Transfer Learning Between U.S. Presidential Elections

Xinran Miao, Jiwei Zhao, Hyunseung Kang

Department of Statistics
University of Wisconsin - Madison

Statistics Student Seminar
April 29, 2024

Disclaimers

None of us are experts in U.S. elections.

But, 2020 and 2024 U.S. presidential elections present a **very unique** opportunity to study transportability/generalizability in U.S. elections.

Comments/Critiques are highly appreciated!

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

- Would online ads against Donald Trump affect voter turnout in five battleground states: AZ, MI, NC, PA, and WI?

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

- Would online ads against Donald Trump affect voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- Ran a stratified randomized experiment in February 2020 to November 2020.

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

- Would online ads against Donald Trump affect voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- Ran a stratified randomized experiment in February 2020 to November 2020.
  - Nearly two million registered voters $(1,999,282)$.

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

- Would online ads against Donald Trump affect voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- Ran a stratified randomized experiment in February 2020 to November 2020.
  - Nearly two million registered voters $(1,999,282)$.
  - Stratified by gender, race, and age groups (**all discrete**).

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

- Would online ads against Donald Trump affect voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- Ran a stratified randomized experiment in February 2020 to November 2020.
  - Nearly two million registered voters ($1,999,282$).
  - Stratified by gender, race, and age groups (**all discrete**).
- Treatment group (85.6% of participants): an average of 754 ads against Trump on social media by Acronym, a left-leaning organization.

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

- Would online ads against Donald Trump affect voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- Ran a stratified randomized experiment in February 2020 to November 2020.
  - Nearly two million registered voters ($1,999,282$).
  - Stratified by gender, race, and age groups (**all discrete**).
- Treatment group (85.6% of participants): an average of 754 ads against Trump on social media by Acronym, a left-leaning organization.
- Control group: no ads from Acronym.

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

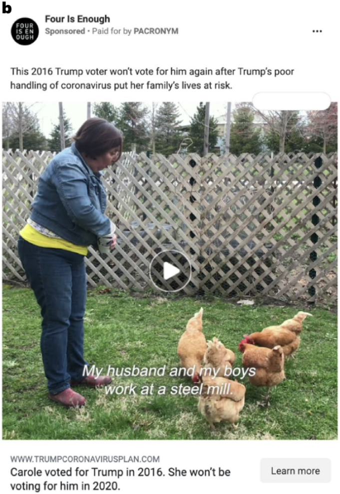
- Would online ads against Donald Trump affect voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- Ran a stratified randomized experiment in February 2020 to November 2020.
  - Nearly two million registered voters ($1,999,282$).
  - Stratified by gender, race, and age groups (**all discrete**).
- Treatment group (85.6% of participants): an average of 754 ads against Trump on social media by Acronym, a left-leaning organization.
- Control group: no ads from Acronym.
- Outcome: voted in 2020 U.S. election? (**binary**)

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

- Would online ads against Donald Trump affect voter turnout in five battleground states: AZ, MI, NC, PA, and WI?
- Ran a stratified randomized experiment in February 2020 to November 2020.
  - Nearly two million registered voters ($1,999,282$).
  - Stratified by gender, race, and age groups (**all discrete**).
- Treatment group (85.6% of participants): an average of 754 ads against Trump on social media by Acronym, a left-leaning organization.
- Control group: no ads from Acronym.
- Outcome: voted in 2020 U.S. election? (**binary**)
- Analysis tool: linear regression.

Example Political Ads in Treatment Group

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



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

Results and Takeaways from Aggarwal et al. (2023)

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

- Negative effect: ad against Trump decreased voter turnout.
  - For example, a Trump supporter may choose not to vote after seeing ads against Trump.
- Positive effect: ad against Trump increased voter turnout.
  - For example, a Biden supporter may be encouraged to vote by ads against Trump.

Results and Takeaways from Aggarwal et al. (2023)

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

- Negative effect: ad against Trump decreased voter turnout.
  - For example, a Trump supporter may choose not to vote after seeing ads against Trump.
- Positive effect: ad against Trump increased voter turnout.
  - For example, a Biden supporter may be encouraged to vote by ads against Trump.

Overall ATE on turnout: -0.06% (SE $= 0.12\%$).

Results and Takeaways from Aggarwal et al. (2023)

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

- Negative effect: ad against Trump decreased voter turnout.
  - For example, a Trump supporter may choose not to vote after seeing ads against Trump.
- Positive effect: ad against Trump increased voter turnout.
  - For example, a Biden supporter may be encouraged to vote by ads against Trump.

Overall ATE on turnout: -0.06% ($\text{SE} = 0.12\%$).

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

Results and Takeaways from Aggarwal et al. (2023)

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

- Negative effect: ad against Trump decreased voter turnout.
  - For example, a Trump supporter may choose not to vote after seeing ads against Trump.
- Positive effect: ad against Trump increased voter turnout.
  - For example, a Biden supporter may be encouraged to vote by ads against Trump.

Overall ATE on turnout: -0.06% ($\text{SE} = 0.12\%$).

- *"...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?

2024 election provides a **unique** opportunity to study this question.

Some similarities between 2020 and 2024:

- Same presumptive candidates from major political parties (Biden, Trump)[1]. Both candidates are seeking a second term[2].
- Nearly similar treatment/ad campaigns (i.e. ad against Trump).
- Recent polls suggest economy is still a major concern for voters.

[1] Last time this occurred was in 1956.
[2] Last time this occurred was in 1892.

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

2024 election provides a **unique** opportunity to study this question.

Some similarities between 2020 and 2024:

- Same presumptive candidates from major political parties (Biden, Trump)[1]. Both candidates are seeking a second term[2].
- Nearly similar treatment/ad campaigns (i.e. ad against Trump).
- Recent polls suggest economy is still a major concern for voters.

Some notable differences:

- 2020: COVID-19, death of George Floyd and racial unrest, etc.
- 2024: abortion, immigration, Ukraine/Israel, etc.

[1] Last time this occurred was in 1956.
[2] Last time this occurred was in 1892.

Our Contribution: Application-Driven Setup & Methods

- "Design" elements:
  - Allows for covariate shift.
  - Does not assume same covariates between 2020 and 2024.
  - Does not assume the transportability.

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

- "Analysis" elements:
  - Simple, design-inspired estimator.
  - One (theoretically) correct approach to bootstrapping in transfer learning.
  - Efficient influence function based estimator.
  - "Demystifying" sensitivity analysis with source data calibration.

Outline

- Notations and review
- Setup: Transfer learning with sensitivity analysis
- Estimators of the target ATE
- Preliminary data analysis on Pennsylvania

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).
- Outcome: $Y \in \{0, 1\}$ where $Y = 1$ is voted.
- Treatment: $A \in \{0, 1\}$ where $A = 1$ is ad against Trump.
- Covariates: $\mathbf{X}$ where
  - Source covariates: $\mathbf{X}$,
  - Target covariates: $\mathbf{V} \subset \mathbf{X}$.
- Potential outcomes: $Y(a) \in \{0, 1\}$, $a \in \{0, 1\}$.
  - $Y(1)$: voted if, contrary to fact, voter got negative ad.
  - $Y(0)$: voted if, contrary to fact, voter did not get negative ad.

