Using Individual-Level Data in Difference-in-Differences for Health Policy Evaluation

Nicholas J. Seewald, PhD

Department of Health Policy and Management
Johns Hopkins Bloomberg School of Public Health

8 March 2023
University of Pennsylvania

Slides are online!

slides.nickseewald.com/upenn2023.pdf

About Me

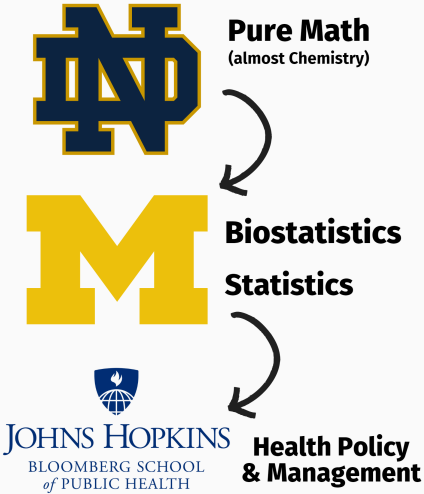
I am a collaborative statistician who builds statistical tools that enable
high quality, impactful science.

About Me

I am a collaborative statistician who builds statistical tools that enable **high quality, impactful science.**

My goal is to develop methods and collaborations that create knowledge to improve health.

How I Got Here



- (Bio)statistics let me combine interests in math and science to improve lives
- Balance between mathematical rigor and cutting-edge science
- “The best thing about being a statistician is that you get to play in everyone’s backyard.” (J.W. Tukey)

My Core Values

Core Tenets

1. Keep people first
2. Build useful, accessible methods
3. Collaboration is key

Core Strategies

1. Deep substantive engagement
2. Bridge gaps between theory and application
3. Disseminate, teach, and train

My Work, Broadly

Causal inference with complex repeated-measures data

Experimental methods:

Sequentially-randomized trials

- Design & analysis of SMARTs with longitudinal outcomes
- Micro-randomized trials

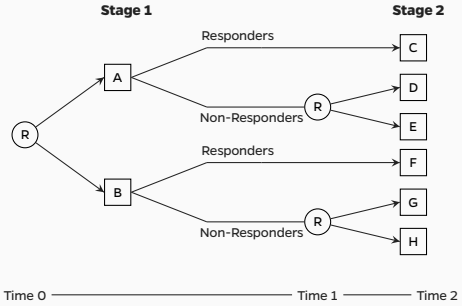
Non-experimental methods:

Health policy evaluation

- How to use multilevel data for state policy evaluation
- Appropriate variance estimation

Sequential Multiple-Assignment Randomized Trials (SMARTs)

- Multistage trials in which some or all participants are randomized more than once, often according to a *tailoring variable*.
- Typically motivated by construction of a high-quality *dynamic treatment regime*.

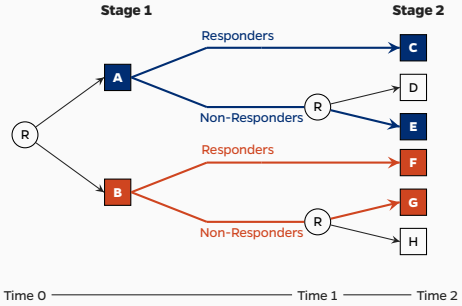


Seewald, Hackworth, and Almirall, (2021), *Principles and Practice of Clinical Trials*.

Seewald et al., (2020), *Statistical Methods in Medical Research*.

Sequential Multiple-Assignment Randomized Trials (SMARTs)

- Multistage trials in which some or all participants are randomized more than once, often according to a *tailoring variable*.
- Typically motivated by construction of a high-quality *dynamic treatment regime*.

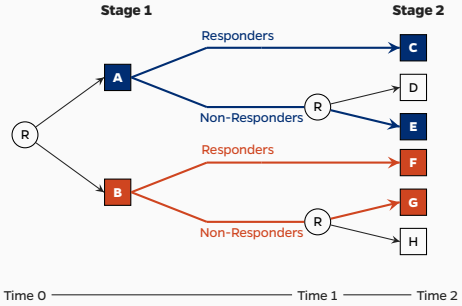


Seewald, Hackworth, and Almirall, (2021), *Principles and Practice of Clinical Trials*.

Seewald et al., (2020), *Statistical Methods in Medical Research*.

Sequential Multiple-Assignment Randomized Trials (SMARTs)

- Sample size considerations for comparing embedded dynamic treatment regimes in a SMART with a longitudinal outcome
- Trade-offs between sample size and measurement occasions subject to budget constraint

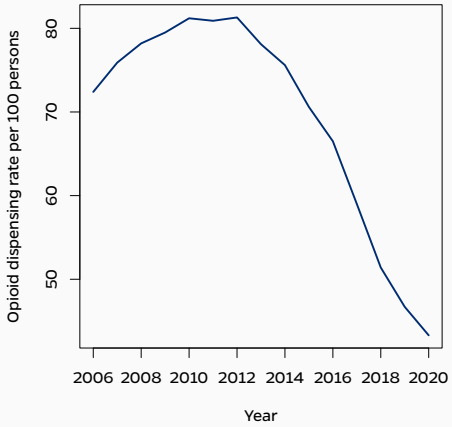


Seewald, Hackworth, and Almirall, (2021), *Principles and Practice of Clinical Trials*.

Seewald et al., (2020), *Statistical Methods in Medical Research*.

Motivating Example: Medical Cannabis Laws and Opioid Prescribing

- **4x** increase in opioid prescribing in U.S. from 1999-2012
  - Opioid prescribing for chronic non-cancer pain has played a meaningful role
- Getting better: prescribing down since 2012, but still ~3x higher than 1999



Dart et al., (2015), *New England Journal of Medicine*.

https://www.cdc.gov/drugoverdose/rxrate-maps/index.html

Do Medical Cannabis Laws Change Opioid Prescribing?

- Cannabis is a potentially effective treatment for chronic non-cancer pain, but evidence is limited.
- Patients with chronic non-cancer pain are eligible to use cannabis under all existing state medical cannabis laws
- Some evidence of substitution among adults with chronic non-cancer pain

Question: What are the effects of state medical cannabis laws on receipt of opioid and non-opioid treatment among patients with chronic non-cancer pain?

Bicket, Stone, and McGinty, (2023), *JAMA Network Open*.

