Transfer Learning Between U.S. Presidential Elections

How can we learn from a 2020 ad campaign to inform 2024 ad campaigns?

Xinran Miao, Jiwei Zhao, Hyunseung Kang

arXiv:2411.01100

University of Wisconsin-Madison

Models, Experiments, and Data Workshop (MEAD)
February 14, 2025, Madison, Wisconsin

Two-Slide Overview

- Question: Would digital ads against President Trump change voter turnout in Pennsylvania (PA), a key swing state, during the 2024 U.S. presidential election?
  - On average, would a voter in Philadelphia county be more (or less) likely to vote when exposed to digital ads against Trump?
- Ideal Approach: Randomized experiment to measure the ad effect (e.g., [Aggarwal et al. 2023])
  - Voters are randomized to receive the negative ad (i.e., treatment) or not (i.e., control)
  - Unfortunately, these experiments are very, very expensive

Two-Slide Overview

- Our Approach: "Carefully" borrow information from past experiments measuring ad effects to estimate the 2024 effect
  - Methodology: transfer learning with sensitivity analysis

Two-Slide Overview

- Our Approach: "Carefully" borrow information from past experiments measuring ad effects to estimate the 2024 effect
  - Methodology: transfer learning with sensitivity analysis
    - Transfer learning: use data from the 2020 voters (i.e., source population) to learn a parameter of 2024 PA voters (i.e., target population) and measured voter demographics in two years

Two-Slide Overview

- Our Approach: "Carefully" borrow information from past experiments measuring ad effects to estimate the 2024 effect
  - Methodology: transfer learning with sensitivity analysis
    - Transfer learning: use data from the 2020 voters (i.e., source population) to learn a parameter of 2024 PA voters (i.e., target population) and measured voter demographics in two years
    - Sensitivity analysis: account for unmeasurable differences between 2020 and 2024 elections (e.g., shift in electoral context)

Two-Slide Overview

- Our Approach: "Carefully" borrow information from past experiments measuring ad effects to estimate the 2024 effect
  - Methodology: transfer learning with sensitivity analysis
    - Transfer learning: use data from the 2020 voters (i.e., source population) to learn a parameter of 2024 PA voters (i.e., target population) and measured voter demographics in two years
    - Sensitivity analysis: account for unmeasurable differences between 2020 and 2024 elections (e.g., shift in electoral context)
  - Other (mostly technical) innovations:
    - A theoretically valid bootstrap procedure to quantify uncertainty in transfer learning
    - Estimators based on the plug-in principle and the efficient influence function
    - Calibration of the sensitivity parameter via data splitting

Two-Slide Overview

- Our Approach: "Carefully" borrow information from past experiments measuring ad effects to estimate the 2024 effect
  - Methodology: transfer learning with sensitivity analysis
    - Transfer learning: use data from the 2020 voters (i.e., source population) to learn a parameter of 2024 PA voters (i.e., target population) and measured voter demographics in two years
    - Sensitivity analysis: account for unmeasurable differences between 2020 and 2024 elections (e.g., shift in electoral context)
  - Other (mostly technical) innovations:
    - A theoretically valid bootstrap procedure to quantify uncertainty in transfer learning
    - Estimators based on the plug-in principle and the efficient influence function
    - Calibration of the sensitivity parameter via data splitting
- Data analysis for Pennsylvania, a "tipping-point" state in the 2024 U.S. presidential election

Disclaimers

None of us are experts in U.S. elections or voter mobilizations.

We approached this empirical question as data scientists. Also, the 2024 provided a **very unique** opportunity to explore transfer learning/generalizability in U.S. elections

Comments/Critiques are highly encouraged:)

Background: 2020 Digital Ad Campaign [Aggarwal et al. 2023]

- What is the effect of online ads against Donald Trump on voter turnout in five battleground states: AZ, MI, NC, PA, and WI?

Background: 2020 Digital Ad Campaign [Aggarwal et al. 2023]

- What is the effect of online ads against Donald Trump on voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- A stratified randomized experiment from February 2020 to November 2020
  - $n = 1,999,282$ registered voters
  - Stratified by gender, race, and age groups (all discrete)

Background: 2020 Digital Ad Campaign [Aggarwal et al. 2023]

- What is the effect of online ads against Donald Trump on voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- A stratified randomized experiment from February 2020 to November 2020
  - $n = 1,999,282$ registered voters
  - Stratified by gender, race, and age groups (all discrete)
- Treatment group ($\approx 85.6\%$): an average of 754 ads against Trump by Acronym, a left-leaning, non-profit organization.

Background: 2020 Digital Ad Campaign [Aggarwal et al. 2023]

- What is the effect of online ads against Donald Trump on voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- A stratified randomized experiment from February 2020 to November 2020
  - $n = 1,999,282$ registered voters
  - Stratified by gender, race, and age groups (all discrete)
- Treatment group ($\approx 85.6\%$): an average of 754 ads against Trump by Acronym, a left-leaning, non-profit organization.
- Control group ($\approx 14.4\%$): no ads from Acronym

Background: 2020 Digital Ad Campaign [Aggarwal et al. 2023]

- What is the effect of online ads against Donald Trump on voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- A stratified randomized experiment from February 2020 to November 2020
  - $n = 1,999,282$ registered voters
  - Stratified by gender, race, and age groups (all discrete)
- Treatment group ($\approx 85.6\%$): an average of 754 ads against Trump by Acronym, a left-leaning, non-profit organization.
- Control group ($\approx 14.4\%$): no ads from Acronym
- Outcome: voted in 2020 U.S. election? (**1 = voted; 0 = not voted**)
  - We don't know which candidate a voter voted for

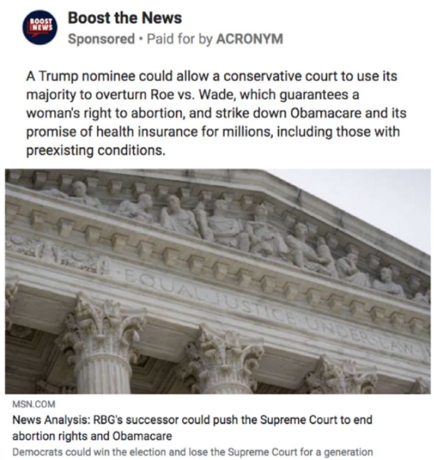
Background: 2020 Digital Ad Campaign [Aggarwal et al. 2023]

- What is the effect of online ads against Donald Trump on voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- A stratified randomized experiment from February 2020 to November 2020
  - $n = 1,999,282$ registered voters
  - Stratified by gender, race, and age groups (all discrete)
- Treatment group ($\approx 85.6\%$): an average of 754 ads against Trump by Acronym, a left-leaning, non-profit organization.
- Control group ($\approx 14.4\%$): no ads from Acronym
- Outcome: voted in 2020 U.S. election? (**1 = voted; 0 = not voted**)
  - We don't know which candidate a voter voted for
- Analysis: linear regression

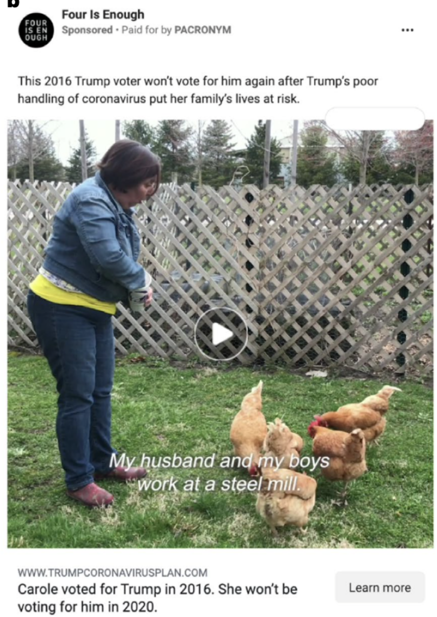
Example Political Ads in Treatment Group

Ads were on Facebook, Instagram, and Outbrain advertising network.

a



b



(a): ads promoting news stories (b): traditional video campaign ads

Results from [Aggarwal et al. (2023)]

The effect was difference in voter turnout between treated and control.

Results from [Aggarwal et al. (2023)]

The effect was difference in voter turnout between treated and control.

- Negative effect: ad against Trump decreased turnout
  - For example, a Trump supporter may be discouraged to vote after the ads against him
- Positive effect: ad against Trump increased turnout
  - For example, a Democratic voter may be encouraged to vote after the ads against Trump

Results from [Aggarwal et al. (2023)]

The effect was difference in voter turnout between treated and control.

- Negative effect: ad against Trump decreased turnout
  - For example, a Trump supporter may be discouraged to vote after the ads against him
- Positive effect: ad against Trump increased turnout
  - For example, a Democratic voter may be encouraged to vote after the ads against Trump

Overall ad effect on turnout: -0.06 percentage points (SE $= 0.12$)

Results from [Aggarwal et al. (2023)]

The effect was difference in voter turnout between treated and control.

- Negative effect: ad against Trump decreased turnout
  - For example, a Trump supporter may be discouraged to vote after the ads against him
- Positive effect: ad against Trump increased turnout
  - For example, a Democratic voter may be encouraged to vote after the ads against Trump

Overall ad effect on turnout: -0.06 percentage points (SE $= 0.12$)

- *"...we can affirm that our overall estimate is effectively equivalent to zero..."*

Results from [Aggarwal et al. (2023)]

The effect was difference in voter turnout between treated and control.

- Negative effect: ad against Trump decreased turnout
  - For example, a Trump supporter may be discouraged to vote after the ads against him
- Positive effect: ad against Trump increased turnout
  - For example, a Democratic voter may be encouraged to vote after the ads against Trump

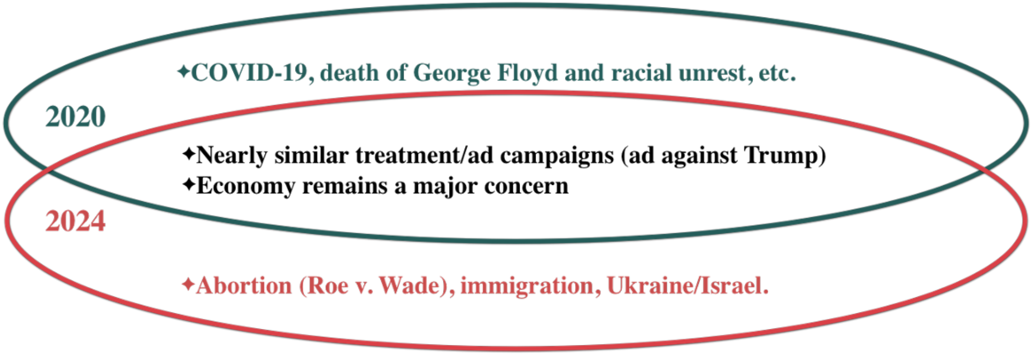
Overall ad effect on turnout: -0.06 percentage points (SE $= 0.12$)

- *"...we can affirm that our overall estimate is effectively equivalent to zero..."*

A quote that grabbed our attention and motivated this work:

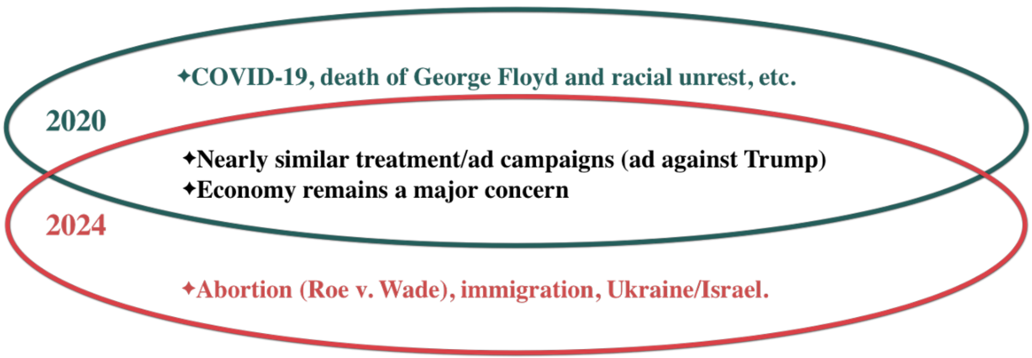
"One reasonable question for our study is how well ***our findings would generalize****...* ***to other electoral contexts****....it could be that the 2020 election was exceptional because of COVID and the idiosyncrasies of the candidates, so perhaps digital advertising would have larger effects in more typical settings...."*

Key Question: Would the Negative Ad Against Trump Remain Ineffective in 2024?