Data Table for Our Setup

	S	$\mathbf{V}$	$\mathbf{X} \setminus \mathbf{V}$	A	$Y(1)$	$Y(0)$	Y
		$\mathbf{X}$					
Source RCT (n_s)	1	✓	✓	1	✓		✓
	⋮	⋮	⋮	⋮	⋮		⋮
	1	✓	✓	1	✓		✓
	1	✓	✓	0		✓	✓
	⋮	⋮	⋮	⋮		⋮	⋮
	1	✓	✓	0		✓	✓
Target (n_t)	0	✓					
	⋮	⋮					
	0	✓					

The goal is to identify and estimate the average treatment effect (ATE) on the target,

$$\theta = \mathbb{E}[Y(1) - Y(0) \mid S = 0].$$

Review of Transportability in an RCT: Assumptions on the Source

Causal assumptions on the source (under stratified RCT):

(A1) SUTVA: Under $S = 1$, if $A = a$, $Y = Y(a)$.

Review of Transportability in an RCT: Assumptions on the Source

Causal assumptions on the source (under stratified RCT):

(A1) SUTVA: Under $S = 1$, if $A = a$, $Y = Y(a)$.

- The observed Y is one realization of the two potential outcomes.

Review of Transportability in an RCT: Assumptions on the Source

Causal assumptions on the source (under stratified RCT):

(A1) 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.,

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

Review of Transportability in an RCT: Assumptions on the Source

Causal assumptions on the source (under stratified RCT):

(A1) 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 an RCT: Assumptions on the Source

Causal assumptions on the source (under stratified RCT):

(A1) 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 an RCT: Assumptions on the Source

Causal assumptions on the source (under stratified RCT):

(A1) 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 A: $0 < \pi(\mathbf{x}) = \text{pr}(A = 1 \mid \mathbf{X} = \mathbf{x}, S = 1) < 1$.

Review of Transportability in an RCT: Assumptions on the Source

Causal assumptions on the source (under stratified RCT):

(A1) 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 A: $0 < \pi(\mathbf{x}) = \mathrm{pr}(A = 1 \mid \mathbf{X} = \mathbf{x}, S = 1) < 1$.

- The treated individuals and untreated individuals have a common support of $\mathbf{X}$.

Review of Transportability in an RCT: Assumptions on the Source

Causal assumptions on the source (under stratified RCT):

(A1) 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 A: $0 < \pi(\mathbf{x}) = \text{pr}(A = 1 \mid \mathbf{X} = \mathbf{x}, S = 1) < 1$.

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

Review of Transportability in an RCT: Assumptions for Transportation

Transportation assumptions:

(A4) Overlap of S: $0 < \text{pr}(S = 1 \mid \mathbf{V} = \mathbf{v}) < 1$ for all $\mathbf{v}$.

Review of Transportability in an RCT: Assumptions for Transportation

Transportation assumptions:

(A4) Overlap of S: $0 < \mathrm{pr}(S = 1 \mid \mathbf{V} = \mathbf{v}) < 1$ for all $\mathbf{v}$.

- The source voters and target voters have a common support of $\mathbf{V}$.

Review of Transportability in an RCT: Assumptions for Transportation

Transportation assumptions:

(A4) Overlap of S: $0 < \text{pr}(S = 1 \mid \mathbf{V} = \mathbf{v}) < 1$ for all $\mathbf{v}$.

- The source voters and target voters have a common support of $\mathbf{V}$.

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

Review of Transportability in an RCT: Assumptions for Transportation

Transportation assumptions:

(A4) Overlap of S: $0 < \mathrm{pr}(S = 1 \mid \mathbf{V} = \mathbf{v}) < 1$ for all $\mathbf{v}$.

- The source voters and target voters have a common support of $\mathbf{V}$.

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

- For any $\mathbf{v}$,
  $\mathrm{pr}(Y(a) = 1 \mid \mathbf{V} = \mathbf{v}, S = 0) = \mathrm{pr}(Y(a) = 1 \mid \mathbf{V} = \mathbf{v}, S = 1)$.

Review of Transportability in an RCT: Assumptions for Transportation

Transportation assumptions:

(A4) Overlap of S: $0 < \text{pr}(S = 1 \mid \mathbf{V} = \mathbf{v}) < 1$ for all $\mathbf{v}$.

- The source voters and target voters have a common support of $\mathbf{V}$.

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

- For any $\mathbf{v}$,
$\text{pr}(Y(a) = 1 \mid \mathbf{V} = \mathbf{v}, S = 0) = \text{pr}(Y(a) = 1 \mid \mathbf{V} = \mathbf{v}, S = 1)$.
- It is often assumed, but cannot be verified.

Review of Transportability in an RCT: Assumptions for Transportation

Transportation assumptions:

(A4) Overlap of S: $0 < \text{pr}(S = 1 \mid \mathbf{V} = \mathbf{v}) < 1$ for all $\mathbf{v}$.

- The source voters and target voters have a common support of $\mathbf{V}$.

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

- For any $\mathbf{v}$,
  $\text{pr}(Y(a) = 1 \mid \mathbf{V} = \mathbf{v}, S = 0) = \text{pr}(Y(a) = 1 \mid \mathbf{V} = \mathbf{v}, S = 1)$.
- It is often assumed, but cannot be verified.
- We assume it for now, but will relax it shortly!

Review: Identification Under (A1)-(A5)

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

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

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

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

What If Transportability (A5) Fails? Sensitivity Analysis

Suppose transportability (A5) does not hold:

$$\underbrace{\mathrm{pr}(Y(1), Y(0) \mid \mathbf{V}, S = 0)}_{\text{unobserved target (2024)}} \neq \underbrace{\mathrm{pr}(Y(1), Y(0) \mid \mathbf{V}, S = 1)}_{\text{observed source (2020)}}, \text{ then}$$

What If Transportability (A5) Fails? Sensitivity Analysis

Suppose transportability (A5) does not hold:

$$\underbrace{\mathrm{pr}(Y(1), Y(0) \mid \mathbf{V}, S=0)}_{\text{unobserved target (2024)}} \neq \underbrace{\mathrm{pr}(Y(1), Y(0) \mid \mathbf{V}, S=1)}_{\text{observed source (2020)}}, \text{ then}$$

$$\boxed{\mathrm{Odd}(Y(a)|\mathbf{V}, S=0)} := \frac{\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=0)}{1-\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=0)}$$

$$\neq \frac{\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=1)}{1-\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=1)} := \boxed{\mathrm{Odd}(Y(a)|\mathbf{V}, S=1)}.$$

What If Transportability (A5) Fails? Sensitivity Analysis

Suppose transportability (A5) does not hold:

$$\underbrace{\mathrm{pr}(Y(1), Y(0) \mid \mathbf{V}, S=0)}_{\text{unobserved target (2024)}} \neq \underbrace{\mathrm{pr}(Y(1), Y(0) \mid \mathbf{V}, S=1)}_{\text{observed source (2020)}}, \text{ then}$$

$$\boxed{\mathrm{Odd}(Y(a)|\mathbf{V}, S=0)} := \frac{\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=0)}{1-\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=0)} \neq \frac{\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=1)}{1-\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=1)} := \boxed{\mathrm{Odd}(Y(a)|\mathbf{V}, S=1)}.$$