Motivating Example: Medical Cannabis Laws and Opioid Prescribing

Previous studies have found mixed results, but have key methodological limitations:

1. No individual-level data
2. General population samples lead to policy endogeneity

Motivating Example: Medical Cannabis Laws and Opioid Prescribing

Previous studies have found mixed results, but have key methodological limitations:

1. No individual-level data
2. General population samples lead to policy endogeneity

Individual-level data lets us identify the population, but adds methodological complexity. My work addresses this.

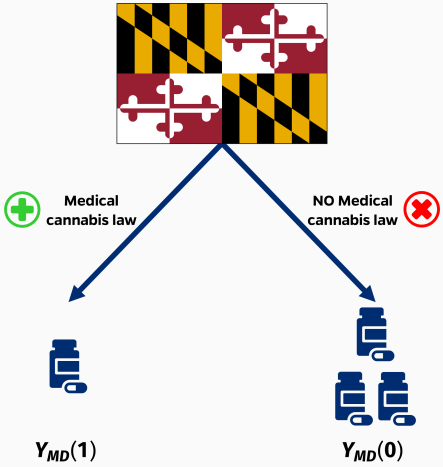
Causal Inference for Policy Evaluation is Hard

- Necessarily limited sample size
- Can't randomize
- Often high variability in "treatment" definitions
  - Lots of implementation science issues and opportunities for statistical work: measurement, identifying which components of a policy work, etc.
- Hard to isolate a particular policy's effects when other policies are in place.

Causal Inference for Policy Evaluation is Hard

- Necessarily limited sample size
- Can't randomize
- Often high variability in "treatment" definitions
  - Lots of implementation science issues and opportunities for statistical work: measurement, identifying which components of a policy work, etc.
- Hard to isolate a particular policy's effects when other policies are in place.
- Huge range of applications for these methods: health services research, education, mental health, substance use...

Formulating a Causal Question



Estimand: Average treatment effect among the treated (ATT)

$$\begin{aligned} \text{ATT} &= \text{E}\left[Y(1) - Y(0) \mid A = 1\right] \\ &= \text{E}\left[Y(1) \mid A = 1\right] - \text{E}\left[Y(0) \mid A = 1\right], \end{aligned}$$

where

- $Y(a)$ is the potential outcome that would be observed for a state had they implemented policy a
- A is "treatment" status (1 = med. cannabis law, 0 = no law)

Estimand: Average treatment effect among the treated (ATT)

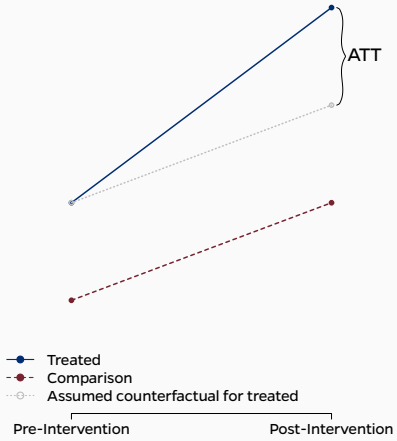
$$
\begin{aligned}
\text{ATT} &= \mathrm{E}\left[Y(1) - Y(0) \mid A = 1\right] \\
&= \mathrm{E}\left[Y(1) \mid A = 1\right] - \underbrace{\mathrm{E}\left[Y(0) \mid A = 1\right]}_{\text{not observable!}},
\end{aligned}
$$

where

- $Y(a)$ is the potential outcome that would be observed for a state had they implemented policy a
- A is "treatment" status (1 = med. cannabis law, 0 = no law)

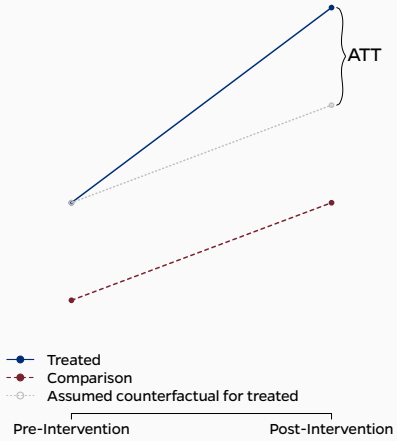
Difference-in-Differences: A Conceptual Introduction

- A commonly-used method for estimating the ATT in policy evaluation
- **Idea:** Compare change in outcome over time between treated and comparison groups
- **Key assumption:** In the absence of treatment, the outcome evolution in the treated group would have looked like the outcome evolution in the comparison group.
  - This is called the *counterfactual parallel trends assumption*



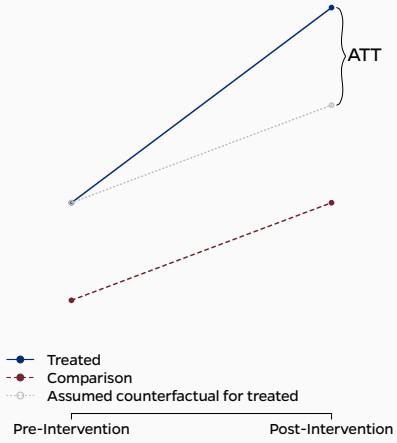
Counterfactual Parallel Trends: Pre/Post Setting

$$\mathrm{E}\left[Y_{\text{post}}(0) - Y_{\text{pre}}(0) \mid A = 1\right]$$
$$= \mathrm{E}\left[Y_{\text{post}}(0) - Y_{\text{pre}}(0) \mid A = 0\right]$$



Counterfactual Parallel Trends: Pre/Post Setting

$$\overbrace{\mathrm{E}\left[Y_{\text{post}}(0) - Y_{\text{pre}}(0) \mid A = 1\right]}^{\text{not observable}} = \underbrace{\mathrm{E}\left[Y_{\text{post}}(0) - Y_{\text{pre}}(0) \mid A = 0\right]}_{\text{observable}}$$



