Handling Correlation in Stacked Difference-in-Differences Estimates with Application to Medical Cannabis Policy

Nicholas J. Seewald, PhD

Department of Biostatistics, Epidemiology, and Informatics
University of Pennsylvania Perelman School of Medicine

17 December 2023
CMStatistics 2023
Joint with E.E. McGinty, K.N. Tormoholen, I. Schmid, E.A. Stuart

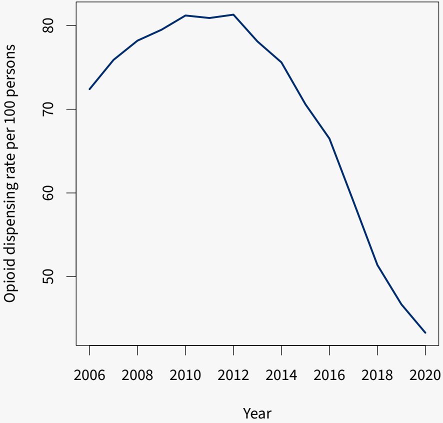
Perelman
SCHOOL OF MEDICINE
UNIVERSITY of PENNSYLVANIA

Slides are online!

slides.nickseewald.com/cmstatistics2023.pdf

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, R. C. et al. (2015). *New England Journal of Medicine*.
https://www.cdc.gov/drugoverdose/rxrate-maps/index.html

Federalism in the United States

"States are the laboratories of democracy."
(Louis Brandeis, *New State Ice Co. vs. Liebmann*)

States in the U.S. have wide latitude to implement or not implement policies and those policies can vary widely. States generally have jurisdiction over things that stay within state lines.

State laws permitting cannabis use are a great example of this.

Do Medical Cannabis Laws Change Opioid Prescribing?

- Cannabis industry & advocates argue medical cannabis for chronic pain could be a partial solution to opioid crisis via substitution
- 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 treatment among patients with chronic non-cancer pain?

Bicket, M. C., Stone, E. M., and McGinty, E. E. (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:

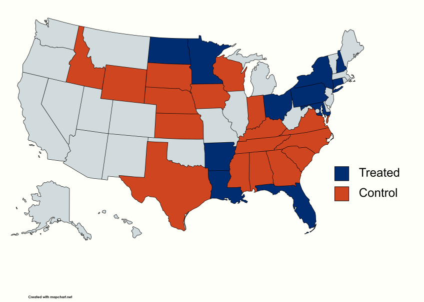
1. No individual-level data
2. General population samples lead to policy endogeneity

Individual-level data lets us identify the population, but adds methodological complexity.

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

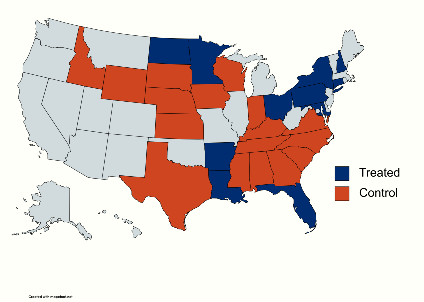


Motivating Example: Medical Cannabis Laws and Opioid Prescribing

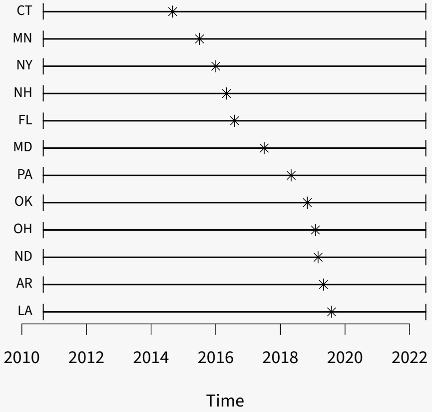
Our sample:

- 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 (an ATT).



Staggered Adoption of Medical Cannabis Laws



States implemented medical cannabis laws at different times

Difference-in-Differences with Multiple Time Periods

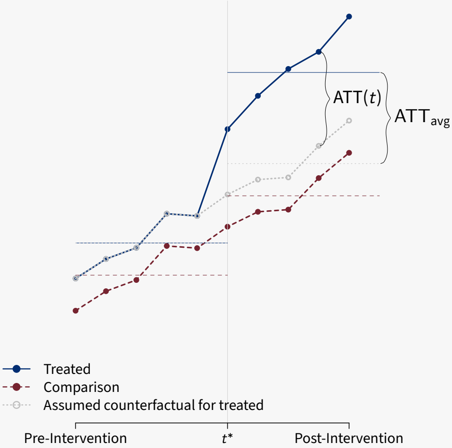
Now, times $t = \{1, \dots, t^*, \dots, T\}$; t^* first measurement after treatment.

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 assumpfbmtion varies with choice of estimand.