For each $Y(a)$, we measure the deviation between the unobserved target the source counterpart via odds ratios (see formulation for a continuous outcome in Appendix):

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

What If Transportability (A5) Fails? Sensitivity Analysis

Suppose transportability (A5) does not hold:

$$\underbrace{\mathrm{pr}(Y(1), Y(0) \mid \mathbf{V}, S=0)}_{\text{unobserved target (2024)}} \neq \underbrace{\mathrm{pr}(Y(1), Y(0) \mid \mathbf{V}, S=1)}_{\text{observed source (2020)}}, \text{ then}$$

$$\boxed{\mathrm{Odd}(Y(a)|\mathbf{V}, S=0)} := \frac{\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=0)}{1-\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=0)}$$

$$\neq \frac{\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=1)}{1-\mathrm{pr}(Y(a)=1 \mid \mathbf{V}, S=1)} := \boxed{\mathrm{Odd}(Y(a)|\mathbf{V}, S=1)}.$$

For each $Y(a)$, we measure the deviation between the unobserved target the source counterpart via odds ratios (see formulation for a continuous outcome in Appendix):

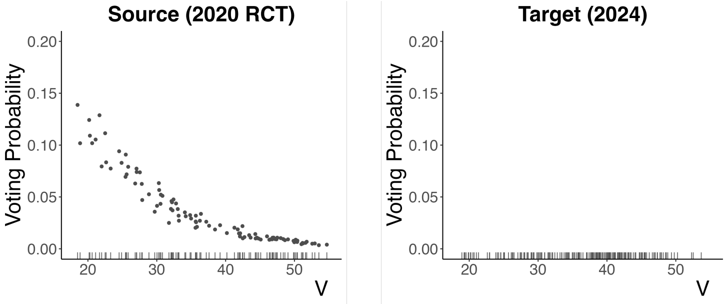
$$\exp(\gamma_a) = \frac{\mathrm{Odd}(Y(a) \mid \mathbf{v}, S=0)}{\mathrm{Odd}(Y(a) \mid \mathbf{v}, S=1)}, \ \gamma_a \in (-\infty, \infty). \tag{1}$$

- When $\gamma_a = 0$, $\exp(\gamma_a) = \exp(0) = 1 = \frac{\mathrm{Odd}(Y(a) \mid \mathbf{v}, S=0)}{\mathrm{Odd}(Y(a) \mid \mathbf{v}, S=1)} \Longrightarrow$ (A5) holds.
- When $\gamma_a \neq 0$, (A5) does not hold; γ_a measures the degree of violation to (A5).

Interpreting Sensitivity Parameter γ_a

Sensitivity model: $\exp(\gamma_a) = \text{Odd}(Y(a) \mid \mathbf{v}, S = 0)/\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)$.

- When $\gamma_a = 0$, the transportability (A5) holds.

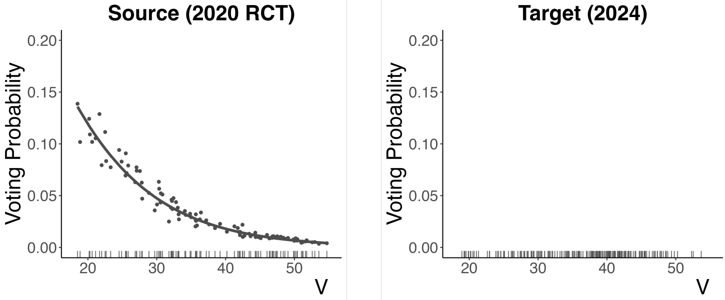


Toy example: $\text{pr}(Y(a) = 1 \mid V, S = 1) = \text{expit}(-0.1V)$.

Interpreting Sensitivity Parameter γ_a

Sensitivity model: $\exp(\gamma_a) = \text{Odd}(Y(a) \mid \mathbf{v}, S = 0)/\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)$.

- When $\gamma_a = 0$, the transportability (A5) holds.

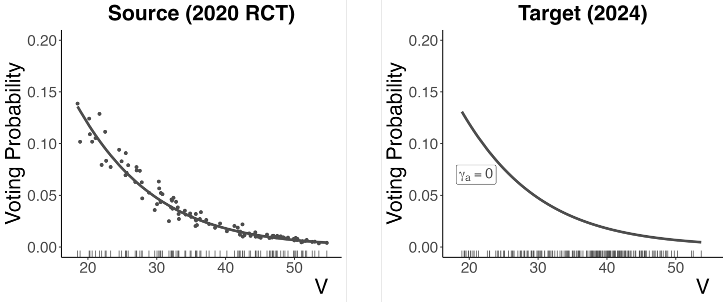


Toy example: $\text{pr}(Y(a) = 1 \mid V, S = 1) = \text{expit}(-0.1V)$.

Interpreting Sensitivity Parameter γ_a

Sensitivity model: $\exp(\gamma_a) = \text{Odd}(Y(a) \mid \mathbf{v}, S=0)/\text{Odd}(Y(a) \mid \mathbf{v}, S=1)$.

- When $\gamma_a = 0$, the transportability (A5) holds.

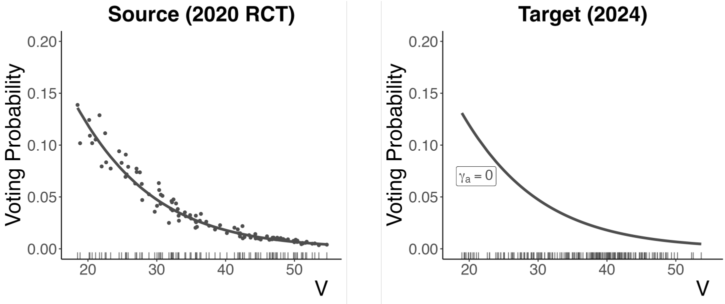


Toy example: $\text{pr}(Y(a) = 1 \mid V, S=1) = \text{expit}(-0.1V)$.

Interpreting Sensitivity Parameter γ_a

Sensitivity model: $\exp(\gamma_a) = \text{Odd}(Y(a) \mid \mathbf{v}, S = 0)/\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)$.

- When $\gamma_a = 0$, the transportability (A5) holds.
- When $\gamma_a \neq 0$, larger $|\gamma_a| \Rightarrow$ larger differences between 2024 and 2020.

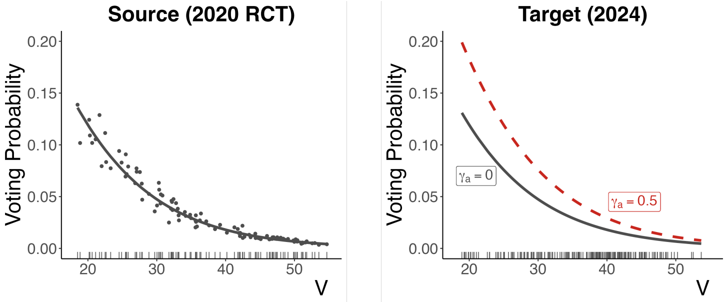


Toy example: $\text{pr}(Y(a) = 1 \mid V, S = 1) = \text{expit}(-0.1V)$.