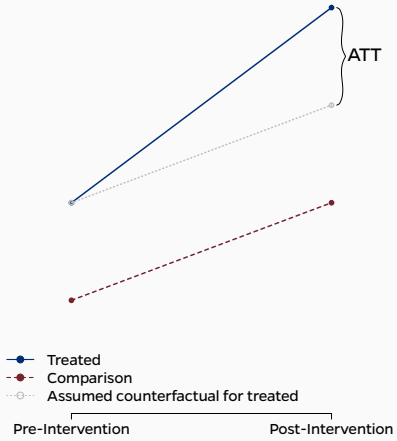
Counterfactual Parallel Trends: Pre/Post Setting

$$\overbrace{\mathrm{E}\left[Y_{\text{post}}(0) - Y_{\text{pre}}(0) \mid A = 1\right]}^{\text{not observable}} = \underbrace{\mathrm{E}\left[Y_{\text{post}}(0) - Y_{\text{pre}}(0) \mid A = 0\right]}_{\text{observable}}$$

Under this assumption,

$$\text{ATT} = \left(\mathrm{E}\left[Y_{\text{post}} \mid A = 1\right] - \mathrm{E}\left[Y_{\text{pre}} \mid A = 1\right]\right) - \left(\mathrm{E}\left[Y_{\text{post}} \mid A = 0\right] - \mathrm{E}\left[Y_{\text{pre}} \mid A = 0\right]\right),$$

Literally a difference in differences!



Difference-in-Differences with Multiple Time Periods

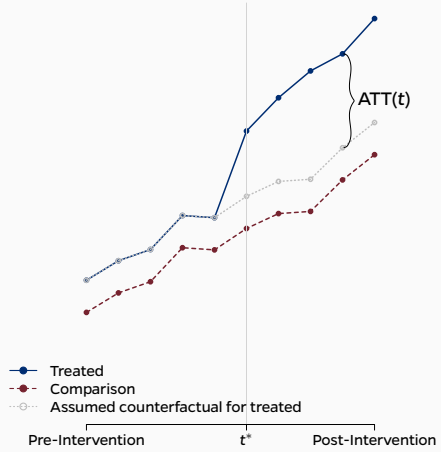
Now, times $t = \{1, \ldots, t^*, \ldots, T\}$; t^* first measurement after treatment.

Alternative estimands:

$$\text{ATT}(t) = \text{E}\left[Y_t(1) - Y_t(0) \mid A = 1\right], \quad t \geq t^*$$

$$\text{ATT}_{\text{avg}} = \text{E}\left[\bar{Y}_{\{t \geq t^*\}}(1) - \bar{Y}_{\{t \geq t^*\}}(0) \mid A = 1\right]$$

Strength of counterfactual parallel trends assumption varies with choice of estimand.



Two-Way Fixed Effects Estimation

A common "modeling" approach to estimate *ATT*:

$$Y_{sit} = \beta_{0,s} + \beta_{1,t} + \boldsymbol{\beta_2} A_{st} + \varepsilon_{sit},$$

where

- $A_{st} = \mathbb{1}\{\text{state } s \text{ treated at time } t\}$
- β_0's are *state fixed effects*
- β_1's are *time fixed effects*

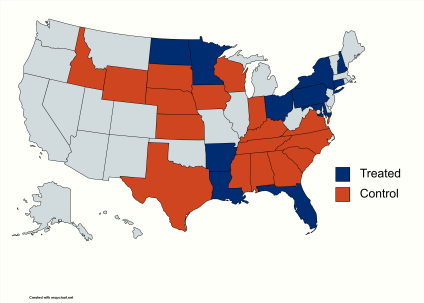
With 1 treated state or "simultaneous adoption",

$$\hat{\beta}_2 \equiv \left(\bar{Y}^{\text{tx}}_{\{t \geq t^*\}} - \bar{Y}^{\text{tx}}_{\{t < t^*\}}\right) - \left(\bar{Y}^{\text{ctrl}}_{\{t \geq t^*\}} - \bar{Y}^{\text{ctrl}}_{\{t < t^*\}}\right)$$

Motivating Example: Medical Cannabis Laws and Opioid Prescribing

Our sample:

- 12 *treated* states that implemented a medical cannabis law between 2012 and 2019 and *do not also have recreational cannabis laws*
- 17 *comparison* states without medical or recreational cannabis laws

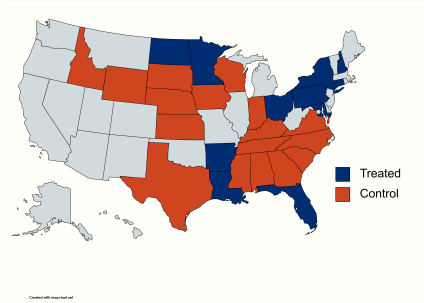


Motivating Example: Medical Cannabis Laws and Opioid Prescribing

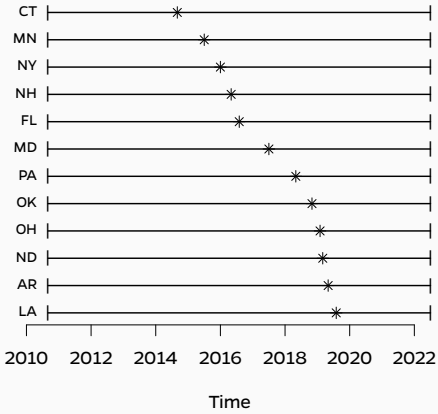
Our sample:

- 12 *treated* states that implemented a medical cannabis law between 2012 and 2019 and *do not also have recreational cannabis laws*
- 17 *comparison* states without medical or recreational cannabis laws

Goal: Estimate the effect of implementing a medical cannabis law on opioid prescribing outcomes, relative to what would have happened in the absence of treatment, among states that implemented such a law.



Medical Cannabis Study: Study Periods



States implemented medical cannabis laws at different times

Two-Way Fixed Effects under Staggered Adoption

$$Y_{sit} = \beta_{0,s} + \beta_{1,t} + \boldsymbol{\beta_2} A_{st} + \varepsilon_{sit}$$

- Not all states implemented medical cannabis policy at the same time.
- Two-way fixed effects can yield a (very) biased overall effect estimate in this setting.
  - Problematic under time-varying treatment effects
  - Estimator inadvertently adjusts for post-treatment information

---

Goodman-Bacon, (2021), *Journal of Econometrics*.

