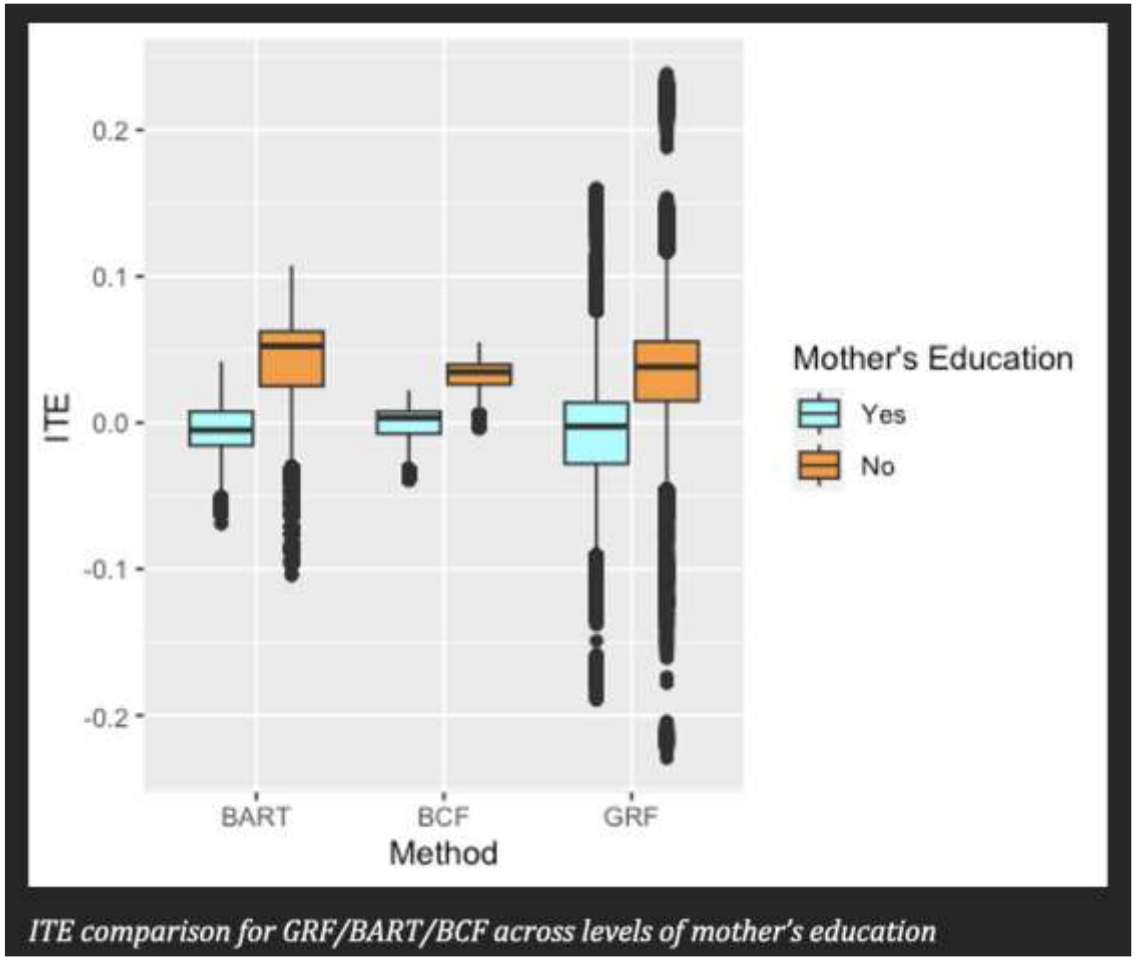


Bayesian Additive Regression Trees (BART)



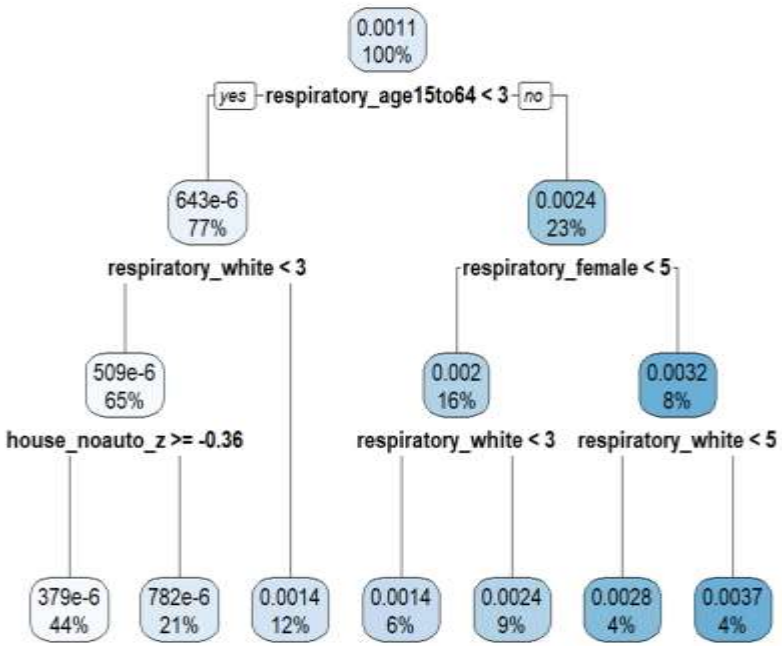
Bayesian Causal Forests (BCF)

MOTHER'S EDUCATION AS AN EXAMPLE

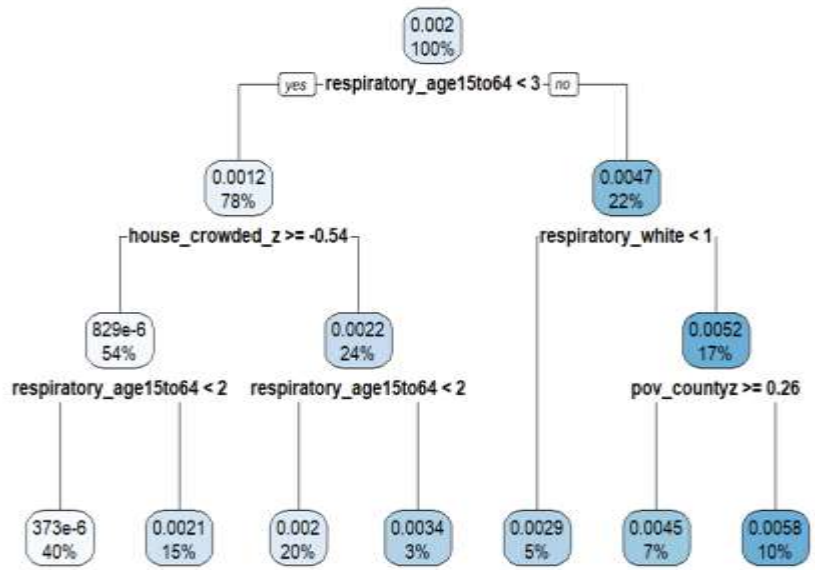


ITE comparison for GRF/BART/BCF across levels of mother's education

USING GENERALIZED RANDOM FORESTS TO CHARACTERIZE VULNERABILITY TO ADVERSE HEALTH OUTCOMES FOLLOWING WILDFIRE SMOKE EXPOSURE IN CALIFORNIA (LETELLIER ET AL.)