Interpreting Sensitivity Parameter γ_a

Sensitivity model: $\exp(\gamma_a) = \text{Odd}(Y(a) \mid \mathbf{v}, S = 0)/\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)$.

- When $\gamma_a = 0$, the transportability (A5) holds.
- When $\gamma_a \neq 0$, larger $|\gamma_a| \Rightarrow$ larger differences between 2024 and 2020.
- Positive $\gamma_1 \Rightarrow$ more turnout in 2024 after receiving ads against Trump compared to that in 2020.

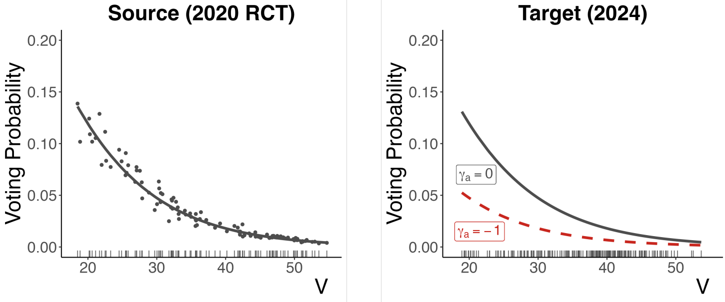


Toy example: $\text{pr}(Y(a) = 1 \mid V, S = 1) = \text{expit}(-0.1V)$.

Interpreting Sensitivity Parameter γ_a

Sensitivity model: $\exp(\gamma_a) = \text{Odd}(Y(a) \mid \mathbf{v}, S = 0)/\text{Odd}(Y(a) \mid \mathbf{v}, S = 1)$.

- When $\gamma_a = 0$, the transportability (A5) holds.
- When $\gamma_a \neq 0$, larger $|\gamma_a| \Rightarrow$ larger differences between 2024 and 2020.
- 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: $\text{pr}(Y(a) = 1 \mid V, S = 1) = \text{expit}(-0.1V)$.

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

Recall $\mu_a(\mathbf{X}) = \mathbb{E}(Y \mid \mathbf{X}, A = a, S = 1)$. Under (A1)-(A4) and the sensitivity model,

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

- When $\gamma_a = 0$, transportability (A5) holds, we return to the previous result:

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

- When $\gamma_a \neq 0$, it is an exponential tilt of $\rho_a(\mathbf{V}) = \mathbb{E}[\mu_a(\mathbf{X}) \mid \mathbf{V}, S = 1]$.

A Simple, Designed-Inspired Estimator and Its Inference

- A simple, design inspired estimation procedure.
  (1) Estimate $\rho_a(\mathbf{V}) = \mathbb{E}[\mu_a(\mathbf{X}) \mid \mathbf{V}, S = 1]$ from source data.

A Simple, Designed-Inspired Estimator and Its Inference

- A simple, design inspired estimation procedure.
  (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 $Y/\widehat{\pi}(\mathbf{X})$ on $\mathbf{V}$ to get $\widehat{\rho}_1(\mathbf{V})$.

A Simple, Designed-Inspired Estimator and Its Inference

- A simple, design inspired estimation procedure.
  (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 $Y/\widehat{\pi}(\mathbf{X})$ on $\mathbf{V}$ to get $\widehat{\rho}_1(\mathbf{V})$.
  (2) Average the exponentially tilted $\widehat{\rho}_a(\mathbf{V})$ among target sample,

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

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

A Simple, Designed-Inspired Estimator and Its Inference

- A simple, design inspired estimation procedure.
  (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 $Y/\widehat{\pi}(\mathbf{X})$ on $\mathbf{V}$ to get $\widehat{\rho}_1(\mathbf{V})$.

  (2) Average the exponentially tilted $\widehat{\rho}_a(\mathbf{V})$ among target sample,

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

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

Since our voter data is discrete, the plug-in estimator is **nonparametric** and **efficient** ([Chamberlain, 1987, Theorem 1]).

A Simple, Designed-Inspired Estimator and Its Inference

- A simple, design inspired estimation procedure.
  (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 $Y/\widehat{\pi}(\mathbf{X})$ on $\mathbf{V}$ to get $\widehat{\rho}_1(\mathbf{V})$.

  (2) Average the exponentially tilted $\widehat{\rho}_a(\mathbf{V})$ among target sample,

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

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

  Since our voter data is discrete, the plug-in estimator is **nonparametric** and **efficient** ([Chamberlain, 1987, Theorem 1]).

- Constructing a $1 - \alpha$ CI of θ (see Appendix for details).
  - At each iteration b, bootstrap the source data and target data separately, construct an estimator $\widehat{\theta}_b^*$ with resampled data.
  - After B iterations, take $\alpha/2$ and $1 - \alpha/2$ quantiles of $\left\{\widehat{\theta}_b^*\right\}_{b=1}^{B}$.

Estimator Based on Efficient Influence Function

For a given γ_a, we have the efficient influence function (EIF) of θ.

- The EIF is **very messy** because (a) $\mathbf{V} \neq \mathbf{X}$ and (b) sensitivity analysis; see Appendix.
- Our EIF recovers [Zeng et al., 2023]'s EIF when $\gamma_a = 0$.

Practically, an EIF-based estimator is useful if $\mathbf{V}$ is continuous.

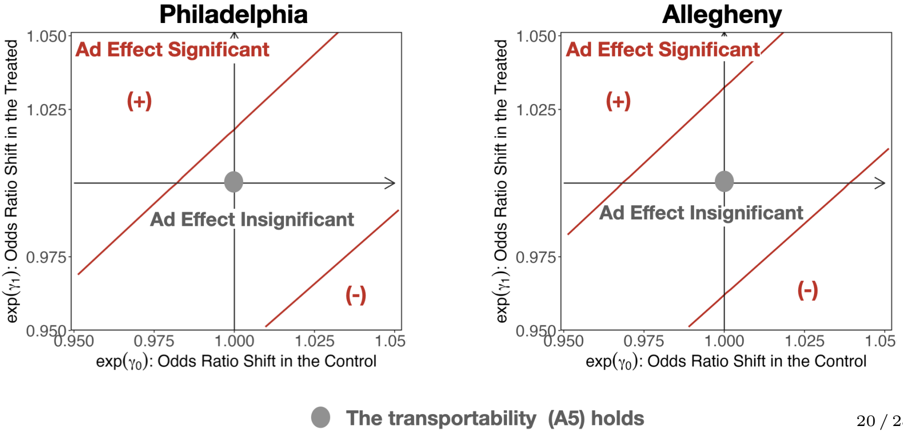
- Four nuisance functions: (i) propensity score $\pi(\mathbf{X})$, (ii) outcome regression $\mu_a(\mathbf{X})$, (iii) projection of outcome regression $\rho(\mathbf{V})$, and (iv) weights between source and target $w(\mathbf{V}) = p(\mathbf{V} \mid S = 0)/p(\mathbf{V} \mid S = 1)$.
- To avoid Donsker conditions, we need cross-fitting in source data.
- The estimator is not doubly robust for $\gamma_a \neq 0$.
- Also, the estimator does not reduce “plug-in bias” from $\rho_a(\mathbf{V})$.

Target Data: Registered Voters in Pennsylvania (PA)

- 4,880,730 registered voters (as of Apr.15, 2024) from 67 counties.
- $\mathbf{V}$: age group, gender, party, an incomplete voting history.
  - $\mathbf{X} \setminus \mathbf{V}$: race, missing voting history.
- We look at each county in Pennsylvania
  - n_t ranges from 1,117 to 685,620; median is 25,182.
- For sensitivity parameters, we use $-0.05 \leq \gamma_a \leq 0.05$ ($0.951 \leq \exp(\gamma_a) \leq 1.051$).
- For inference, we use the simple plug-in estimator with bootstrap.

Example of Sensitivity Contours

- When $\gamma_1 = \gamma_0 = 0$ (transportability (A5) holds; 2024 $\approx$ 2020), all effects remain insignificant.
- When $\gamma_1\gamma_0 \neq 0$ ((A5) does not hold; 2024 $\neq$ 2020), the effect can be significant under 0.05 level.



The Most and Least Sensitive Counties

We calculate the smallest $\exp(|\gamma_1 - \gamma_0|)$ that turns the ATE significant.

- A **lower** value indicates a smaller difference between 2024 and 2020 can make the ad to significantly affect the voter turnout $\Longrightarrow$ **more sensitive**.
- A **higher** value indicates only a large difference between 2024 and 2020 can make the ad to significantly affect the voter turnout $\Longrightarrow$ **less sensitive**.

Table 1: The most and least sensitive counties.

	Positive ad effects		Negative ad effects	
	County	$\exp(\lvert\gamma_1 - \gamma_0\rvert)$	County	$\exp(\lvert\gamma_1 - \gamma_0\rvert)$
Most Sensitive	Philadelphia	1.018	Bedford	1.002
	Monroe	1.028	Fulton	1.002
Least Sensitive	Bedford	1.010	Allegheny	1.041
	Fulton	1.105	Philadelphia	1.062
	Clinton	-	Clinton	-

Calibrating Sensitivity Parameters

In sensitivity analysis, there is always a question about what is a "large", "small", or a "plausible" sensitivity parameter γ_a.

Calibrating Sensitivity Parameters

In sensitivity analysis, there is always a question about what is a "large", "small", or a "plausible" sensitivity parameter γ_a.

We present one solution to this question by creating **dis-similar** partitions of the source data.

Calibrating Sensitivity Parameters

In sensitivity analysis, there is always a question about what is a "large", "small", or a "plausible" sensitivity parameter γ_a.

We present one solution to this question by creating **dis-similar** partitions of the source data.

(1) Partition source into two "dissimilar" subpopulations (i.e. subpopulations are not similar after adjusting for covariates).
   - For example, in [Aggarwal et al., 2023], we created [PA,MI,WI] ("blue collar states") and [NC,AZ] ("not blue collar states").

(2) Use one partition as source, the other as target.

Calibrating Sensitivity Parameters

In sensitivity analysis, there is always a question about what is a "large", "small", or a "plausible" sensitivity parameter γ_a.

We present one solution to this question by creating **dis-similar** partitions of the source data.

(1) Partition source into two "dissimilar" subpopulations (i.e. subpopulations are not similar after adjusting for covariates).
   - For example, in [Aggarwal et al., 2023], we created [PA,MI,WI] ("blue collar states") and [NC,AZ] ("not blue collar states").

(2) Use one partition as source, the other as target.
   - For each (γ_1, γ_0) pair, we obtain the CI of the transported ATE (transported CI).

Calibrating Sensitivity Parameters

In sensitivity analysis, there is always a question about what is a "large", "small", or a "plausible" sensitivity parameter γ_a.

We present one solution to this question by creating **dis-similar** partitions of the source data.

(1) Partition source into two "dissimilar" subpopulations (i.e. subpopulations are not similar after adjusting for covariates).
   - For example, in [Aggarwal et al., 2023], we created [PA,MI,WI] ("blue collar states") and [NC,AZ] ("not blue collar states").

(2) Use one partition as source, the other as target.
   - For each (γ_1, γ_0) pair, we obtain the CI of the transported ATE (transported CI).
   - Because the RCT has been run on this target, we have the CI of the target ATE (oracle CI).

Calibrating Sensitivity Parameters

In sensitivity analysis, there is always a question about what is a "large", "small", or a "plausible" sensitivity parameter γ_a.

We present one solution to this question by creating **dis-similar** partitions of the source data.

(1) Partition source into two "dissimilar" subpopulations (i.e. subpopulations are not similar after adjusting for covariates).
   - For example, in [Aggarwal et al., 2023], we created [PA,MI,WI] ("blue collar states") and [NC,AZ] ("not blue collar states").

(2) Use one partition as source, the other as target.
   - For each (γ_1, γ_0) pair, we obtain the CI of the transported ATE (transported CI).
   - Because the RCT has been run on this target, we have the CI of the target ATE (oracle CI).

(3) Let the subset $\mathcal{C} \subset \mathbb{R}^2$ be the range of (γ_1, γ_0) pairs where the corresponding transported CI overlaps with the oracle CI.

Calibrating Sensitivity Parameters

In sensitivity analysis, there is always a question about what is a "large", "small", or a "plausible" sensitivity parameter γ_a.

We present one solution to this question by creating **dis-similar** partitions of the source data.

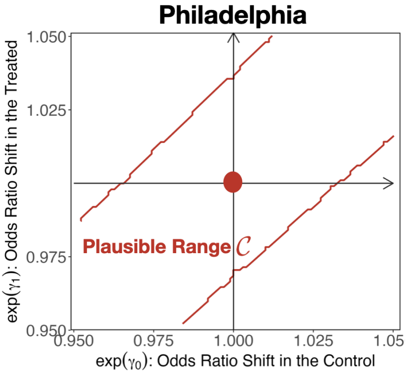
(1) Partition source into two "dissimilar" subpopulations (i.e. subpopulations are not similar after adjusting for covariates).
   - For example, in [Aggarwal et al., 2023], we created [PA,MI,WI] ("blue collar states") and [NC,AZ] ("not blue collar states").

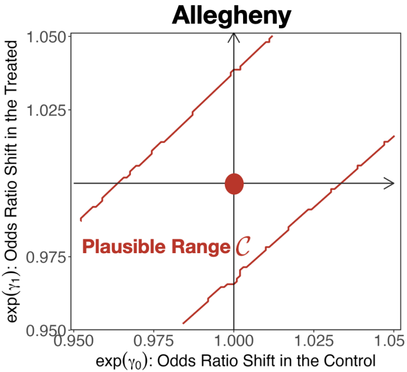
(2) Use one partition as source, the other as target.
   - For each (γ_1, γ_0) pair, we obtain the CI of the transported ATE (transported CI).
   - Because the RCT has been run on this target, we have the CI of the target ATE (oracle CI).

(3) Let the subset $\mathcal{C} \subset \mathbb{R}^2$ be the range of (γ_1, γ_0) pairs where the corresponding transported CI overlaps with the oracle CI.

The γ_a's from PA that overlap with the subset $\mathcal{C}$ are "plausible".

Examples of Calibrations





● **The transportability (A5) holds**

Calibrated Results for Positive Ad Effects

Figure 1: The smallest $\exp(\gamma_1 - \gamma_0)$ turns ATE significant when $\gamma_1 > \gamma_0$.

- Nine counties could have a positive ad effect.