Difference-in-Differences is Booming

- Issues with two-way fixed effects have led to massive growth in difference-in-differences methods/estimators (see, e.g., Roth et al. 2021)
- Work is primarily done by econometricians
- Lots of questions remain:
  - Little to no head-to-head methods comparison
  - Different (sometimes unclear) data requirements
  - Subtleties arise in implementation
  - Lack of translational work

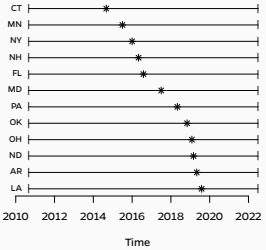
Callaway and Sant'Anna, (2021), *Journal of Econometrics*; Cengiz et al., (2019), *The Quarterly Journal of Economics*; de Chaisemartin and D'Haultfœuille, (2018), *The Review of Economic Studies*; Roth et al., (2022), *arXiv:2201.01194 [econ, stat]*.

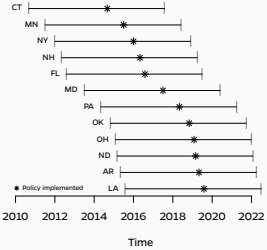
Difference-in-Differences is Booming

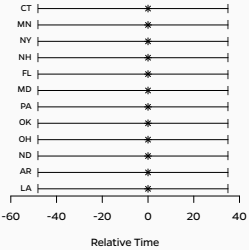
- Issues with two-way fixed effects have led to massive growth in difference-in-differences methods/estimators (see, e.g., Roth et al. 2021)
- Work is primarily done by econometricians
- Lots of questions remain:
  - Little to no head-to-head methods comparison
  - Different (sometimes unclear) data requirements
  - Subtleties arise in implementation
  - Lack of translational work
  - **This is what I do best!**

Callaway and Sant'Anna, (2021), *Journal of Econometrics*; Cengiz et al., (2019), *The Quarterly Journal of Economics*; de Chaisemartin and D'Haultfœuille, (2018), *The Review of Economic Studies*; Roth et al., (2022), *arXiv:2201.01194 [econ, stat]*.

Stacked Difference-in-Differences / Serial Trial Emulation







Start with full data

Anchor time

Estimate and aggregate

Hernán and Robins, (2016), *American Journal of Epidemiology*; Ben-Michael, Feller, and Stuart, (2021), *Epidemiology*.

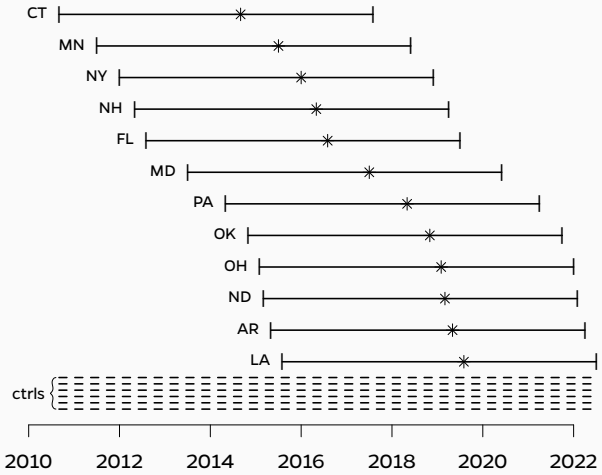
Medical Cannabis Study: State Cohorts

Data are individual-level commercial health insurance claims from $N = 583,820$ unique individuals in 29 states.

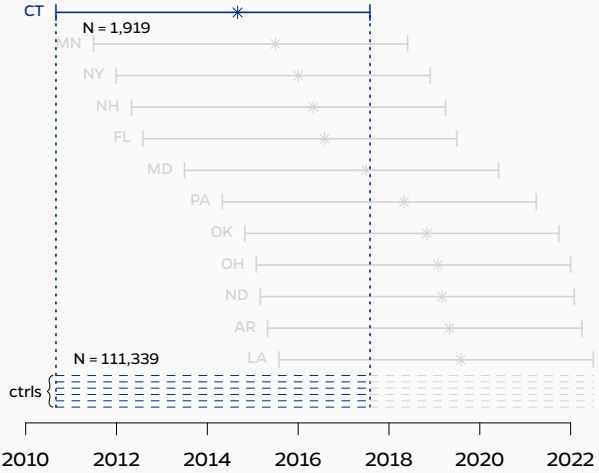
For each treatment state, we build a *cohort* of individuals in that state and the control states over the study period.

- Individuals included if they have a chronic non-cancer pain diagnosis in the pre-law period **and** are continuously enrolled in commercial health insurance for the full study period.