Ideal approach: running a randomized experiment in 2024.

Key Question: Would the Negative Ad Against Trump Remain Ineffective in 2024?



Ideal approach: running a randomized experiment in 2024.

However, ad campaigns are expensive:

- [Aggarwal et al. 2023]'s experiment costed $8.9 million USD
- A recent super political action committee for the Democratic Party has already spent $450 million USD for the 2024 election [Schleifer and Goldmacher 2024]

Carefully Learning from the Past: Transfer Learning with Sensitivity Analysis

Our approach: **transfer learning with sensitivity analysis**.

- It is cheaper and faster

We use transfer learning (domain adaptation/distributional shift) to learn from past ad campaigns in order to inform ad campaigns in 2024.

We use sensitivity analysis to quantify our belief that 2020 and 2024 are not the same and provide a **plausible range of ad effects in 2024** based on how similar/different 2020 and 2024 are.

Some Prior Works

The closest related literature is transportability/generalizability.

We are definitely not the first nor only ones to incorporate sensitivity analysis in transportability/generalizability. A very small, partial list:

- Linear outcome sensitivity model: [Nguyen et al. 2017; Dahabreh, Petito, et al. 2020; Dahabreh, Robins, S. J.-P. Haneuse, et al. 2023; Zeng et al. 2023]
- Exponential tilting sensitivity model: [Dahabreh, Robins, S. J. Haneuse, Robertson, et al. 2022]
- Marginal sensitivity model for domain adaptation: [Nie, Imbens, and Wager 2021]
- Omitted variable bias approach with weighted estimators: [M. Y. Huang 2024]

Our goal is to tailor these methods to address our key question.

Our Setup

- We treat the 2020 RCT from [Aggarwal et al. 2023] as the source population and the 2024 PA voters as the target population
  - Source: ∼2 million registered voters in five states in 2020
  - Target: ∼4.9 million registered PA voters in 2024

Our Setup

- We treat the 2020 RCT from [Aggarwal et al. 2023] as the source population and the 2024 PA voters as the target population
  - Source: ∼2 million registered voters in five states in 2020
  - Target: ∼4.9 million registered PA voters in 2024
- We allow voter demographics to shift between 2020 and 2024
- We allow voter turnout from ad to change between 2020 and 2024:

$$\begin{aligned} &\mathrm{pr}(\text{voted if given ad} \mid \text{target (i.e. } \textbf{2024 PA voters}\text{), voter demographics}) \\ \neq &\mathrm{pr}(\text{voted if given ad} \mid \text{source (i.e. } \textbf{2020 RCT}\text{), voter demographics}) \end{aligned}$$

Our Setup

- We treat the 2020 RCT from [Aggarwal et al. 2023] as the source population and the 2024 PA voters as the target population
  - Source: $\sim$2 million registered voters in five states in 2020
  - Target: $\sim$4.9 million registered PA voters in 2024
- We allow voter demographics to shift between 2020 and 2024
- We allow voter turnout from ad to change between 2020 and 2024:

$$\begin{aligned} &\text{pr}(\text{voted if given ad} \mid \text{target (i.e. } \textbf{2024 PA voters}\text{), voter demographics}) \\ \neq &\text{pr}(\text{voted if given ad} \mid \text{source (i.e. } \textbf{2020 RCT}\text{), voter demographics}) \end{aligned}$$

- Source and target did not measure the same covariates.
  - Example: voters' race was measured in the 2020 RCT but it was not measured in 2024
  - Generally, RCTs have detailed covariates, which leads to source covariates $\supseteq$ target covariates
  - Most works assume source covariates $=$ target covariates
- Our data structure is closest to [Zeng et al. 2023]
- Goal: study the ad effect on voter turnout in 2024 PA voters

Our Methodology

- Use sensitivity analysis [Robins, Rotnitzky, and Scharfstein 2000] to quantify unmeasured differences between 2020 and 2024
- Use regression (most popular & familiar tool for analysis) and (theoretically correct) bootstrap, tailored for transfer learning
- Estimator based on the efficient influence function

Our Methodology

- Use sensitivity analysis [Robins, Rotnitzky, and Scharfstein 2000] to quantify unmeasured differences between 2020 and 2024
- Use regression (most popular & familiar tool for analysis) and (theoretically correct) bootstrap, tailored for transfer learning
- Estimator based on the efficient influence function
- Calibrate sensitivity parameters with a simple, **dissimilar partition** of source population
  - Use a biased sampling to restrict sensitivity parameters into a plausible range

Notation and Causal Assumptions

Notation and Causal Assumptions

- Population type: $S \in \{0, 1\}$ where $S = 1$ is source (e.g. 2020) and $S = 0$ is target (e.g. 2024 PA voters)
- Outcome: $Y \in \{0, 1\}$ where $Y = 1$ is voted
- Treatment: $A \in \{0, 1\}$ where $A = 1$ is ad against Trump
- Covariates: $\mathbf{X} \in \mathbb{R}^p$ where
  - Source covariates: $\mathbf{X}$
  - Target covariates: $\mathbf{V} \subset \mathbf{X}$
    $\mathbf{V} =$ (age group, gender, party, part of voting history)
    $\mathbf{X} = \mathbf{V}$ + (race, missing voting history)
- Potential outcomes: $Y(a) \in \{0, 1\}$, $a \in \{0, 1\}$
  - $Y(1)$: voted if, potentially contrary to fact, voter received ads
  - $Y(0)$: voted if, potentially contrary to fact, voter didn't receive ads
- Causal estimand: $\theta = \mathbb{E}[Y(1) - Y(0) \mid S = 0]$
  - Average treatment effect; the ad effect on turnout among 2024 PA voters

Data Table for Our Setup

		$\mathbf{X}$					
	S	$\mathbf{V}$	$\mathbf{X} \setminus \mathbf{V}$	A	$Y(1)$	$Y(0)$	Y
Source RCT (i.e. 2020)	1	$\checkmark$	$\checkmark$	1	$\checkmark$		$\checkmark$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	1	$\checkmark$	$\checkmark$	1	$\checkmark$		$\checkmark$
	1	$\checkmark$	$\checkmark$	0		$\checkmark$	$\checkmark$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	1	$\checkmark$	$\checkmark$	0		$\checkmark$	$\checkmark$
Target (i.e. 2024 PA voters)	0	$\checkmark$					
	$\vdots$	$\vdots$					
	0	$\checkmark$					

The goal is to identify and estimate $\theta = \mathbb{E}[Y(1) - Y(0) \mid S = 0]$.

Review Of Transportability In RCT: Assumptions On The Source ($S = 1$)

Causal assumptions on the source (under stratified RCT):

(A1) Stable unit treatment value assumption (SUTVA):
Under $S = 1$, if $A = a$, $Y = Y(a)$

- The observed Y is one realization of the two potential outcomes
- There are no multiple versions of treatment, e.g.,

$$Y(700 \text{ ads}) = Y(800 \text{ ads}) = Y(1)$$

- There is no interference: a voter's voting result cannot be affected by other voters' treatments

Review Of Transportability In RCT: Assumptions On The Source ($S = 1$)

Causal assumptions on the source (under stratified RCT):

(A1) Stable unit treatment value assumption (SUTVA):
Under $S = 1$, if $A = a$, $Y = Y(a)$

- The observed Y is one realization of the two potential outcomes
- There are no multiple versions of treatment, e.g.,

$$Y(700 \text{ ads}) = Y(800 \text{ ads}) = Y(1)$$

- There is no interference: a voter's voting result cannot be affected by other voters' treatments

(A2) Randomized treatment: $A \perp Y(1), Y(0) \mid \mathbf{X}, S = 1$

Review Of Transportability In RCT: Assumptions On The Source ($S = 1$)

Causal assumptions on the source (under stratified RCT):

(A1) Stable unit treatment value assumption (SUTVA):
Under $S = 1$, if $A = a$, $Y = Y(a)$

- The observed Y is one realization of the two potential outcomes
- There are no multiple versions of treatment, e.g.,

$$Y(700 \text{ ads}) = Y(800 \text{ ads}) = Y(1)$$

- There is no interference: a voter's voting result cannot be affected by other voters' treatments

(A2) Randomized treatment: $A \perp Y(1), Y(0) \mid \mathbf{X}, S = 1$

(A3) Overlap of treatment: $0 < \pi(\mathbf{x}) = \mathrm{pr}(A = 1 \mid \mathbf{X} = \mathbf{x}, S = 1) < 1$

- The treated individuals and untreated individuals have a common support of $\mathbf{X}$
- The propensity score $\pi(\mathbf{x})$ is known in an RCT

Review Of Transportability In RCT: Assumptions On The Source ($S = 1$)

Causal assumptions on the source (under stratified RCT):

(A1) Stable unit treatment value assumption (SUTVA):
Under $S = 1$, if $A = a$, $Y = Y(a)$

- The observed Y is one realization of the two potential outcomes
- There are no multiple versions of treatment, e.g.,

$$Y(700 \text{ ads}) = Y(800 \text{ ads}) = Y(1)$$

- There is no interference: a voter's voting result cannot be affected by other voters' treatments

(A2) Randomized treatment: $A \perp Y(1), Y(0) \mid \mathbf{X}, S = 1$

(A3) Overlap of treatment: $0 < \pi(\mathbf{x}) = \text{pr}(A = 1 \mid \mathbf{X} = \mathbf{x}, S = 1) < 1$

- The treated individuals and untreated individuals have a common support of $\mathbf{X}$
- The propensity score $\pi(\mathbf{x})$ is known in an RCT

(A2)-(A3) are satisfied with the 2020 RCT by [Aggarwal et al. 2023].

Review Of Transportability In RCT: Assumptions For Transportation

(A4) Overlap between source and target: $0 < \text{pr}(S = 1 \mid \mathbf{V} = \mathbf{v}) < 1$ for all $\mathbf{v}$

- The source voters and target voters have a common support of $\mathbf{V}$
- It can be empirically verified

Review Of Transportability In RCT: Assumptions For Transportation

(A4) Overlap between source and target: $0 < \text{pr}(S = 1 \mid \mathbf{V} = \mathbf{v}) < 1$ for all $\mathbf{v}$

- The source voters and target voters have a common support of $\mathbf{V}$
- It can be empirically verified

(A5) Transportability: $Y(1), Y(0) \perp S \mid \mathbf{V}$

- $\underbrace{\text{pr}(Y(a) = 1 \mid \mathbf{V}, S = 0)}_{\text{2024 PA voters}} = \underbrace{\text{pr}(Y(a) = 1 \mid \mathbf{V}, S = 1)}_{\text{2020 RCT}}$
- It is often assumed, but cannot be verified
- We assume it for this short review, but will relax it shortly

See [Tipton and Peck 2017; Dahabreh, Robins, S. J. Haneuse, and Hernán 2019; Egami and Hartman 2023; Degtiar and Rose 2023] for details.

Review: Identification Under (A1)-(A5)

Let $\mu_a = \mathbb{E}[Y \mid \mathbf{X}, A = a, S = 1]$. Under (A1)-(A5), the target ATE θ is identified:

$$\begin{aligned}\theta &= \mathbb{E}[Y(1) - Y(0) \mid S = 0] \\ &= \underbrace{\mathbb{E}[\underbrace{\mathbb{E}[\underbrace{\mu_1(\mathbf{X}) - \mu_0(\mathbf{X})}_{\text{CATE}(\mathbf{X}) \text{ in source}} \mid \mathbf{V}, S = 1]}_{\text{CATE}(\mathbf{V}) \text{ in source}} \mid S = 0]}_{\text{Reweigh CATE}(\mathbf{V}) \text{ to target}}\end{aligned}$$

For reference, when $\mathbf{X} = \mathbf{V}$, we have

$$\theta = \mathbb{E}[\mathbb{E}[\mu_1(\mathbf{X}) - \mu_0(\mathbf{X}) \mid S = 0].$$

In other words, when source and target covariates differ, we have more nuisance parameters (i.e. projection of CATE($\mathbf{X}$) onto $\mathbf{V}$); see recent work by [Zeng et al. 2023].

What If Transportability (A5) Fails? Sensitivity Analysis

Suppose transportability (A5) does not hold:

$$\underbrace{\mathrm{pr}(Y(a) \mid \mathbf{V}, S=0)}_{\text{2024 (i.e., target)}} \neq \underbrace{\mathrm{pr}(Y(a) \mid \mathbf{V}, S=1)}_{\text{2020 (i.e., source)}}$$

For each $Y(a)$, we measure the deviation between the two probabilities via γ_a:

$$\exp(\gamma_a) = \frac{\mathrm{Odd}(Y(a) \mid \mathbf{v}, S=0)}{\mathrm{Odd}(Y(a) \mid \mathbf{v}, S=1)}, \ \gamma_a \in (-\infty, \infty) \tag{1}$$

$$\mathrm{Odd}(Y(a) \mid \mathbf{v}, s) = \frac{\mathrm{pr}(Y(a)=1 \mid \mathbf{V}=\mathbf{v}, S=s)}{1-\mathrm{pr}(Y(a)=1 \mid \mathbf{V}=\mathbf{v}, S=s)}, \ s \in \{0,1\} \tag{2}$$

What If Transportability (A5) Fails? Sensitivity Analysis

Suppose transportability (A5) does not hold:

$$\underbrace{\text{pr}(Y(a) \mid \mathbf{V}, S=0)}_{\text{2024 (i.e., target)}} \neq \underbrace{\text{pr}(Y(a) \mid \mathbf{V}, S=1)}_{\text{2020 (i.e., source)}}$$

For each $Y(a)$, we measure the deviation between the two probabilities via γ_a:

$$\exp(\gamma_a) = \frac{\text{Odd}(Y(a) \mid \mathbf{v}, S=0)}{\text{Odd}(Y(a) \mid \mathbf{v}, S=1)}, \; \gamma_a \in (-\infty, \infty) \tag{1}$$

$$\text{Odd}(Y(a) \mid \mathbf{v}, s) = \frac{\text{pr}(Y(a)=1 \mid \mathbf{V}=\mathbf{v}, S=s)}{1-\text{pr}(Y(a)=1 \mid \mathbf{V}=\mathbf{v}, S=s)}, \; s \in \{0,1\} \tag{2}$$

Broadly speaking, the unobservable part ($S=0$; i.e., 2024) differs from the observable part ($S=1$; i.e., 2020) by γ_a.

What If Transportability (A5) Fails? Sensitivity Analysis

Suppose transportability (A5) does not hold:

$$\underbrace{\mathrm{pr}\big(Y(a) \mid \mathbf{V}, S=0\big)}_{\text{2024 (i.e., target)}} \neq \underbrace{\mathrm{pr}\big(Y(a) \mid \mathbf{V}, S=1\big)}_{\text{2020 (i.e., source)}}$$

For each $Y(a)$, we measure the deviation between the two probabilities via γ_a:

$$\exp(\gamma_a) = \frac{\mathrm{Odd}\big(Y(a) \mid \mathbf{v}, S=0\big)}{\mathrm{Odd}\big(Y(a) \mid \mathbf{v}, S=1\big)}, \ \gamma_a \in (-\infty, \infty) \qquad (1)$$

$$\mathrm{Odd}\big(Y(a) \mid \mathbf{v}, s\big) = \frac{\mathrm{pr}\big(Y(a)=1 \mid \mathbf{V}=\mathbf{v}, S=s\big)}{1-\mathrm{pr}\big(Y(a)=1 \mid \mathbf{V}=\mathbf{v}, S=s\big)}, \ s \in \{0,1\} \qquad (2)$$

Broadly speaking, the unobservable part ($S=0$; i.e., 2024) differs from the observable part ($S=1$; i.e., 2020) by γ_a.

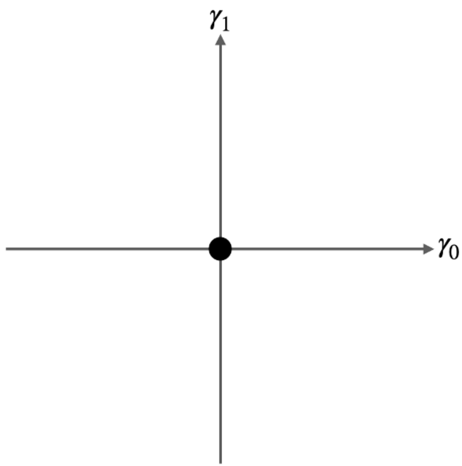
- $\gamma_a = 0$ means $\mathrm{pr}\big(Y(a) \mid \mathbf{V}, S=0\big) = \mathrm{pr}\big(Y(a) \mid \mathbf{V}, S=1\big)$; i.e., transportability
- In general, a large $|\gamma_a| \Rightarrow$ large difference between 2020 and 2024
- Since the red part is unobserved, γ_a cannot be estimated. Instead, γ_a is a sensitivity parameter chosen to quantify the difference between the target and the source

Remarks

Interpreting Sensitivity Parameter γ_a

- (γ_0, γ_1) quantifies the shift in $(Y(0), Y(1))$ between 2024 and 2020 via an odds ratio model

$$\exp(\gamma_a) = \frac{\text{Odd}(Y(a) \mid \mathbf{v}, S = 0)}{\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)}, \ \gamma_a \in (-\infty, \infty).$$



γ_0	γ_1
0	0

Transported ATE (Ad Effect 2024)

$\theta(\gamma_0, \gamma_1)$

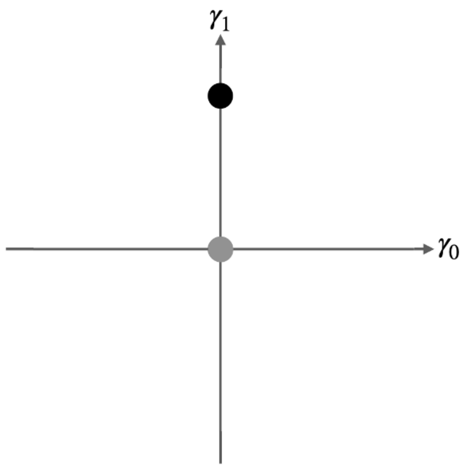
$2024 \approx 2020$ • 0.38

0 (%)

Interpreting Sensitivity Parameter γ_a

- (γ_0, γ_1) quantifies the shift in $(Y(0), Y(1))$ between 2024 and 2020 via an odds ratio model

$$\exp(\gamma_a) = \frac{\text{Odd}(Y(a) \mid \mathbf{v}, S = 0)}{\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)}, \quad \gamma_a \in (-\infty, \infty).$$



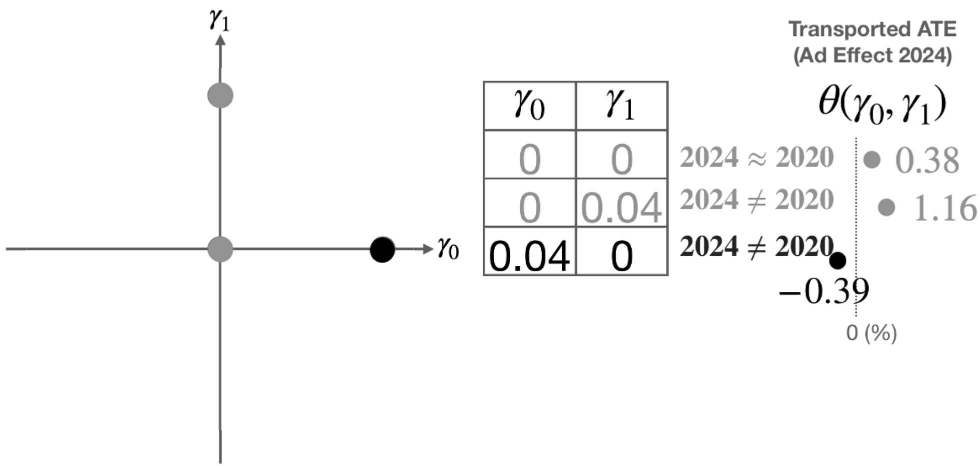
γ_0	γ_1		Transported ATE (Ad Effect 2024) $\theta(\gamma_0, \gamma_1)$
0	0	$2024 \approx 2020$	0.38
0	0.04	$2024 \neq 2020$	1.16

0 (%)

Interpreting Sensitivity Parameter γ_a

- (γ_0, γ_1) quantifies the shift in $(Y(0), Y(1))$ between 2024 and 2020 via an odds ratio model

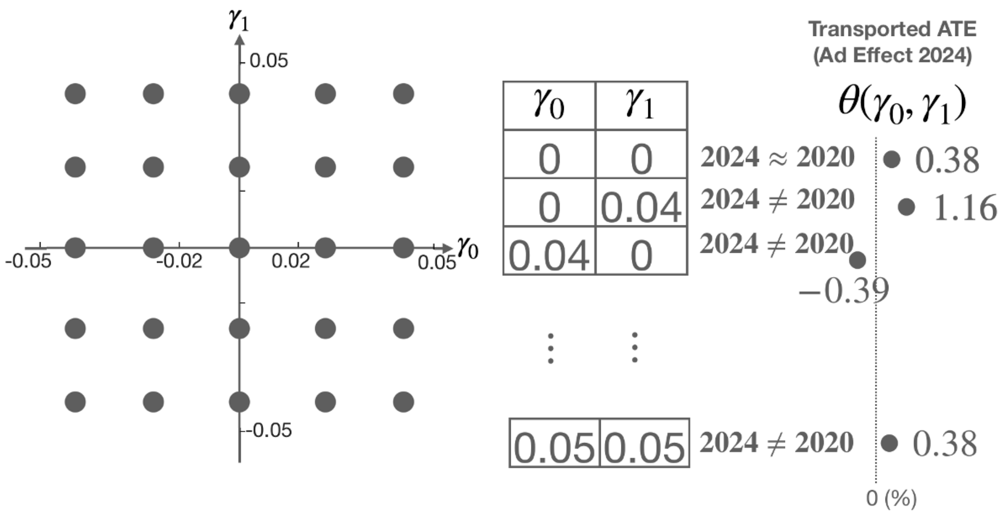
$$\exp(\gamma_a) = \frac{\text{Odd}(Y(a) \mid \mathbf{v}, S = 0)}{\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)}, \; \gamma_a \in (-\infty, \infty).$$



Interpreting Sensitivity Parameter γ_a

- (γ_0, γ_1) quantifies the shift in $(Y(0), Y(1))$ between 2024 and 2020 via an odds ratio model

$$\exp(\gamma_a) = \frac{\text{Odd}(Y(a) \mid \mathbf{v}, S = 0)}{\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)}, \quad \gamma_a \in (-\infty, \infty).$$



Identification Under (A1)-(A4) + Sensitivity Model

Again, let $\mu_a(\mathbf{X}) = \mathrm{E}(Y \mid \mathbf{X}, A = a, S = 1)$. Under (A1)-(A4) and the sensitivity model, we have

$$\mathrm{E}[Y(a) \mid S = 0] = \mathrm{E}\left[\frac{\mathrm{E}\{\exp(\gamma_a)\mu_a(\mathbf{X}) \mid \mathbf{V}, S = 1\}}{\mathrm{E}\{\exp(\gamma_a)\mu_a(\mathbf{X}) + 1 - \mu_a(\mathbf{X}) \mid \mathbf{V}, S = 1\}}\middle| S = 0\right].$$

- It is an exponential tilt of $\rho_a(\mathbf{V}) = \mathbb{E}[\mu_a(\mathbf{X}) \mid \mathbf{V}, S = 1]$
- If $\gamma_a = 0$ (i.e. transportability (A5) holds), we return to the previous result:

$$\mathbb{E}[Y(a) \mid S = 0] = \mathbb{E}[\mathbb{E}[\mu_a(\mathbf{X}) \mid \mathbf{V}, S = 1] \mid S = 0]$$

Estimation: Simple, Plug-In Estimator

Identification leads to a simple, plug-in estimator:

1. Estimate $\rho_a(\mathbf{V}) = \mathbb{E}[\mu_a(\mathbf{X}) \mid \mathbf{V}, S = 1]$ from source data
   - An example: for $a = 1$, regress $Y/\hat{\pi}(\mathbf{X})$ on $\mathbf{V}$ to get $\hat{\rho}_1(\mathbf{V})$

```
rho1 <- lm(Y ~ AgeGroup + Gender + Party,
           data = source |> filter(a == 1), weights = 1 / pi)

rho0 <- lm(Y ~ AgeGroup + Gender + Party,
           data = source |> filter(a == 0), weights = 1 / (1 - pi))
```

   - Because source is an RCT, $\rho_a(\mathbf{V})$ can be consistently estimated

Estimation: Simple, Plug-In Estimator

Identification leads to a simple, plug-in estimator:

1. Estimate $\rho_a(\mathbf{V}) = \mathbb{E}[\mu_a(\mathbf{X}) \mid \mathbf{V}, S = 1]$ from source data
   - An example: for $a = 1$, regress $Y/\hat{\pi}(\mathbf{X})$ on $\mathbf{V}$ to get $\hat{\rho}_1(\mathbf{V})$

```
rho1 <- lm(Y ~ AgeGroup + Gender + Party,
           data = source |> filter(a == 1), weights = 1 / pi)

rho0 <- lm(Y ~ AgeGroup + Gender + Party,
           data = source |> filter(a == 0), weights = 1 / (1 - pi))
```

   - Because source is an RCT, $\rho_a(\mathbf{V})$ can be consistently estimated
2. Average the exponentially tilted $\rho_a(\mathbf{V})$ among target sample

$$\widehat{\mathbb{E}}[Y(a) \mid S = 0] = \frac{1}{n_t} \sum_{i \in \text{Target}} \frac{\exp(\gamma_a)\widehat{\rho}_a(\mathbf{V}_i)}{\exp(\gamma_a)\hat{\rho}_a(\mathbf{V}_i) + 1 - \hat{\rho}_a(\mathbf{V}_i)}$$

$$\widehat{\theta}(\gamma_1, \gamma_0) = \widehat{\mathbb{E}}[Y(1) \mid S = 0] - \widehat{\mathbb{E}}[Y(0) \mid S = 0]$$

```
rho1pred <- predict(rho1, data = target)
rho0pred <- predict(rho0, data = target)
theta <- exp(gamma1) * rho1pred / ((exp(gamma1) * rho1pred + 1 - rho1pred) -
         exp(gamma0) * rho0pred / ((exp(gamma0) * rho0pred + 1 - rho0pred)
```

Bootstrapping for Transfer Learning

In general, bootstrapping is easy and convenient for estimating SEs or confidence intervals (CIs).
Here, we lay out one (theoretically valid) bootstrap for transfer learning with sensitivity analysis.

In general, bootstrapping is easy and convenient for estimating SEs or confidence intervals (CIs).

Here, we lay out one (theoretically valid) bootstrap for transfer learning with sensitivity analysis.

In each bootstrap iteration $b \in \{1, \cdots, B\}$:

1. Resample source data with replacement of size n_s, obtain data $\mathcal{D}_{\mathcal{S}}^*$
2. Resample target data with replacement of size n_t, obtain data $\mathcal{D}_{\mathcal{T}}^*$.
3. With $\mathcal{D}_{\mathcal{S}}^*$ and $\mathcal{D}_{\mathcal{T}}^*$, construct the ATE estimator $\widehat{\theta}_b^*$ from above.

Take $\alpha/2$ and $1-\alpha/2$ quantiles of $\left\{\widehat{\theta}_b^*\right\}_{b=1}^B$ as a $1-\alpha$ CI of θ.

In general, bootstrapping is easy and convenient for estimating SEs or confidence intervals (CIs).
Here, we lay out one (theoretically valid) bootstrap for transfer learning with sensitivity analysis.
In each bootstrap iteration $b \in \{1, \cdots, B\}$:

1. Resample source data with replacement of size n_s, obtain data $\mathcal{D}_{\mathcal{S}}^*$
2. Resample target data with replacement of size n_t, obtain data $\mathcal{D}_{\mathcal{T}}^*$.
3. With $\mathcal{D}_{\mathcal{S}}^*$ and $\mathcal{D}_{\mathcal{T}}^*$, construct the ATE estimator $\widehat{\theta}_b^*$ from above.

Take $\alpha/2$ and $1 - \alpha/2$ quantiles of $\left\{\widehat{\theta}_b^*\right\}_{b=1}^B$ as a $1 - \alpha$ CI of θ.
Theorem: If $\rho_a(\mathbf{V})$ is smooth enough and Donsker condition holds, the above procedure yields a valid $1 - \alpha$ CI of θ.
The smoothness + Donsker conditions hold for our discrete voter data.

Estimator Based on Efficient Influence Function

For given γ_a, we have the efficient influence function (EIF) of θ and it directs an EIF-based estimator.

- Four nuisance functions: (i) propensity score $\pi(\mathbf{X})$, (ii) outcome regression $\mu_a(\mathbf{X})$, (iii) projection of outcome regression $\rho_a(\mathbf{V})$, and (iv) population weights $w(\mathbf{V}) = \mathrm{pr}(\mathbf{V} \mid S = 0)/\mathrm{pr}(\mathbf{V} \mid S = 1)$
- The efficient influence function (EIF) is

$$\begin{aligned}\mathrm{EIF}(\mathbf{O}_i, \theta_a(\gamma_a)) =& \frac{S_i w(\mathbf{V}_i)}{\mathrm{pr}(S_i = 1)} \frac{\exp(\gamma_a)}{[\exp(\gamma_a)\rho_a(\mathbf{V}_i) + 1 - \rho_a(\mathbf{V}_i)]^2} \left[\left\{ \frac{A_i}{\pi(\mathbf{X}_i)} + \frac{1 - A_i}{1 - \pi(\mathbf{X}_i)} \right\} \{Y_i - \mu_a(\mathbf{X}_i)\} \right. \\ &\left. + \mu_a(\mathbf{X}_i) - \rho_a(\mathbf{V}_i) \right] + \frac{1 - S_i}{\mathrm{pr}(S_i = 0)} \left[\frac{\exp(\gamma_a)\rho_a(\mathbf{V}_i)}{\exp(\gamma_a)\rho_a(\mathbf{V}) + 1 - \rho_a(\mathbf{V}_i)} - \theta_a(\gamma_a) \right]\end{aligned}$$

- Our EIF recovers [Zeng et al. 2023]'s EIF when $\gamma_a = 0$:

$$\mathrm{EIF}(\mathbf{O}_i, \theta_a(\gamma_a)) = \frac{S_i w(\mathbf{V}_i)}{\mathrm{pr}(S_i = 1)} \left[\left\{ \frac{A_i}{\pi(\mathbf{X}_i)} + \frac{1 - A_i}{1 - \pi(\mathbf{X}_i)} \right\} \{Y_i - \mu_a(\mathbf{X}_i)\} + \mu_a(\mathbf{X}_i) - \rho_a(\mathbf{V}_i) \right] + \frac{1 - S_i}{\mathrm{pr}(S_i = 0)} \rho_a(\mathbf{V}_i)$$

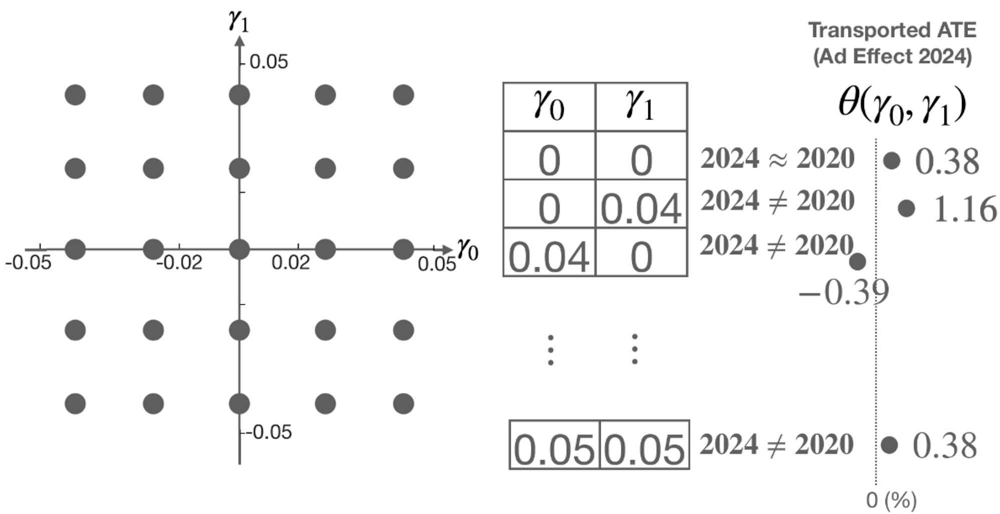
Estimator Based on Efficient Influence Function

Some remarks:

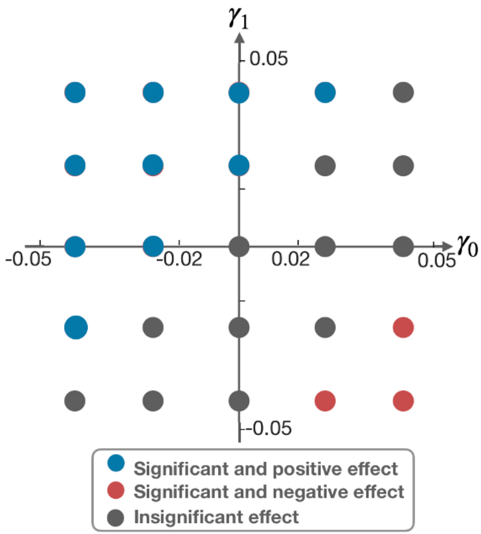
- Practically speaking, the estimator is useful if $\mathbf{V}$ is continuous
- To avoid parametric assumptions, we need cross-fitting
- The estimator is not doubly robust when $\gamma_a \neq 0$; see Appendix D of [Dahabreh, Robins, S. J. Haneuse, Robertson, et al. 2022] for a related comment when $\mathbf{V} = \mathbf{X}$
- EIF-based estimator does not reduce the "plug-in bias" from $\rho_a(\mathbf{V})$ when $\gamma_a \neq 0$

Details are deferred to the Appendix.

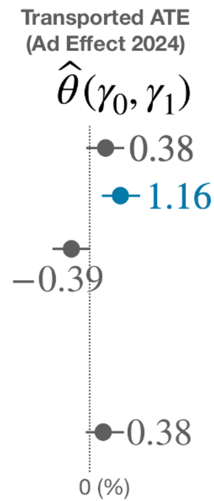
Making Inferences under the Sensitivity Model



- Calculate p-values using the proposed plug-in estimator with bootstrap

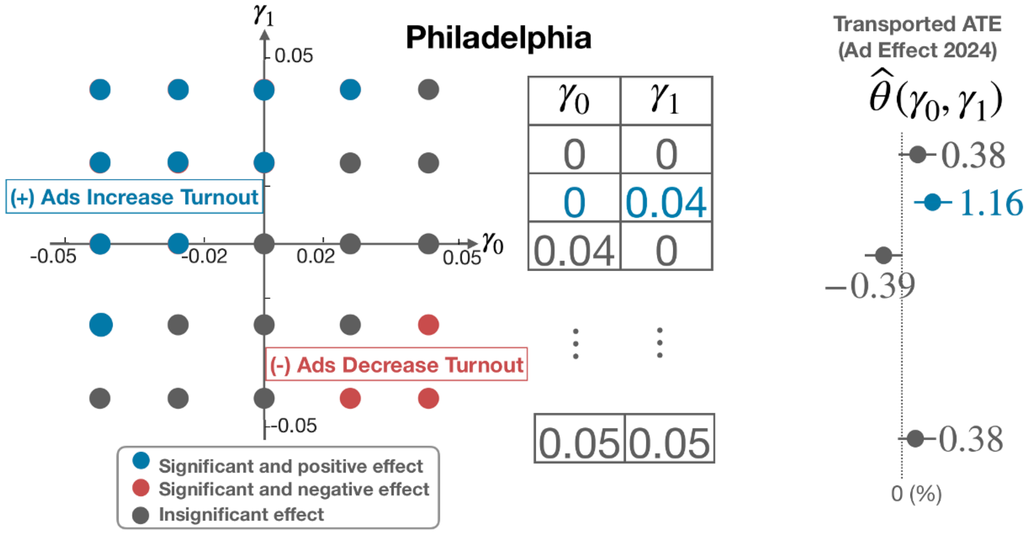


γ_0	γ_1
0	0
0	0.04
0.04	0
⋮	⋮
0.05	0.05



Making Inferences under the Sensitivity Model

- Calculate p-values using the proposed plug-in estimator with bootstrap



Calibrating Sensitivity Parameters

- In sensitivity analysis, there is always a question about what is a "large", "small", or a "plausible" sensitivity parameter γ_a
- Another common way is to omit an observed covariate (e.g., [Hsu and Small 2013; Cinelli and Hazlett 2020; Ek and Zachariah 2023; M. Y. Huang 2024]).
  - It treats the omitted variable as the confounder and quantify the induced unmeasured difference.
  - This can lead to a misleading understanding of the magnitude of unmeasured confounding (Section 6.2 of [Cinelli and Hazlett 2020])
  - We avoid this issue by using the same covariates $\mathbf{V}$ for sensitivity analysis and calibration
- We present one solution to this question by creating **dis-similar** partitions of the source data[1]

[1] See (M. Huang 2024) for a related idea to assess overlap of S (i.e. our (A4))

A Three-Step Calibration Procedure

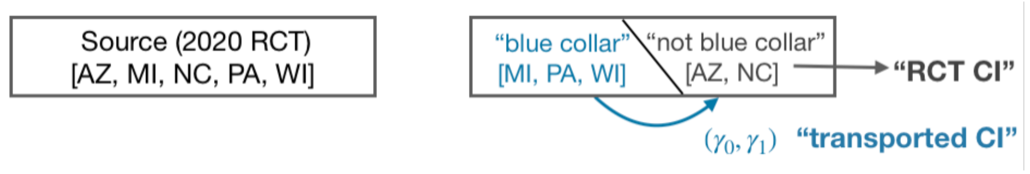
1. Partition source into blue-collar states and non-blue-collar states (i.e., two "dissimilar" partitions that are not similar w.r.t. ad effect)
    - It is meant to create unmeasurable differences between the two partitions

A Three-Step Calibration Procedure

1. Partition source into blue-collar states and non-blue-collar states (i.e., two "dissimilar" partitions that are not similar w.r.t. ad effect)
   - It is meant to create unmeasurable differences between the two partitions
2. Treat one partition as "proxy source", the other partition as "proxy target"

A Three-Step Calibration Procedure

1. Partition source into blue-collar states and non-blue-collar states (i.e., two "dissimilar" partitions that are not similar w.r.t. ad effect)
   - It is meant to create unmeasurable differences between the two partitions
2. Treat one partition as "proxy source", the other partition as "proxy target"
   a. Use RCT to estimate ATE and its CI in proxy target partition (i.e. **"RCT CI"**)
   b. For each (γ_0, γ_1), use our method to estimate ATE and its CI in target partition (i.e. **"transported CI"**)



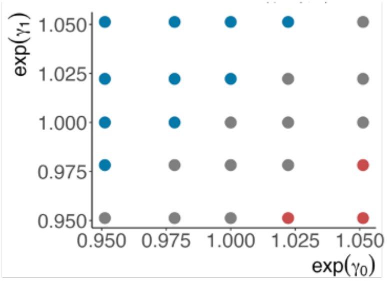
A Three-Step Calibration Procedure

1. Partition source into blue-collar states and non-blue-collar states (i.e., two "dissimilar" partitions that are not similar w.r.t. ad effect)
   - It is meant to create unmeasurable differences between the two partitions
2. Treat one partition as "proxy source", the other partition as "proxy target"
   a Use RCT to estimate ATE and its CI in proxy target partition (i.e. **"RCT CI"**)
   b For each (γ_0, γ_1), use our method to estimate ATE and its CI in target partition (i.e. **"transported CI"**)

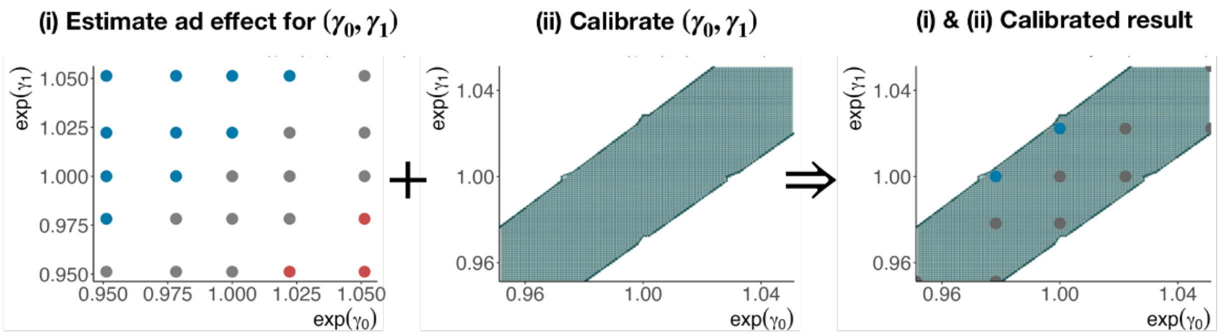
3. Plausible range $\mathcal{C}_1$ = subset of (γ_0, γ_1) where CIs from both overlap
   - $(\gamma_0, \gamma_1) \in \mathcal{C}_1$ represent plausible, unmeasured differences between the two partitions since they correctly transport from the proxy source to match the ATE of the proxy target, up to sampling errors
   - We switch roles of proxy source and proxy target and obtain $\mathcal{C}_2$. The final plausible range is $\mathcal{C} = \mathcal{C}_1 \cap \mathcal{C}_2$

Example: Calibrating Sensitivity Parameters in Philadelphia

(i) Estimate ad effect for (γ_0, γ_1)

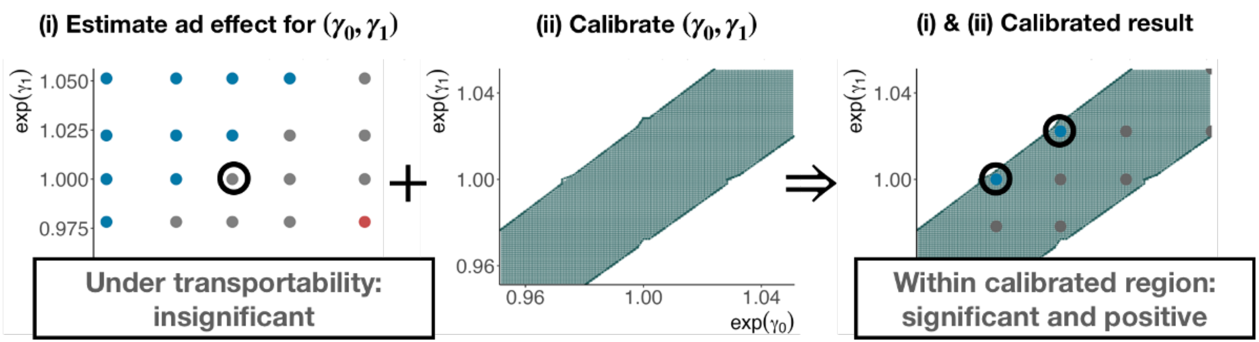


Example: Calibrating Sensitivity Parameters in Philadelphia



We focus on the change of conclusions before and after sensitivity analysis:

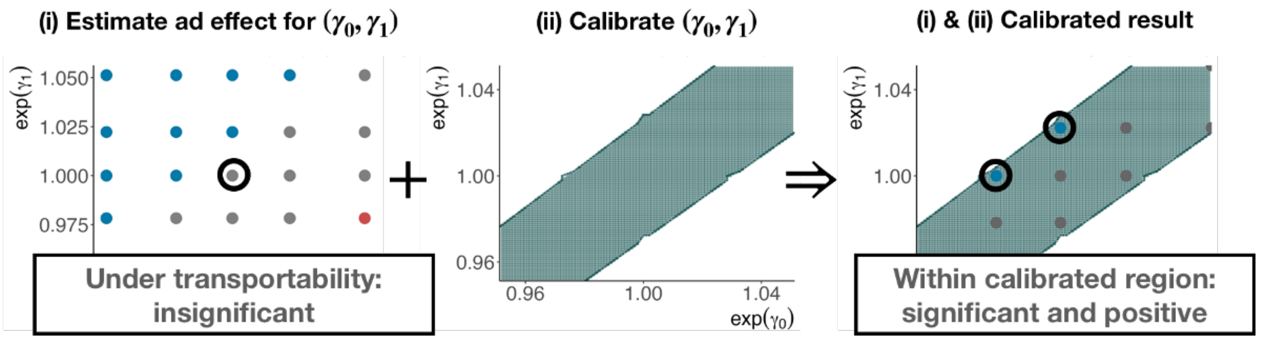
Philadelphia is *sensitive* to a significant and positive ad effect.



Example: Calibrating Sensitivity Parameters in Philadelphia

We focus on the change of conclusions before and after sensitivity analysis:

Philadelphia is *sensitive* to a significant and positive ad effect.



The calibrated result depends on the way of partitioning.

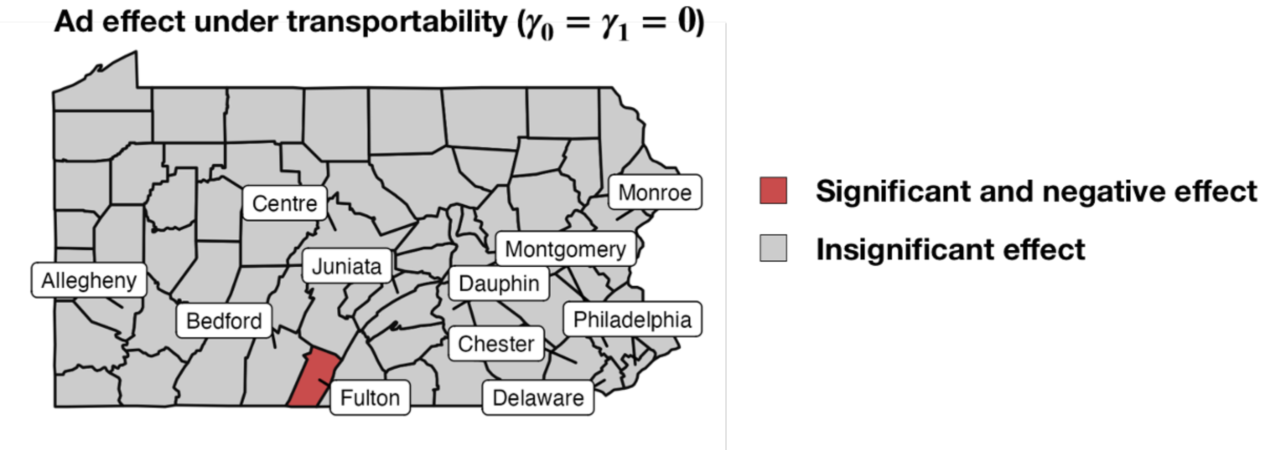
- Another example: using race to partition the source
- Proxy source: white; proxy target: non-white

FAQs

Data Analysis: Registered Voters in Pennsylvania (PA)

- Target population: 4,880,730 registered voters (as of April 15, 2024) from 67 counties in PA
- **V**: age group, gender, party, an incomplete voting history
- **X** : **V** + race and missing voting history
- We performed a county-by-county analysis and a subgroup analysis for groups defined through gender, urbanicity, and education attainment in the neighborhood
- For sensitivity parameters, we use $-0.05 \leq \gamma_a \leq 0.05$ (odd ratio: $0.951 \leq \exp(\gamma_a) \leq 1.051$)
- For inference, we use the simple plug-in estimator with bootstrap

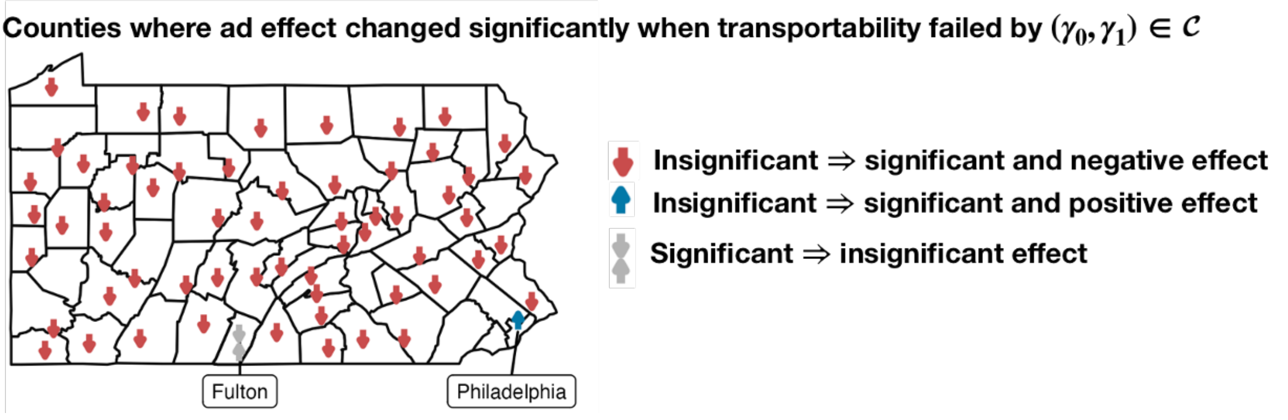
- When transportability holds ($\gamma_1 = \gamma_0 = 0$), the digital ads against Trump decreases turnout in Fulton county
  - Trump had 86.03% of votes in Fulton in 2024 (highest margin among counties in PA)



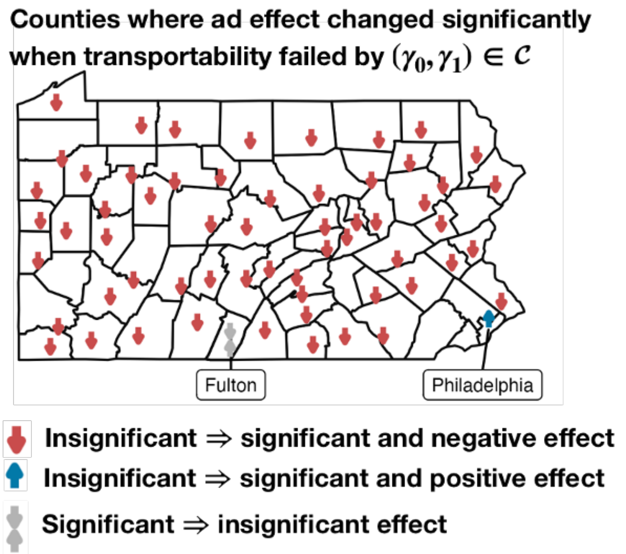
Data Analysis: Ad Effect Across Counties

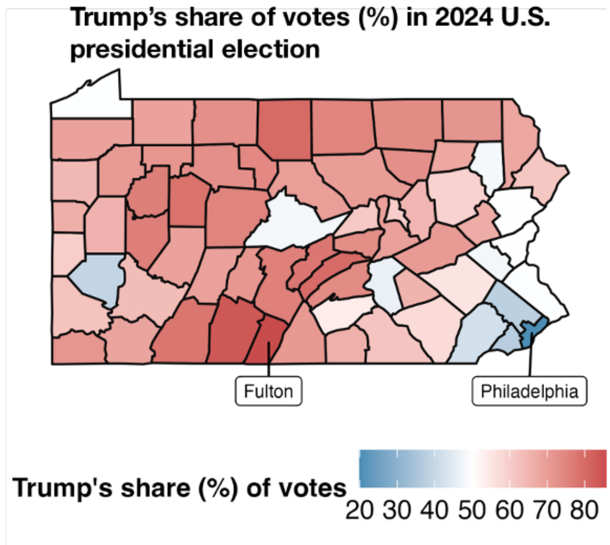
When $\gamma_0\gamma_1 \neq 0$, we applied the calibrated sensitivity analysis.

- Some conclusions may change within the calibrated region $\mathcal{C}$ (i.e., *sensitive* effects)
  - Philadelphia county is sensitive to a significant and positive effect
  - 59 counties is sensitive to a significant and negative effect
- 6 counties remain to have an insignificant effect (i.e., *insensitive* effects)



Data Analysis: Ad Effect Across Counties

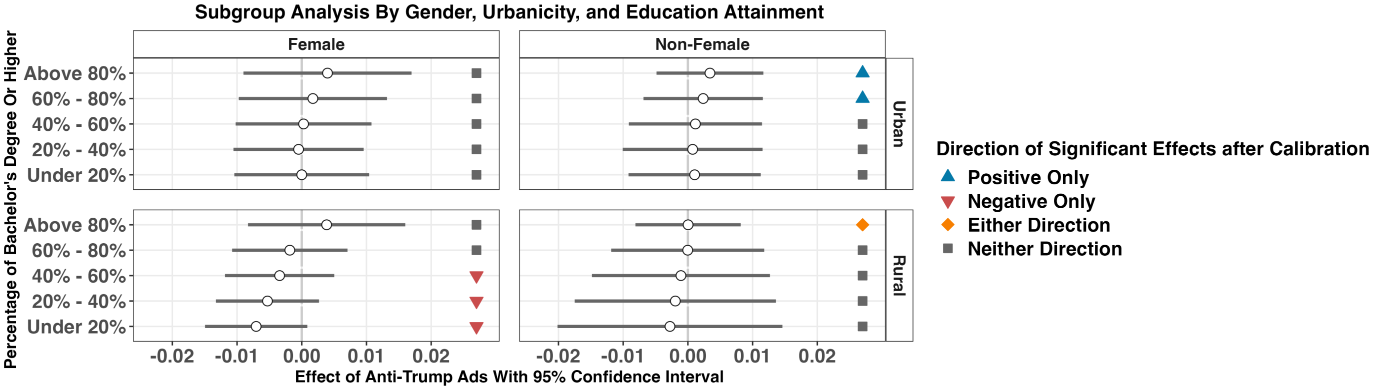




The direction of the ad effect after calibration roughly corresponds to Trump's share of votes in the 2024 U.S. presidential election.

Data Analysis: Ad Effect In Subgroups

- We consider 20 subgroups defined by a three-way interaction between
  - gender (female & male)
  - urbanicity of the census tract (rural & urban, obtained from U.S. Census)
  - % of Bachelor's degree or higher in the same zip-code area (five levels, obtained from U.S. Census)



Summary

- Motivation: From [Aggarwal et al. 2023], would ads against Trump in 2020 remain ineffective in 2024?
- Our approach: transfer learning with sensitivity analysis
  - Setup: (a) source is from RCT, (b) $\mathbf{V} \neq \mathbf{X}$,
  - Analysis: (a) simple plug-in estimator with bootstrap SE/CIs, (b) EIF-based approach, (c) calibration of sensitivity parameter γ_a with source data
- Application: would ads against Trump be effective for PA voters in 2024?

  County-by-county analysis:
  - Under transportability, the ad effect is significant and negative in Fulton county
  - With sensitivity analysis, after calibration, the direction of significant ad effects roughly corresponds to Trump's share of votes in the 2020

  Subgroup analysis:
  - Under transportability, ad effects are insignificant
  - Ads can decrease voter turnout among female voters living in rural areas with low college education and increase voter turnout among non-female voters living in urban areas with high college education

- Thank you for listening. Comments are highly appreciated!
  - arXiv:2411.01100
  - xinran.miao@wisc.edu
- Acknowledgements

  - UW-Madison: Xiaobin Zhou, Ang Yu, Adeline Y. Lo, Xindi Lin, Jingqi Duan, Sameer Deshpande, Elaine Chiu, and statistics student seminar participants on April 29, 2024
  - Chan Park (UIUC), Melody Huang (Yale), Ying Jin (Harvard), OCIS seminar participants on April 30, 2024, ACIC participants on May 15, 2024

FAQs About Data

- Is the voting data self-reported?
  No. See [Aggarwal et al. 2023] for details.
- Does the data contain which candidate the voter voted for?
  No.
- Why is your voter data discrete?
  We're not sure. Perhaps, this is done to preserve some privacy?
- Is party registration measured accurately?
  Yes and no. [Aggarwal et al. 2023] and current voter registration data documentation discuss some reasons for errors.
- How was the treatment randomized?
  The randomization was stratified within gender, race, and age groups with the intention of increasing the propensity for women, black, and young people. The average treatment probability was 85.6%.
- Why is there a high proportion of treated individuals?
  We're not sure. Perhaps Acronym wanted to deliver the ads against Trump to as many voters as possible?

FAQs About Data

- Was the randomization done through Facebook/Meta?
  No. Our understanding from [Aggarwal et al. 2023] is that the participants were randomized before the advertising company delivered the ads.
- Were ads delivered?
  Yes and no. 60% of the treatment group participants were identified and served ads. The analysis was intention-to-treat. See [Aggarwal et al. 2023] for details. We followed [Aggarwal et al. 2023] and considered intention-to-treat effect.

◂ Back to Background

FAQs About Framework And Assumptions

- Is SUTVA violated?
  Great question! It is possible, especially if
  - Different doses of ads: $Y(700 \text{ ads}) \neq Y(800 \text{ ads}) \neq Y(1)$
  - Different ads in 2020 and 2024: $Y(\text{ad in } 2020) \neq Y(\text{ad in } 2024)$
  - Voters talk to each other due to ads: $Y(\text{my trt}, \text{your trt}) \neq Y(\text{my trt})$
  - There are carry-over effects from 2020 ad campaign into 2024

  But, we also picked 2020 and 2024 to minimize SUTVA violations as Trump remains the Republican candidates between the two years.
- Is your population infinite?
  Excellent question! Our framework implicitly assumes that the units in the target and the source data are sampled from an infinite population of voters. But, it may be more appropriate to treat the PA voter registration database as a finite population or a large sample from a finite population. See [Jin and Rothenhäusler 2024] for transfer learning when the target sample is fixed. We are also happy to discuss more.

FAQs About Framework And Assumptions

- Is the data from 2020 independent from 2024?
  Excellent question! Theoretically, this concern is a bit tricky to address, especially if 2024 is a fixed, census-level data. We're happy to talk more about this.
  Application-wise, we repeated the analysis after excluding voters in PA, WI, MI from the source data. Results exhibit similar patterns but are less powerful (see Appendix of the paper). Although this ensure independence between the source and target, it also increases the difference between the source and target. We believe that in general the source and target should be as similar as possible, so we decided to report the results when using all data in [Aggarwal et al. 2023] as the source.
- What about treating this data as longitudinal?
  Excellent idea, but this requires measuring same voter over time. This data is not easy to get.

◂ Back to Framework

FAQs About Sensitivity Analysis

- Can you give us some understanding on the odds ratio?
  Yes. The turnout in PA is 76.5% in 2020 and 76.6% in 2024. The odds ratio is roughly 1.0055 (logarithm: 0.0055).
- Can the sensitivity model depend on covariates?
  Yes. While this introduces more complexity in interpreting the sensitivity parameters, it could be useful if there is a priori knowledge about how the ad effects in 2024 and 2020 differ with respect to measured covariates. For example, for $A = 0$, we can define

$$\exp(\gamma_{\text{Democrat?}} I(\text{Democrat?})), \quad \exp(\gamma_{\text{Republican?}} I(\text{Republican?}))$$

  if we believe the change between 2020 and 2024 is different for Democrats and Republicans. We call this a **local sensitivity model.**

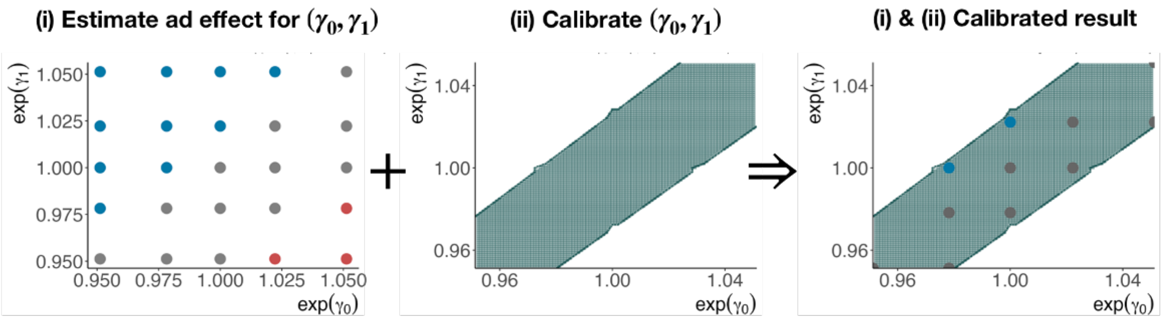
◂ Back to Sensitivity Analysis

FAQs About Analysis

- Is your conclusion valid after Biden dropped out?
  Great question! Our original analysis plan assumed that President Biden is the Democratic Party's nominee for the presidency. While we believe the interpretations from our analysis about ads will still be plausible since Trump is the nominee for the Republication party and the digital ad campaign consists of negative ads against Trump, we caution readers from over-interpreting the results.
  Notably, our calibration procedure based on blue-collar and non blue-collar states could under-estimate the dramatic shift in electoral context after Biden dropped out of the race and the consequences of this unprecedented event in American politics.
- Is your conclusion sensitive to data quality from 2024 voter registration data (i.e. target data)?
  Yes. Unfortunately, high quality target data is expensive.

◂ Back to Summary

- Can you explain the shape of your calibrated region?
  Sure. The off-diagonal area represents case when $|\gamma_1 - \gamma_0|$ is large, i.e., when the change in treated voters differs a lot from the change in control voters from 2020 to 2024.
  - It can be unrealistic when $\gamma_1 > 0$ is very high and $\gamma_0 < 0$ is very low
  - The calibration rules out such extreme scenarios

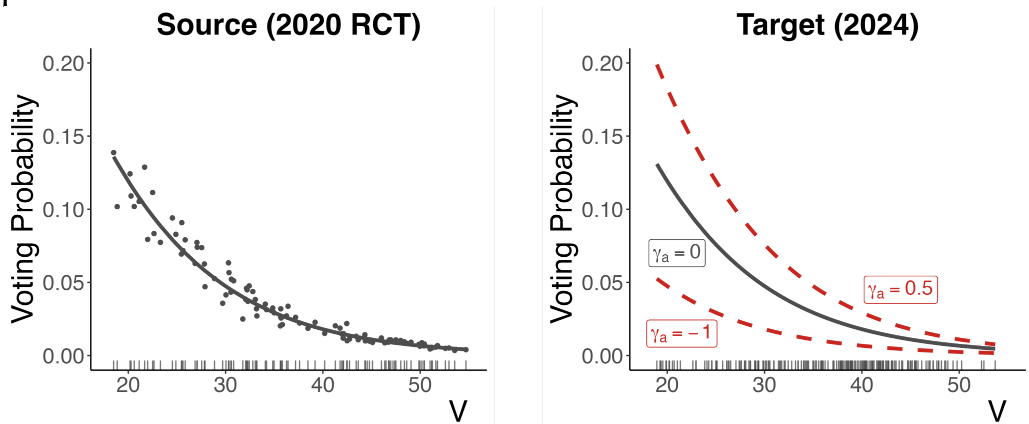


Back to Calibration

Interpreting Sensitivity Parameter γ_a

$$\exp(\gamma_a) = \frac{\text{Odd}(Y(a) \mid \mathbf{v}, S = 0)}{\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)}$$

- When $\gamma_a = 0$, transportability (A5) holds
- Positive $\gamma_1 \Rightarrow$ more turnout in 2024 after receiving ads against Trump compared to that in 2020
- Negative $\gamma_1 \Rightarrow$ less turnout in 2024 after receiving ads against Trump compared to that in 2020



Toy example: $\text{pr}(Y(a) = 1 \mid V, S = 1) = \text{expit}(-0.1V)$. Back to Sensitivity

Some Remarks On The Sensitivity Model

$$\exp(\gamma_a) = \frac{\text{Odd}(Y(a) \mid \mathbf{v}, S=0)}{\text{Odd}(Y(a) \mid \mathbf{v}, S=1)}$$

Some remarks:

- The sensitivity model does not place any observable restrictions on the data [Robins, Rotnitzky, and Scharfstein 2000; Franks, D'Amour, and Feller 2020]
- A pseudo-R^2 version of γ_a is in Proposition 3 of [Franks, D'Amour, and Feller 2020].
- We can also reparametrize the sensitivity model in terms of $P(S=1 \mid Y(1), Y(0), \mathbf{V}=\mathbf{v})$; see Appendix and [Carroll et al. 1997].
- The sensitivity model can depend on covariates (e.g. $\exp(\gamma_{\mathbf{v}}^\intercal \mathbf{v} + \ldots)$; "local" sensitivity analysis)
- Some works that use this model: [Robins, Rotnitzky, and Scharfstein 2000; Franks, D'Amour, and Feller 2020; Scharfstein et al. 2021; Dahabreh, Robins, S. J. Haneuse, Robertson, et al. 2022]
- There is a **long and healthy** discussion about what constitutes a "good" model for sensitivity analysis [Robins 2002; Rosenbaum 2002]

Back to Sensitivity Analysis

Alternate Parametrization of The Sensitivity Model

- The sensitivity model $\exp(\gamma_a) = \dfrac{\text{Odd}(Y(a) \mid \mathbf{v}, S=0)}{\text{Odd}(Y(a) \mid \mathbf{v}, S=1)}$ implies the following partially linear logistic regression model [Carroll et al. 1997]:

$$\text{pr}(S=1 \mid Y(a)=y, \mathbf{V}=\mathbf{v}) = \text{expit}\left(-\gamma_a y - \eta_a(\mathbf{v})\right)$$
$$\eta_a(\mathbf{v}) = \log\left(\frac{\text{pr}(S=0)}{\text{pr}(S=1)} \frac{w(\mathbf{v})}{\mathrm{E}\{\exp(\gamma_a Y(a)) \mid \mathbf{v}, S=1\}}\right)$$
$$w(\mathbf{V}) = p(\mathbf{V} \mid S=0)/p(\mathbf{V} \mid S=1)$$

- The joint sensitivity model implies the following partially linear logistic regression model:

$$\text{pr}(S=1 \mid Y(1)=y_1, Y(0)=y_0, \mathbf{V}=\mathbf{v}) = \text{expit}\left(-\gamma_1 y_1 - \gamma_0 y_0 - \eta(\mathbf{v})\right)$$
$$\eta(\mathbf{v}) = \log\left(\frac{\text{pr}(S=0)}{\text{pr}(S=1)} \frac{w(\mathbf{v})}{\mathrm{E}\{\exp(\gamma_1 Y(1) + \gamma_0 Y(0)) \mid \mathbf{V}, S=1\}}\right)$$

Back to Sensitivity

EIF-Based Cross-Fitting Estimator

We follow the modern trend in causal inference where we use cross-fitting [Chernozhukov et al. 2017; Kennedy 2022] and the EIF to estimate $\theta_a(\gamma_a)$.

1. Randomly partition the source and target sample indices $\mathcal{I}_s$ and $\mathcal{I}_t$ into K disjoint sets, $\mathcal{I}_s^{(k)}$ and $\mathcal{I}_t^{(k)}$, respectively, for $k = 1, 2, \cdots, K$ with some pre-specified integer K.
2. For each k, estimate nuisance functions with data in $\mathcal{I} \setminus \mathcal{I}^{(k)}$ and denote them as $\widehat{\pi}^{(k)}(\mathbf{x})$, $\widehat{\mu}_a^{(k)}(\mathbf{x})$, $\widehat{w}^{(k)}(\mathbf{v})$ and $\widehat{\rho}_a^{(k)}(\mathbf{v})$. Plug them into the "uncentered" EIF and evaluate it with the data in $\mathcal{I}^{(k)}$, i.e.,

$$\widehat{\theta}^{(k)}_{\text{EIF,a}}(\gamma_a) = \frac{1}{|\mathcal{I}_s^{(k)}|} \sum_{i \in \mathcal{I}_s^{(k)}} \frac{\exp(\gamma_a)\widehat{w}^{(k)}(\mathbf{V}_i)}{[\exp(\gamma_a)\widehat{\rho}_a^{(k)}(\mathbf{V}_i) + 1 - \widehat{\rho}_a^{(k)}(\mathbf{V}_i)]^2} \left[\left\{ \frac{A_i}{\widehat{\pi}^{(k)}(\mathbf{X}_i)} + \frac{1 - A_i}{1 - \widehat{\pi}^{(k)}(\mathbf{X}_i)} \right\} \{Y_i - \widehat{\mu}_a^{(k)}(\mathbf{X}_i)\} + \widehat{\mu}_a^{(k)}(\mathbf{X}_i) - \widehat{\rho}_a^{(k)}(\mathbf{V}_i) \right] + \frac{1}{|\mathcal{I}_t^{(k)}|} \sum_{i \in \mathcal{I}_t^{(k)}} \frac{\exp(\gamma_a)\widehat{\rho}_a^{(k)}(\mathbf{V}_i)}{\exp(\gamma_a)\widehat{\rho}_a^{(k)}(\mathbf{V}_i) + 1 - \widehat{\rho}_a^{(k)}(\mathbf{V}_i)}.$$

3. Take an average of $\widehat{\theta}^{(k)}_{\text{EIF,a}}(\gamma_a)$ to arrive at the EIF-based cross-fitting estimator of $\theta_{\text{EIF,a}}(\gamma_a)$, which we denote as $\widehat{\theta}_{\text{EIF,a}}(\gamma_a) = \sum_{k=1}^K \widehat{\theta}^{(k)}_{\text{EIF,a}}(\gamma_a)/K$.

◂ Back to EIF

EIF-Based Estimation

Theorem 1

Suppose regularity assumptions hold. Then we have
(i) [Conditional Double Robustness]. Suppose $\widehat{\rho}_a^{(k)}$ is a consistent estimator of $\rho_a^{(k)}$. Then, $\widehat{\theta}_{\mathrm{EIF,a}}(\gamma_a) \to_p \theta_a(\gamma_a)$ if

$$\|\widehat{\pi}^{(k)}(\mathbf{X}_i) - \pi^{(k)}(\mathbf{X}_i)\| \cdot \|\widehat{\mu}_a^{(k)}(\mathbf{X}_i) - \mu_a^{(k)}(\mathbf{X}_i)\| = o_p(1), \tag{3}$$

(ii) [Asymptotic normality and Semiparametric Efficiency] Suppose $\widehat{\rho}_a^{(k)}$, $\widehat{\mu}_a^{(k)}$, $\widehat{w}^{(k)}$, and $\widehat{\pi}^{(k)}$ are consistent estimators with the following rates:

$$\|\widehat{\pi}^{(k)}(\mathbf{X}_i) - \pi^{(k)}(\mathbf{X}_i)\| \cdot \|\widehat{\mu}_a^{(k)}(\mathbf{X}_i) - \mu_a^{(k)}(\mathbf{X}_i)\| = o_p(n^{-1/2}), \tag{4a}$$

$$\|\widehat{w}^{(k)}(\mathbf{V}_i) - w^{(k)}(\mathbf{V}_i)\| \cdot \|\widehat{\rho}_a^{(k)}(\mathbf{V}_i) - \rho_a^{(k)}(\mathbf{V}_i)\| = o_p(n^{-1/2}), \text{ and} \tag{4b}$$

$$\|\widehat{\rho}_a^{(k)}(\mathbf{V}_i) - \rho_a^{(k)}(\mathbf{V}_i)\|^2 = o_p(n^{-1/2}). \tag{4c}$$

Then, $\sqrt{n}\left\{\widehat{\theta}_{\mathrm{EIF,a}}(\gamma_a) - \theta_a(\gamma_a)\right\} \to_d N\left(0, \sigma^2_{\mathrm{EIF,a}}(\gamma_a)\right)$ where $\sigma^2_{\mathrm{EIF,a}}(\gamma_a) = \mathbb{E}[\{\mathrm{EIF}(\mathbf{O}_i, \theta_a(\gamma_a))\}^2]$.

Theorem 2 (EIF-Based Estimation, Continued)

(iii) [Consistent Estimator of Standard Error] Suppose the same assumptions in (ii) hold. Then, $\widehat{\sigma}^2_{\mathrm{EIF,a}}(\gamma_a) \rightarrow_p \sigma^2_{\mathrm{EIF,a}}(\gamma_a)$*, where*
$\widehat{\sigma}^2_{\mathrm{EIF,a}}(\gamma_a) = K^{-1} \sum_{k=1}^{K} \frac{1}{\left|\mathcal{I}^{(k)}\right|} \sum_{i \in \mathcal{I}^{(k)}} \left\{ \widehat{\mathrm{EIF}}^{(k)}(\mathbf{O}_i, \widehat{\theta}_{\mathrm{EIF,a}}(\gamma_a)) \right\}^2$ *and*
$\widehat{\mathrm{EIF}}^{(k)}(\mathbf{O}_i, \widehat{\theta}_{\mathrm{EIF,a}}(\gamma_a))$ *is the empirical counterpart of* $\mathrm{EIF}^{(k)}(\mathbf{O}_i, \widehat{\theta}_{\mathrm{EIF,a}}(\gamma_a))$ *with plug-in estimates of the nuisance parameters* $\widehat{\pi}^{(k)}$, $\widehat{\rho}_a^{(k)}$, $\widehat{w}^{(k)}$, *and* $\widehat{\mu}_a^{(k)}$.

Back to EIF

Voter Demographics Between Source and Target Population

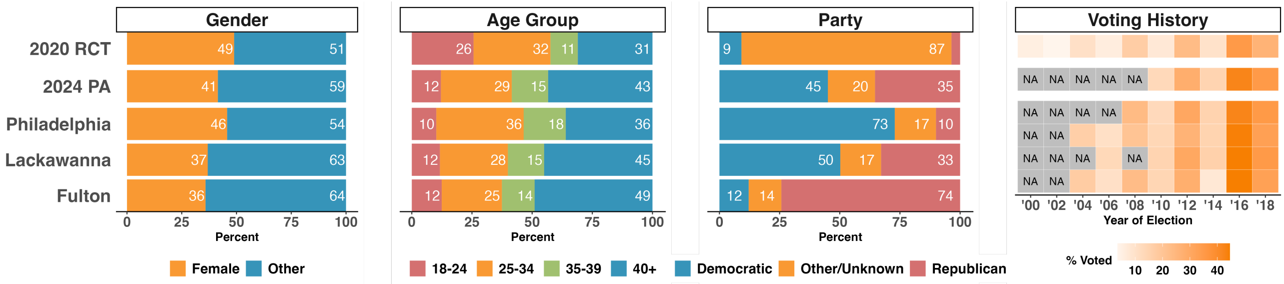


Figure. Registered voter demographics.

Back to Data Application

References I

- Aggarwal, Minali et al. (2023). **"A 2 million-person, campaign-wide field experiment shows how digital advertising affects voter turnout".** In: *Nature Human Behaviour* 7.3, pp. 332–341.
- Carroll, Raymond J et al. (1997). **"Generalized partially linear single-index models".** In: *Journal of the American Statistical Association* 92.438, pp. 477–489.
- Chernozhukov, Victor et al. (2017). **"Double/Debiased/Neyman machine learning of treatment effects".** In: *American Economic Review* 107.5, pp. 261–265.
- Cinelli, Carlos and Chad Hazlett (2020). **"Making sense of sensitivity: Extending omitted variable bias".** In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 82.1, pp. 39–67.
- Dahabreh, Issa J, Lucia C Petito, et al. (2020). **"Toward causally interpretable meta-analysis: transporting inferences from multiple randomized trials to a new target population".** In: *Epidemiology* 31.3, pp. 334–344.
- Dahabreh, Issa J, James M Robins, Sebastien J-PA Haneuse, et al. (2023). **"Sensitivity analysis using bias functions for studies extending inferences from a randomized trial to a target population".** In: *Statistics in Medicine* 42.13, pp. 2029–2043.
- Dahabreh, Issa J, James M Robins, Sebastien JP Haneuse, and Miguel A Hernán (2019). **"Generalizing causal inferences from randomized trials: counterfactual and graphical identification".** In: *arXiv preprint arXiv:1906.10792*.

References II

- Dahabreh, Issa J, James M Robins, Sebastien JP Haneuse, Sarah E Robertson, et al. (2022). **"Global sensitivity analysis for studies extending inferences from a randomized trial to a target population".** In: *arXiv preprint arXiv:2207.09982.*
- Degtiar, Irina and Sherri Rose (2023). **"A review of generalizability and transportability".** In: *Annual Review of Statistics and Its Application* 10, pp. 501–524.
- Egami, Naoki and Erin Hartman (2023). **"Elements of external validity: Framework, design, and analysis".** In: *American Political Science Review* 117.3, pp. 1070–1088.
- Ek, Sofia and Dave Zachariah (2023). **"Externally Valid Policy Evaluation Combining Trial and Observational Data".** In: *arXiv preprint arXiv:2310.14763.*
- Franks, Alexander M, Alexander D'Amour, and Avi Feller (2020). **"Flexible sensitivity analysis for observational studies without observable implications".** In: *Journal of the American Statistical Association* 115.532, pp. 1730–1746.
- Hsu, Jesse Y and Dylan S Small (2013). **"Calibrating sensitivity analyses to observed covariates in observational studies".** In: *Biometrics* 69.4, pp. 803–811.
- Huang, Melody (2024). **"Overlap violations in external validity".** In: *arXiv preprint arXiv:2403.19504.*
- Huang, Melody Y (2024). **"Sensitivity analysis for the generalization of experimental results".** In: *Journal of the Royal Statistical Society Series A: Statistics in Society.*

References III

- Jin, Ying and Dominik Rothenhäusler (2024). **“Tailored inference for finite populations: conditional validity and transfer across distributions”.** In: *Biometrika* 111.1, pp. 215–233.
- Kennedy, Edward H (2022). **“Semiparametric doubly robust targeted double machine learning: a review”.** In: *arXiv preprint arXiv:2203.06469*.
- Nguyen, Trang Quynh et al. (2017). **“Sensitivity analysis for an unobserved moderator in RCT-to-target-population generalization of treatment effects”.** In: *The Annals of Applied Statistics* 11.1, pp. 225–247.
- Nie, Xinkun, Guido Imbens, and Stefan Wager (2021). **“Covariate balancing sensitivity analysis for extrapolating randomized trials across locations”.** In: *arXiv preprint arXiv:2112.04723*.
- Robins, James M (2002). **“[Covariance Adjustment in Randomized Experiments and Observational Studies]: Comment”.** In: *Statistical Science* 17.3, pp. 309–321.
- Robins, James M, Andrea Rotnitzky, and Daniel O Scharfstein (2000). **“Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models”.** In: *Statistical Models in Epidemiology, the Environment, and Clinical Trials*. Springer, pp. 1–94.
- Rosenbaum, Paul R (2002). **“Covariance adjustment in randomized experiments and observational studies”.** In: *Statistical Science* 17.3, pp. 286–327.
- Scharfstein, Daniel O et al. (2021). **“Semiparametric sensitivity analysis: Unmeasured confounding in observational studies”.** In: *arXiv preprint arXiv:2104.08300*.

References IV

Schleifer, Theodore and Shane Goldmacher (Oct. 2024). ***Inside the Secretive $700 Million Ad-Testing Factory for Kamala Harris.*** `https://www.nytimes.com/2024/10/17/us/elections/future-forward-kamala-harris-ads.html`. Accessed: 2024-10-19.

Tipton, Elizabeth and Laura R Peck (2017). **"A design-based approach to improve external validity in welfare policy evaluations".** In: *Evaluation Review* 41.4, pp. 326–356.

Zeng, Zhenghao et al. (2023). **"Efficient Generalization and Transportation".** In: *arXiv preprint arXiv:2302.00092*.