Difference-in-Differences for Ordinal Outcomes

Soichiro Yamauchi

Harvard University

Applied Statistics Workshop, IQSS
April 1, 2020

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

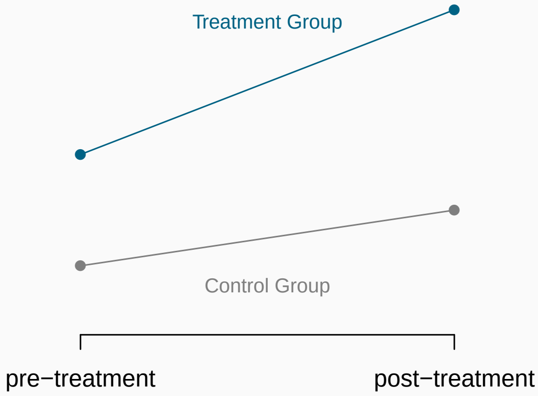
Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment

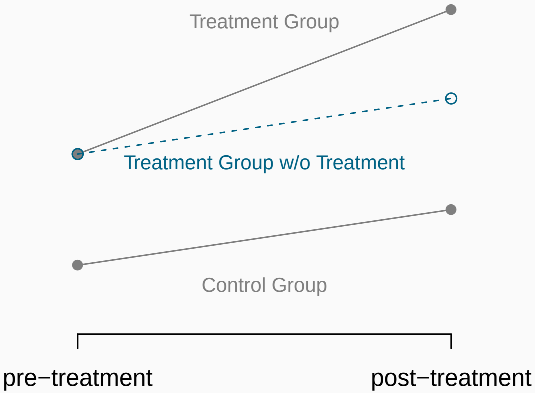
Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption
  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment



Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption
  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment



Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment

  $\rightsquigarrow$ Relies on the differences between two potential outcomes: Linearity

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment

  $\rightsquigarrow$ Relies on the differences between two potential outcomes: Linearity

- In social science, many outcomes are measured on an ordinal scale (e.g., survey questions) $\rightsquigarrow$ "differences" are not well defined

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment

  $\rightsquigarrow$ Relies on the differences between two potential outcomes: Linearity

- In social science, many outcomes are measured on an ordinal scale (e.g., survey questions) $\rightsquigarrow$ “differences” are not well defined
- Problems in common practices:

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment

  $\rightsquigarrow$ Relies on the differences between two potential outcomes: Linearity

- In social science, many outcomes are measured on an ordinal scale (e.g., survey questions) $\rightsquigarrow$ "differences" are not well defined
- Problems in common practices:
  - Treat as a continuous variable $\rightsquigarrow$ Difficult to interpret + linearity

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment

  $\rightsquigarrow$ Relies on the differences between two potential outcomes: Linearity

- In social science, many outcomes are measured on an ordinal scale (e.g., survey questions) $\rightsquigarrow$ "differences" are not well defined
- Problems in common practices:
  - Treat as a continuous variable $\rightsquigarrow$ Difficult to interpret + linearity
  - Dichotomize the outcome $\rightsquigarrow$ *Multiple* distinct parallel trends assumptions

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment

  $\rightsquigarrow$ Relies on the differences between two potential outcomes: Linearity

- In social science, many outcomes are measured on an ordinal scale (e.g., survey questions) $\rightsquigarrow$ "differences" are not well defined
- Problems in common practices:
  - Treat as a continuous variable $\rightsquigarrow$ Difficult to interpret + linearity
  - Dichotomize the outcome $\rightsquigarrow$ *Multiple* distinct parallel trends assumptions
  - Ordered probit/logit $\rightsquigarrow$ Identification assumptions are not explicitly stated

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment

  $\rightsquigarrow$ Relies on the differences between two potential outcomes: Linearity

- In social science, many outcomes are measured on an ordinal scale (e.g., survey questions) $\rightsquigarrow$ "differences" are not well defined
- Problems in common practices:
  - Treat as a continuous variable $\rightsquigarrow$ Difficult to interpret + linearity
  - Dichotomize the outcome $\rightsquigarrow$ *Multiple* distinct parallel trends assumptions
  - Ordered probit/logit $\rightsquigarrow$ Identification assumptions are not explicitly stated

- Propose: A latent variable framework for DiD for the ordinal outcomes

Difference-in-Differences Design in Observational Studies

- Difference-in-differences for causal inference in observational studies
- Adjust for the time-invariant confounders by utilizing the past outcome
- Key identification assumption: parallel trends assumption

  $\rightsquigarrow$ Identical *trends* across the treated & the control without the treatment

  $\rightsquigarrow$ Relies on the differences between two potential outcomes: Linearity

- In social science, many outcomes are measured on an ordinal scale (e.g., survey questions) $\rightsquigarrow$ "differences" are not well defined
- Problems in common practices:
  - Treat as a continuous variable $\rightsquigarrow$ Difficult to interpret + linearity
  - Dichotomize the outcome $\rightsquigarrow$ *Multiple* distinct parallel trends assumptions
  - Ordered probit/logit $\rightsquigarrow$ Identification assumptions are not explicitly stated

- Propose: A latent variable framework for DiD for the ordinal outcomes
- Application: Revisit a recent debate on the relationship between the mass shootings and the attitude toward gun control

Difference-in-Differences for Ordinal Outcomes

Contributions: New identification strategy & diagnostic tool

Difference-in-Differences for Ordinal Outcomes

Contributions: New identification strategy & diagnostic tool

- Introduce a latent variable framework

Difference-in-Differences for Ordinal Outcomes

Contributions: New identification strategy & diagnostic tool

- Introduce a latent variable framework
  - Extend the latent utility representation of the standard probit/logit

Difference-in-Differences for Ordinal Outcomes

Contributions: New identification strategy & diagnostic tool

- Introduce a latent variable framework
  - Extend the latent utility representation of the standard probit/logit

- Apply the assumption by Athey & Imbens (2006) on the latent variable scale

Difference-in-Differences for Ordinal Outcomes

Contributions: New identification strategy & diagnostic tool

- Introduce a latent variable framework
  - Extend the latent utility representation of the standard probit/logit

- Apply the assumption by Athey & Imbens (2006) on the latent variable scale
  - Assumes temporal changes in quantiles are identical across two groups

Difference-in-Differences for Ordinal Outcomes

Contributions: New identification strategy & diagnostic tool

- Introduce a latent variable framework
  - Extend the latent utility representation of the standard probit/logit

- Apply the assumption by Athey & Imbens (2006) on the latent variable scale
  - Assumes temporal changes in quantiles are identical across two groups

  $\leadsto$ Avoid imposing the linearity assumption in the standard DiD

Difference-in-Differences for Ordinal Outcomes

Contributions: New identification strategy & diagnostic tool

- Introduce a latent variable framework
  - Extend the latent utility representation of the standard probit/logit

- Apply the assumption by Athey & Imbens (2006) on the latent variable scale
  - Assumes temporal changes in quantiles are identical across two groups

  $\rightsquigarrow$ Avoid imposing the linearity assumption in the standard DiD

- Derive a diagnostic with one additional pre-treatment period

Difference-in-Differences for Ordinal Outcomes

Contributions: New identification strategy & diagnostic tool

- Introduce a latent variable framework
  - Extend the latent utility representation of the standard probit/logit
- Apply the assumption by Athey & Imbens (2006) on the latent variable scale
  - Assumes temporal changes in quantiles are identical across two groups

  $\rightsquigarrow$ Avoid imposing the linearity assumption in the standard DiD
- Derive a diagnostic with one additional pre-treatment period
  - Analogous to the pre-treatment trend check in the standard DiD

Difference-in-Differences for Ordinal Outcomes

Contributions: New identification strategy & diagnostic tool

- Introduce a latent variable framework
  - Extend the latent utility representation of the standard probit/logit

- Apply the assumption by Athey & Imbens (2006) on the latent variable scale
  - Assumes temporal changes in quantiles are identical across two groups

  $\rightsquigarrow$ Avoid imposing the linearity assumption in the standard DiD

- Derive a diagnostic with one additional pre-treatment period
  - Analogous to the pre-treatment trend check in the standard DiD

  $\rightsquigarrow$ Equivalence based test to assess the plausibility of the assumption

Mass Shootings and Attitudes toward Gun Control

- Recent debate on the topic (Barney & Schaffner, 2019; Hartman& Newman, 2019; Newman & Hartman, 2019)

Mass Shootings and Attitudes toward Gun Control

- Recent debate on the topic (Barney & Schaffner, 2019; Hartman& Newman, 2019; Newman & Hartman, 2019)
- Proximity to the shootings as a treatment (dichotomized by 100 miles)

Mass Shootings and Attitudes toward Gun Control

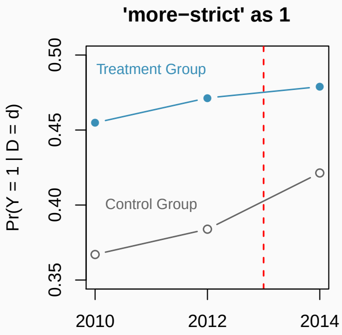
- Recent debate on the topic (Barney & Schaffner, 2019; Hartman& Newman, 2019; Newman & Hartman, 2019)
- Proximity to the shootings as a treatment (dichotomized by 100 miles)
- Ordinal survey outcome: `less-strict`, `keep-the-same` and `more-strict`

```
In general, do you feel that laws covering the sale of firearms
should be made more strict, less strict, or kept as they are?
(0) Less Strict; (1) Kept As They Are; (2) More Strict.
```

Mass Shootings and Attitudes toward Gun Control

- Recent debate on the topic (Barney & Schaffner, 2019; Hartman& Newman, 2019; Newman & Hartman, 2019)
- Proximity to the shootings as a treatment (dichotomized by 100 miles)
- Ordinal survey outcome: `less-strict`, `keep-the-same` and `more-strict`

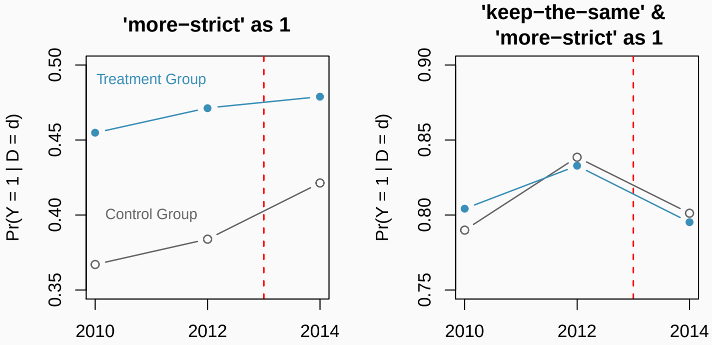
```
In general, do you feel that laws covering the sale of firearms
should be made more strict, less strict, or kept as they are?
```



Mass Shootings and Attitudes toward Gun Control

- Recent debate on the topic (Barney & Schaffner, 2019; Hartman& Newman, 2019; Newman & Hartman, 2019)
- Proximity to the shootings as a treatment (dichotomized by 100 miles)
- Ordinal survey outcome: `less-strict`, `keep-the-same` and `more-strict`

```
In general, do you feel that laws covering the sale of firearms
should be made more strict, less strict, or kept as they are?
```



continuous

DiD for Ordinal Outcomes: Setup

- Observed outcome: $Y_{it} \in \{0, \ldots, J-1\}$ for $i = 1, \ldots, n$ and $t \in \{0, 1\}$

DiD for Ordinal Outcomes: Setup

- Observed outcome: $Y_{it} \in \{0, \ldots, J-1\}$ for $i = 1, \ldots, n$ and $t \in \{0, 1\}$
- Binary treatment: $D_i \in \{0, 1\}$

DiD for Ordinal Outcomes: Setup

- Observed outcome: $Y_{it} \in \{0, \ldots, J-1\}$ for $i = 1, \ldots, n$ and $t \in \{0, 1\}$
- Binary treatment: $D_i \in \{0, 1\}$
- Potential outcome: $Y_{it}(d)$ for $d \in \{0, 1\}$

DiD for Ordinal Outcomes: Setup

- Observed outcome: $Y_{it} \in \{0, \ldots, J-1\}$ for $i = 1, \ldots, n$ and $t \in \{0, 1\}$
- Binary treatment: $D_i \in \{0, 1\}$
- Potential outcome: $Y_{it}(d)$ for $d \in \{0, 1\}$
- Estimand: Differences in choice probabilities for the treated

$$\zeta_j = \Pr(Y_{i1}(1) = j \mid D_i = 1) - \Pr(Y_{i1}(0) = j \mid D_i = 1)$$

DiD for Ordinal Outcomes: Setup

- Observed outcome: $Y_{it} \in \{0, \ldots, J-1\}$ for $i = 1, \ldots, n$ and $t \in \{0, 1\}$
- Binary treatment: $D_i \in \{0, 1\}$
- Potential outcome: $Y_{it}(d)$ for $d \in \{0, 1\}$
- Estimand: Differences in choice probabilities for the treated

$$\zeta_j = \Pr(Y_{i1}(1) = j \mid D_i = 1) - \Pr(Y_{i1}(0) = j \mid D_i = 1)$$

- Example: Difference in prob. of choosing `more-strict` under two conditions

DiD for Ordinal Outcomes: Setup

- Observed outcome: $Y_{it} \in \{0, \ldots, J-1\}$ for $i = 1, \ldots, n$ and $t \in \{0, 1\}$
- Binary treatment: $D_i \in \{0, 1\}$
- Potential outcome: $Y_{it}(d)$ for $d \in \{0, 1\}$
- Estimand: Differences in choice probabilities for the treated

$$\zeta_j = \Pr(Y_{i1}(1) = j \mid D_i = 1) - \Pr(Y_{i1}(0) = j \mid D_i = 1)$$

- Example: Difference in prob. of choosing `more-strict` under two conditions
- $\Pr(Y_{i1}(1) = j \mid D_i = 1)$ is observed from the data: $\Pr(Y_{i1} = j \mid D_i = 1)$

DiD for Ordinal Outcomes: Setup

- Observed outcome: $Y_{it} \in \{0, \ldots, J-1\}$ for $i = 1, \ldots, n$ and $t \in \{0, 1\}$
- Binary treatment: $D_i \in \{0, 1\}$
- Potential outcome: $Y_{it}(d)$ for $d \in \{0, 1\}$
- Estimand: Differences in choice probabilities for the treated

$$\zeta_j = \Pr(Y_{i1}(1) = j \mid D_i = 1) - \Pr(Y_{i1}(0) = j \mid D_i = 1)$$

- Example: Difference in prob. of choosing `more-strict` under two conditions
- $\Pr(Y_{i1}(1) = j \mid D_i = 1)$ is observed from the data: $\Pr(Y_{i1} = j \mid D_i = 1)$
- Need to identify $\Pr(Y_{i1}(0) = j \mid D_i = 1)$ with additional assumptions

Latent Variable Formulation

- Ordinal outcome: $Y_{dt} \sim Y_{it}(0) \mid D_i = d$

Latent Variable Formulation

- Ordinal outcome: $Y_{dt} \sim Y_{it}(0) \mid D_i = d$

  $\rightsquigarrow Y_{11}$ is the counterfactual outcome

Latent Variable Formulation

- Ordinal outcome: $Y_{dt} \sim Y_{it}(0) \mid D_i = d$

  $\rightsquigarrow Y_{11}$ is the counterfactual outcome
- Latent "utility" generating the ordinal outcome: $Y^*_{dt} \in \mathbb{R}$

Latent Variable Formulation

- Ordinal outcome: $Y_{dt} \sim Y_{it}(0) \mid D_i = d$

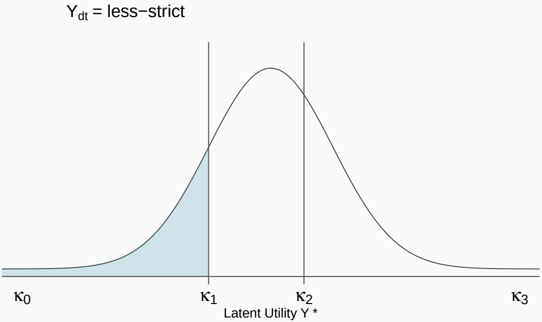
  $\rightsquigarrow$ Y_{11} is the counterfactual outcome
- Latent "utility" generating the ordinal outcome: $Y^*_{dt} \in \mathbb{R}$
- Index model: Mapping Y^*_{dt} to Y_{dt}

$$Y_{dt} = \begin{cases} 0 & \text{if} \quad \kappa_1 \geq Y^*_{dt} \geq \kappa_0 \\ j & \text{if} \quad \kappa_{j+1} \geq Y^*_{dt} \geq \kappa_j \\ J-1 & \text{if} \quad \kappa_J \geq Y^*_{dt} \geq \kappa_{J-1} \end{cases}$$

Latent Variable Formulation

- Ordinal outcome: $Y_{dt} \sim Y_{it}(0) \mid D_i = d$
  $\rightsquigarrow Y_{11}$ is the counterfactual outcome
- Latent "utility" generating the ordinal outcome: $Y^*_{dt} \in \mathbb{R}$
- Index model: Mapping Y^*_{dt} to Y_{dt}

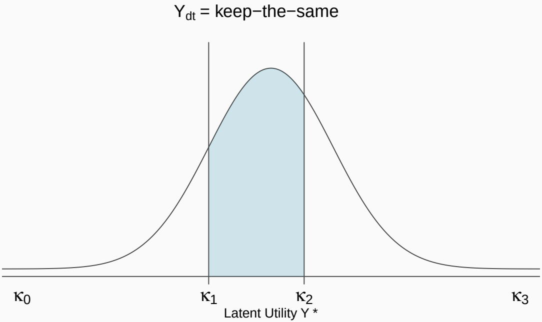
$$Y_{dt} = \begin{cases} 0 & \text{if} \quad \kappa_1 \geq Y^*_{dt} \geq \kappa_0 \\ j & \text{if} \quad \kappa_{j+1} \geq Y^*_{dt} \geq \kappa_j \\ J-1 & \text{if} \quad \kappa_J \geq Y^*_{dt} \geq \kappa_{J-1} \end{cases}$$



Latent Variable Formulation

- Ordinal outcome: $Y_{dt} \sim Y_{it}(0) \mid D_i = d$
  $\rightsquigarrow Y_{11}$ is the counterfactual outcome
- Latent "utility" generating the ordinal outcome: $Y^*_{dt} \in \mathbb{R}$
- Index model: Mapping Y^*_{dt} to Y_{dt}

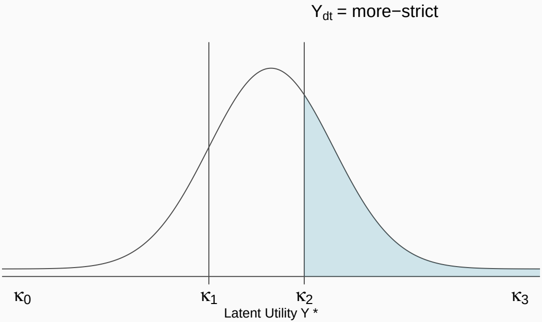
$$Y_{dt} = \begin{cases} 0 & \text{if} \quad \kappa_1 \geq Y^*_{dt} \geq \kappa_0 \\ j & \text{if} \quad \kappa_{j+1} \geq Y^*_{dt} \geq \kappa_j \\ J-1 & \text{if} \quad \kappa_J \geq Y^*_{dt} \geq \kappa_{J-1} \end{cases}$$



Latent Variable Formulation

- Ordinal outcome: $Y_{dt} \sim Y_{it}(0) \mid D_i = d$
  $\rightsquigarrow Y_{11}$ is the counterfactual outcome
- Latent "utility" generating the ordinal outcome: $Y^*_{dt} \in \mathbb{R}$
- Index model: Mapping Y^*_{dt} to Y_{dt}

$$Y_{dt} = \begin{cases} 0 & \text{if} \quad \kappa_1 \geq Y^*_{dt} \geq \kappa_0 \\ j & \text{if} \quad \kappa_{j+1} \geq Y^*_{dt} \geq \kappa_j \\ J-1 & \text{if} \quad \kappa_J \geq Y^*_{dt} \geq \kappa_{J-1} \end{cases}$$



Latent Variable Formulation

- Ordinal outcome: $Y_{dt} \sim Y_{it}(0) \mid D_i = d$

  $\rightsquigarrow Y_{11}$ is the counterfactual outcome
- Latent "utility" generating the ordinal outcome: $Y^*_{dt} \in \mathbb{R}$
- Index model: Mapping Y^*_{dt} to Y_{dt}

$$Y_{dt} = \begin{cases} 0 & \text{if} \quad \kappa_1 \geq Y^*_{dt} \geq \kappa_0 \\ j & \text{if} \quad \kappa_{j+1} \geq Y^*_{dt} \geq \kappa_j \\ J-1 & \text{if} \quad \kappa_J \geq Y^*_{dt} \geq \kappa_{J-1} \end{cases}$$

- Location-scale family: Imposing distribution on Y^*_{dt}

$$Y^*_{dt} \sim \underbrace{\mu_{dt}}_{\text{location}} + \underbrace{\sigma_{dt}}_{\text{scale}} U$$

where U belongs to a parametric family (e.g., normal, logistic, t-dist.)

Main Result

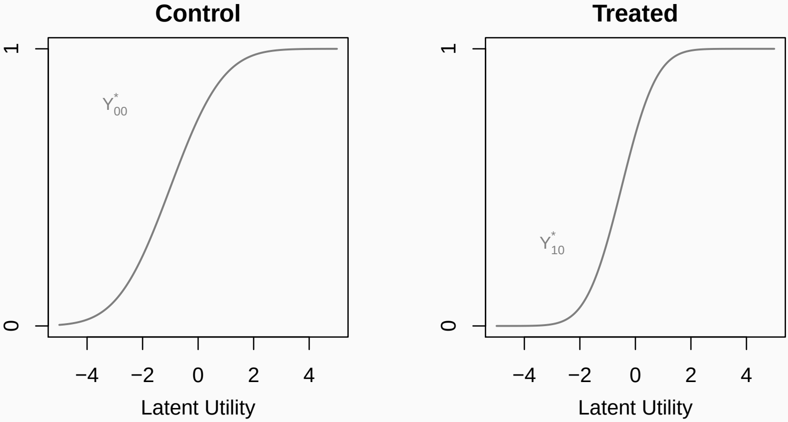
- Distributional parallel-trends assumption (Athey & Imbens 2006)

$$F_{Y^*_{00}}(F^{-1}_{Y^*_{01}}(v)) = F_{Y^*_{10}}(F^{-1}_{Y^*_{11}}(v)) \quad \forall v \in [0, 1]$$

Main Result

- Distributional parallel-trends assumption (Athey & Imbens 2006)

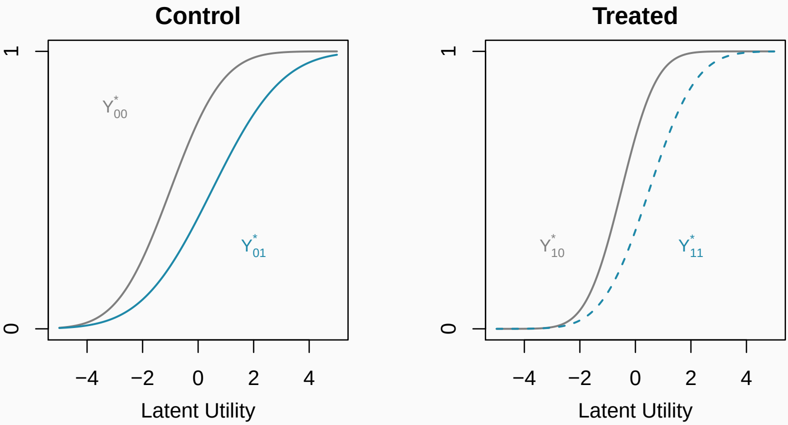
$$\underbrace{F_{Y^*_{00}}(F^{-1}_{Y^*_{01}}(v))}_{=q_0(v)} = \underbrace{F_{Y^*_{10}}(F^{-1}_{Y^*_{11}}(v))}_{=q_1(v)} \quad \forall v \in [0,1]$$



Main Result

- Distributional parallel-trends assumption (Athey & Imbens 2006)

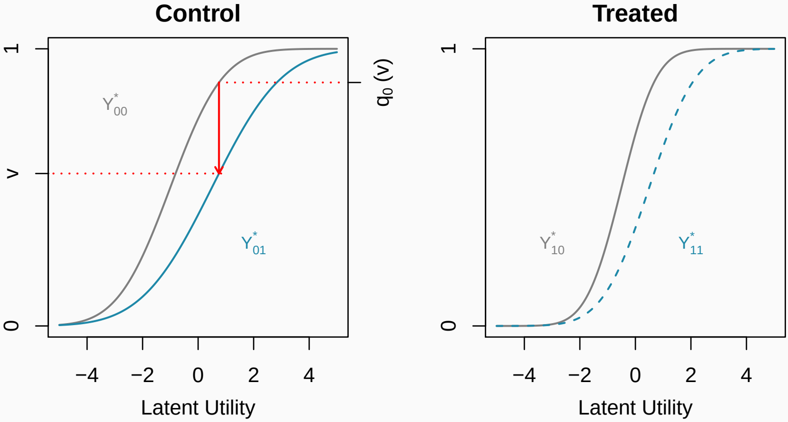
$$\underbrace{F_{Y^*_{00}}(F^{-1}_{Y^*_{01}}(v))}_{=q_0(v)} = \underbrace{F_{Y^*_{10}}(F^{-1}_{Y^*_{11}}(v))}_{=q_1(v)} \quad \forall v \in [0,1]$$



Main Result

- Distributional parallel-trends assumption (Athey & Imbens 2006)

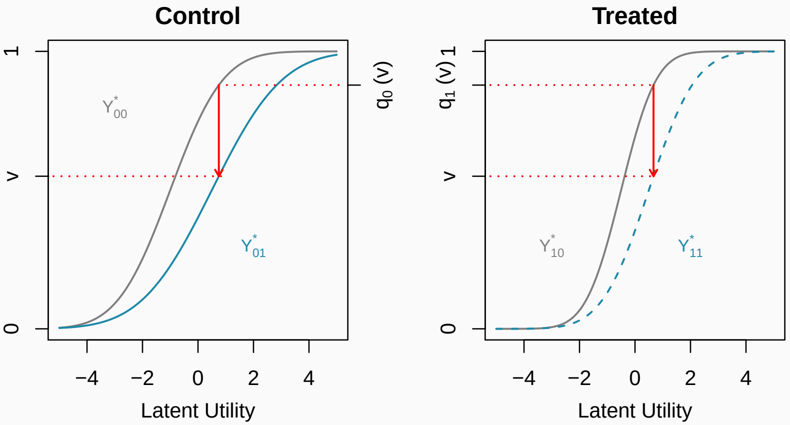
$$\underbrace{F_{Y^*_{00}}(F^{-1}_{Y^*_{01}}(v))}_{=q_0(v)} = \underbrace{F_{Y^*_{10}}(F^{-1}_{Y^*_{11}}(v))}_{=q_1(v)} \quad \forall v \in [0,1]$$



Main Result

- Distributional parallel-trends assumption (Athey & Imbens 2006)

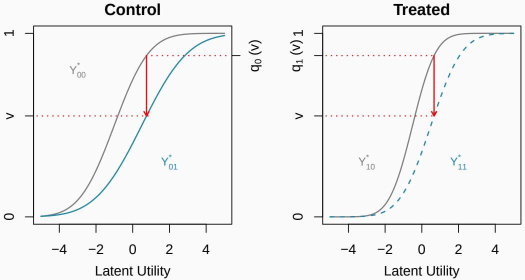
$$\underbrace{F_{Y^*_{00}}(F^{-1}_{Y^*_{01}}(v))}_{=q_0(v)} = \underbrace{F_{Y^*_{10}}(F^{-1}_{Y^*_{11}}(v))}_{=q_1(v)} \quad \forall v \in [0,1]$$



Main Result

- Distributional parallel-trends assumption (Athey & Imbens 2006)

$$\underbrace{F_{Y^*_{00}}(F^{-1}_{Y^*_{01}}(v))}_{=q_0(v)} = \underbrace{F_{Y^*_{10}}(F^{-1}_{Y^*_{11}}(v))}_{=q_1(v)} \quad \forall v \in [0,1]$$



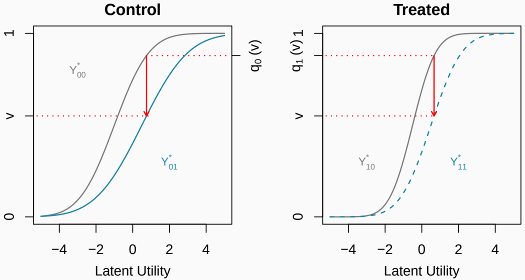
- Proposition: the distribution of the counterfactual latent variable given by

$$\mu_{11} = \mu_{10} + \frac{\mu_{01} - \mu_{00}}{\sigma_{00}/\sigma_{10}}, \quad \text{and} \quad \sigma_{11} = \frac{\sigma_{10}\sigma_{01}}{\sigma_{00}}$$

Main Result

- Distributional parallel-trends assumption (Athey & Imbens 2006)

$$\underbrace{F_{Y^*_{00}}(F^{-1}_{Y^*_{01}}(v))}_{=q_0(v)} = \underbrace{F_{Y^*_{10}}(F^{-1}_{Y^*_{11}}(v))}_{=q_1(v)} \quad \forall v \in [0,1]$$



- Proposition: the distribution of the counterfactual latent variable given by

$$\mu_{11} = \mu_{10} + \frac{\mu_{01} - \mu_{00}}{\sigma_{00}/\sigma_{10}}, \quad \text{and} \quad \sigma_{11} = \frac{\sigma_{10}\sigma_{01}}{\sigma_{00}}$$

- When variances are constant $\sigma_{dt} = \sigma$, recovers the usual parallel trends form

$$\mu_{11} - \mu_{10} = \mu_{01} - \mu_{00}$$

Estimation

- Impose a distribution on U (base distribution): $U \sim \mathcal{N}(0, 1)$

Estimation

- Impose a distribution on U (base distribution): $U \sim \mathcal{N}(0, 1)$

  $\rightsquigarrow$ Estimate parameters by MLE (e.g., variant of the ordered probit)

Estimation

- Impose a distribution on U (base distribution): $U \sim \mathcal{N}(0, 1)$

  $\rightsquigarrow$ Estimate parameters by MLE (e.g., variant of the ordered probit)
- Plug-in estimator for the counter-factual distribution

$$\widehat{\mu}_{11} \leftarrow \widehat{\mu}_{10} + \frac{\widehat{\mu}_{01} - \widehat{\mu}_{00}}{\widehat{\sigma}_{00}/\widehat{\sigma}_{10}}, \quad \text{and} \quad \widehat{\sigma}_{11} \leftarrow \frac{\widehat{\sigma}_{10}\widehat{\sigma}_{01}}{\widehat{\sigma}_{00}}$$

Estimation

- Impose a distribution on U (base distribution): $U \sim \mathcal{N}(0,1)$
  $\rightsquigarrow$ Estimate parameters by MLE (e.g., variant of the ordered probit)
- Plug-in estimator for the counter-factual distribution

$$\widehat{\mu}_{11} \leftarrow \widehat{\mu}_{10} + \frac{\widehat{\mu}_{01} - \widehat{\mu}_{00}}{\widehat{\sigma}_{00}/\widehat{\sigma}_{10}}, \quad \text{and} \quad \widehat{\sigma}_{11} \leftarrow \frac{\widehat{\sigma}_{10}\widehat{\sigma}_{01}}{\widehat{\sigma}_{00}}$$

- Obtain causal estimates: $\widehat{\boldsymbol{\zeta}} = (\widehat{\zeta}_1, \ldots, \widehat{\zeta}_{J-1})^\top$

$$\widehat{\zeta}_j = \underbrace{\frac{1}{n_1}\sum_{i=1}^{n} D_i \mathbf{1}\{Y_{i1} = j\}}_{=\widehat{\Pr}(Y_{i1}(1)=j|D_i=1)} - \underbrace{\Big\{\Phi\big((\widehat{\kappa}_{j+1} - \widehat{\mu}_{11})/\widehat{\sigma}_{11}\big) - \Phi\big((\widehat{\kappa}_j - \widehat{\mu}_{11})/\widehat{\sigma}_{11}\big)\Big\}}_{=\widehat{\Pr}(Y_{i1}(0)=j|D_i=1)}$$

Estimation

- Impose a distribution on U (base distribution): $U \sim \mathcal{N}(0,1)$

  $\rightsquigarrow$ Estimate parameters by MLE (e.g., variant of the ordered probit)
- Plug-in estimator for the counter-factual distribution

$$\widehat{\mu}_{11} \leftarrow \widehat{\mu}_{10} + \frac{\widehat{\mu}_{01} - \widehat{\mu}_{00}}{\widehat{\sigma}_{00}/\widehat{\sigma}_{10}}, \quad \text{and} \quad \widehat{\sigma}_{11} \leftarrow \frac{\widehat{\sigma}_{10}\widehat{\sigma}_{01}}{\widehat{\sigma}_{00}}$$

- Obtain causal estimates: $\widehat{\boldsymbol{\zeta}} = (\widehat{\zeta}_1, \ldots, \widehat{\zeta}_{J-1})^\top$

$$\widehat{\zeta}_j = \underbrace{\frac{1}{n_1}\sum_{i=1}^{n} D_i \mathbf{1}\{Y_{i1} = j\}}_{=\widehat{\Pr}(Y_{i1}(1)=j \mid D_i=1)} - \underbrace{\left\{\Phi\Big((\widehat{\kappa}_{j+1} - \widehat{\mu}_{11})/\widehat{\sigma}_{11}\Big) - \Phi\Big((\widehat{\kappa}_{j} - \widehat{\mu}_{11})/\widehat{\sigma}_{11}\Big)\right\}}_{=\widehat{\Pr}(Y_{i1}(0)=j \mid D_i=1)}$$

Estimation

- Impose a distribution on U (base distribution): $U \sim \mathcal{N}(0, 1)$

  $\rightsquigarrow$ Estimate parameters by MLE (e.g., variant of the ordered probit)
- Plug-in estimator for the counter-factual distribution

$$\widehat{\mu}_{11} \leftarrow \widehat{\mu}_{10} + \frac{\widehat{\mu}_{01} - \widehat{\mu}_{00}}{\widehat{\sigma}_{00}/\widehat{\sigma}_{10}}, \quad \text{and} \quad \widehat{\sigma}_{11} \leftarrow \frac{\widehat{\sigma}_{10}\widehat{\sigma}_{01}}{\widehat{\sigma}_{00}}$$

- Obtain causal estimates: $\widehat{\boldsymbol{\zeta}} = (\widehat{\zeta}_1, \ldots, \widehat{\zeta}_{J-1})^\top$

$$\widehat{\zeta}_j = \underbrace{\frac{1}{n_1}\sum_{i=1}^{n} D_i \mathbf{1}\{Y_{i1} = j\}}_{=\widehat{\Pr}(Y_{i1}(1)=j|D_i=1)} - \underbrace{\left\{\Phi\left((\widehat{\kappa}_{j+1} - \widehat{\mu}_{11})/\widehat{\sigma}_{11}\right) - \Phi\left((\widehat{\kappa}_{j} - \widehat{\mu}_{11})/\widehat{\sigma}_{11}\right)\right\}}_{=\widehat{\Pr}(Y_{i1}(0)=j|D_i=1)}$$

Estimation

- Impose a distribution on U (base distribution): $U \sim \mathcal{N}(0,1)$

  $\rightsquigarrow$ Estimate parameters by MLE (e.g., variant of the ordered probit)
- Plug-in estimator for the counter-factual distribution

$$\widehat{\mu}_{11} \leftarrow \widehat{\mu}_{10} + \frac{\widehat{\mu}_{01} - \widehat{\mu}_{00}}{\widehat{\sigma}_{00}/\widehat{\sigma}_{10}}, \quad \text{and} \quad \widehat{\sigma}_{11} \leftarrow \frac{\widehat{\sigma}_{10}\widehat{\sigma}_{01}}{\widehat{\sigma}_{00}}$$

- Obtain causal estimates: $\widehat{\boldsymbol{\zeta}} = (\widehat{\zeta}_1, \dots, \widehat{\zeta}_{J-1})^\top$

$$\widehat{\zeta}_j = \underbrace{\frac{1}{n_1}\sum_{i=1}^{n} D_i \mathbf{1}\{Y_{i1} = j\}}_{=\widehat{\Pr}(Y_{i1}(1)=j|D_i=1)} - \underbrace{\left\{\Phi\left((\widehat{\kappa}_{j+1} - \widehat{\mu}_{11})/\widehat{\sigma}_{11}\right) - \Phi\left((\widehat{\kappa}_j - \widehat{\mu}_{11})/\widehat{\sigma}_{11}\right)\right\}}_{=\widehat{\Pr}(Y_{i1}(0)=j|D_i=1)}$$

- Obtain variance estimates by the block-bootstrap

▸ Cutoff ▸ Mean-Variance

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

  ```
  In general, do you feel that laws covering the sale of firearms
  should be made more strict, less strict, or kept as they are?

  (0) Less Strict; (1) Kept As They Are; (2) More Strict.
  ```

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

  ```
  In general, do you feel that laws covering the sale of firearms
  should be made more strict, less strict, or kept as they are?

  (0) Less Strict; (1) Kept As They Are; (2) More Strict.
  ```

- Respondents are “treated” if living within 100 miles from the shootings

Defintion of Mass Shootings

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

```
In general, do you feel that laws covering the sale of firearms
should be made more strict, less strict, or kept as they are?

(0) Less Strict; (1) Kept As They Are; (2) More Strict.
```

- Respondents are “treated” if living within 100 miles from the shootings
  - 16 mass-shootings coded at the zip-code level

Defintion of Mass Shootings

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

```
In general, do you feel that laws covering the sale of firearms
should be made more strict, less strict, or kept as they are?

(0) Less Strict; (1) Kept As They Are; (2) More Strict.
```

- Respondents are “treated” if living within 100 miles from the shootings
  - 16 mass-shootings coded at the zip-code level
  - Approx. 30% of respondents (out of 16620) are treated

Defintion of Mass Shootings

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

```
In general, do you feel that laws covering the sale of firearms
should be made more strict, less strict, or kept as they are?

(0) Less Strict; (1) Kept As They Are; (2) More Strict.
```

- Respondents are “treated” if living within 100 miles from the shootings
  - 16 mass-shootings coded at the zip-code level
  - Approx. 30% of respondents (out of 16620) are treated

- Subgroups

Defintion of Mass Shootings

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

```
In general, do you feel that laws covering the sale of firearms
should be made more strict, less strict, or kept as they are?

(0) Less Strict; (1) Kept As They Are; (2) More Strict.
```

- Respondents are "treated" if living within 100 miles from the shootings
  - 16 mass-shootings coded at the zip-code level
  - Approx. 30% of respondents (out of 16620) are treated
- Subgroups
  - `Partisanship`: 3-point scale party self-identification

Defintion of Mass Shootings

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

  `In general, do you feel that laws covering the sale of firearms should be made more strict, less strict, or kept as they are?`

  `(0) Less Strict; (1) Kept As They Are; (2) More Strict.`

- Respondents are “treated” if living within 100 miles from the shootings
  - 16 mass-shootings coded at the zip-code level
  - Approx. 30% of respondents (out of 16620) are treated

- Subgroups
  - `Partisanship`: 3-point scale party self-identification
  - `Prior-exposure`: Living in areas with mass shootings in the past 10 years

Defintion of Mass Shootings

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

```
In general, do you feel that laws covering the sale of firearms
should be made more strict, less strict, or kept as they are?

(0) Less Strict; (1) Kept As They Are; (2) More Strict.
```

- Respondents are "treated" if living within 100 miles from the shootings
  - 16 mass-shootings coded at the zip-code level
  - Approx. 30% of respondents (out of 16620) are treated

- Subgroups
  - `Partisanship`: 3-point scale party self-identification
  - `Prior-exposure`: Living in areas with mass shootings in the past 10 years

- Previous studies used:

▸ Defintion of Mass Shootings

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

  ```
  In general, do you feel that laws covering the sale of firearms
  should be made more strict, less strict, or kept as they are?

  (0) Less Strict; (1) Kept As They Are; (2) More Strict.
  ```

- Respondents are “treated” if living within 100 miles from the shootings
  - 16 mass-shootings coded at the zip-code level
  - Approx. 30% of respondents (out of 16620) are treated

- Subgroups
  - `Partisanship`: 3-point scale party self-identification
  - `Prior-exposure`: Living in areas with mass shootings in the past 10 years

- Previous studies used:
  - Ordinal logit with RE (NH19 and HN19)

Defintion of Mass Shootings

Revisiting the Empirical Application

- Two-wave (2010-12) panel from Cooperative Congressional Election Study:

  `In general, do you feel that laws covering the sale of firearms should be made more strict, less strict, or kept as they are?`

  `(0) Less Strict; (1) Kept As They Are; (2) More Strict.`

- Respondents are “treated” if living within 100 miles from the shootings
  - 16 mass-shootings coded at the zip-code level
  - Approx. 30% of respondents (out of 16620) are treated

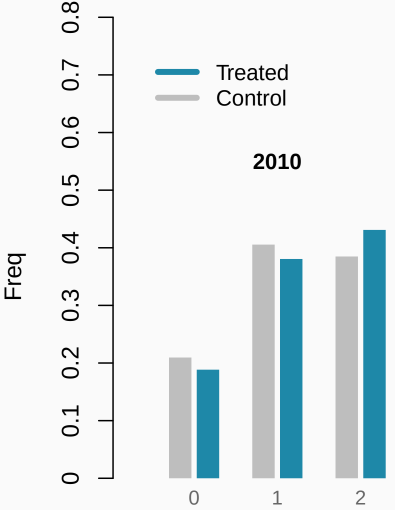
- Subgroups
  - `Partisanship`: 3-point scale party self-identification
  - `Prior-exposure`: Living in areas with mass shootings in the past 10 years

- Previous studies used:
  - Ordinal logit with RE (NH19 and HN19)
  - Linear two-way FE (BS19)

Defintion of Mass Shootings

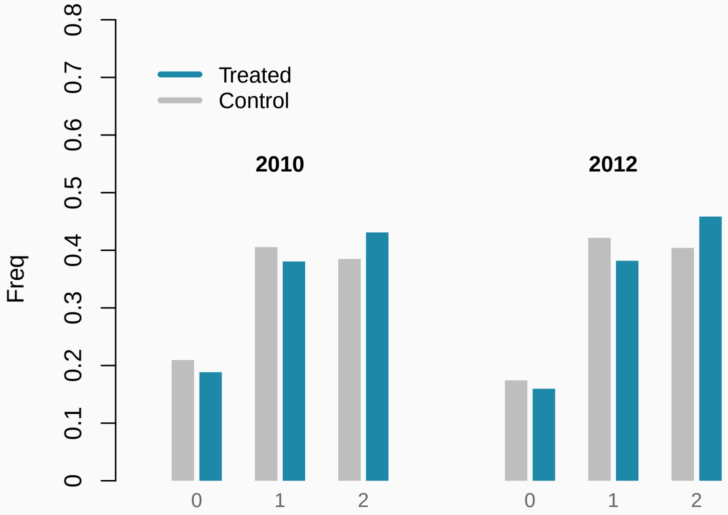
Distribution of Outcome in 2010 & 2012

0: `less-strict`; 1: `keep-the-same`; 2: `more-strict`



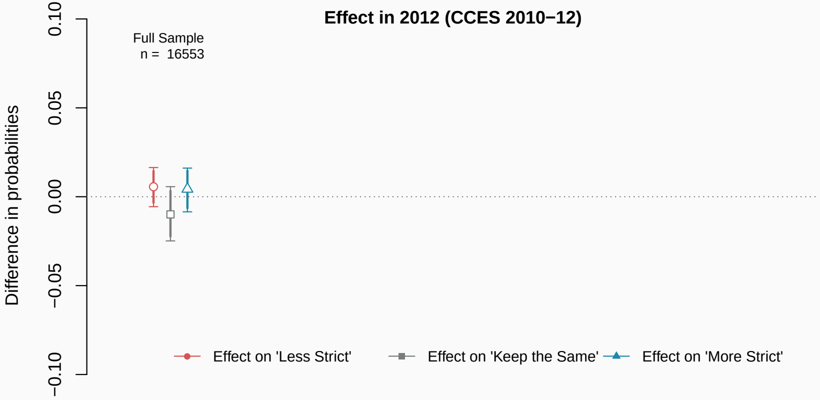
Distribution of Outcome in 2010 & 2012

0: `less-strict`; 1: `keep-the-same`; 2: `more-strict`



Results

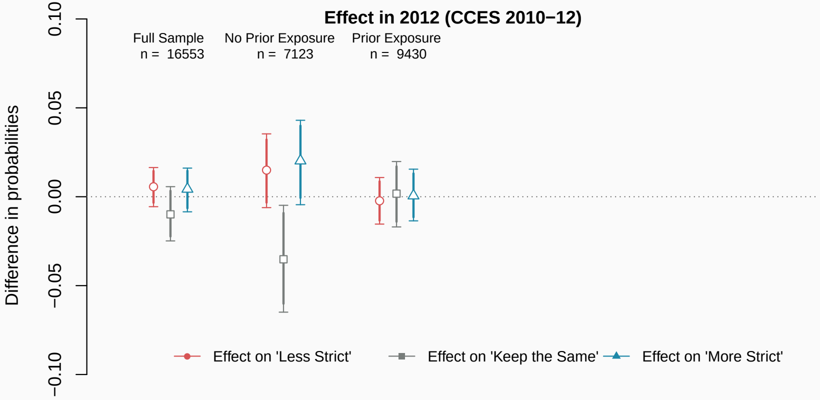
- Estimate $\zeta_j = \Pr(Y_{i1}(1) = j \mid D_i = 1) - \Pr(Y_{i1}(0) = j \mid D_i = 1)$
- Block-bootstrap at the zip-code level



25-miles PID-7 Cumulative effect Bounds

Results

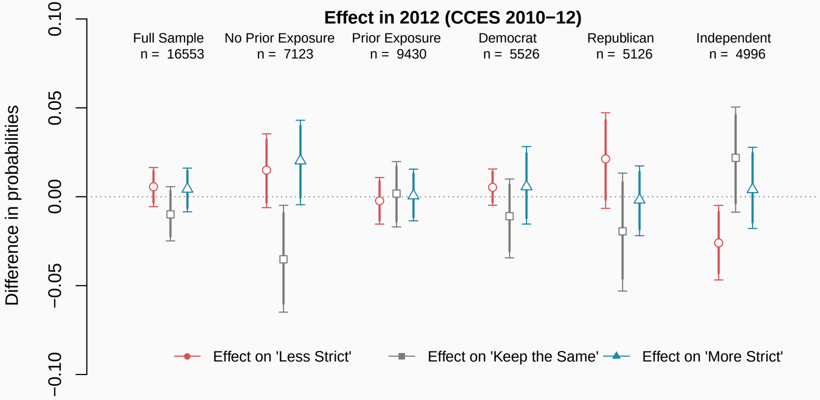
- Estimate $\zeta_j = \Pr(Y_{i1}(1) = j \mid D_i = 1) - \Pr(Y_{i1}(0) = j \mid D_i = 1)$
- Block-bootstrap at the zip-code level