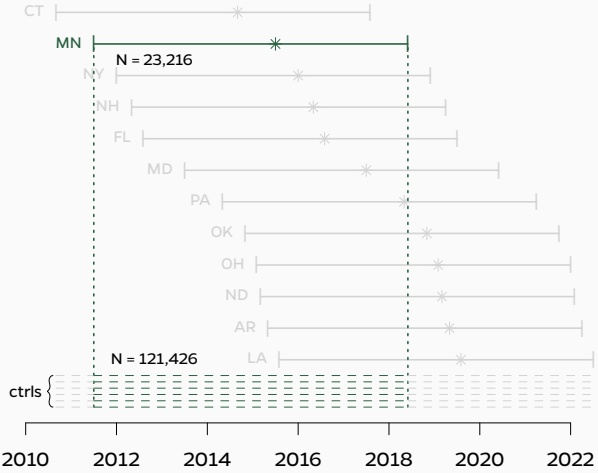
Medical Cannabis Study: State Cohorts



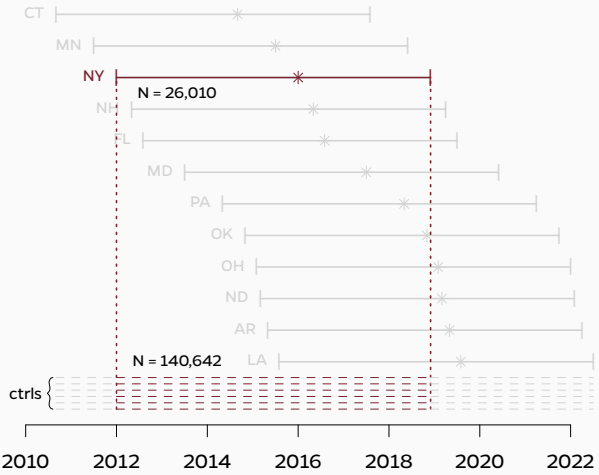
Medical Cannabis Study: State Cohorts



Medical Cannabis Study: State Cohorts



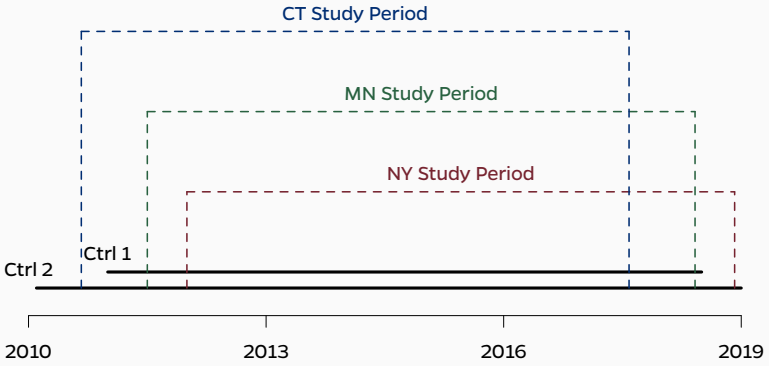
Medical Cannabis Study: State Cohorts



Shared Control Individuals

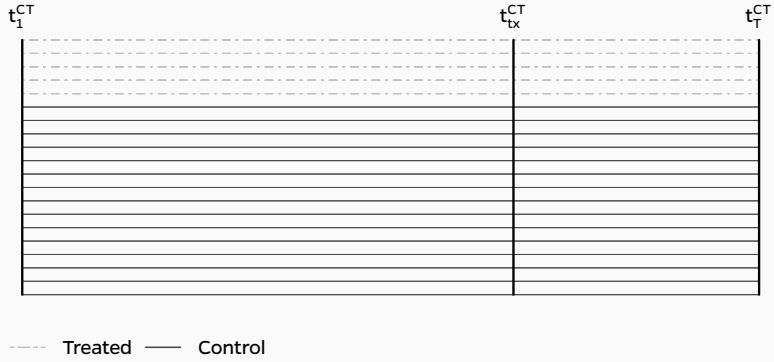
- Individuals in control states might appear in multiple cohorts.
  - "Ctrl 1" is in MN cohort only but "Ctrl 2" is in CT, MN, and NY cohorts

This induces correlation between treatment effect estimates for different cohorts!



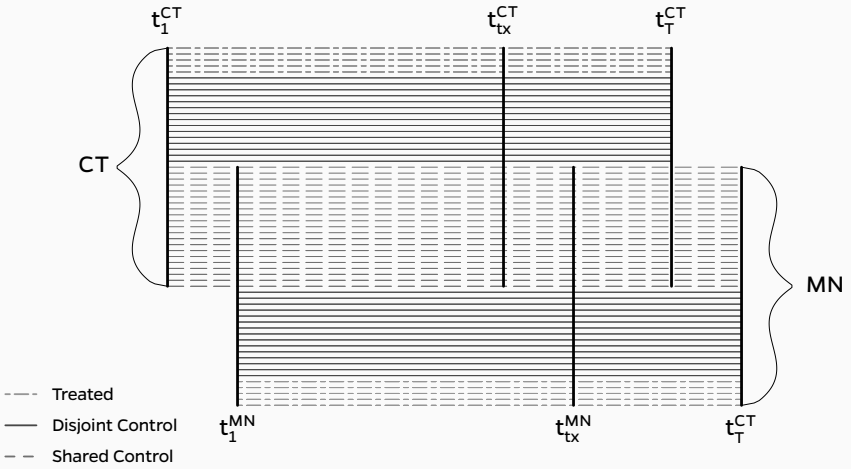
Shared Control Individuals

One cohort:



Shared Control Individuals

Two cohorts:



Handling Correlation Induced by Shared Control Indviduals

Goal: Improved inference on overall ATT averaged across treated units.

- ATT estimates remain unbiased under usual assumptions
- Failure to account for shared control individuals can lead to *incorrect inference*

Big Idea: Incorporate pairwise correlation between estimates into inverse-variance weighted average

Covariance between Diff-in-Diff Effect Estimates

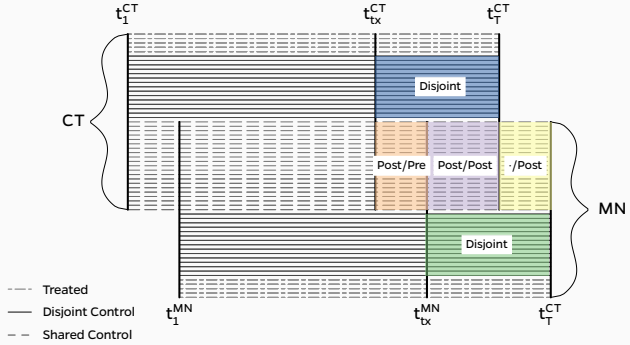
With only one treated unit, we could estimate ATT for cohort C as

$$\widehat{\text{ATT}}_C = \left(\bar{Y}^{\text{tx}}_{s,\text{post}} - \bar{Y}^{\text{tx}}_{s,\text{pre}}\right) - \left(\bar{Y}^{\text{ctrl}}_{s,\text{post}} - \bar{Y}^{\text{ctrl}}_{s,\text{pre}}\right)$$

Assuming states are independent,

$$\begin{aligned}\text{Cov}\left(\widehat{\text{ATT}}_{C_1}, \widehat{\text{ATT}}_{C_2}\right) = \text{Cov}\left(\bar{Y}^{\text{ctrl}}_{C_1,\text{post}}, \bar{Y}^{\text{ctrl}}_{C_2,\text{post}}\right) + \text{Cov}\left(\bar{Y}^{\text{ctrl}}_{C_1,\text{pre}}, \bar{Y}^{\text{ctrl}}_{C_2,\text{pre}}\right) \\ - \text{Cov}\left(\bar{Y}^{\text{ctrl}}_{C_1,\text{post}}, \bar{Y}^{\text{ctrl}}_{C_2,\text{pre}}\right) - \text{Cov}\left(\bar{Y}^{\text{ctrl}}_{C_1,\text{pre}}, \bar{Y}^{\text{ctrl}}_{C_2,\text{post}}\right)\end{aligned}$$