Sacramento Valley



San Francisco Bay

Variable Importance*	
Individual level factors	
Age 0 to 14	22.0%
Age 15 to 64	10.3%
Female	7.5%
Hispanic	6.7%
Male	4.5%
Non-Hispanic White	4.0%
Area level factors	
Unemployment rate	3.9%
A/C prevalence	3.6%
Housing age	3.3%
Black/White dissimilarity index	3.2%

*The top 10 variables with the highest variable importance are presented.

EXTERNAL VALIDITY AND TARGET POPULATION

- Internal validity refers to the validity of a causal effect estimate in the study sample → BIAS
- "**No causal effect is fully specified unless we define a target population for that causal effect**" (Maldonado & Greenland 2002)
- Target population:
  - The population in whom we ultimately want causal knowledge
  - Which population do we target when we make recommendations?
- External validity refers to the similarity between the study population (sample) and the target population
- In this context, we distinguish 2 types of external validity:
  - Generalizability
  - Transportability

Maldonado, G., & Greenland, S. (2002). Estimating causal effects. International journal of epidemiology, 31(2), 422-429.

THE IMPORTANCE OF EFFECT MODIFIERS

- The main reason for which an effect estimate in a given study may not be generalizable (to the target population) or transportable is because of a **differential distribution of effect modifiers**
- And effect (measure) modifiers are not uncommon (on a given scale)

REVISITING IDENTIFICATION ASSUMPTIONS FOR EXTERNAL VALIDITY

- External Exchangeability:
  - No differences in distributions of effect measure modifiers between the study sample and the target population
- External Positivity:
  - For example, if we have no Hispanic women over 65 in the study sample and the target population includes such population subgroup, we have an issue of external positivity
  - Such issue (as for internal validity positivity) can be addressed by model extrapolation but with some strong (and sometimes unverifiable) assumptions.
- External validity implications of the consistency violation:
  - If the exposure of interest is not consistent, the inference may not be generalizable to the target population
  - Ex: Hernán, M. A., & Taubman, S. L. (2008). *Does obesity shorten life? The importance of well-defined interventions to answer causal questions. International journal of obesity, 32(3), S8-S14*.
  - Different interventions can reduce BMI: Physical activity, Diet, Surgery
  - If there are EMMs and the distribution of EMMs differs between the study sample and the target population, results will not be generalizable.

MULTI-STAGE APPROACHES FOR DRIVERS OF SPATIAL OR TEMPORAL EFFECT HETEROGENEITY

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
  - Treated units received the intervention at different times
- 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 (Acemoglu et al. 2013; Dube and Zipperer, 2015)

MULTIPLE TREATED UNITS AND OVERALL EFFECT

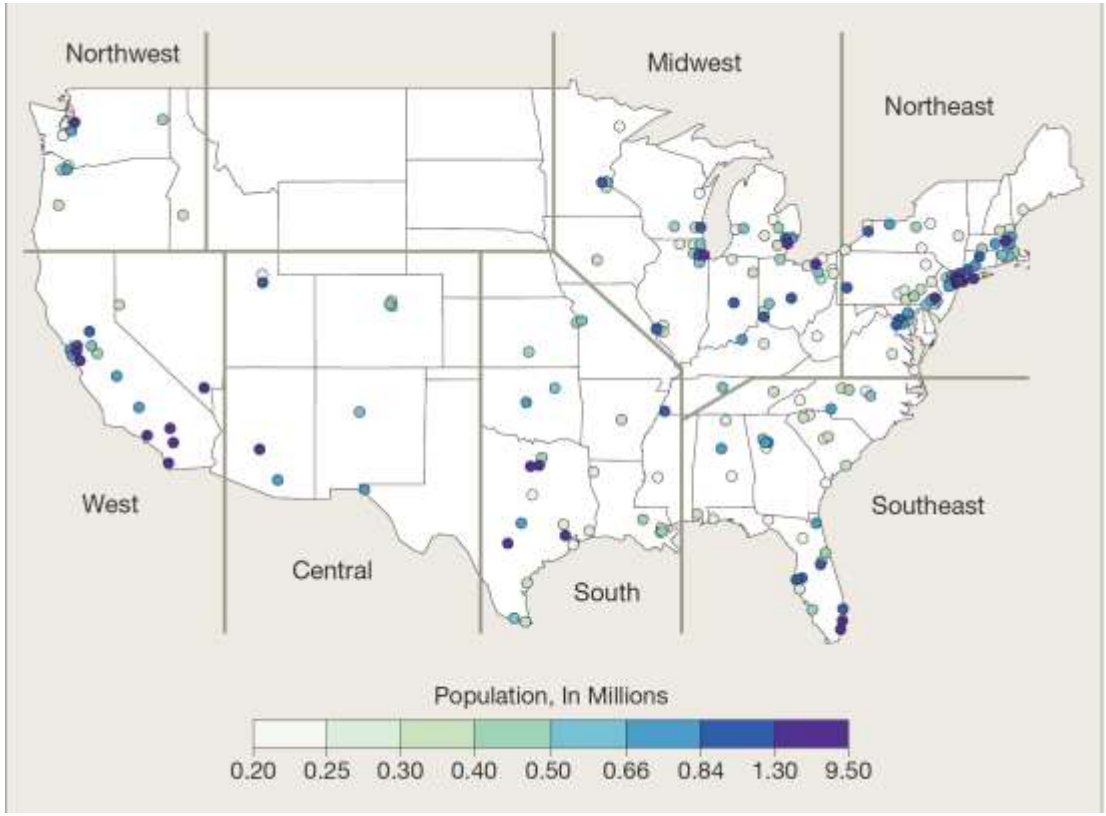
- Assuming **homogeneous concentration-response functions (i.e. association/effect) in all geographical units**, we could pool results across them to increase precision

❖ Time-series analysis: multilevel modeling
  - ❖ Bayesian hierarchical modeling
  - ❖ Meta-analysis with random intercept for each geographical unit

❖ Case-crossover design could skip this step
  - ❖ Including all risk sets in one model
  - ❖ Or stratifying by spatial units (case time series design)

- What if heterogeneity exists?

❖ **<u>This is effect modification</u>**



Counties included in one study of air pollution (Dominici et al. 2006)

MULTIPLE TREATED UNITS AT DIFFERENT TIMES

- A very active area of research
  - Goodman-Bacon (2018) proposed a solution based on a weighted average of all possible DID estimators (using some groups multiple times) in the sample of interest. This approach requires an additional identifying assumption of time-invariant treatment effects.
  - Callaway and Sant'Anna (2020) proposed an analytical solution for such case where there are more than two time periods and units that can become treated at different points in time while relaxing the time-invariant treatment effects assumption

WHAT DO WE MEAN BY STAGGERED INTERVENTIONS?

