Partial Identification of Dose Responses with Hidden Confounders

Myrl G. Marmarelis, Elizabeth Haddad, Andrew Jesson,
Neda Jahanshad, Aram Galstyan, Greg Ver Steeg

39th Conference on Uncertainty in Artificial Intelligence, 2023

Confounders?

2023

Even a Little Alcohol Can Harm Your Health

Recent research makes it clear that any amount of drinking can be detrimental. Here's why you may want to cut down on your consumption beyond Dry January.

2015

Wine May Help With Diabetes

2015

Drink to Your Health (in Moderation), the Science Says

2008

New Hints Seen That Red Wine May Slow Aging

2008

Red Wine May Curb Fat Cells

2008

Regimens: Wine May Help Keep Liver Healthy

2004

Alcohol's Benefits Extend to Hypertension

2002

Evidence Mounting That Moderate Drinking Is Healthful

1994

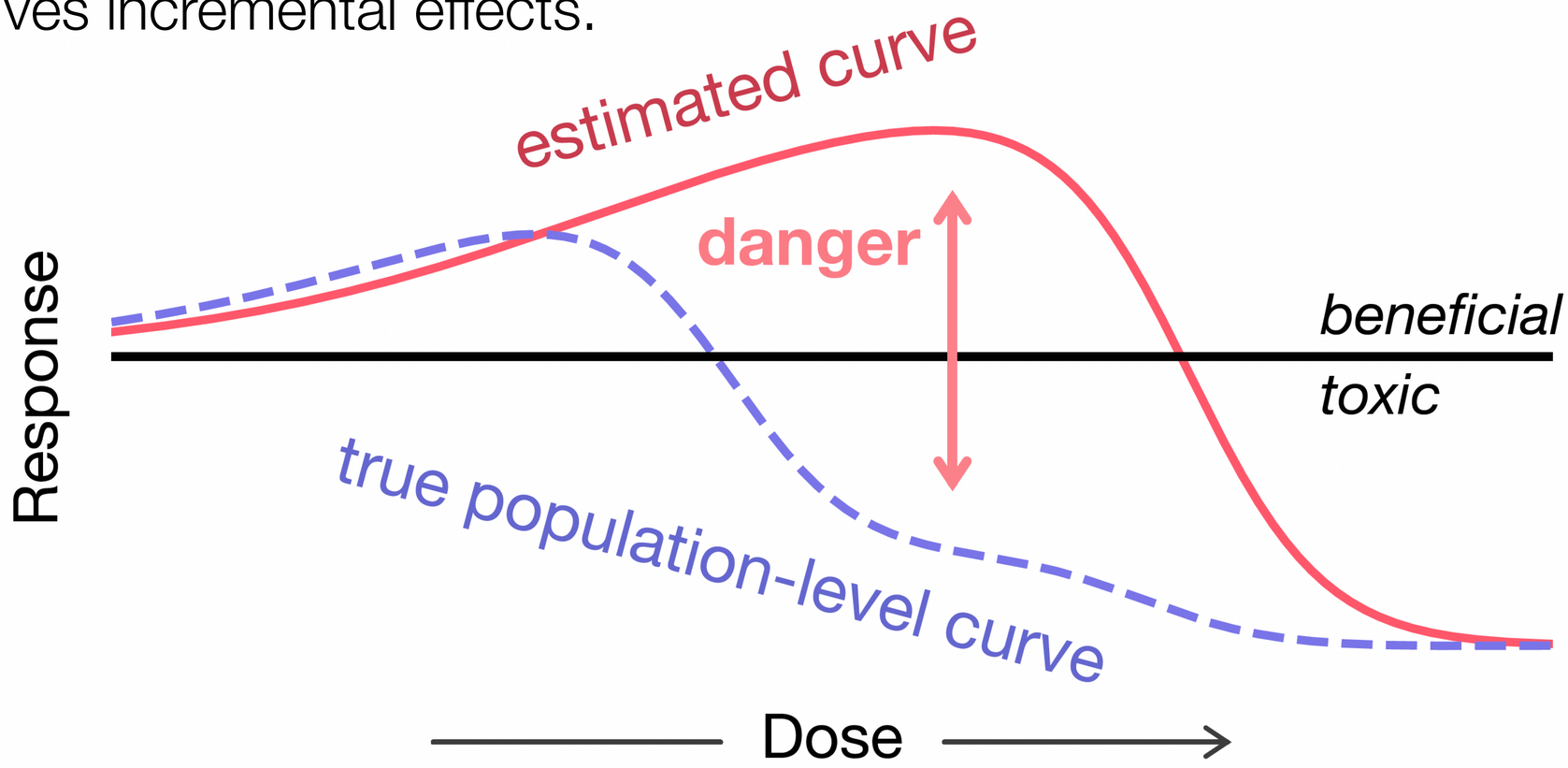
Wine for the Heart: Over All, Risks May Outweigh Benefits

Why Partial Identification […] with Hidden Confounders

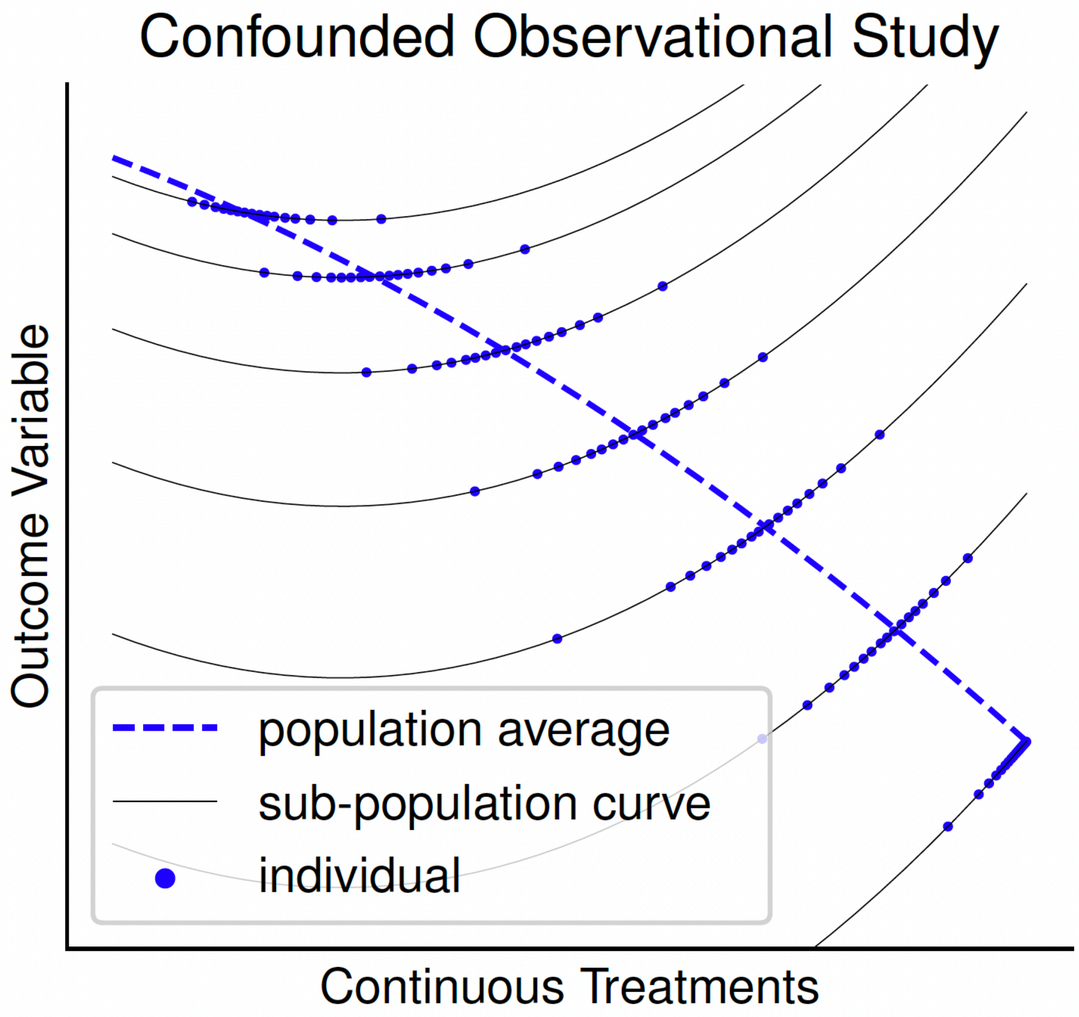
- Typical predictions are *descriptive.*
- Causal inferences are *prescriptive*.
  - We aim to predict the outcome of an intervention.
  - If a study produces *actionable* insights, then it is claiming to make a causal inference, whether explicitly or not.
- If we know there might be confounders (endogeneity), then point identification of causal outcomes is impossible.
- Partial identification is our best bet: produce a *set* of outcomes admitted by the causal setting.

Why Dose Responses

- “Amount” of treatment is important in many problems.
- Derivative of the curve gives incremental effects.



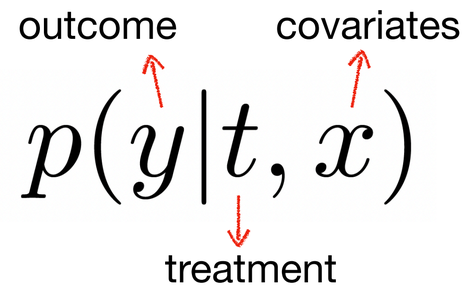
Related: Simpson's Paradox



Our Proposed Method

Potential Outcomes Setup

- Say we have an outcome prediction model. Assume we learned it perfectly.



- This predicts the potential outcome at *t*. $p(y_t|t,x) = p(y|t,x)$
- “Potential outcomes” are the different treatment outcomes after controlling for everything else that is relevant in the problem, i.e. confounders.
- The dose response is $\mathbb{E}[Y_t \mid X = x]$ as a function of t.

The Ignorability Assumption

- No hidden confounding!

potential outcomes

covariates

$$\{(Y_t)_{t\in\mathcal{T}} \perp\!\!\!\perp T\} \mid X$$

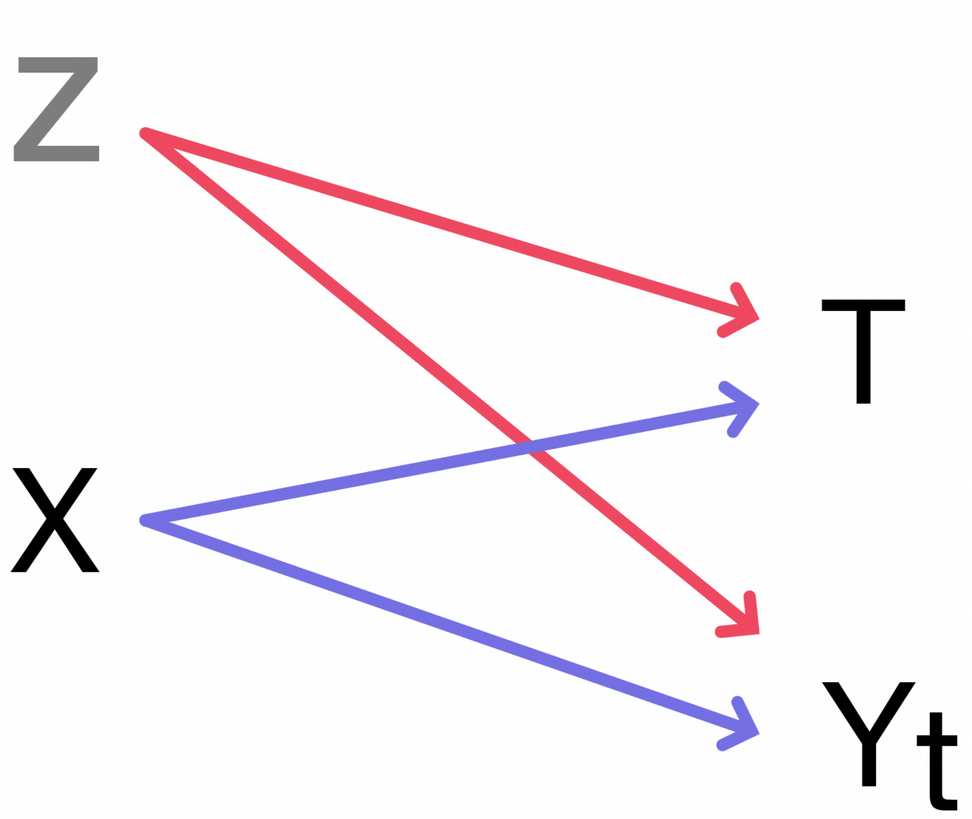
treatment assignment

Hidden Confounding

- If the ignorability assumption held, then assigned treatment wouldn't affect the potential outcome, conditioned on observed confounders.
- In that case,

$$p(y_t|x) = p(y_t|t,x) = p(y|t,x)$$

- However, a hidden confounder could *ruin* this via a backdoor path.
- The graph to the right gives one such example with the red arrows.



Continuous Treatments

propensity

$$p(y_t|x) = \int_{\mathcal{T}} p(y_t|\,\tau, x)p(\tau|x)\,\mathrm{d}\tau$$

the potential outcome

Continuous Treatments

propensity

$$p(y_t|x) = \int_{\tau} p(y_t|\,\tau, x)p(\tau|x)\,\mathrm{d}\tau$$

the potential outcome

$$p(y_t|\,\tau, x)$$

"what is this person's potential outcome at *t* given that their assigned treatment is *tau*"

the counterfactual

The Problem with Continuous Treatments

a) Confounded Outcomes for Binary Treatments

$$P[Y_t] = P[Y_t \mid T = t] \times P[T = t] + P[Y_t \mid T = 1-t] \times P[T = 1-t]$$

counterfactual

The Problem with Continuous Treatments

a) Confounded Outcomes for Binary Treatments

$$P[Y_t] = P[Y_t|\ T = t\ ] \times P[T = t] + P[Y_t|\ T = 1\text{-}t\ ] \times P[T = 1\text{-}t]$$

counterfactual

b) Confounded Outcomes for Continuous Treatments

$$P[Y_t] = \int_{\mathcal{T}}$$

observable $P[Y_t|\ T = t\ ]$ (*infinitesimal support*)

completely unknown counterfactuals

$T \neq t$

$T = t$

possible treatment assignments T

- Infinite unobservable counterfactuals!
- The integrand cannot be identified almost anywhere.

The Problem with Continuous Treatments

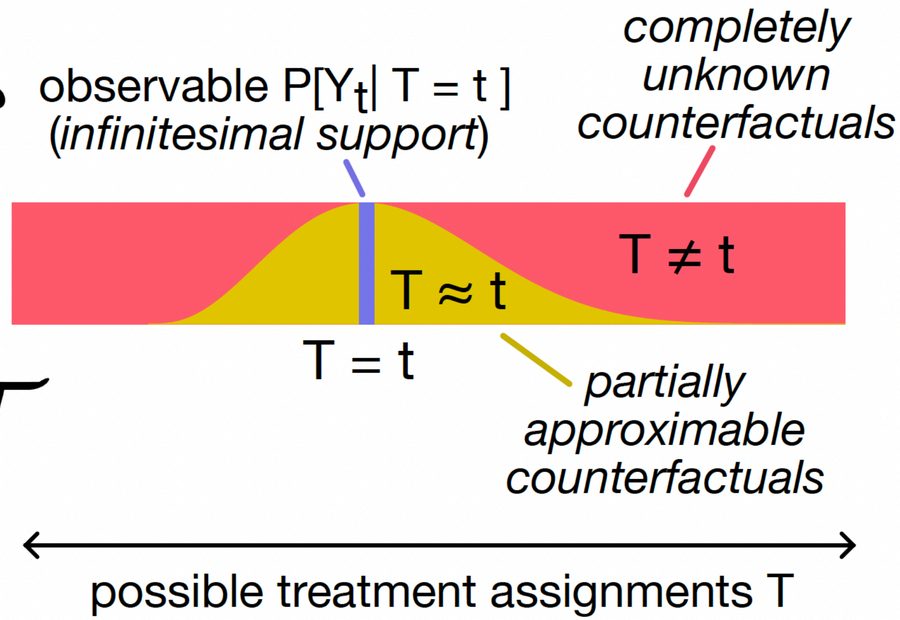
a) Confounded Outcomes for Binary Treatments

$$P[Y_t] = P[Y_t|\, T = t\,] \times P[T = t] + P[Y_t|\, T = 1\text{-}t\,] \times P[T = 1\text{-}t]$$

counterfactual

b) Confounded Outcomes for Continuous Treatments

$$P[Y_t] = \int_{\mathcal{T}}$$



- Infinite unobservable counterfactuals!
- The integrand cannot be identified almost anywhere.
- We need an approximation.

We Know Nothing!

- First step to the solution is extrapolating from the point that we can observe.

the counterfactual the prediction

$$p(y_t|\tau, x) = p(y_t|t, x) + (\tau - t)\partial_\tau p(y_t|\tau, x)|_{\tau=t} + \frac{(\tau - t)^2}{2}\partial_\tau^2 p(y_t|\tau, x)|_{\tau=t} + \mathcal{O}(\tau - t)^3$$

extrapolations

Now We Know Something

- Second step is to specify where that extrapolation can be trusted, and how much.

(from before)

$$p(y_t|x) = \int_{\mathcal{T}} p(y_t|\,\tau, x)p(\tau|x)\,\mathrm{d}\tau$$

Now We Know Something

- Second step is to specify where that extrapolation can be trusted, and how much.

(from before)

$$p(y_t|x) = \int_{\mathcal{T}} p(y_t|\,\tau, x)p(\tau|x)\,\mathrm{d}\tau$$

(use weights and split)

$$\approx \int_{\mathcal{T}} \underbrace{w_t(\tau)\tilde{p}(y_t|\tau, x)p(\tau|x)\,\mathrm{d}\tau}_{(A)\text{ the approximated quantity}}$$

$$+ \int_{\mathcal{T}} \underbrace{[1 - w_t(\tau)]p(\tau|y_t, x)p(y_t|x)\,\mathrm{d}\tau}_{(B)\text{ by Bayes' rule}}$$

Solution Outline

- We find that

$$p(y_t|x) \approx \frac{\int_{\mathcal{T}} w_t(\tau)\tilde{p}(y_t|\,\tau, x)p(\tau|x)\,\mathrm{d}\tau}{\int_{\mathcal{T}} w_t(\tau)p(\tau|\,y_t, x)\,\mathrm{d}\tau}.$$

- We'll figure out what to do with the "trust weights" later.
- We still have two unknowns: the approximation, and the denominator.

Finally Introducing the Sensitivity Model, δMSM

- For binary treatments, the classical MSM bounds the Radon-Nikodym derivative of the two counterfactuals, i.e. the ratio of their densities.

Finally Introducing the Sensitivity Model, δMSM

- For binary treatments, the classical MSM bounds the Radon-Nikodym derivative of the two counterfactuals, i.e. the ratio of their densities.
- We follow a similar route but take it to the infinitesimal limit:

$$\omega_\delta(y_t|\,\tau, x) := \frac{p(y_t|\,\tau+\delta, x)}{p(y_t|\,\tau, x)}$$

ratio of nearby counterfactuals

Finally Introducing the Sensitivity Model, δMSM

- For binary treatments, the classical MSM bounds the Radon-Nikodym derivative of the two counterfactuals, i.e. the ratio of their densities.
- We follow a similar route but take it to the infinitesimal limit:

$$\omega_\delta(y_t|\,\tau, x) := \frac{p(y_t|\,\tau+\delta, x)}{p(y_t|\,\tau, x)}$$

$$= \left[\frac{p(\tau+\delta|x)}{p(\tau|x)}\right]^{-1} \left[\frac{p(\tau+\delta|\,y_t, x)}{p(\tau|\,y_t, x)}\right]$$

by Bayes' rule.

nominal propensities

complete propensities

Definition of the δMSM

For treatments $t \in \mathcal{T} \subseteq \mathbb{R}$, where $\mathcal{T}$ is connected, and violation-of-ignorability factor $\Gamma \geq 1$, the δMSM requires

$$\left| \frac{\partial}{\partial \tau} \log \frac{p(\tau | y_t, x)}{p(\tau | x)} \right| \leq \log \Gamma$$

Definition of the δMSM

For treatments $t \in \mathcal{T} \subseteq \mathbb{R}$, where $\mathcal{T}$ is connected, and violation-of-ignorability factor $\Gamma \geq 1$, the δMSM requires

$$\left| \frac{\partial}{\partial \tau} \log \frac{p(\tau | y_t, x)}{p(\tau | x)} \right| \leq \log \Gamma$$

ratio of complete and nominal propensities

"how much is this person's treatment assignment informed by a potential outcome (through backdoor paths)"

Necessary Assumptions for Hidden Confounding

- First Assumption: δMSM holds with some Γ.
- Second Assumption: we need an “anchor point,” designated as zero treatment.

$$p(\tau = 0 \mid y_t, x) = p(\tau = 0 \mid x) \text{ for all } x, t, \text{ and } y_t$$

Necessary Assumptions for Hidden Confounding

- First Assumption: δMSM holds with some Γ.
- Second Assumption: we need an “anchor point,” designated as zero treatment.

$$p(\tau = 0 \mid y_t, x) = p(\tau = 0 \mid x) \text{ for all } x, t, \text{ and } y_t$$

- This is necessary to solve the integrals.
- What does the second assumption mean for our partial identification?
  - Informally, hidden confounders “matter less” at near-zero treatment values.

Combining the Ingredients

partial identification factor

$$\tilde{p}(y_t|x) = d(t|y_t, x)^{-1} p(y_t|t, x)$$

approx. potential outcome

outcome prediction

Combining the Ingredients

partial identification factor

$$\tilde{p}(y_t|x) = d(t|y_t, x)^{-1} p(y_t|t, x)$$

approx. potential outcome

outcome prediction

- Admissible probability densities are governed by

$$d(t|y_t, x) \in \left[\underline{d}(t|y_t, x), \overline{d}(t|y_t, x) \right]$$

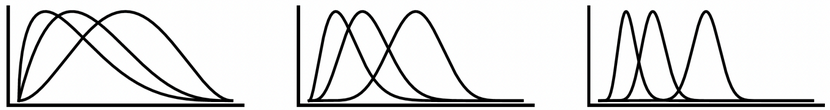
which can be **solved in closed form!** Note: $\underline{d} \leq 1 \leq \overline{d}$

Settling the Trust Weights

- Accuracy of the extrapolation depends on the continuity of the counterfactual density with respect to treatment assignment.
- Narrower treatment propensity densities $p(\tau|x)$ suggest worse extrapolations.

- Therefore, we parametrize the weights $w_t(\tau)$ to have the same narrowness (and form) as the nominal propensities, but always centered at *t* of course.

- We found solutions for various exponential families.



One Last Thing: Relaxing the Second Assumption

- For Beta-distributed treatments, we symmetrify the anchor point assumption.

$$\begin{cases} p(\tau = 0 \mid y_t, x) = p(\tau = 0 \mid x) & \text{w.p.} \quad t, \\ p(\tau = 1 \mid y_t, x) = p(\tau = 1 \mid x) & \text{w.p.} \quad 1 - t \end{cases}$$

- New interpretation: the more distant the potential outcome, the less informative it is about treatment assignment.

Relaxed Anchor Points, Illustrated

- **Are you the kind of person that drinks a lot of wine?** $(\tau = 1)$
  - **Depends on your health outcome from drinking a lot of wine.** (y_1)
  - Depends on your health outcome from drinking no wine. (y_0)
- **Are you the kind of person that drinks no wine?** $(\tau = 0)$
  - Depends on your health outcome from drinking a lot of wine. (y_1)
  - **Depends on your health outcome from drinking no wine.** (y_0)

Benchmark Results

Benchmarks	brain		blood		pbmc		mftc			
	mean	(std.)	mean	(std.)	mean	(std.)	mean	(std.)	% best	ratio to best
δMSM (ours)	**138**	(120)	**141**	(129)	**138**	(121)	**144**	(124)	**78.4**	**1.03** (0.08)
CMSM	186	(153)	188	(156)	205	(169)	182	(145)	7.8	1.81 (2.15)
uniform	158	(137)	162	(146)	157	(136)	167	(141)	4.8	1.20 (0.10)
binary MSM	211	(128)	213	(131)	222	(127)	214	(127)	9.0	2.57 (2.34)

- Partial-identification costs of 90% coverage of the average dose responses.
- Semi-synthetic confounders are random projections of original data.
- Random quadratic forms describe the potential outcome.
- 500 experiments per dataset and method.

Conclusion — So What

paper

code

- **We described the first sensitivity model for continuous treatments that**
  - **changes with the propensity (& in a sensical way)**
  - **always admits valid potential outcome densities.**

- Extensive semi-synthetic benchmarks show consistently superior performance to baseline sensitivity models.

myrlm@isi.edu