Covariances with Shared Control Individuals

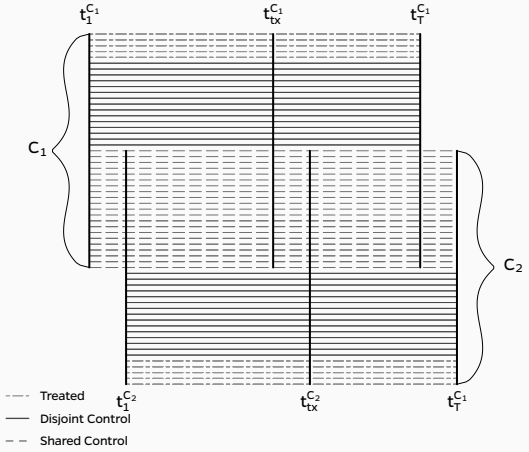


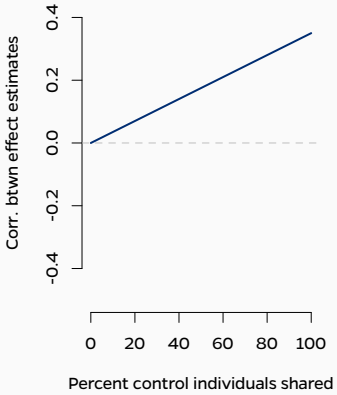
$$\text{Cov}\left(\bar{Y}^{\text{ctrl}}_{\text{CT,post}}, \bar{Y}^{\text{ctrl}}_{\text{MN,post}}\right) \text{“}=\text{”} \text{Cov}\left(\bar{Y}_{\text{CT Disjoint}} + \bar{Y}_{\text{Post/Pre}} + \bar{Y}_{\text{Post/Post}}, \bar{Y}_{\text{MN Disjoint}} + \bar{Y}_{\text{Post/Post}} + \bar{Y}_{\cdot\text{/Post}}\right)$$

When Does This Matter?

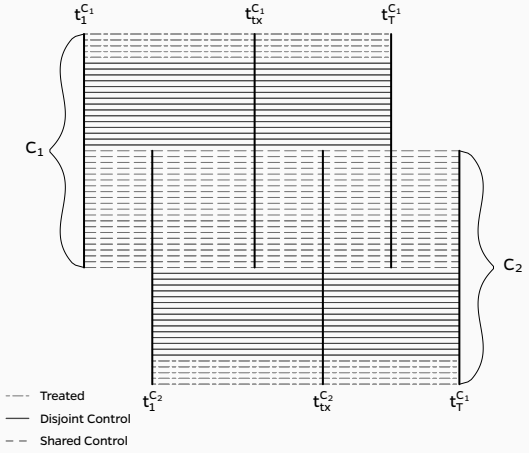
- Setting / simplifying assumptions:
  - Exchangeable within-person correlation
  - Interest is in ATT_{avg}
  - Individuals are independent of people who live in other states
- Correlation between effect estimates depends on:
  - duration of pre- and post-treatment periods
  - delay between study period start times
  - proportion of shared control individuals
  - within- and between-person correlations