- When a given policy/treatment is affecting multiple units <u>but at different times points</u>

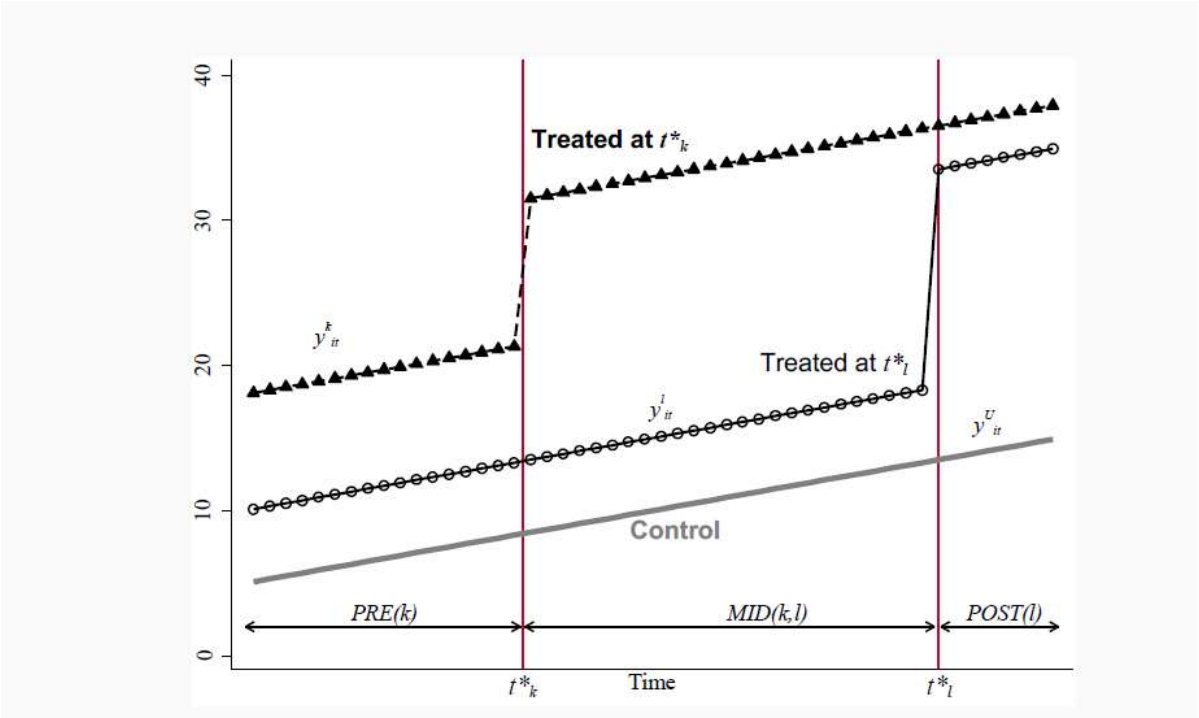


Figure from Goodman-Bacon, 2021

WHY DO TRADITIONAL METHODS FAIL IN SUCH SETTINGS?

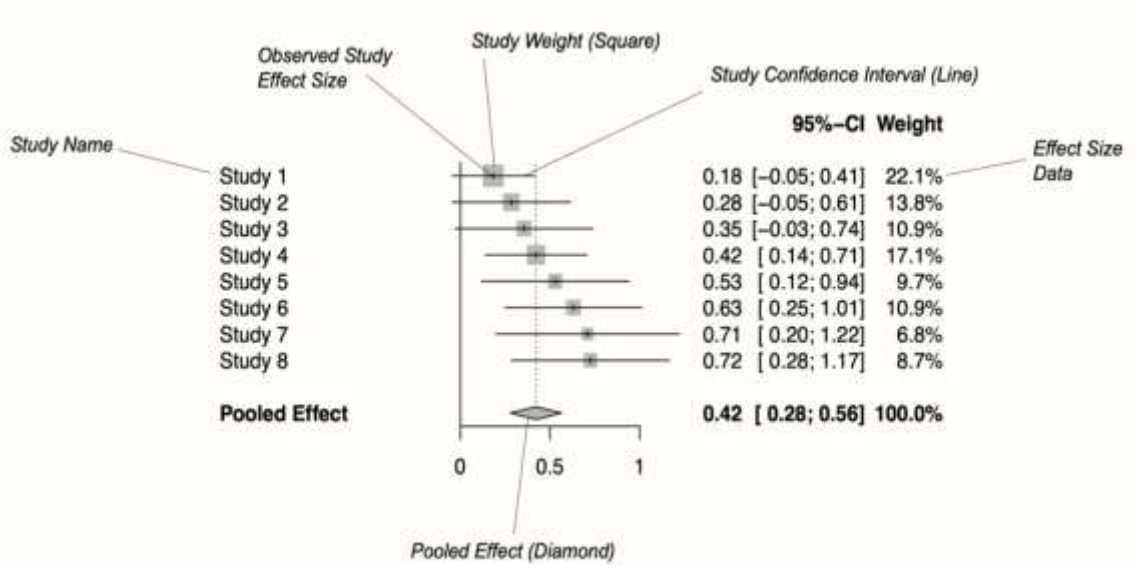
- Simple settings for DID and SCM do not apply here for a few reasons:
  - Each treated unit requires a suitable control group specific the timing of the policy/treatment
  - Units that receive the policy/treatment later may be used as control unit earlier in the study period
    - Some potential statistical dependence issues
  - **We do not want to assume that a given policy/treatment has the same effect across units and over time**

ANALYTICAL APPROACHES TO DEAL WITH STAGGERED INTERVENTIONS

- Multiple approaches have been proposed in the past few years
  - A very active area: De Chaisemartin, C., & d'Haultfoeuille, X. (2022)
- We follow the 2-stage approach proposed by Callaway and Sant'Anna (2020)
- This approach requires:
  - To have –enough- never treated units
  - To assume independence regarding the timing of the implementation over treated units
  - Similar assumptions as traditional settings

A QUICK NOTE ON META-ANALYSES AND META-REGRESSIONS

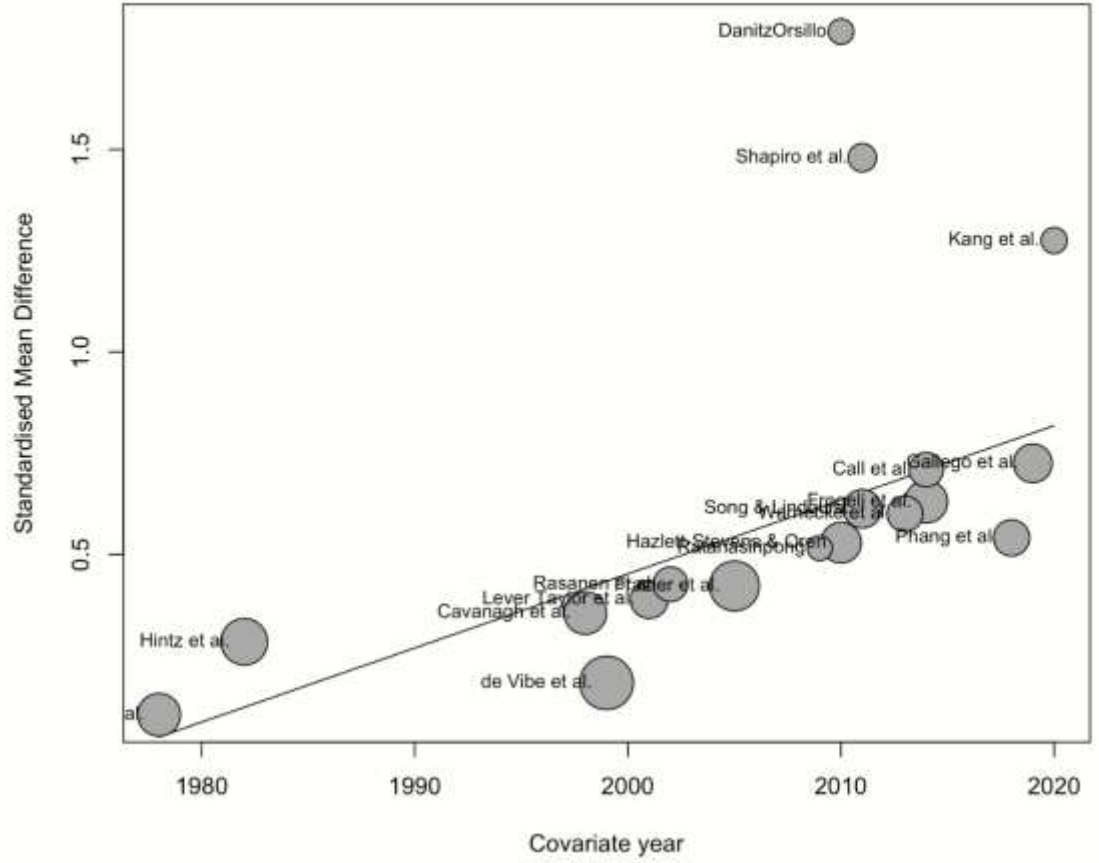
- Meta-analyses are typically used to pool effect estimates from multiple studies
- This is a form of multilevel model, in which participants are nested within studies
- We obtain a pooled effect estimate by applying a weighting procedure (usually based on the inverse of the standard error)
- We can also quantify the level of heterogeneity across studies
- If there is some heterogeneity, we then can conduct a meta-regression in which the dependent variable is the study-specific effect estimate and independent variables study-specific characteristics (e.g. study period, population composition etc…)