Two-Way Fixed Effects Estimation

A common "modeling" approach to estimate ATT:

$$Y_{sit} = \underbrace{\beta_{0,s}}_{\text{state fixed effects}} + \underbrace{\beta_{1,t}}_{\text{time fixed effects}} + \underbrace{\beta_2 A_{st}}_{\text{treatment}} + \varepsilon_{sit},$$

With 1 treated state or "simultaneous adoption",

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

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

Two-Way Fixed Effects Estimation

A common "modeling" approach to estimate ATT:

$$Y_{sit} = \underbrace{\beta_{0,s}}_{\text{state fixed effects}} + \underbrace{\beta_{1,t}}_{\text{time fixed effects}} + \underbrace{\beta_2 A_{st}}_{\text{treatment}} + \varepsilon_{sit},$$

With 1 treated state or "simultaneous adoption",

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

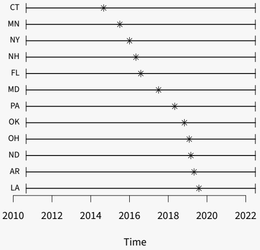
- Two-way fixed effects can yield a (very) biased overall effect estimate under staggered adoption if there's a time-varying treatment effect.
  - Estimator inadvertently adjusts for post-treatment information

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

Two-Way Fixed Effects under Staggered Adoption

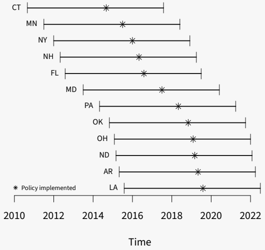
$$Y_{sit} = \beta_{0,s} + \beta_{1,t} + \beta_{\mathbf{2}} A_{st} + \varepsilon_{sit}$$

Stacked Difference-in-Differences / Serial Trial Emulation



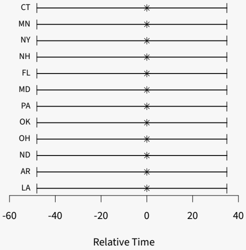
Start with full data

→



Anchor time

→



Estimate and aggregate

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

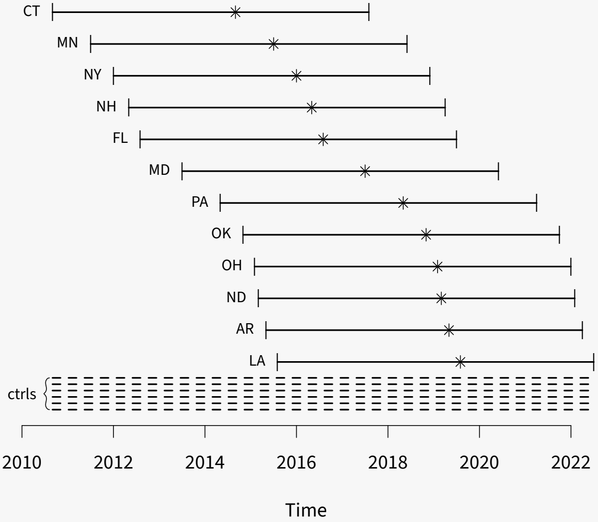
Medical Cannabis Study: State Cohorts

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

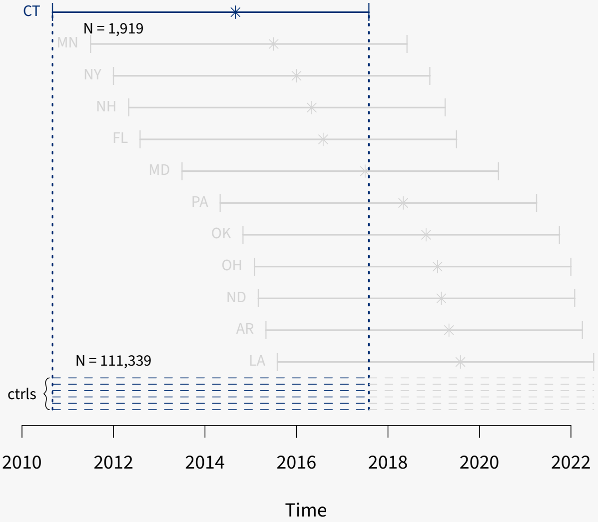
For each treatment state, we build a *cohort* of individuals in that state and the control states over the study period.

- 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.

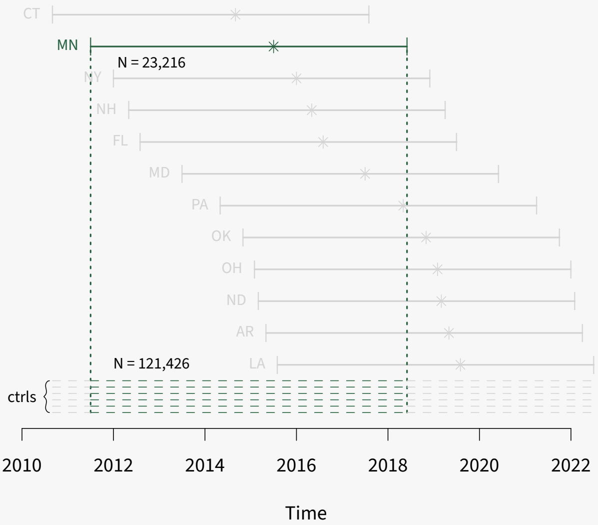
Medical Cannabis Study: State Cohorts



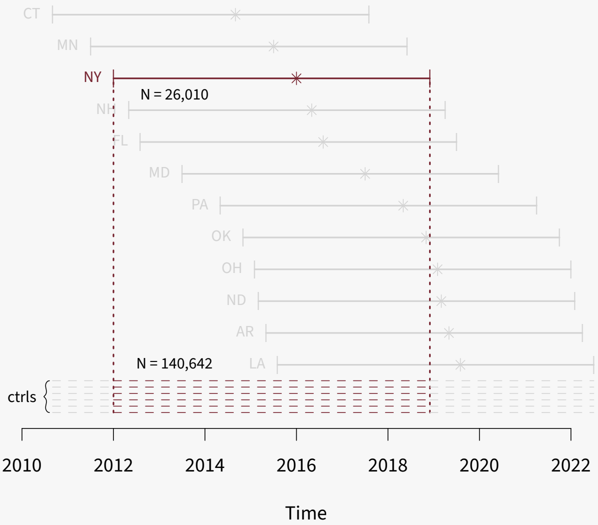
Medical Cannabis Study: State Cohorts



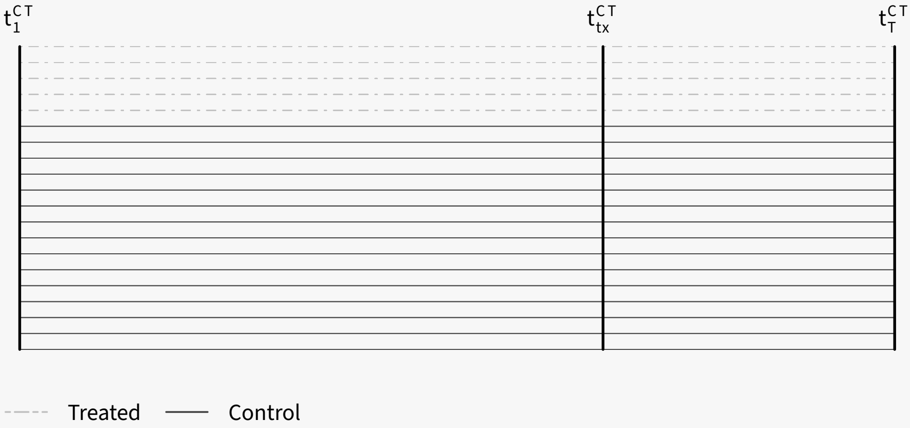
Medical Cannabis Study: State Cohorts



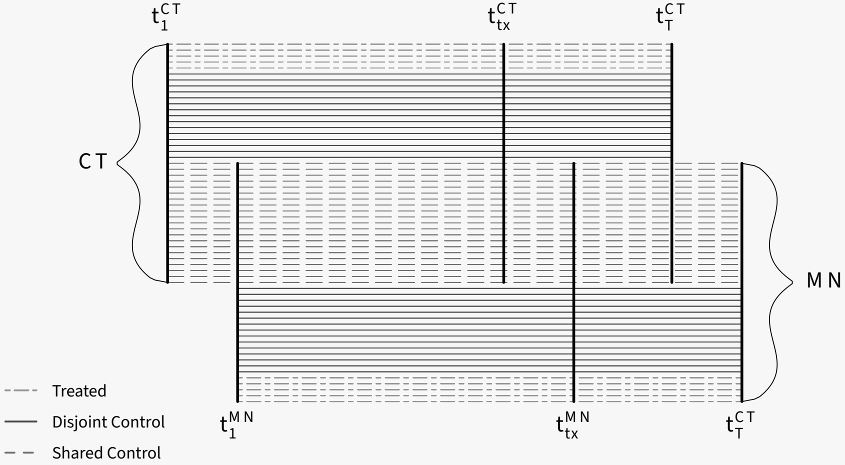
Medical Cannabis Study: State Cohorts



Cohort Schematic



Shared Control Individuals



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 a generalized least squares-esque weighting procedure

Covariance between Diff-in-Diff Effect Estimates

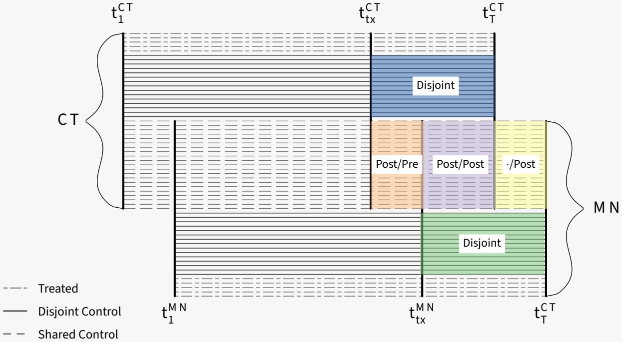
With only one treated unit, we could estimate ATT for cohort C as