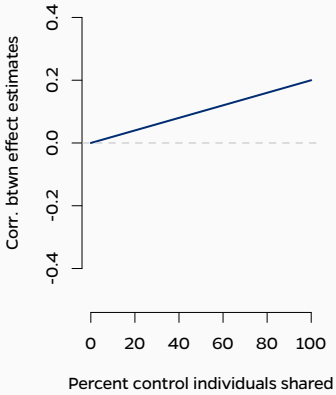
Correlation Due to Shared Controls



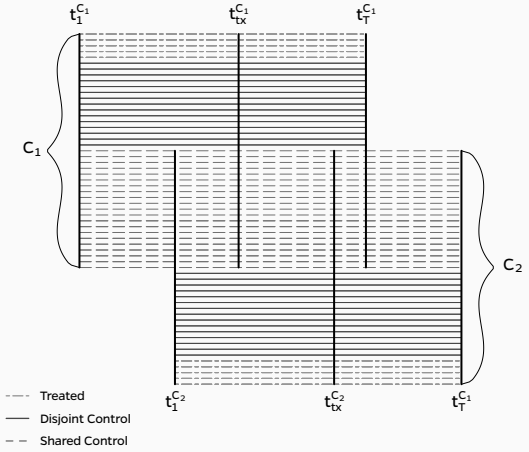


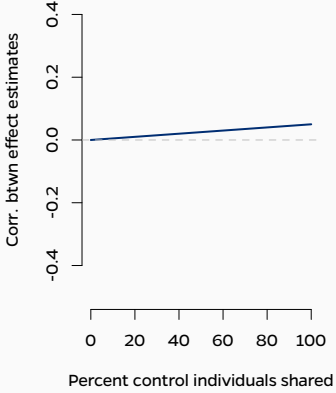
Correlation Due to Shared Controls



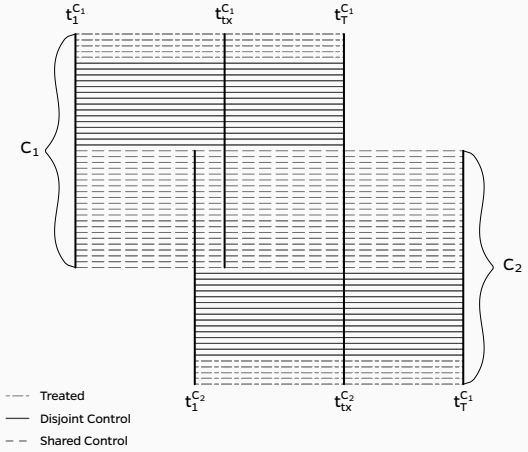


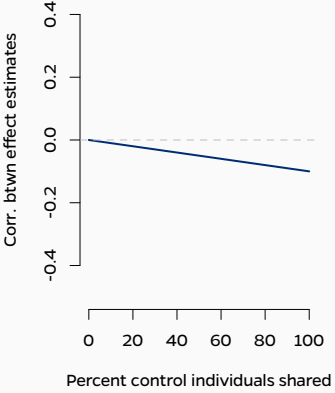
Correlation Due to Shared Controls



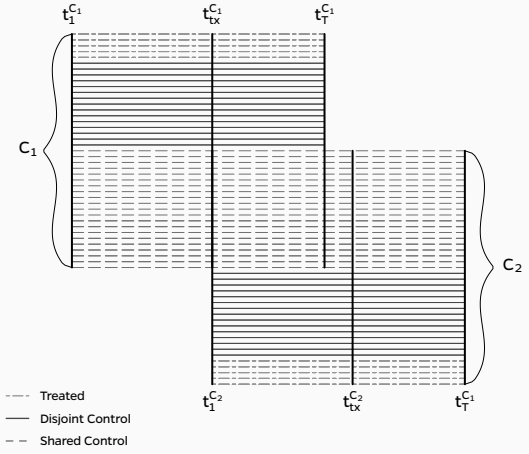


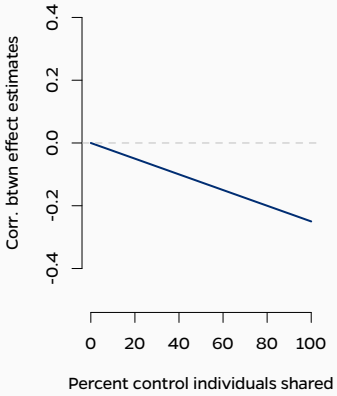
Correlation Due to Shared Controls



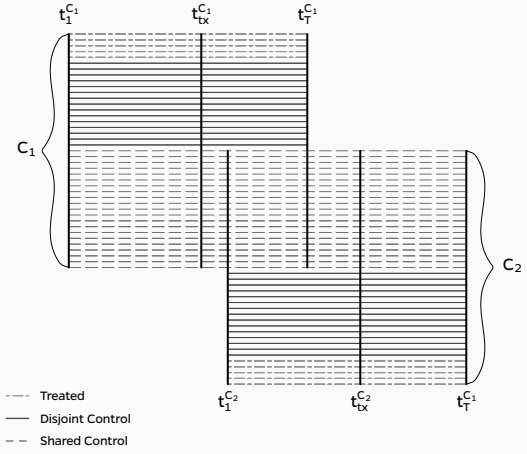


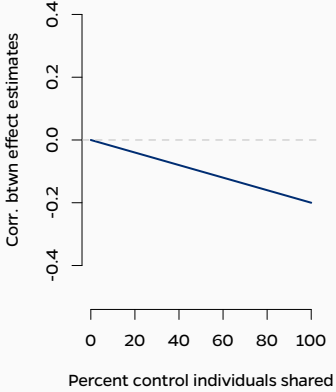
Correlation Due to Shared Controls



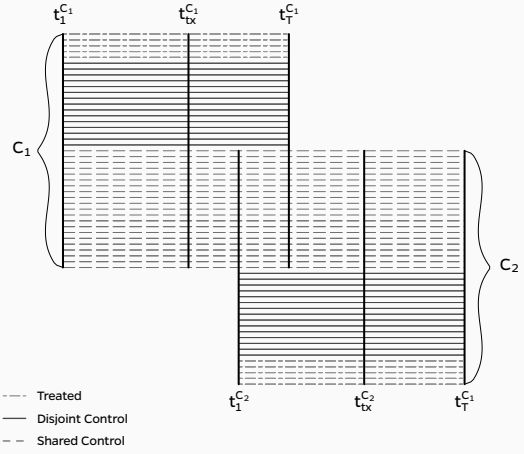


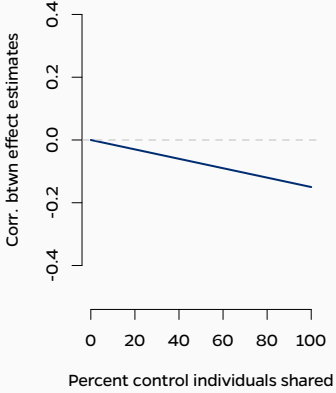
Correlation Due to Shared Controls



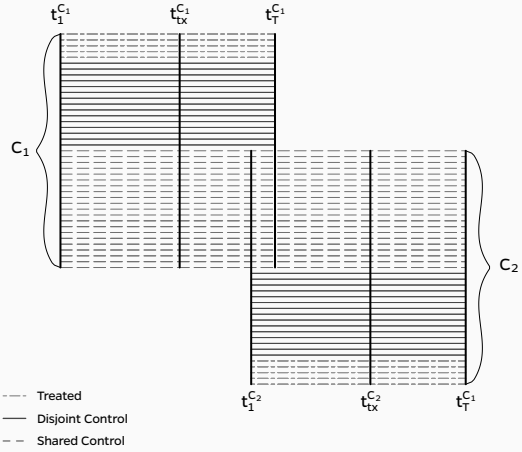


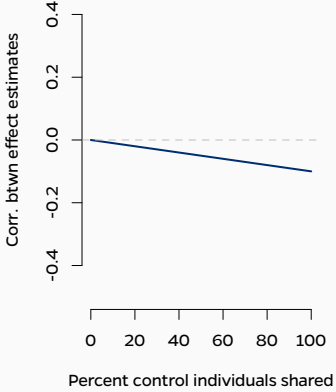
Correlation Due to Shared Controls



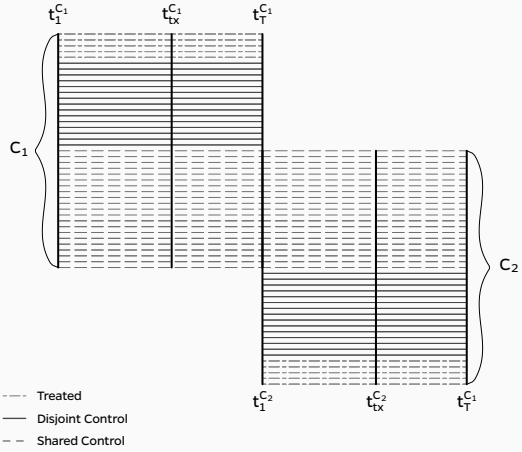


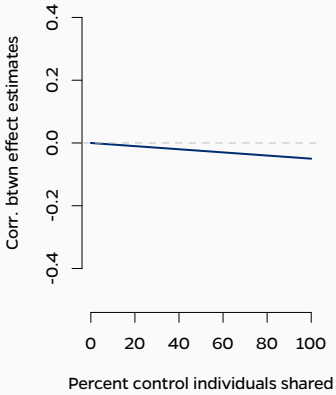
Correlation Due to Shared Controls



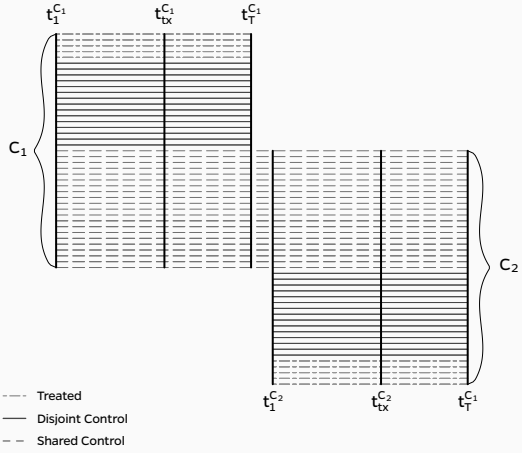


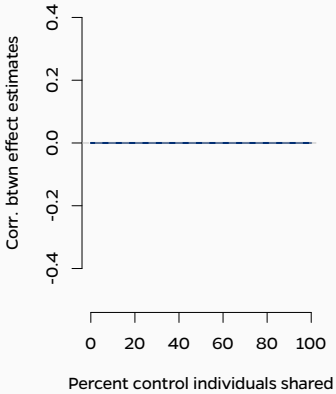
Correlation Due to Shared Controls





Correlation Due to Shared Controls





Inverse Variance Weighted Averaging

Estimating correlations (covariances) lets us construct a covariance matrix Σ for all state-specific ATTs.

Then,

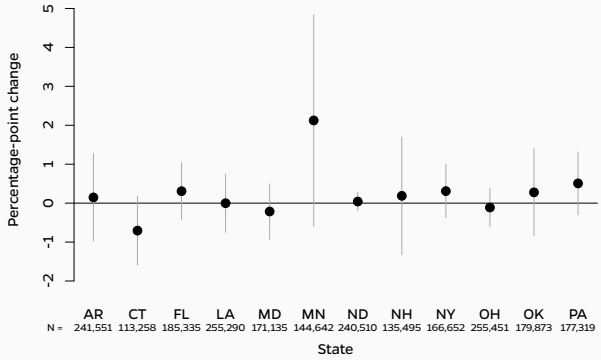
$$\widehat{\text{ATT}}_{\text{overall}} = \frac{1}{\sum_s (1/\sigma_s^2)} \sum_s \widehat{\text{ATT}}_s / \sigma_s^2$$

and

$$\text{Var}\left(\widehat{\text{ATT}}_{\text{overall}}\right) = \frac{1}{\left(\boldsymbol{v}^\top \boldsymbol{v}\right)^2} \boldsymbol{v}^\top \boldsymbol{\Sigma} \boldsymbol{v},$$

where $\boldsymbol{v}^\top = \left(1/\sigma_1, \ldots, 1/\sigma_S\right)$

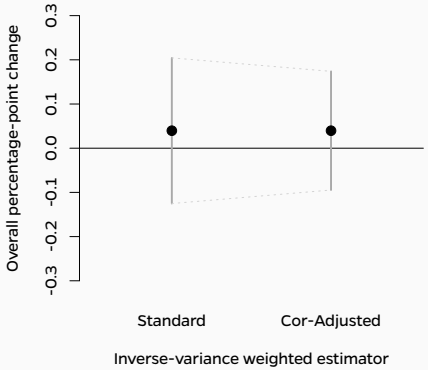
Medical Cannabis Study: Results



Change in proportion of chronic noncancer pain patients receiving *any opioid prescription*, per month, attributable to state medical cannabis law in first 3 years of implementation

Medical Cannabis Study: Results

- In this case, accounting for between-estimate correlation gives *smaller* SE (here, by 18.5%)
- State-level policy evaluations are (often) notoriously underpowered – this could be a step in the right direction!



Schell, Griffin, and Morral, (2018).

Conclusions

- Individual-level data is useful for identifying populations of interest in policy evaluation, but introduces methodological complexity.