[3] From Ballotpedia

Calibrated Results for Positive Ad Effects

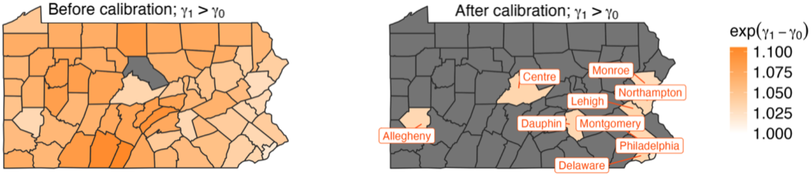


Figure 1: The smallest $\exp(\gamma_1 - \gamma_0)$ turns ATE significant when $\gamma_1 > \gamma_0$.

- Nine counties could have a positive ad effect.
  - Receiving ads against Trump may increase their turnout.

[3] From Ballotpedia

Calibrated Results for Positive Ad Effects

Figure 1: The smallest $\exp(\gamma_1 - \gamma_0)$ turns ATE significant when $\gamma_1 > \gamma_0$.

- Nine counties could have a positive ad effect.
  - Receiving ads against Trump may increase their turnout.
  - They are mostly near urban-ish (Philadelphia, Pittsburgh, suburbs of Philadelphia) or college towns (i.e. Centre county).

[3] From Ballotpedia

Calibrated Results for Positive Ad Effects

Figure 1: The smallest $\exp(\gamma_1 - \gamma_0)$ turns ATE significant when $\gamma_1 > \gamma_0$.

- Nine counties could have a positive ad effect.
    - Receiving ads against Trump may increase their turnout.
    - They are mostly near urban-ish (Philadelphia, Pittsburgh, suburbs of Philadelphia) or college towns (i.e. Centre county).
    - Northampton is a "pivot" county[3] (Biden won by 1,233 votes; 172,065 voters voted).

[3] From Ballotpedia

Calibrated Results for Positive Ad Effects

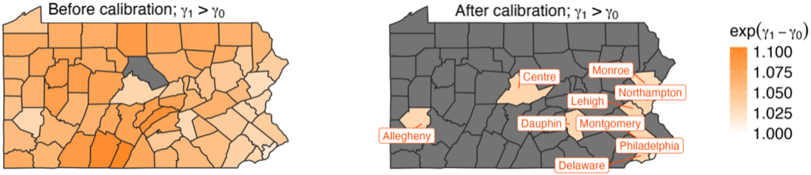


Figure 1: The smallest $\exp(\gamma_1 - \gamma_0)$ turns ATE significant when $\gamma_1 > \gamma_0$.

- Nine counties could have a positive ad effect.
  - Receiving ads against Trump may increase their turnout.
  - They are mostly near urban-ish (Philadelphia, Pittsburgh, suburbs of Philadelphia) or college towns (i.e. Centre county).
  - Northampton is a "pivot" county[3] (Biden won by 1,233 votes; 172,065 voters voted).
- Biden has won all of the nine counties in the 2020 election.

[3] From Ballotpedia

Calibrated Results for Positive Ad Effects

Figure 1: The smallest $\exp(\gamma_1 - \gamma_0)$ turns ATE significant when $\gamma_1 > \gamma_0$.

- Nine counties could have a positive ad effect.
  - Receiving ads against Trump may increase their turnout.
  - They are mostly near urban-ish (Philadelphia, Pittsburgh, suburbs of Philadelphia) or college towns (i.e. Centre county).
  - Northampton is a "pivot" county[3] (Biden won by 1,233 votes; 172,065 voters voted).
- Biden has won all of the nine counties in the 2020 election.
  - Biden leaners may be encouraged to vote by ads against Trump; this aligns with analyses in [Aggarwal et al., 2023] stratified by Trump supporting score.

[3]From Ballotpedia

Calibrated Results for Negative Ad Effects

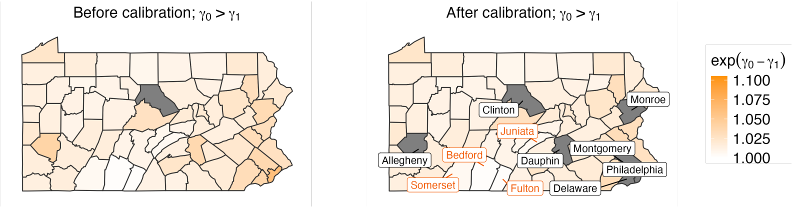


Figure 2: The smallest $\exp(\gamma_0 - \gamma_1)$ inducing significance when $\gamma_0 > \gamma_1$.

- Most counties could have a negative ad effect.

Calibrated Results for Negative Ad Effects

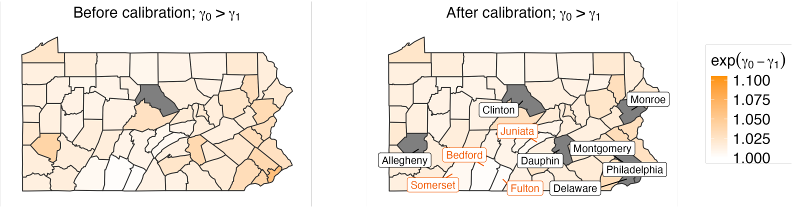


Figure 2: The smallest $\exp(\gamma_0 - \gamma_1)$ inducing significance when $\gamma_0 > \gamma_1$.

- Most counties could have a negative ad effect.
  - Receiving ads against Trump may discourage them to vote.

Calibrated Results for Negative Ad Effects

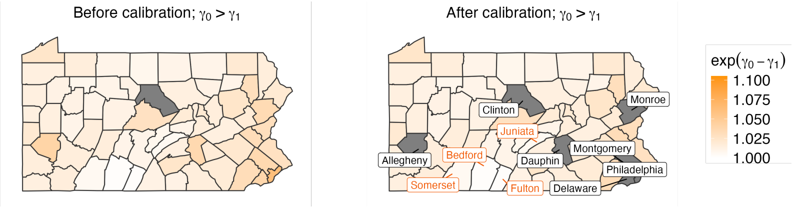


Figure 2: The smallest $\exp(\gamma_0 - \gamma_1)$ inducing significance when $\gamma_0 > \gamma_1$.

- Most counties could have a negative ad effect.
  - Receiving ads against Trump may discourage them to vote.
  - The most sensitive counties are Fulton and Bedford; smallest $\exp(\gamma_0 - \gamma_1) = 1.002$.

Calibrated Results for Negative Ad Effects

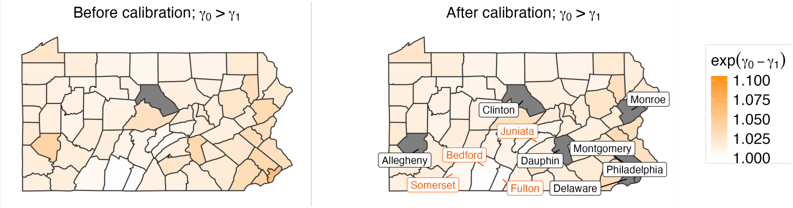


Figure 2: The smallest $\exp(\gamma_0 - \gamma_1)$ inducing significance when $\gamma_0 > \gamma_1$.

- Most counties could have a negative ad effect.
  - Receiving ads against Trump may discourage them to vote.
  - The most sensitive counties are Fulton and Bedford; smallest $\exp(\gamma_0 - \gamma_1) = 1.002$.
  - They have the largest margin for Trump (85.41% for Trump in Fulton; 83.39% for Trump in Bedford) in 2020 U.S. presidential election.

Calibrated Results for Negative Ad Effects

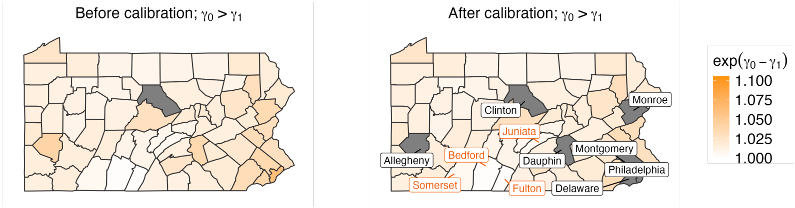


Figure 2: The smallest $\exp(\gamma_0 - \gamma_1)$ inducing significance when $\gamma_0 > \gamma_1$.

- Most counties could have a negative ad effect.
  - Receiving ads against Trump may discourage them to vote.
  - The most sensitive counties are Fulton and Bedford; smallest $\exp(\gamma_0 - \gamma_1) = 1.002$.
  - They have the largest margin for Trump (85.41% for Trump in Fulton; 83.39% for Trump in Bedford) in 2020 U.S. presidential election.
- Three swing counties may have ad effects in either direction: Centre, Lehigh, Northampton.

Some Preliminary Takeaways from PA

- If transportability (A5) holds (i.e. 2020 $\approx$ 2024), **all** counties will have near zero ad effects in 2024.
- If (A5) fails, a few counties could have positive ad effects, whereas most could have negative ad effects in 2024.
  - The direction largely depends on their leaning towards Trump/Biden (Republican/Democrat).
  - Counties with mostly Trump leaners are likely to vote less, whereas counties with Biden leaners will vote more.
  - The direction can go either way in swing counties.

Summary and Ongoing Work

- Motivation: From [Aggarwal et al., 2023], would the negative ad against Trump in 2020 remain ineffective in 2024?
- Our approach: transfer learning with sensitivity analysis
  - Setup: (a) source is from RCT, (b) $\mathbf{V} \neq \mathbf{X}$, (c) data is discrete.
  - Analysis: (a) simple plug-in estimator with bootstrap SE/CIs, (b) EIF-based approach, (c) calibration of sensitivity parameters with source data.
- Preliminary analysis of Pennsylvania.
- Ongoing work
  - Repeat analysis with other states (WI, NC and GA).
  - Use 2022 U.S. midterm elections to improve $\hat{\theta}$ and to improve calibration.

Acknowledgements

- Thank you to Xiaobin Zhou for finding this data :)
- Thank you all for coming. Comments are highly appreciated!
- Thank you to Steven Moen for finding typos in an earlier version!

Part I

Appendix

References I

Aggarwal, M., Allen, J., Coppock, A., Frankowski, D., Messing, S., Zhang, K., Barnes, J., Beasley, A., Hantman, H., and Zheng, S. (2023).

A 2 million-person, campaign-wide field experiment shows how digital advertising affects voter turnout.

Nature Human Behaviour, 7(3):332–341.

Chamberlain, G. (1987).

Asymptotic efficiency in estimation with conditional moment restrictions.

Journal of econometrics, 34(3):305–334.

References II

Dahabreh, I. J., Petito, L. C., Robertson, S. E., Hernán, M. A., and Steingrimsson, J. A. (2020).

Toward causally interpretable meta-analysis: transporting inferences from multiple randomized trials to a new target population.

Epidemiology, 31(3):334–344.

Dahabreh, I. J., Robins, J. M., Haneuse, S. J., Robertson, S. E., Steingrimsson, J. A., and Hernán, M. A. (2022).

Global sensitivity analysis for studies extending inferences from a randomized trial to a target population.

arXiv preprint arXiv:2207.09982.

References III

Dahabreh, I. J., Robins, J. M., Haneuse, S. J.-P., Saeed, I., Robertson, S. E., Stuart, E. A., and Hernán, M. A. (2023).

Sensitivity analysis using bias functions for studies extending inferences from a randomized trial to a target population.

Statistics in Medicine.

Franks, A., D'Amour, A., and Feller, A. (2019).

Flexible sensitivity analysis for observational studies without observable implications.

Journal of the American Statistical Association.

References IV

Huang, M. Y. (2024).

Sensitivity analysis for the generalization of experimental results.

Journal of the Royal Statistical Society Series A: Statistics in Society.

Nguyen, T. Q., Ebnesajjad, C., Cole, S. R., and Stuart, E. A. (2017).

Sensitivity analysis for an unobserved moderator in rct-to-target-population generalization of treatment effects.

The Annals of Applied Statistics, pages 225–247.

References V

Nie, X., Imbens, G., and Wager, S. (2021).

Covariate balancing sensitivity analysis for extrapolating randomized trials across locations.

arXiv preprint arXiv:2112.04723.

Robins, J. M., Rotnitzky, A., and Scharfstein, D. O. (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*, pages 1–94. Springer.

References VI

Scharfstein, D. O., Nabi, R., Kennedy, E. H., Huang, M.-Y., Bonvini, M., and Smid, M. (2021).

Semiparametric sensitivity analysis: Unmeasured confounding in observational studies.

arXiv preprint arXiv:2104.08300.

Zeng, Z., Kennedy, E. H., Bodnar, L. M., and Naimi, A. I. (2023).

Efficient generalization and transportation.

arXiv preprint arXiv:2302.00092.

FAQs

- Is the calibration reasonable?
  Great question! We're also experimenting with the "right" way to assess whether a give γ value is extreme or not.

- Why is your voter data all discrete?
  We're not sure and this surprised us too. 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.

- Is your conclusion sensitive to data quality from 2024 voter registration data (i.e. target data)?
  Yes. Unfortunately, high quality target data is expensive.

- What about treating this data as longitudinal?
  Excellent idea, but this requires measuring same voter over time.

Some Prior Works

The closest related literature is transportability/generalizability.

We are definitely not the only ones to incorporate sensitivity analysis in transportability/generalizability. A very small, partial list:

- Linear, outcome sensitivity model: [Nguyen et al., 2017, Dahabreh et al., 2020, Dahabreh et al., 2023, Zeng et al., 2023]
- Exponential tilting sensitivity model: [Dahabreh et al., 2022]
- Marginal sensitivity model for transfer learning functionals: [Nie et al., 2021]
- Omitted variable bias approach with weighted estimators: [Huang, 2024]

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

Two-Parameter Sensitivity Model and Some Remarks

To jointly characterize $Y(1), Y(0)$, we use the following model

$$\begin{aligned}&\mathrm{pr}(Y(1) = y_1, Y(0) = y_0 \mid \mathbf{V} = \mathbf{v}, S = 0)\\ \propto\, &\exp(\gamma_1 y_1 + \gamma_0 y_0) \cdot \mathrm{pr}(Y(1) = y_1, Y(0) = y_0 \mid \mathbf{V} = \mathbf{v}, S = 1)\end{aligned}$$

Some remarks:

- The sensitivity model does not place any observable restrictions on the observed data [Robins et al., 2000, Franks et al., 2019]
- A pseudo-R^2 version of γ_a is in Proposition 3 of [Franks et al., 2019].
- The sensitivity model can depend on covariates ($\exp(\gamma_{\mathbf{v}}^{\intercal}\mathbf{v} + ...)$)
- Some works that use this model: [Robins et al., 2000, Franks et al., 2019, Scharfstein et al., 2021, Dahabreh et al., 2022]
- There is a long and healthy debate about what constitutes a "good" sensitivity analysis.