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

Assuming states are independent,

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

Covariances with Shared Control Individuals



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

Simplifying Assumptions

- Same pre- and post-treatment durations for all treated states
- Pairwise independence of states
- Block-exchangeable correlation structure in outcomes within each state:

$$\Sigma_\gamma := \mathrm{Var}\left(\boldsymbol{Y}_\gamma\right) = \begin{pmatrix} \begin{matrix} 1 & \rho_\gamma & \cdots & \rho_\gamma \\ \rho_\gamma & 1 & \cdots & \rho_\gamma \\ \vdots & \vdots & \ddots & \vdots \\ \rho_\gamma & \rho_\gamma & \cdots & 1 \end{matrix} & \cdots & \begin{matrix} \phi_\gamma & \psi_\gamma & \cdots & \psi_\gamma \\ \psi_\gamma & \phi_\gamma & \cdots & \psi_\gamma \\ \vdots & \vdots & \ddots & \vdots \\ \psi_\gamma & \psi_\gamma & \cdots & \phi_\gamma \end{matrix} \\ \vdots & \ddots & \vdots \\ \begin{matrix} \phi_\gamma & \psi_\gamma & \cdots & \psi_\gamma \\ \psi_\gamma & \phi_\gamma & \cdots & \psi_\gamma \\ \vdots & \vdots & \ddots & \vdots \\ \psi_\gamma & \psi_\gamma & \cdots & \phi_\gamma \end{matrix} & \cdots & \begin{matrix} 1 & \rho_\gamma & \cdots & \rho_\gamma \\ \rho_\gamma & 1 & \cdots & \rho_\gamma \\ \vdots & \vdots & \ddots & \vdots \\ \rho_\gamma & \rho_\gamma & \cdots & 1 \end{matrix} \end{pmatrix} \sigma_\gamma^2$$

Between-Estimate Covariance in Stacked Diff-in-Diff

Here's some math, to prove I can do it:

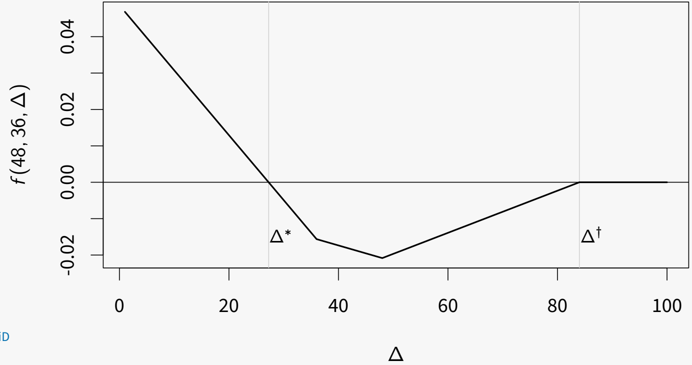
$$\mathrm{Cov}\left(\widehat{\mathrm{ATT}}_\gamma, \widehat{\mathrm{ATT}}_\nu\right) = \frac{f\left(T_{\mathrm{pre}}, T_{\mathrm{post}}, \Delta\right)}{N_\gamma^{\mathrm{ctrl}} N_\nu^{\mathrm{ctrl}}} \sum_{\zeta \in \text{ctrl states}} \sigma_\zeta^2 \Bigg[\underbrace{N_\gamma(\zeta) N_\nu(\zeta)}_{\#\text{ctrls per state } \zeta} \underbrace{\left(\phi_\zeta - \psi_\zeta\right)}_{\text{diff. in btwn-person corrs}}$$

$$+ \underbrace{N_{\gamma \cap \nu}(\zeta)}_{\#\text{shared ctrls}} \left(1 - \rho_\zeta - \left(\phi_\zeta - \psi_\zeta\right)\right) \Bigg],$$

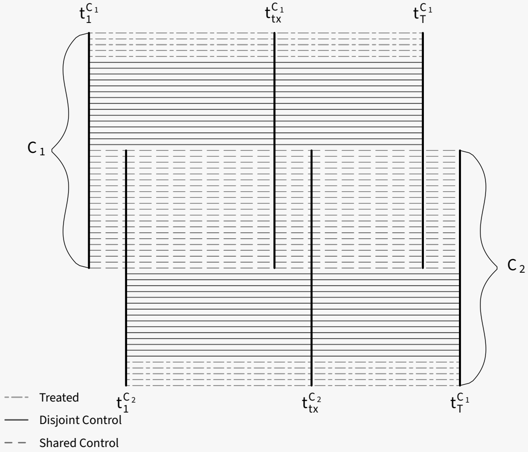
Summand is strictly positive under (quite weak) assumption that $\phi_\zeta > \phi_\zeta > \psi_\zeta$.