- When using individual-level data that might be shared across cohorts in stacked diff-in-diff, it may be important to account for correlation between estimates
- A closed-form formula for induced correlation is available for select analyses

My Future Work in Health Policy Evaluation

1. Are difference-in-differences analyses using individual-level longitudinal data more efficient than those using aggregated data?
    - Strangely, probably not. *Maybe* in complicated confounding scenarios.
    - Wide-ranging simulation project (novel in policy context)

My Future Work in Health Policy Evaluation

1. Are difference-in-differences analyses using individual-level longitudinal data more efficient than those using aggregated data?
    - Strangely, probably not. *Maybe* in complicated confounding scenarios.
    - Wide-ranging simulation project (novel in policy context)
2. Which pieces of a policy's implementation work in what contexts to produce an effect?
    - What combinations or sequences of implementation strategies yield best outcomes?
    - Blend methods from policy evaluation and dynamic treatment regimes literatures
    - Proposed as part of "methods core" for NIMH P50 renewal (sub. May 2023)

About Me

I am a collaborative statistician who builds statistical tools that enable
high quality, impactful science.

Acknowledgements

Mentors

- Daniel Almirall, PhD
- Kelley Kidwell, PhD
- Beth McGinty, PhD*
- Susan A. Murphy, PhD
- Elizabeth A. Stuart, PhD *

Collaborators

- Gail L. Daumit, MD, MHS
- Pedja Klasnja, PhD
- Laura Samuel, PhD
- Ian Schmid, ScM *
- Shawna N. Smith, PhD
- Elizabeth Stone, MSPH
- Kayla Tormohlen, PhD *

NIDA R01DA049789 (PI: McGinty)

* Co-author on presented work

Extra slides

Why is the correlation negative?

Remember the covariance between two estimated ATTs:

$$\text{Cov}\left(\widehat{\text{ATT}}_{C_1}, \widehat{\text{ATT}}_{C_2}\right) = \text{Cov}\left(\bar{Y}^{\text{ctrl}}_{C_1,\text{post}}, \bar{Y}^{\text{ctrl}}_{C_2,\text{post}}\right) + \text{Cov}\left(\bar{Y}^{\text{ctrl}}_{C_1,\text{pre}}, \bar{Y}^{\text{ctrl}}_{C_2,\text{pre}}\right) - \text{Cov}\left(\bar{Y}^{\text{ctrl}}_{C_1,\text{post}}, \bar{Y}^{\text{ctrl}}_{C_2,\text{pre}}\right) - \text{Cov}\left(\bar{Y}^{\text{ctrl}}_{C_1,\text{pre}}, \bar{Y}^{\text{ctrl}}_{C_2,\text{post}}\right)$$

In this setting, one cohort's post period is the other's pre: