Policy Evaluation in the Real World:
Experiences with Translating Cutting-Edge Methods for an Applied Audience

Nicholas J. Seewald, Ph.D.

nseewal1@jhu.edu

Joint with Beth McGinty, Ph.D. and Elizabeth A. Stuart, Ph.D.

25 May 2022
American Causal Inference Conference 2022
Berkeley, CA

JOHNS HOPKINS
BLOOMBERG SCHOOL
of PUBLIC HEALTH

These slides are online!

I just tweeted a link to the slides:
`@nickseewald`

A NOTE:
Some of this might feel obvious!
But I hope you'll find it useful to have some vague feelings given structure.

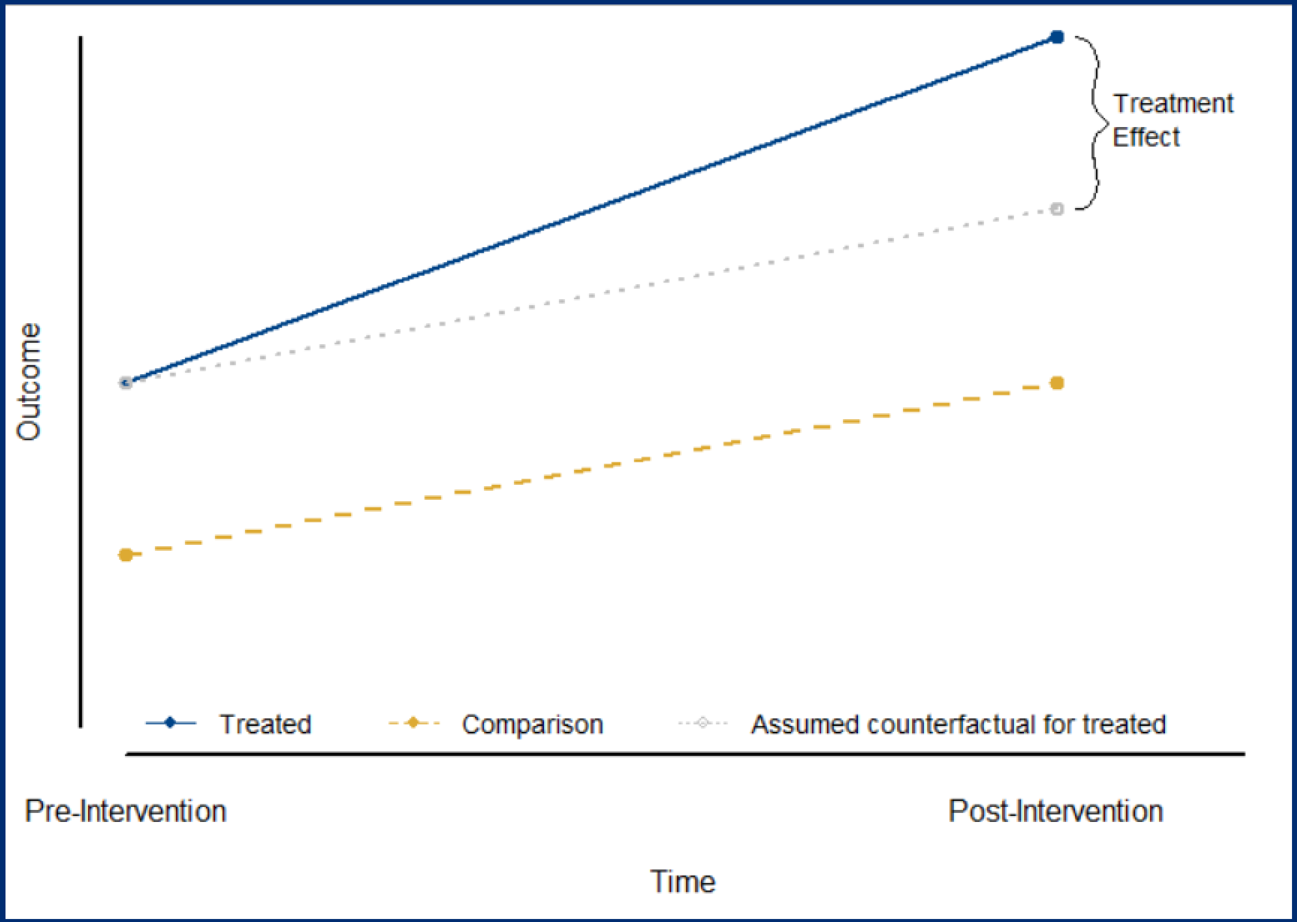
`https://slides.nickseewald.com/acic2022.pdf`

Difference-in-Differences

Compare change in outcome over time between treated and comparison groups.

Under assumption that treated group *would look like comparison group in absence of treatment*, can estimate causal treatment effect.

The above assumption is called "parallel trends".



Assumption: Parallel Trends

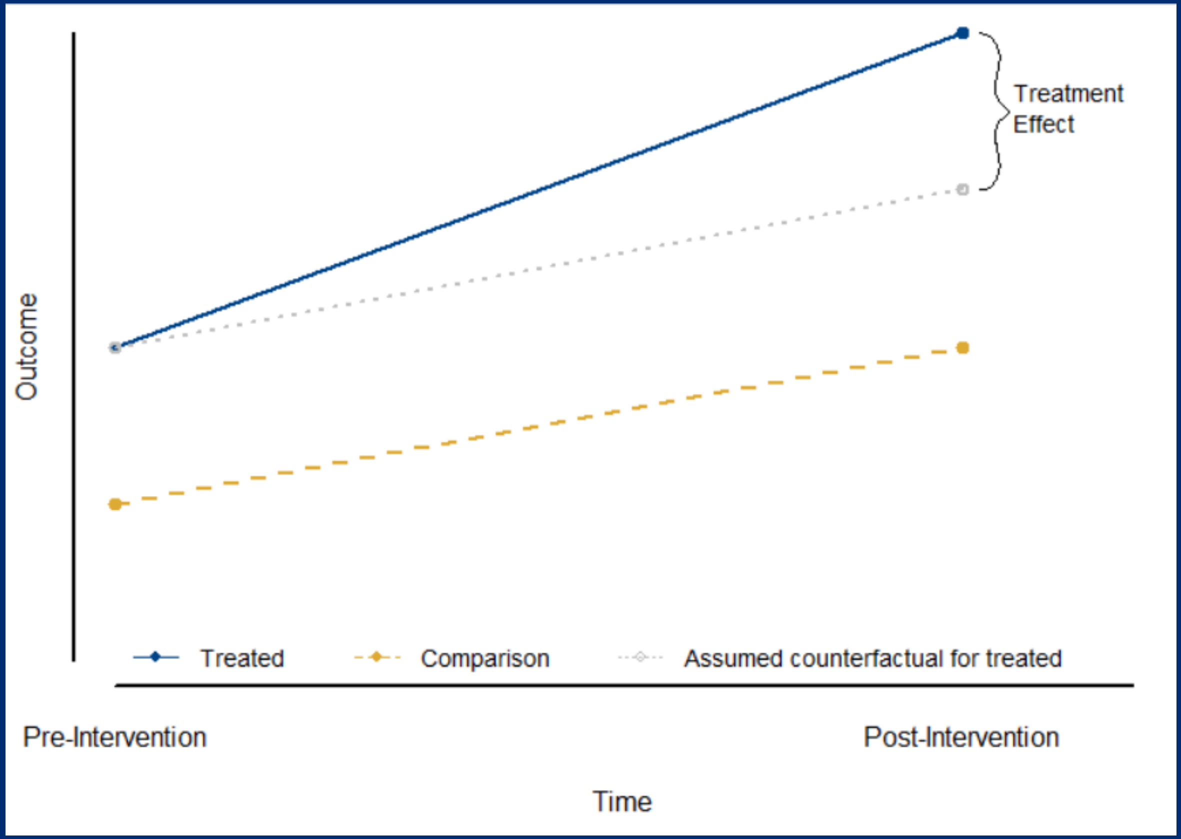
- In a 2-period setting,

  $$E[\,Y_{\text{post}}(0) - Y_{\text{pre}}(0) \mid \text{treated}\,] = E[Y_{\text{post}}(0) - Y_{\text{pre}}(0) \mid \text{comparison}]$$

- In a multi-period setting, more options, e.g.,

  $$E[\,\bar{Y}_{t\geq t^*}(0) - \bar{Y}_{t<t^*}(0) \mid \text{treated}\,] = E[\,\bar{Y}_{t\geq t^*}(0) - \bar{Y}_{t<t^*}(0) \mid \text{comparison}\,]$$

- Untestable, but good agreement over longer pre periods strengthens plausibility.



New Methods for Causal Inference in Policy Evaluation

- Lots of new methods for causal inference in health policy evaluation lately!
- Big focus on relaxing assumptions in classical difference-in-differences (e.g., strict parallel trends)
- "Fixing" problems with classical approaches introduces complexity
  - Complexity is difficult to communicate, impedes adoption of new methods.

How can we translate these cutting-edge methods for an applied audience?

Opioid Prescribing Cap Laws

Annals of Internal Medicine — ORIGINAL RESEARCH

Effects of State Opioid Prescribing Laws on Use of Opioid and Other Pain Treatments Among Commercially Insured U.S. Adults

Emma E. McGinty, PhD, MS; Mark C. Bicket, MD, PhD; Nicholas J. Seewald, PhD; Elizabeth A. Stuart, PhD; G. Caleb Alexander, MD, MS; Colleen L. Barry, PhD, MPP; Alexander D. McCourt, JD, PhD, MPH; and Lainie Rutkow, JD, PhD, MPH

- Laws limit dose and/or duration of opioid prescriptions
- Concern about legislating clinical judgment, impact on patients with chronic noncancer pain.
  - May lead to rapid tapering, discontinuation, etc., without appropriate substitution
- Do these laws change prescribing behavior?

Research Question

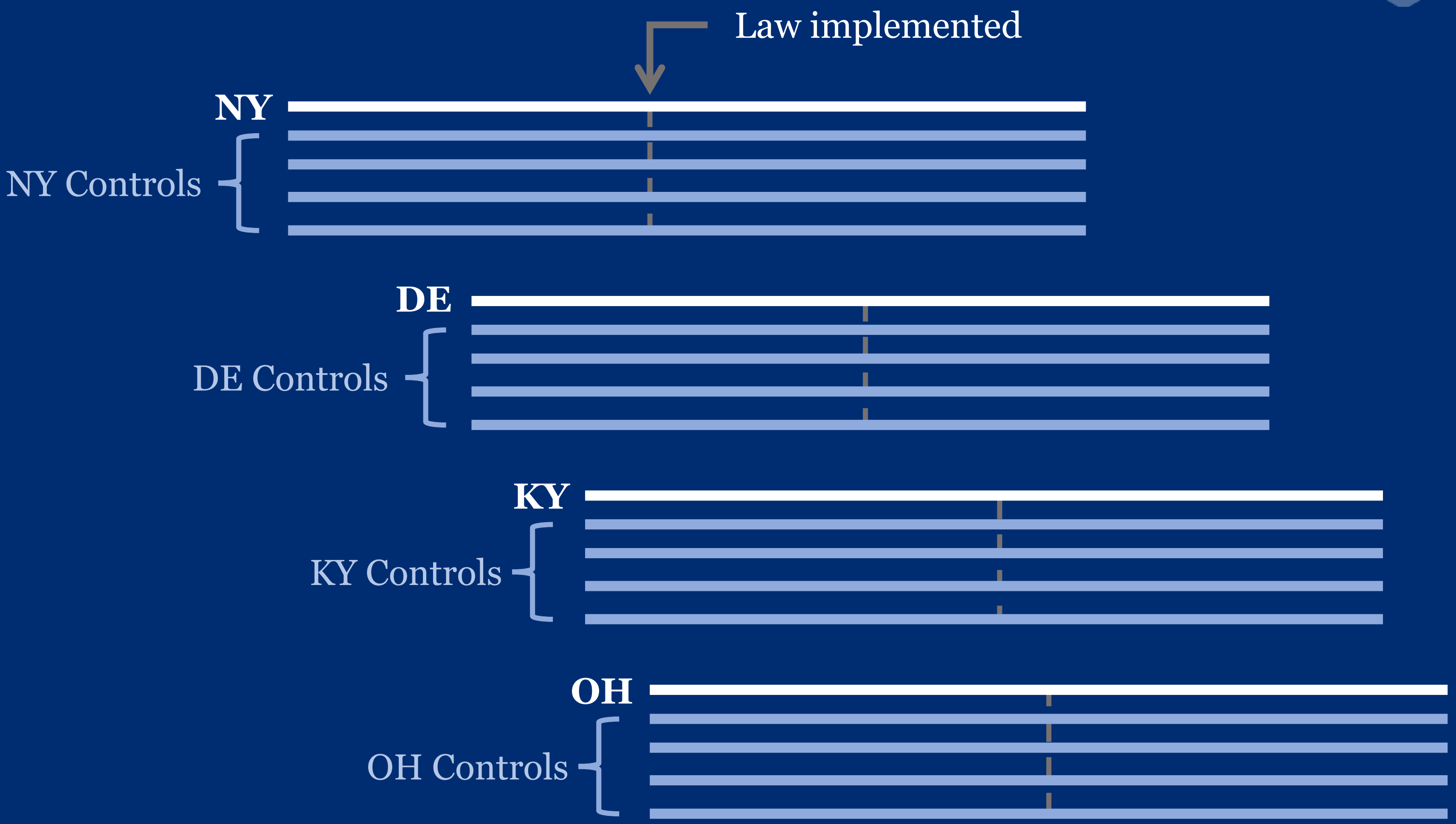
Question: What is the effect of implementing a state opioid prescribing cap law on receipt of opioid prescriptions among commercially insured adults in that state, relative to the expected levels of opioid prescribing absent the law?

Want *state-specific* effects: gives idea of effect heterogeneity, can match up with corresponding qualitative work.

Study Design

4 states implemented opioid prescribing cap law and no other laws related to opioid prescribing in 4-year period.

Each “treated” state has comparison group that did not implement prescribing cap law nor change any other opioid laws in 4-year period.



Analysis Plan

- Original plan was straightforward difference-in-differences analysis
  - BUT: Clear pre-treatment trends in differences
  - Suggests parallel trends violated
- Needed a different approach!

Synthetic Controls

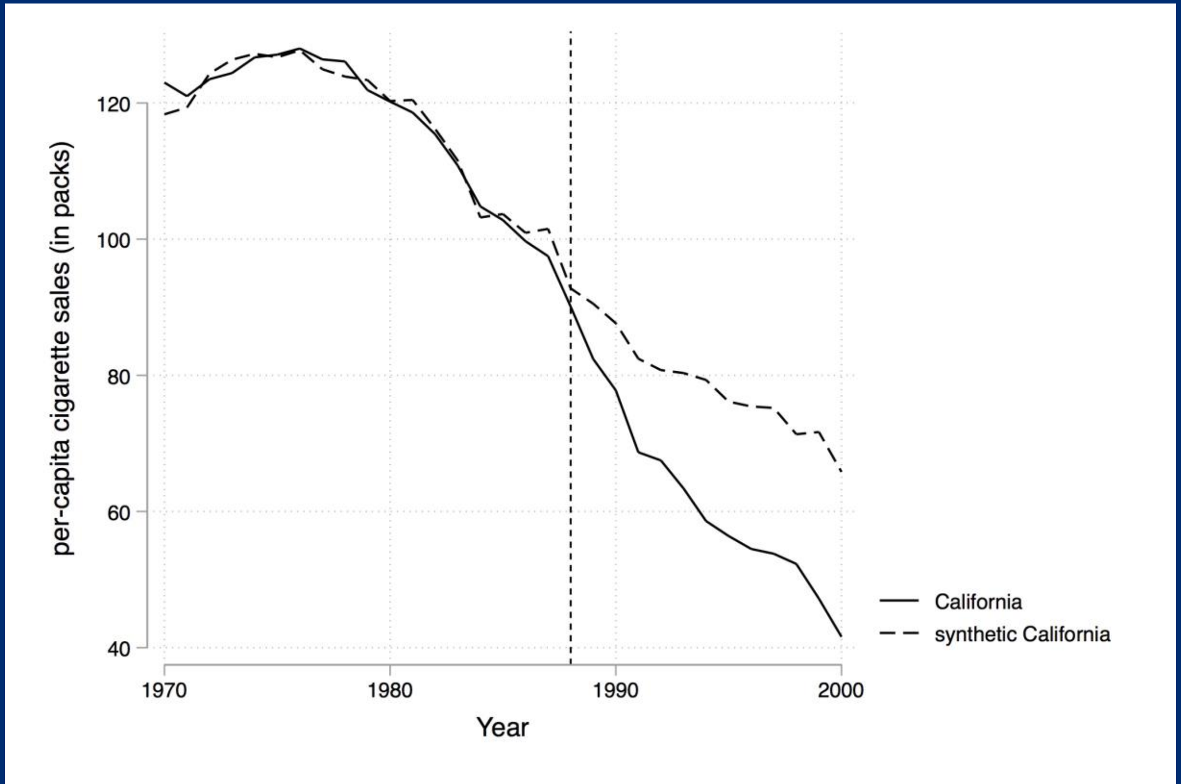


Image from Scott Cunningham, https://mixtape.scunning.com/synthetic-control.html

- **Key idea:** construct a convex combination of untreated states' outcomes to match the treated state's pre-treatment outcomes, then carry that forward into the post-treatment period.
- "Extrapolated" synthetic control outcomes estimate what would have happened in treated state in absence of treatment.
- **CON: Excellent pre-treatment fit is not always possible.**

Augmented Synthetic Controls

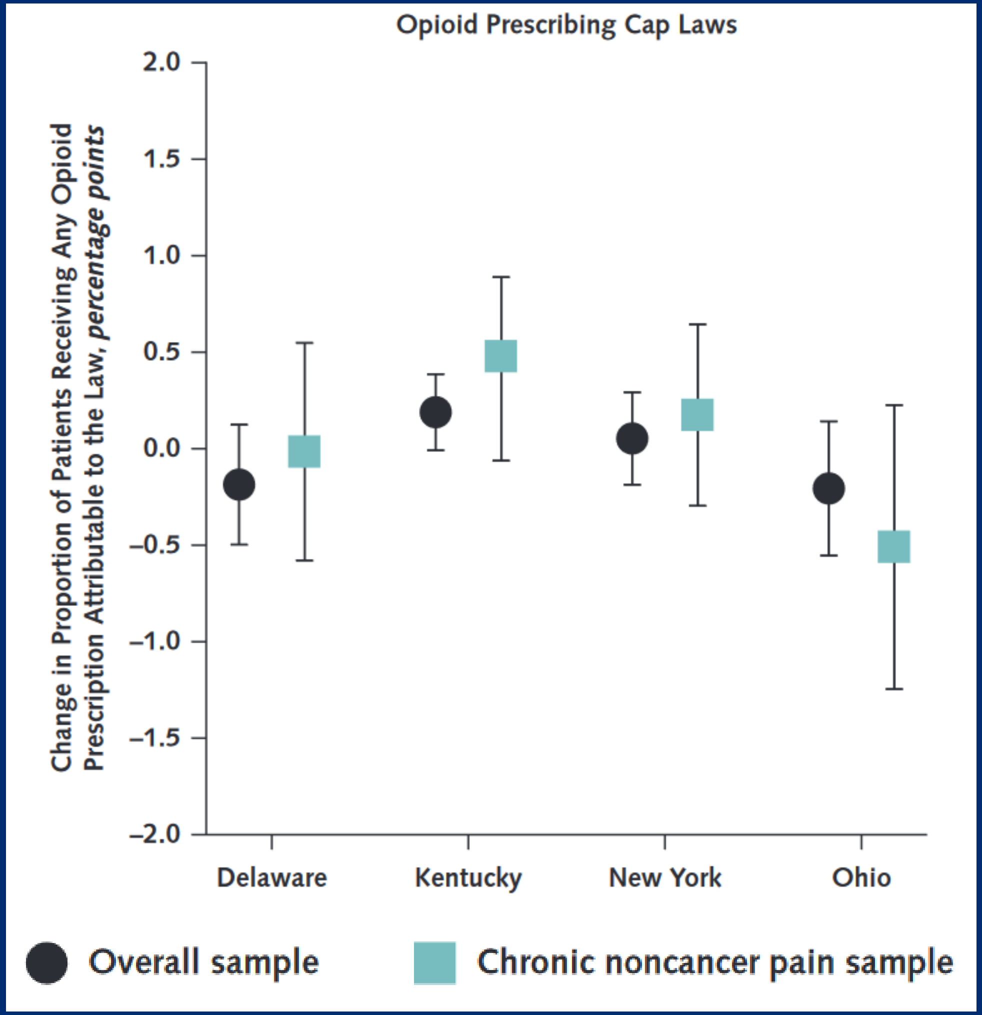
(Ben-Michael et al., *JASA* 2021)

- Augments synthetic control method with an outcome model and allows more flexibility in weights on comparison states
  - Outcome model estimates bias due to imperfect pre-treatment fit
  - De-bias synthetic controls using estimate
- Like synthetic controls, able to *estimate effects in single treated states*

Results

"After analyzing the insurance claims from both groups, the researchers found that state laws were associated with a less than one percent change in the proportion of patients receiving an opioid prescription."

JHSPH Press Release
https://tinyurl.com/yvbwsubs



Communication Strategies for Using Cutting-Edge Methods

Based on experiences with Annals of Internal Medicine

Communication Strategies

1. Focus on the research question
2. Justify additional complexity
3. Be clear and precise
4. Be ready to talk about the “C” word
5. Anticipate extensions
6. Develop tutorials

Focus on the research question

"You need a careful statement of the precise question that synthetic control methods try to answer."

- **Question:** What is the effect of implementation of a state opioid prescribing cap law on receipt of opioid prescriptions among commercially insured adults in that state, relative to the expected levels of opioid prescribing absent the law?
- Note the specificity: this defines our estimand!
  - Effect of [well-defined treatment] on [well-defined outcome], relative to [comparator]
- The estimand guides the choice of methods
  - We want state-specific effects, so we don't need a method that pools multiple effect estimates
  - Use only methods that can answer your research question!

Justify additional complexity

- Answer the question, "Why do you need to use *this* [new/more complex/'nonstandard'] method?"

- Policy evaluation is quite solidified around a set of "standard" methods
  - Things like difference-in-differences, two-way fixed effects, etc.
- Using unfamiliar methods may invite skepticism
  - *Crucial* to lean on research question
  - Explain why the standard approaches won't work

Justify additional complexity

"You should start with the problem you are trying to solve: that conventional approaches really apply only to the case of two groups and two periods, but that you have issues of 24 + 24 times (if we understand correctly) and staggered start times (time zero) [...]. For these reasons, you are using an alternative approach that overcomes these shortcoming [sic]."

- Sort of right, sort of wrong: an opportunity to clarify!
  - Reviewer wanted us to justify complexity, and gave it a shot *for us* (!!!)
- We used augmented synthetic controls to get *better balance* between treatment and comparison states in the pre-treatment period.
- Our study avoids issues with staggered adoption *by design*
  - Careful avoidance of states with confounding laws, focus on state-specific effects

Be clear and precise

- Avoid jargon
  - You might not expect what reviewers see as jargon! (e.g., ATT)
- Focus on the applied audience: motivate the method well.
  - *"Focus more on saying what you are doing and why for the Annals audience; refer the economists to JASA."*
- Keep in mind who your reviewers might be
  - We expected an economist, someone more clinical, and a statistical reviewer
  - Can be tricky to balance appropriate amount of detail: might err toward caution and wait for R&R

Be clear and precise

- Be very clear about your estimand.
- Spend time **crafting** language to describe effect estimates
  - This is hard: must balance precision and clarity
- We settled on "change in proportion of patients receiving any opioid prescription, per month, attributable to the law in its first two years of implementation".
  - Needed to capture (1) change over time, (2) difference between treated state and synthetic control, (3) outcome measurement time scale, (4) causation

Be ready to talk about the "C" word

"[A]void implying causality given the observational study design."

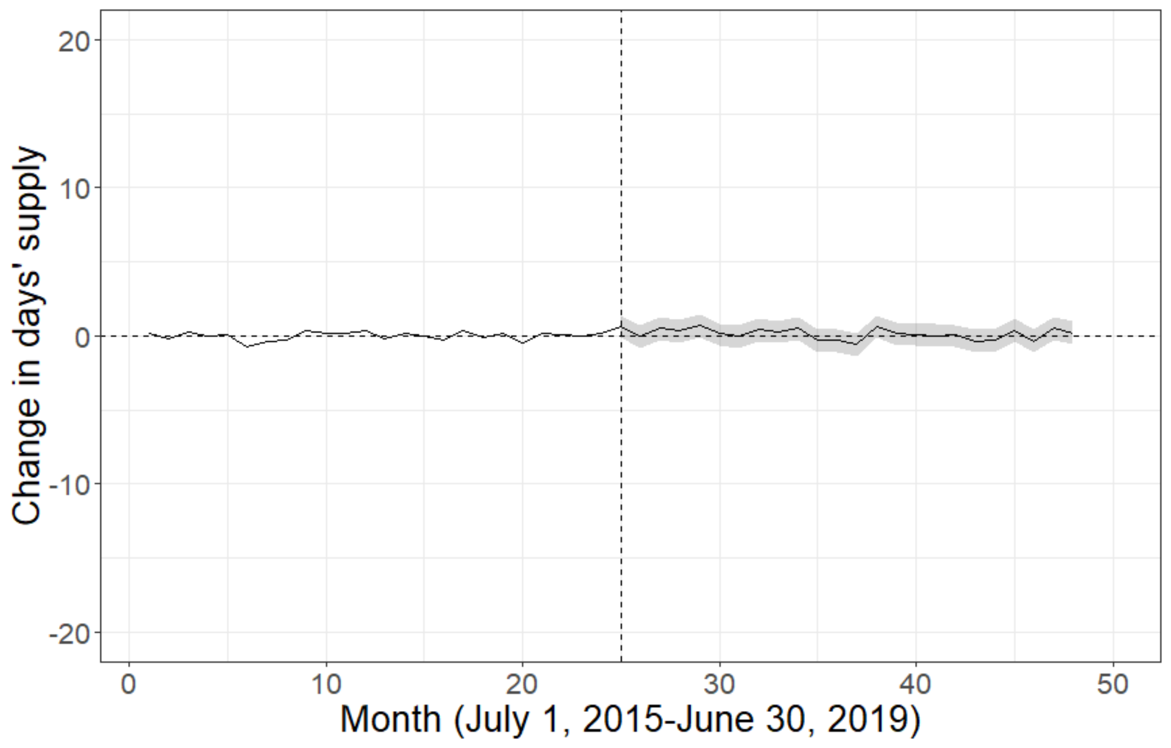
- Cue rants, etc. etc.
- Pick your battles!
- Sneaky backdoor: "attributable to the law"

Anticipate extensions

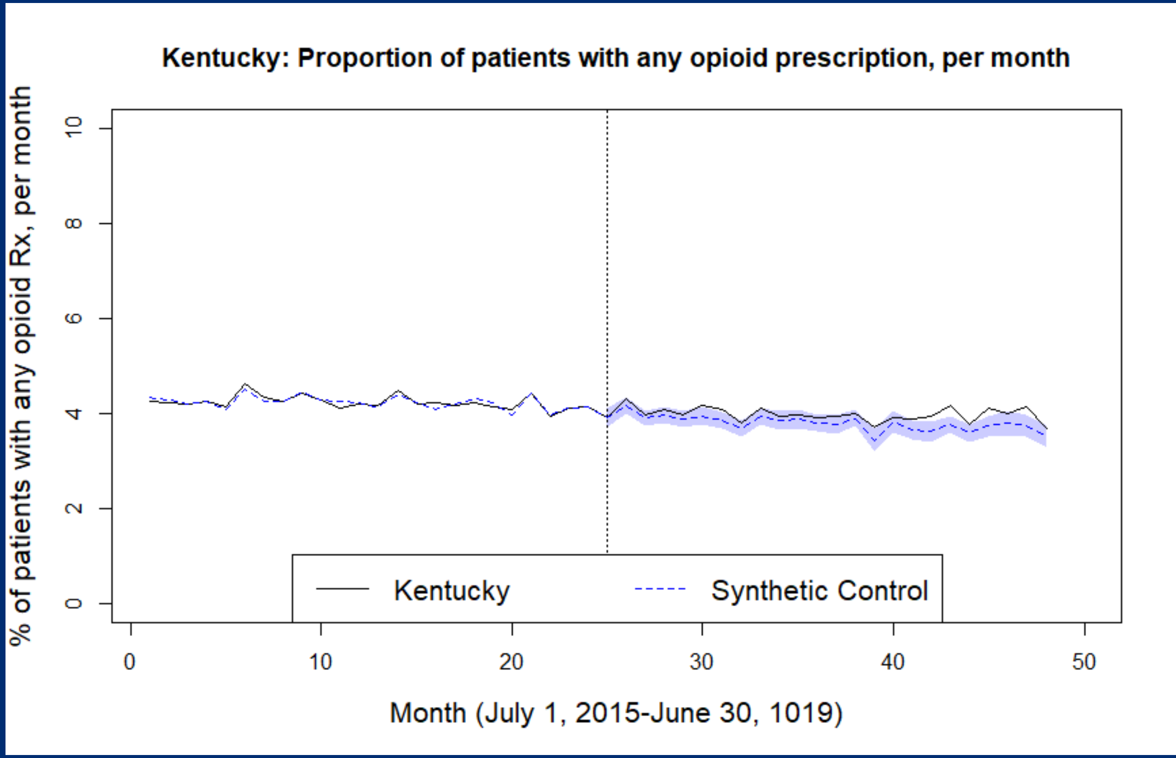
- A key advantage of augmented synthetic controls (vs. synthetic controls) is the capacity for inference
  - Using an outcome model gets us p-values and confidence intervals (friendly!)
  - But maybe not for every quantity reviewers/editors expect
- We extended `augsynth` R package output in two ways: (1) inference for aggregated estimates, (2) plotting

Anticipate extensions

Default `augsynth`: ATT scale



Our extension: outcome scale



Consider developing tutorials

- Getting your new method in the hands of practitioners and applied scientists can be hard. Make it easier by building explanatory companion material aimed at an applied audience.
  - Software packages
  - Annotated example code (vignettes)
  - Tutorial paper in applied journal
  - YouTube videos
  - Blogs
  - Use your imagination!

Acknowledgments

- National Institute on Drug Abuse [R01DA044987, R01DA049789]
- ACIC/NSF New Researchers travel grant
- Editors & reviewers of *Annals of Internal Medicine*

The content of this presentation is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.