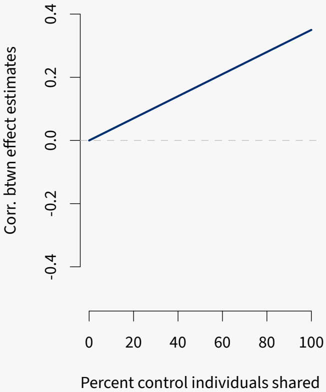
Sign of Between-Estimate Covariance Depends on Δ

$$f\left(T_{\text{pre}}, T_{\text{post}}, \Delta\right) = \frac{1}{T_{\text{pre}}^2 T_{\text{post}}^2} \cdot \Big[T_{\text{pre}}^2 \max\left(T_{\text{post}} - \Delta, 0\right) + T_{\text{post}}^2 \max\left(T_{\text{pre}} - \Delta, 0\right)$$
$$- T_{\text{pre}} T_{\text{post}} \min\Big(T_{\text{pre}}, T_{\text{post}}, \Delta, \max\left(T_{\text{pre}} + T_{\text{post}} - \Delta, 0\right)\Big)\Big].$$

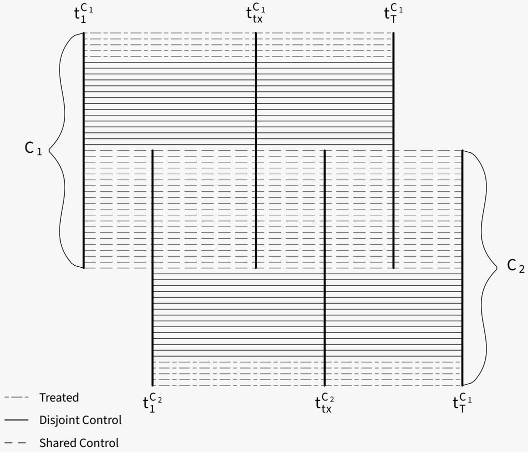


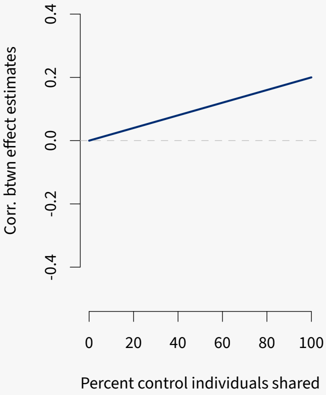
Correlation Due to Shared Controls



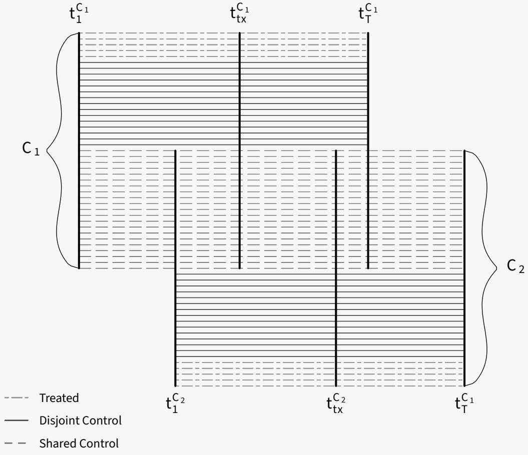


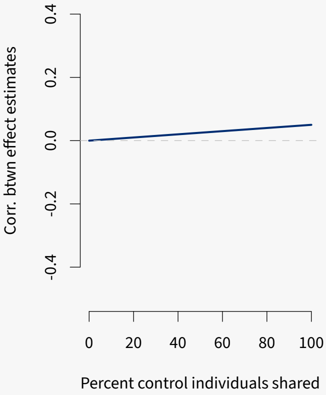
Correlation Due to Shared Controls



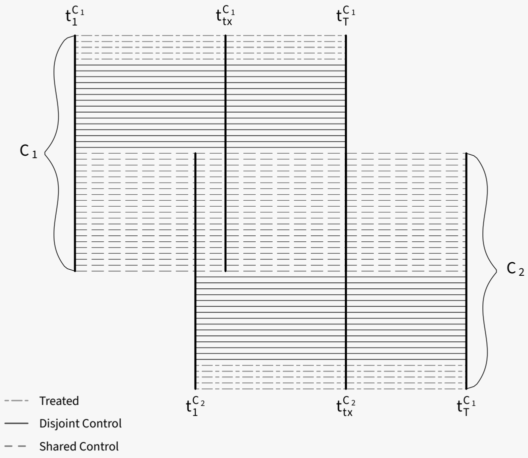


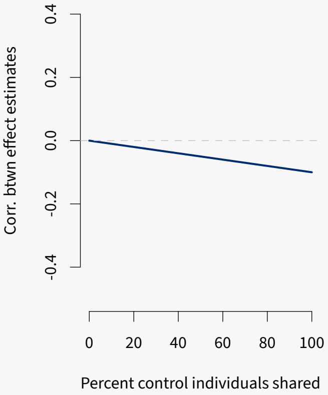
Correlation Due to Shared Controls



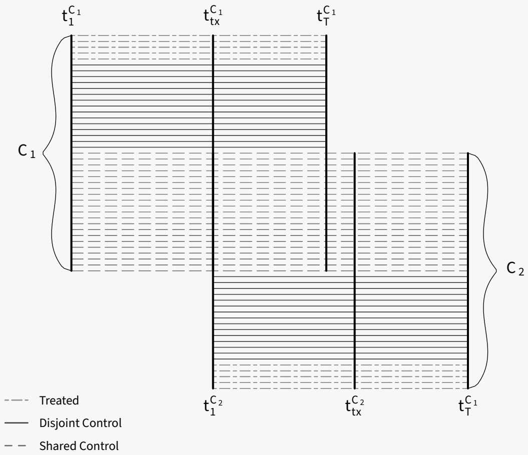


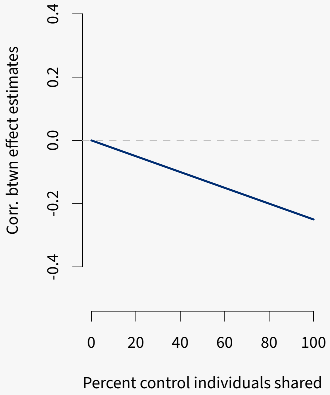
Correlation Due to Shared Controls



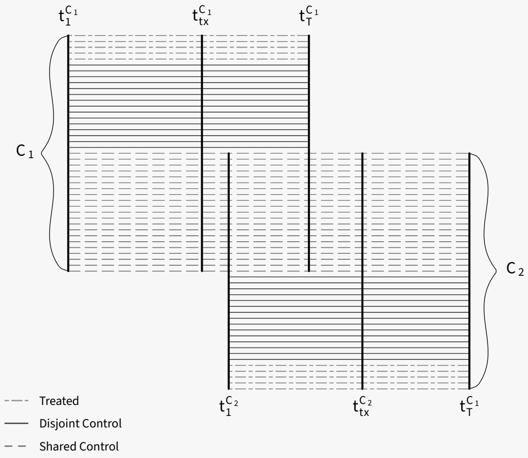


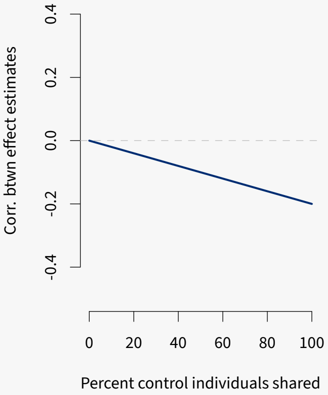
Correlation Due to Shared Controls



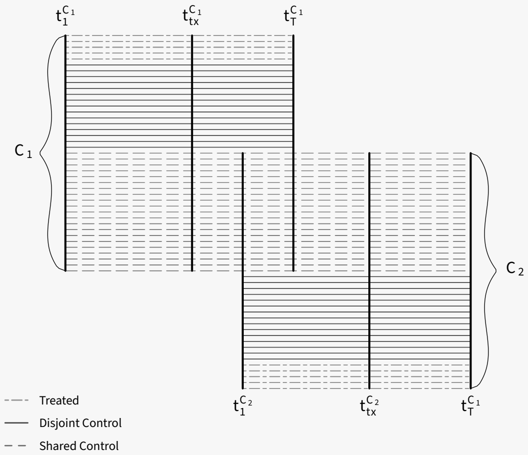


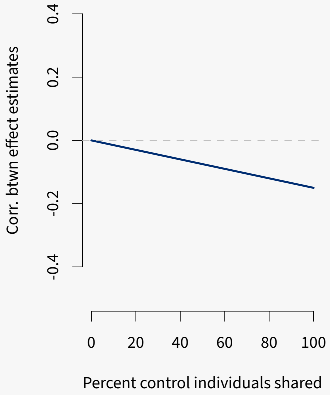
Correlation Due to Shared Controls



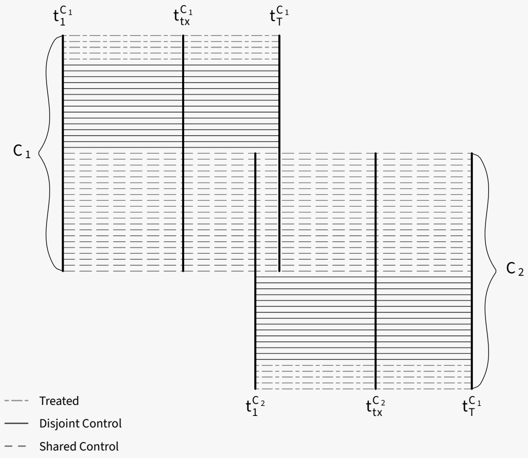


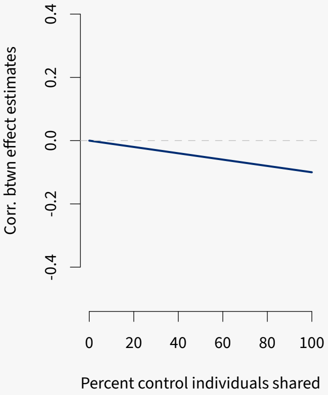
Correlation Due to Shared Controls



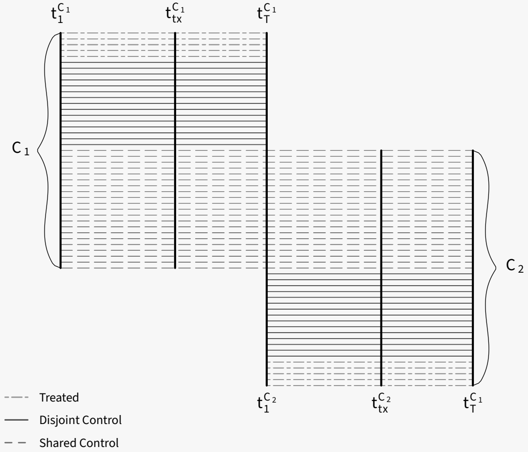


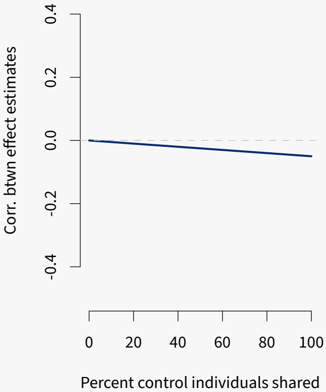
Correlation Due to Shared Controls



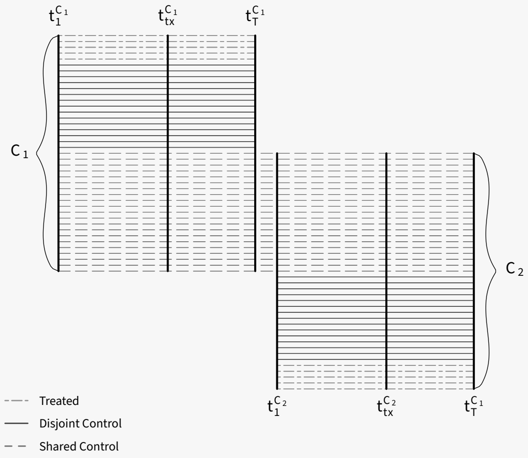


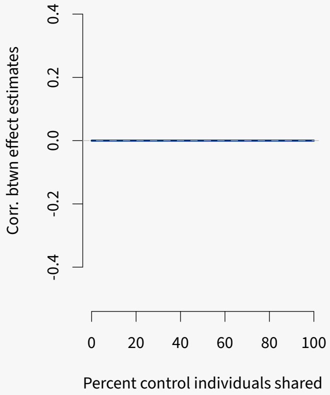
Correlation Due to Shared Controls





Correlation Due to Shared Controls





Aggregation: Inverse-Variance Weighting

If estimates are uncorrelated, could use inverse-variance weighted averaging to aggregate. Let $\boldsymbol{V}$ be a diagonal matrix with entries variances of the $\widehat{\text{ATT}}$s. Then,