https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/forest.html

HOW DO META-REGRESSION WORK?

- An intuitive way to describe meta-regression is a weighted regression based on the study-specific variance
  - If all studies have the same variance: meta-regression = simple regression
- We then obtain a slope for each independent variable, and we can understand the drivers of effect estimates heterogeneity

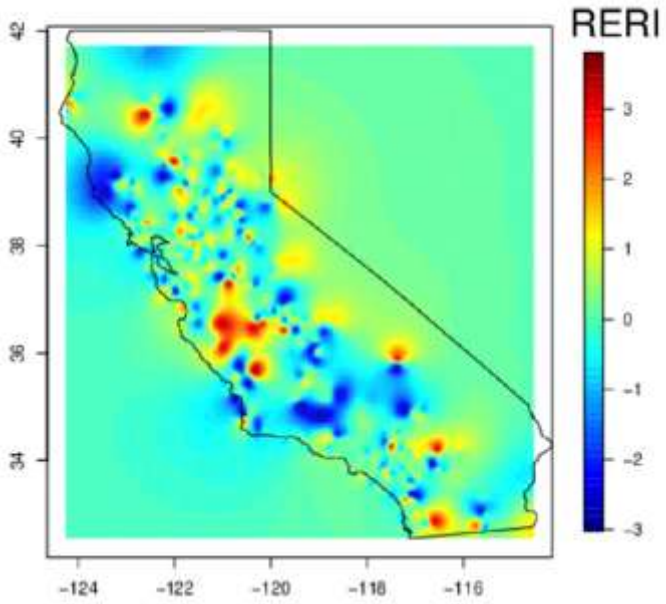


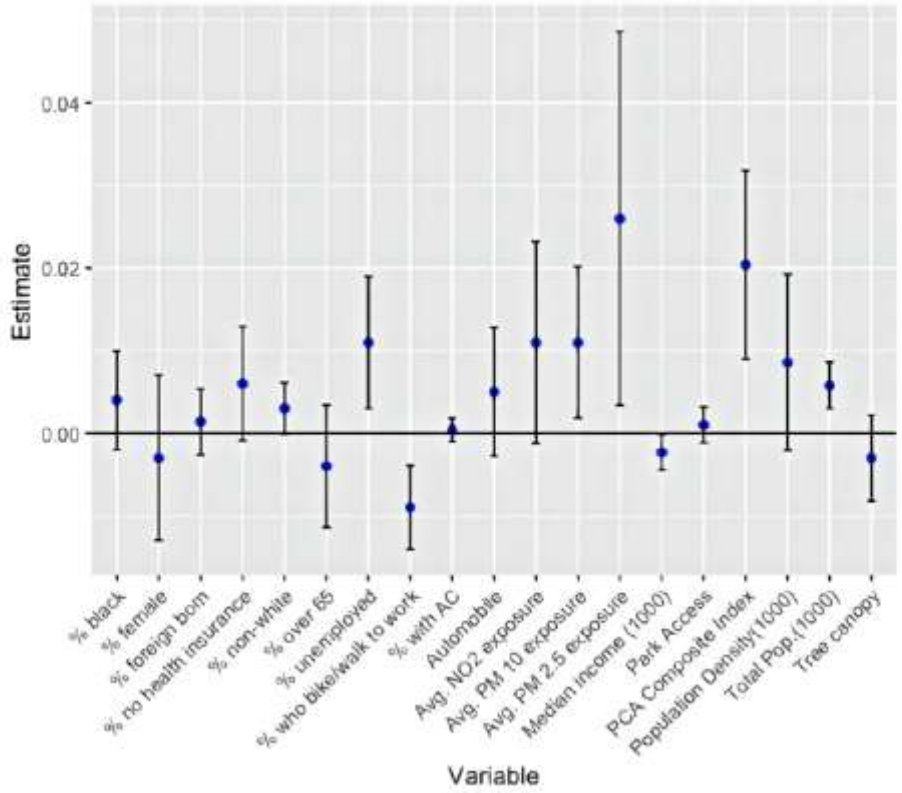
https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/metareg.html

AN EXAMPLE

Spatial variation in the joint effect of extreme heat events and ozone on respiratory hospitalizations in California

Lara Schwarz[a,b,1,2], Kristen Hansen[b,1,2], Anna Alari[c], Sindana D. Ilango[d], Nelson Bernal[e], Rupa Basu[f], Alexander Gershunov[g], and Tarik Benmarhnia[b,g]



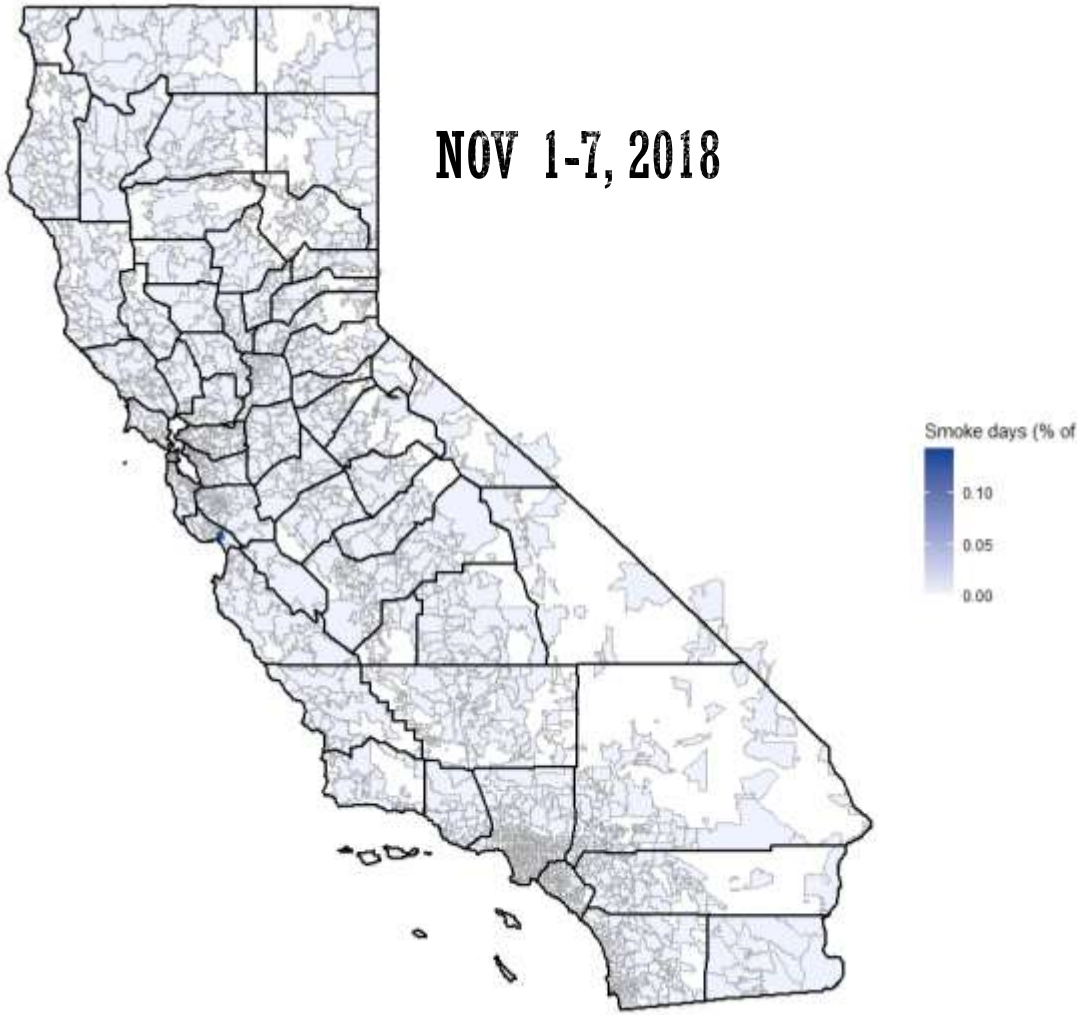


A CASE STUDY: WILDFIRE SMOKE AND COMPOUNDED RISKS FROM RESPIRATORY INFECTIOUS DISEASES

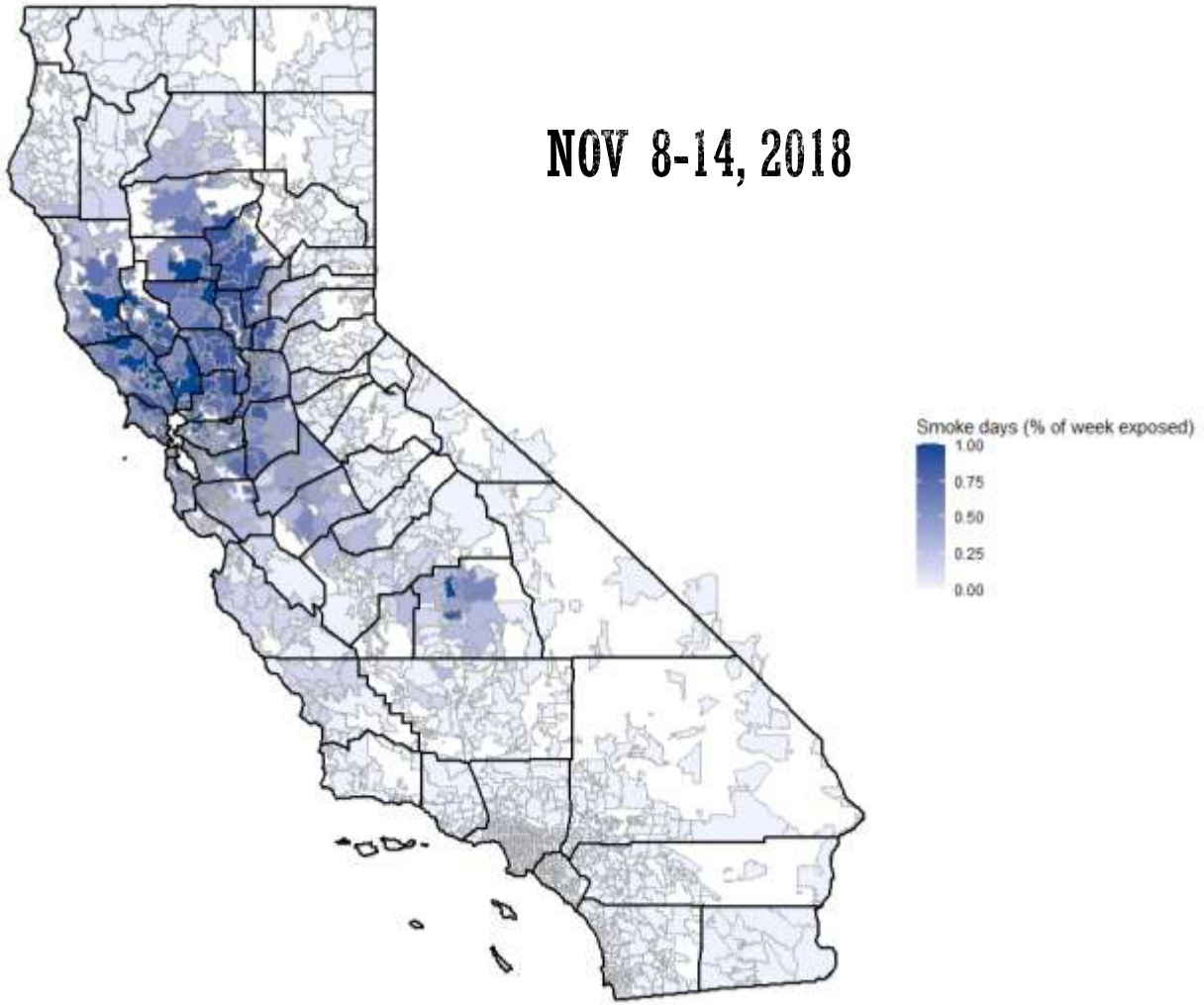
- Objectives:
  - Assess the county-specific effects of wildfire smoke on all respiratory hospitalizations using a large wildfire event in November 2018 as a case study
    - Using generalized synthetic control methods
  - Evaluate how population-level influenza frailty modifies such effects
    - Using a random effect meta-regression

TREATED ZIP CODES WITH SMOKE PM2.5>0

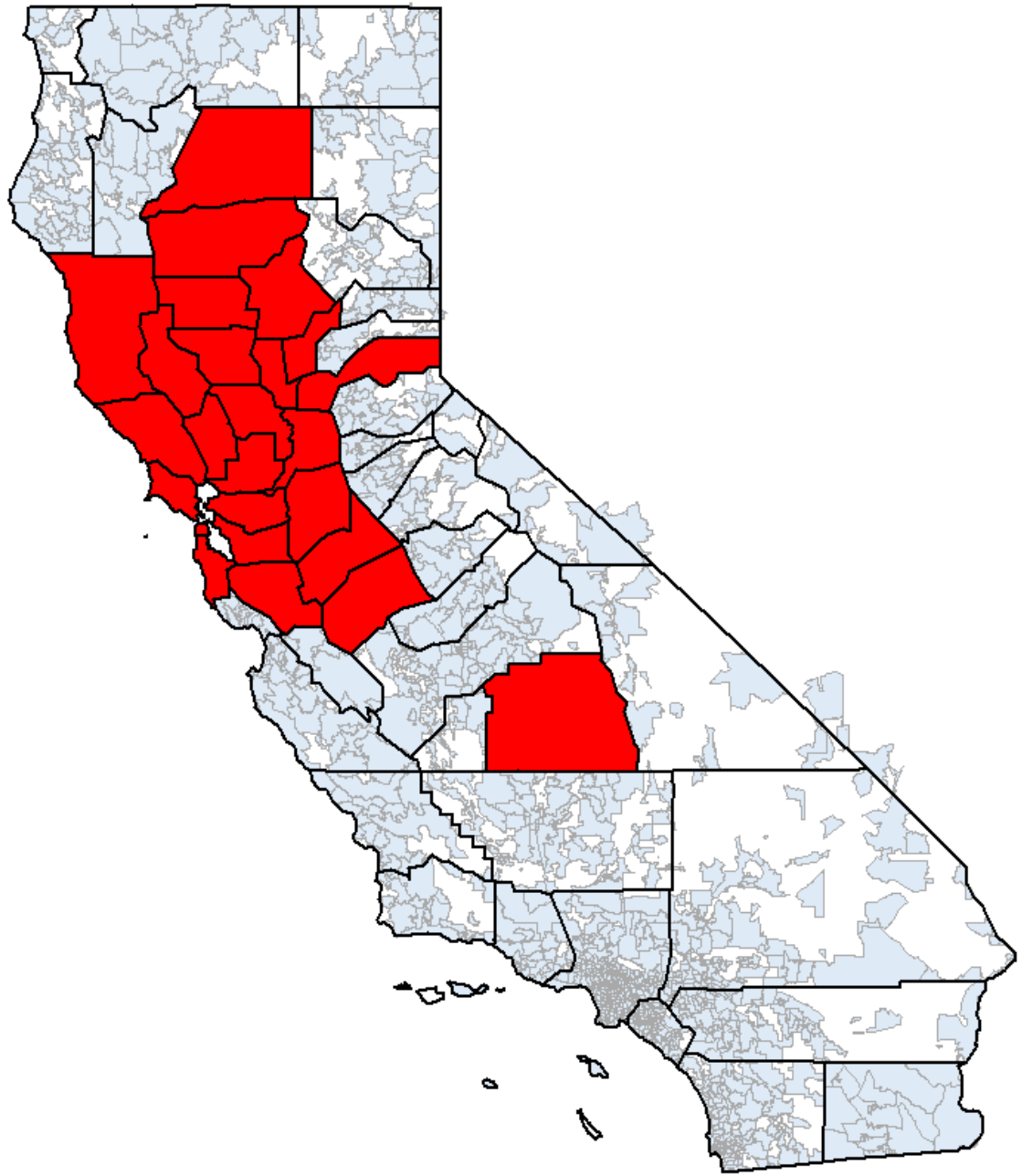
NOV 1-7, 2018



NOV 8-14, 2018



25 COUNTIES EXPOSED DURING NOV 8-14 WEEK



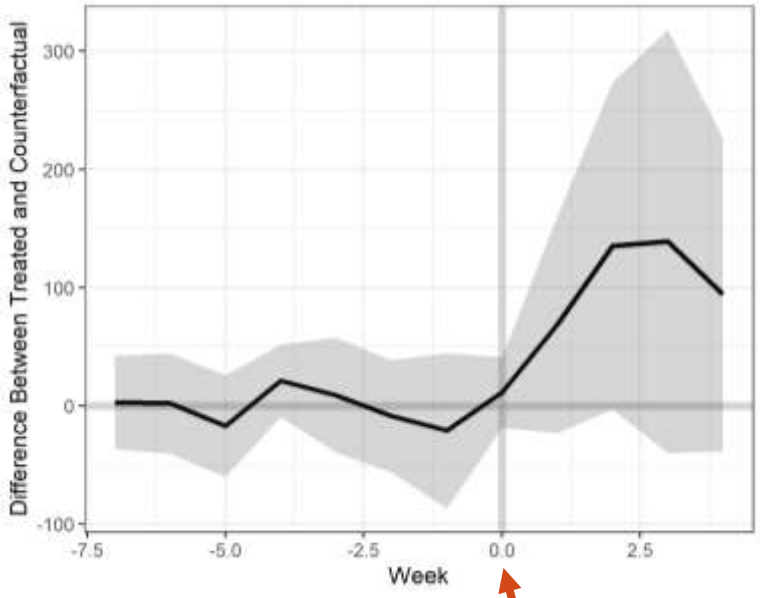
ESTIMATING INFLUENZA BURDEN Z-SCORE

- OSHPD flu data was processed at the daily zip code level
  - ICD-9 codes: 487, 488 ICD-10 codes: J09, J10, J11
- Weekly counts were summarized at the County level
- Weekly z-score computed for every County
  - Restricted to 2010-2019, flu season based on CDC (October to May)
- Average flu z-score for weeks 42-44 (**3 weeks prior** to wildfire smoke start) for each County

*OSHPD: Office of Statewide Health Planning and Development Patient Discharge Data

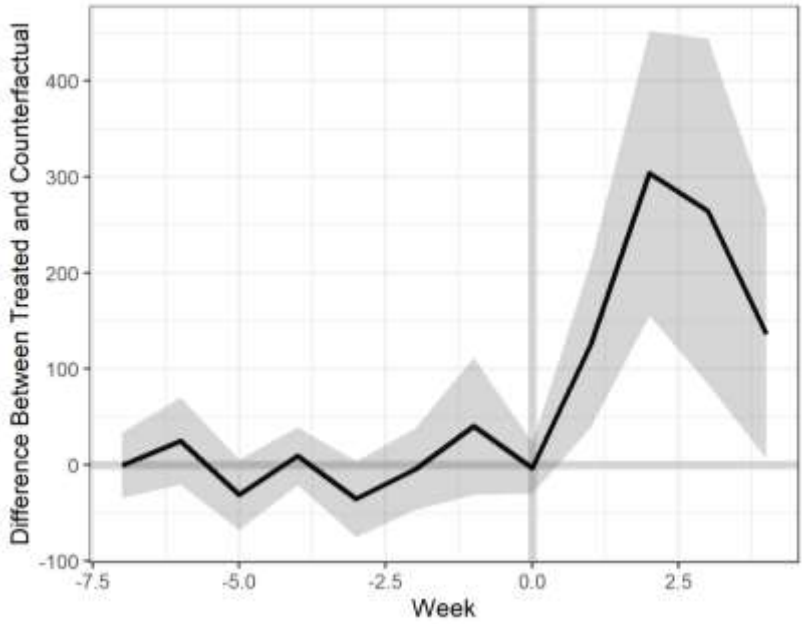
RESULTS (1/2)

SAN JOAQUIN COUNTY

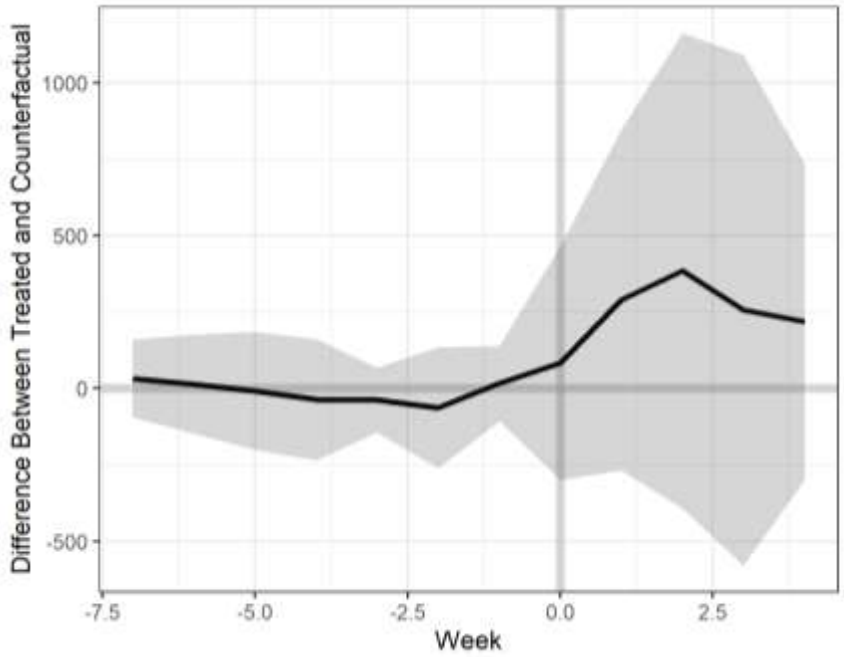


WILDFIRE START

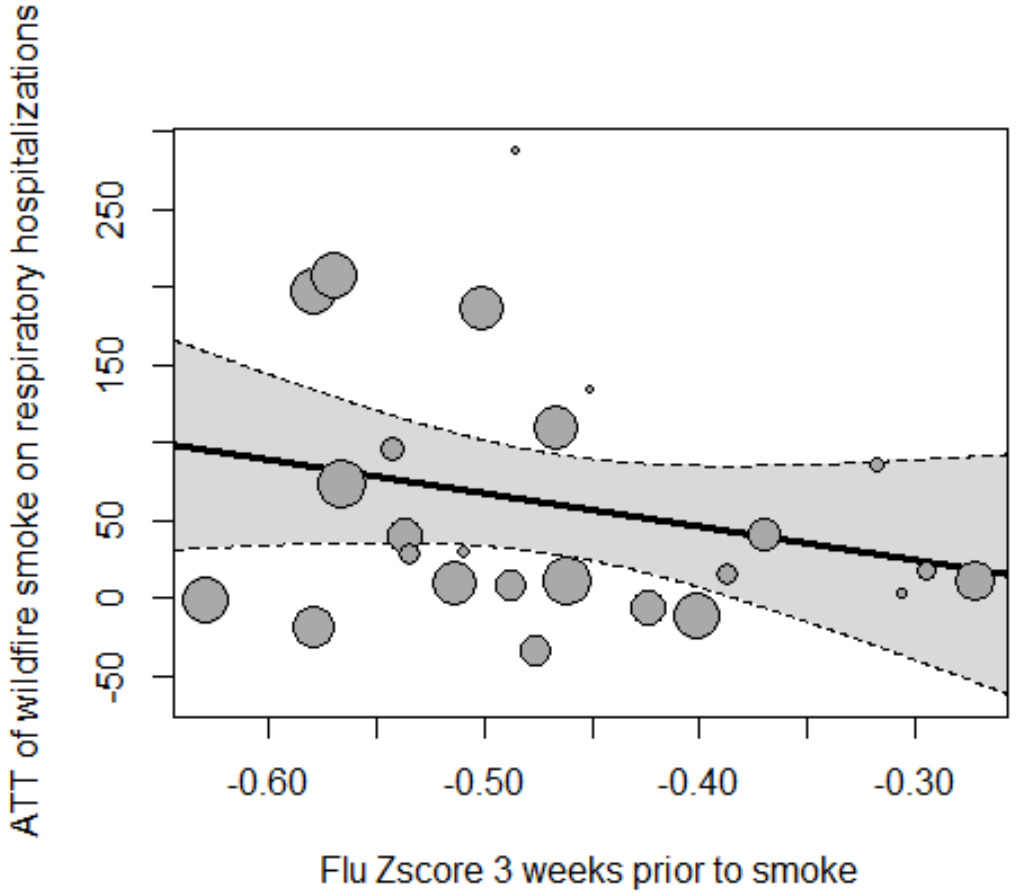
SANTA CLARA COUNTY



ALAMEDA COUNTY



RESULTS (2/2)



RESULTS FROM THE RANDOM-EFFECT META-REGRESSION MODEL

β = -215.42 (95% CI:, -550.72, -119.88)

TIME-VARYING HETEROGENEOUS HEALTH EFFECTS: EXTREME HEAT AS AN EXAMPLE

- Most of the heat-health literature focused on time-fixed effect modifiers
  - E.g. age, gender, SES, built environment features…
- Some literature about spatially compounded effects with air pollutants
  - A systematic review focused on compound exposures found consistent evidence for synergistic effects for heat and air pollution, especially for O3 and PM2.5

Anenberg SC, Haines S, Wang E, Nassikas N, Kinney PL. 2020. Synergistic health effects of air pollution, temperature, and pollen exposure: a systematic review of epidemiological evidence. Environ Health. 19:130

- **Time-varying** effect modifiers **<u>are also important</u>**
  - Seasonality, cascading events, recovery period and policy evaluation

MEDIATION ANALYSES TO BETTER UNDERSTAND THE MECHANISMS IN CLIMATE AND HEALTH: HEAT AND OZONE AS A CASE STUDY

Alari A, Schwarz L, Chen C, Hansen K, Chaix B, Benmarhnia T. *The role of ozone as a mediator in the relation between heat waves and mortality in 15 French urban agglomerations*. American Journal of Epidemiology, 192(6), 949-962.

THE LINKS BETWEEN EXTREME HEAT AND OZONE IN RELATION TO POPULATION HEALTH

- Extreme heat and ozone have been shown to impact respiratory and CVD outcomes
- Ozone is a mediator in the relationship between heat and health
  - Ozone is generated by some precursors (NOx, NO2..), sunlight and <u>when temperatures are high</u>
  - Multiple papers have proposed such mechanism, but little empirical evidence to date

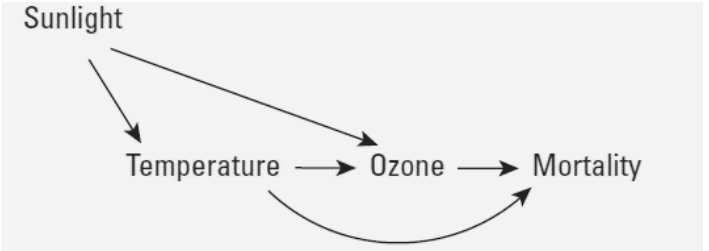
Does Air Pollution Confound Studies of Temperature?