Alternative Formulation: Exponential Tilting Model

The selection odds model (1) can be equivalently written as

$$p(y(a) \mid \mathbf{V}, S=0) \propto \exp\{\gamma_a y(a)\} \cdot p(y(a) \mid \mathbf{v}, S=1). \tag{2}$$

Under (A1)-(A4) and (2), we can identify $\mathbb{E}(Y(a) \mid S=0)$ as

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

Bootstrapping for Transfer Learning

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 $\rho(\mathbf{V})$ 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.

[Return to main slides.]

Voter Demographics Between Source and Target Population

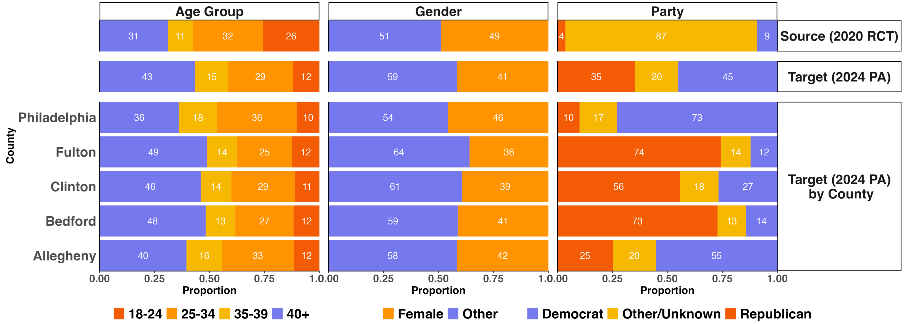


Figure 3: Registered voter demographics.

Preliminary Result for PA: A Larger **V** Robustifies Conclusions

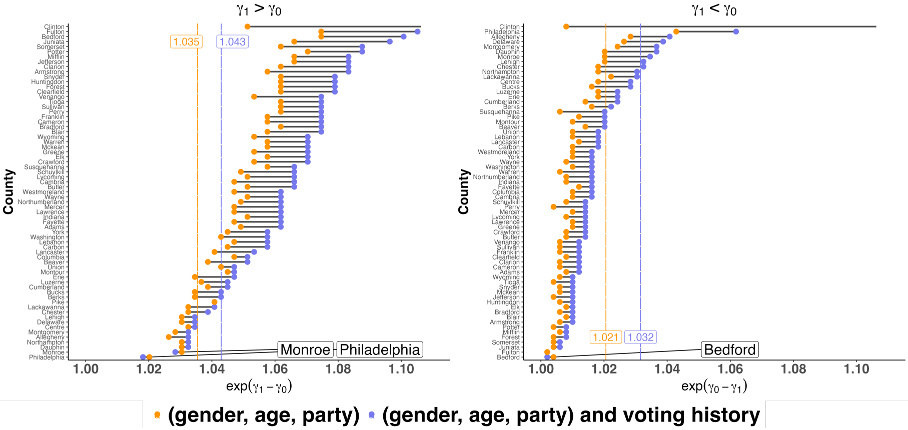


Figure 4: The smallest $\exp(|\gamma_1 - \gamma_0|)$ that makes $\hat{\theta}(\gamma_1, \gamma_0)$ significant. Left considers the case where $\gamma_1 > \gamma_0$ and right considers the case when $\gamma_1 < \gamma_0$.

EIF-Based Estimator

$$\widehat{\theta}_{\mathrm{EIF},a} =$$

$$\frac{1}{n_s}\sum_{i\in\mathrm{Source}}\widehat{w}(\mathbf{v}_i)\left(\left\{\frac{a_i}{\widehat{\pi}(\mathbf{x}_i)}+\frac{1-a_i}{1-\widehat{\pi}(\mathbf{x}_i)}\right\}\left[\frac{e^{\gamma_a y_i}y_i}{e^{\gamma_a}\widehat{\rho}(\mathbf{v}_i)+1-\widehat{\rho}(\mathbf{v}_i)}-\frac{e^{\gamma_a}\widehat{\mu}_a(\mathbf{x}_i)}{e^{\gamma_a}\widehat{\rho}(\mathbf{v}_i)+1-\widehat{\rho}(\mathbf{v}_i)}\right.\right.$$

$$\left.-\frac{e^{\gamma_a y_i}\widehat{\rho}_a(\mathbf{v}_i)}{[e^{\gamma_a}\widehat{\rho}_a(\mathbf{v}_i)+1-\rho_a(\mathbf{v}_i)]^2}+\frac{\{e^{\gamma_a}\widehat{\mu}_a(\mathbf{x}_i)+1-\widehat{\mu}_a(\mathbf{x}_i)\}e^{\gamma_a}\widehat{\rho}(\mathbf{v}_i)}{[e^{\gamma_a}\widehat{\rho}_a(\mathbf{v}_i)+1-\rho_a(\mathbf{v}_i)]^2}\right]$$

$$\left.+\frac{\widehat{\mu}_a(\mathbf{x}_i)\{e^{\gamma_a}\widehat{\rho}(\mathbf{v}_i)+1-\widehat{\rho}(\mathbf{v}_i)\}-\widehat{\rho}_a(\mathbf{v}_i)\{e^{\gamma_a}\widehat{\mu}_a(\mathbf{x}_i)+1-\widehat{\mu}_a(\mathbf{x}_i)\}}{[e^{\gamma_a}\widehat{\rho}_a(\mathbf{v}_i)+1-\rho_a(\mathbf{v}_i)]^2}\right)$$

$$+\frac{1}{n_t}\sum_{i\in\mathrm{Target}}\frac{e^{\gamma_a}\widehat{\rho}_a(\mathbf{v}_i)}{e^{\gamma_a}\widehat{\rho}_a(\mathbf{v}_i)+1-\widehat{\rho}_a(\mathbf{v}_i)}.$$

When $\gamma_a = 0$, it collapses to [Zeng et al., 2023]:

$$=\frac{1}{n_s}\sum_{i\in\mathrm{Source}}\widehat{w}(\mathbf{v}_i)\left(\frac{a_i}{\widehat{\pi}(\mathbf{x}_i)}+\frac{1-a_i}{1-\widehat{\pi}(\mathbf{x}_i)}\right)[y_i-\widehat{\mu}_a(\mathbf{x}_i)]+$$

$$\frac{1}{n_s}\sum_{i\in\mathrm{Source}}\widehat{w}(\mathbf{v}_i)[\widehat{\mu}_a(\mathbf{x}_i)-\widehat{\rho}_a(\mathbf{v}_i)]+\frac{1}{n_t}\sum_{i\in\mathrm{Target}}\widehat{\rho}(\mathbf{v}_i).$$

[Return to main slides.]