$$\widehat{\text{ATT}}_{\text{ivw}} := \left(\mathbf{1}^\top \boldsymbol{V}^{-1}\mathbf{1}\right)^{-1} \mathbf{1}\boldsymbol{V}^{-1}\widehat{\mathbf{ATT}}_{\text{tx}} = \frac{1}{\sum_{s\in\text{tx states}} v_{ss}^{-1}} \sum_{s\in\text{tx states}} v_{ss}^{-1}\widehat{\text{ATT}}_s,$$

This has variance

$$\text{Var}\left(\widehat{\text{ATT}}_{\text{ivw}}\right) = \left(\mathbf{1}^\top \boldsymbol{V}^{-1}\mathbf{1}\right)^{-1} = \frac{1}{\sum_{s\in\text{tx states}} 1/v_{ss}}.$$

This does **not** account for between-estimate correlation!

Aggregation: GLS-Based Strategy

Now consider $\boldsymbol{W} = \text{Cov}\left(\textbf{ATT}\right)$.

Then,

$$\widehat{\text{ATT}}_{\text{gls}} = \left(\mathbf{1}^\top \boldsymbol{W}^{-1}\mathbf{1}\right)^{-1} \mathbf{1}\boldsymbol{W}^{-1}\widehat{\textbf{ATT}}_{\text{tx}}$$

and

$$\text{Var}\left(\widehat{\text{ATT}}_{\text{gls}}\right) = \left(\mathbf{1}^\top \boldsymbol{W}^{-1}\mathbf{1}\right)^{-1}.$$

Lin, D.-Y. and Sullivan, P. F. (2009). *Am J Hum Genet*.

Aggregation: GLS-Based Strategy

Now consider $\boldsymbol{W} = \text{Cov}(\textbf{ATT})$.

Then,

$$\widehat{\text{ATT}}_{\text{gls}} = \left(\mathbf{1}^\top \boldsymbol{W}^{-1}\mathbf{1}\right)^{-1} \mathbf{1}\boldsymbol{W}^{-1}\widehat{\textbf{ATT}}_{\text{tx}}$$

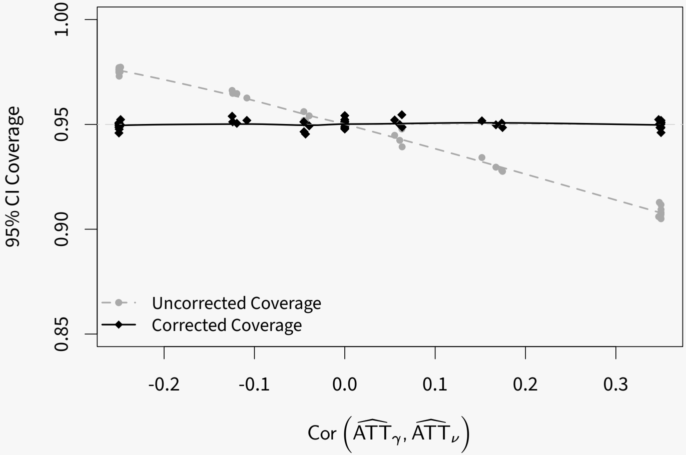
and

$$\text{Var}\left(\widehat{\text{ATT}}_{\text{gls}}\right) = \left(\mathbf{1}^\top \boldsymbol{W}^{-1}\mathbf{1}\right)^{-1}.$$

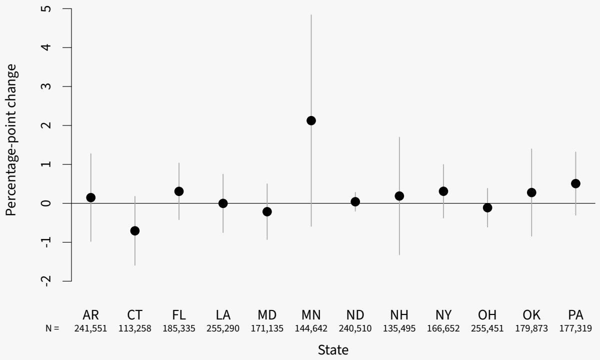
For 2 treated states, $\text{Var}\left(\widehat{\text{ATT}}_{\text{gls}}\right) > \text{Var}\left(\widehat{\text{ATT}}_{\text{ivw}}\right)$, *unless* between-estimate correlation is positive and sufficiently small.

Lin, D.-Y. and Sullivan, P. F. (2009). *Am J Hum Genet*.

Correlation Correction Yields Nominal Coverage for $\widehat{ATT}_{\text{avg}}$



Medical Cannabis Laws Study: Results



State-specific effects of medical cannabis laws on proportion of chronic noncancer pain patients receiving *any opioid prescription*, on average in a given month in first 3 years of law implementation

Between-Estimate Correlation

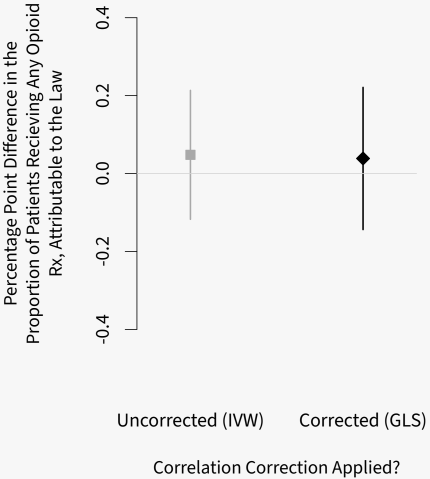
Correlation between state-specific estimates of the percentage point difference in proportion of patients prescribed any opioid, attributable to medical cannabis laws, in a given month in the first 3 years of law implementation.

- Correlations generally small in magnitude, but as high as 0.19.

	MN	NY	NH	FL	MD	PA	OK	OH	ND	AR	LA
CT	0.08	0.04	0.02	0.01	-0.02	-0.03	-0.03	-0.03	-0.05	-0.02	-0.02
MN		0.12	0.09	0.06	0.02	-0.02	-0.04	-0.05	-0.08	-0.04	-0.05
NY			0.1	0.08	0.04	0.01	-0.03	-0.04	-0.06	-0.03	-0.04
NH				0.08	0.05	0.02	-0.01	-0.02	-0.04	-0.03	-0.03
FL					0.05	0.03	0	-0.01	-0.02	-0.01	-0.03
MD						0.08	0.04	0.03	0.04	0.02	0.01
PA							0.07	0.06	0.09	0.04	0.03
OK								0.08	0.13	0.06	0.06
OH									0.19	0.09	0.09
ND										0.15	0.15
AR											0.09

Medical Cannabis Laws Study: Results

- In this case, accounting for between-estimate correlation gives $\sim$10% *larger* SE



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

Acknowledgements

- NIDA R01DA049789 (PI: McGinty)
- Co-Authors: Beth McGinty, Kayla Tormohlen, Ian Schmid, Elizabeth Stuart

Under review at *JASA*

Preprint at `arXiv:2311.18093`

`seewaldn@pennmedicine.upenn.edu`

`nickseewald.com`