Jessie P. Buckley,[a] Jonathan M. Samet,[b] and David B. Richardson[a]

The Role of Ambient Ozone in Epidemiologic Studies of Heat-Related Mortality

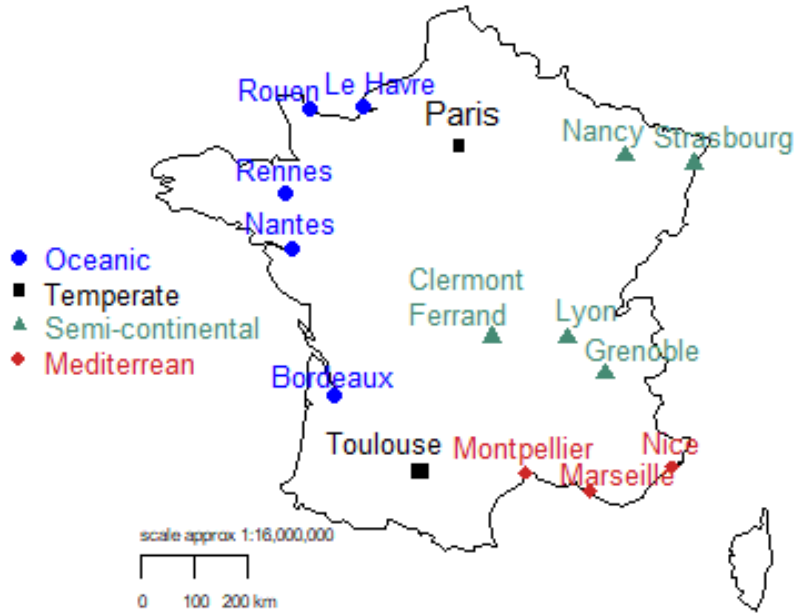
Colleen E. Reid,[1] Jonathan M. Snowden,[2,3] Caitlin Kontgis,[4] and Ira B. Tager[2]

[1]Department of Environmental Health Sciences, and [2]Division of Epidemiology, School of Public Health, University of California, Berkeley, Berkeley, California, USA; [3]Department of Obstetrics and Gynecology, Oregon Health and Science University, Portland, Oregon, USA; [4]Department of Geography, Center for Sustainability and the Global Environment, University of Wisconsin–Madison, Madison, Wisconsin, USA

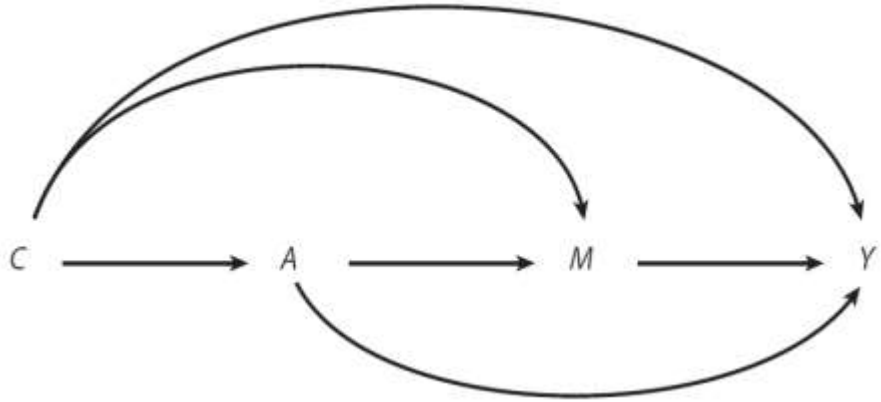


THE ROLE OF OZONE AS A MEDIATOR IN THE RELATION BETWEEN HEAT WAVES AND MORTALITY IN 15 FRENCH URBAN AGGLOMERATIONS

- Aims of this paper:
  - To decompose the total effect of heat wave on mortality into natural direct and indirect effects via increasing ozone levels
  - Compare the proportion mediation across French agglomerations
- Data Sources
  - 15 major French urban agglomerations, summer period (years 2000 to 2015)
  - We used the Official Météo-France Heat Wave definition for each city
    - We also analyzed several alternative HW definitions
  - Ozone and No2 obtained from air pollution monitors in each city
  - Mortality: daily counts for non-accidental deaths
    - We also considered respiratory and CVD deaths



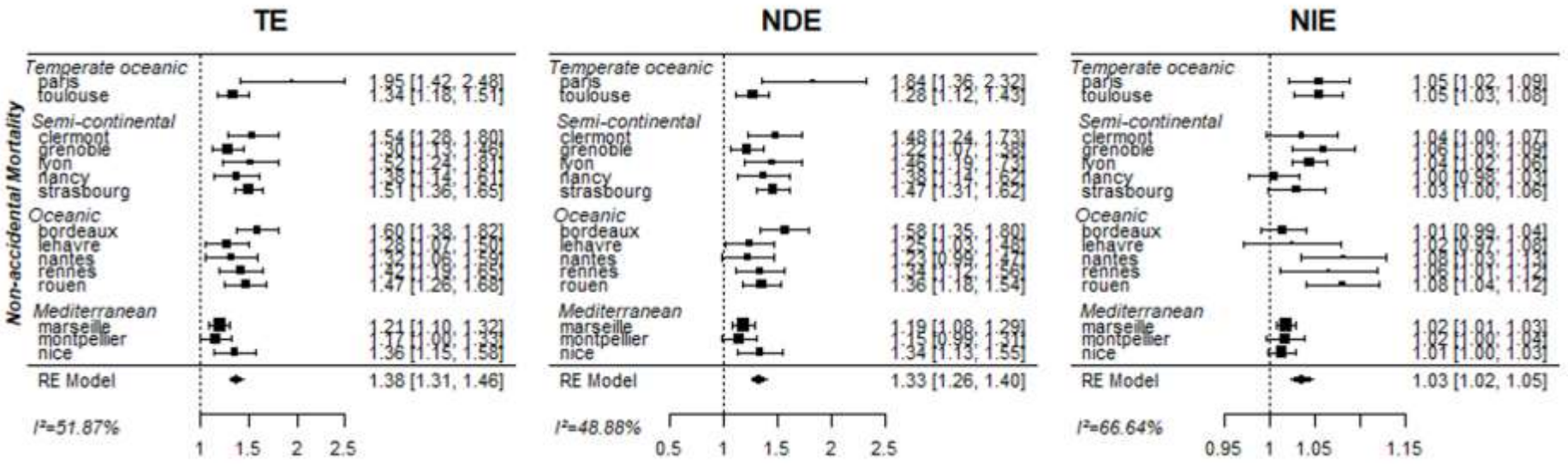
ANALYTICAL APPROACH



- We conducted a causal mediation analysis through a regression-based approach coupled with DLM
  - We estimated total effects, natural direct and indirect effects
  - With 4 identification assumptions
  - We did not identify any E-M interaction, so NDE and CDE were similar
- Two sequential regressions
  - $E[Y|hw, o3, c] = exp(\theta_0 + \theta_1 hw + \theta_2 o3 + \sum \theta_n c_n)$
  - $E[o3|hw, c] = \beta_0 + \beta_1 hw + \sum \beta_n c$
- *Then*
  - $NDE = \exp(\theta_1)$
  - $NIE = \exp(\theta_2 \beta_1)$
  - $PM = RR_{NDE}(RR_{NIE} - 1)/(RR_{NDE} * RR_{NIE} - 1)$

https://github.com/aalari/Heat-wave_O3_Mediation

MAIN RESULTS FOR NON-ACCIDENTAL MORTALITY



Urban Agglomeration	
Temperate-Oceanic climate	**PM (%)**
Paris	11%
Toulouse	20%
Semi-continental climate	
Clermont-Ferrand	10%
Grenoble	25%
Lyon	13%
Nancy	2%
Strasbourg	8%
Oceanic climate	
Bordeaux	4%
Le-Havre	11%
Nantes	31%
Rennes	20%
Rouen	23%
Mediterranean climate	
Marseille	10%
Montpellier	11%
Nice	5%

Acknowledgements

Yiqun Ma, Chen Chen and Lara Schwarz and the Climate Epidemiology Group

Sigal Maya and the Workshop Organizing Committee

Sheri Weiser, Arianne Teherani and the P20 team

THANK YOU

YIM022@UCSD.EDU
TBENMARHNIA@UCSD.EDU