▸ 25-miles ▸ PID-7 ▸ Cumulative effect ▸ Bounds

Results

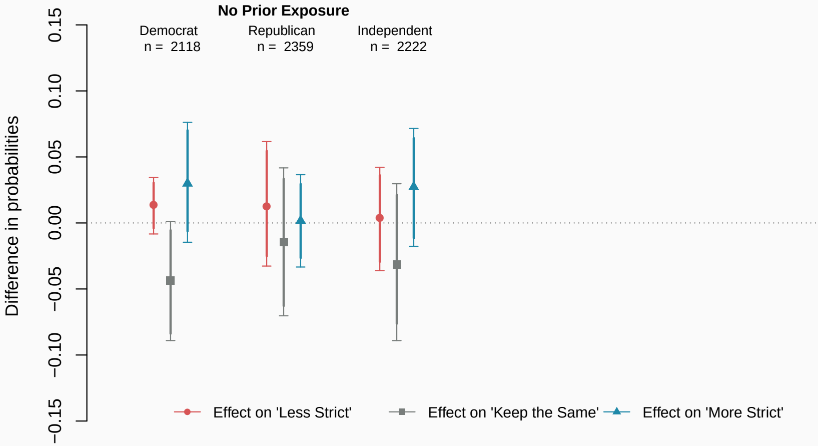
- Estimate $\zeta_j = \Pr(Y_{i1}(1) = j \mid D_i = 1) - \Pr(Y_{i1}(0) = j \mid D_i = 1)$
- Block-bootstrap at the zip-code level



25-miles PID-7 Cumulative effect Bounds

Results

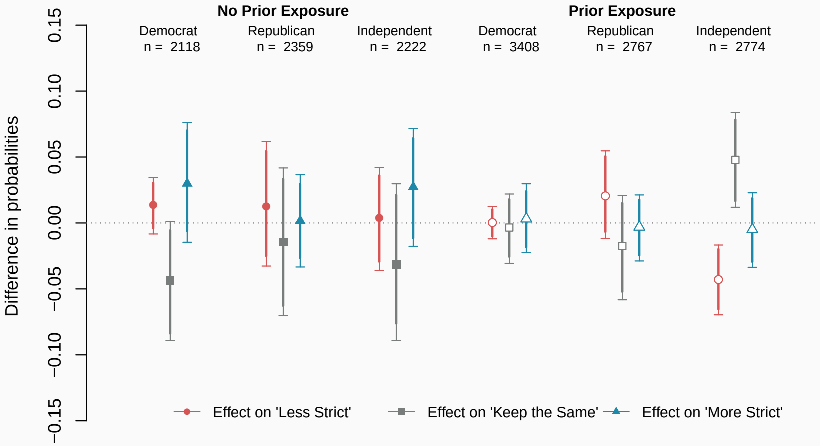
- Estimate $\zeta_j = \Pr(Y_{i1}(1) = j \mid D_i = 1) - \Pr(Y_{i1}(0) = j \mid D_i = 1)$
- Block-bootstrap at the zip-code level



▸ 25-miles ▸ PID-7 ▸ Cumulative effect ▸ Bounds

Results

- Estimate $\zeta_j = \Pr(Y_{i1}(1) = j \mid D_i = 1) - \Pr(Y_{i1}(0) = j \mid D_i = 1)$
- Block-bootstrap at the zip-code level



25-miles PID-7 Cumulative effect Bounds

Assessing the Distributional Parallel Trends Assumption

- Additional pre-treatment time-periods $\rightsquigarrow$ Assessment of the distributional PT

Assessing the Distributional Parallel Trends Assumption

- Additional pre-treatment time-periods $\rightsquigarrow$ Assessment of the distributional PT
- If the assumption holds for the pre-treatment, we have $\tilde{q}_1(v) - \tilde{q}_0(v) = 0$

$$\tilde{q}_d(v) = F_{Y^*_{d0}}(F^{-1}_{Y^*_{d1}}(v))$$

Assessing the Distributional Parallel Trends Assumption

- Additional pre-treatment time-periods $\rightsquigarrow$ Assessment of the distributional PT
- If the assumption holds for the pre-treatment, we have $\tilde{q}_1(v) - \tilde{q}_0(v) = 0$

$$\tilde{q}_d(v) = F_{Y^*_{d0}}(F^{-1}_{Y^*_{d1}}(v))$$

- Test the following non-equivalence null (i.e., H_0: Assumption does not hold)

$$H_0 : \max_{v \in [0,1]} |\tilde{q}_1(v) - \tilde{q}_0(v)| > \delta \quad \text{vs} \quad H_1 : \max_{v \in [0,1]} |\tilde{q}_1(v) - \tilde{q}_0(v)| \leq \delta$$

Assessing the Distributional Parallel Trends Assumption

- Additional pre-treatment time-periods $\rightsquigarrow$ Assessment of the distributional PT
- If the assumption holds for the pre-treatment, we have $\tilde{q}_1(v) - \tilde{q}_0(v) = 0$

$$\tilde{q}_d(v) = F_{Y^*_{d0}}(F^{-1}_{Y^*_{d1}}(v))$$

- Test the following non-equivalence null (i.e., H_0: Assumption does not hold)

$$H_0: \max_{v \in [0,1]} |\tilde{q}_1(v) - \tilde{q}_0(v)| > \delta \quad \text{vs} \quad H_1: \max_{v \in [0,1]} |\tilde{q}_1(v) - \tilde{q}_0(v)| \leq \delta$$

$\rightsquigarrow$ Two one-sided tests (TOST)

$$H_0^+: \max_{v \in [0,1]} \{\tilde{q}_1(v) - \tilde{q}_0(v)\} > \delta \quad \text{and} \quad H_0^-: \max_{v \in [0,1]} \{\tilde{q}_1(v) - \tilde{q}_0(v)\} < -\delta$$

Assessing the Distributional Parallel Trends Assumption

- Additional pre-treatment time-periods $\rightsquigarrow$ Assessment of the distributional PT
- If the assumption holds for the pre-treatment, we have $\tilde{q}_1(v) - \tilde{q}_0(v) = 0$

$$\tilde{q}_d(v) = F_{Y^*_{d0}}(F^{-1}_{Y^*_{d1}}(v))$$

- Test the following non-equivalence null (i.e., H_0: Assumption does not hold)

$$H_0: \max_{v\in[0,1]} |\tilde{q}_1(v) - \tilde{q}_0(v)| > \delta \quad \text{vs} \quad H_1: \max_{v\in[0,1]} |\tilde{q}_1(v) - \tilde{q}_0(v)| \leq \delta$$

$\rightsquigarrow$ Two one-sided tests (TOST)

$$H_0^+: \max_{v\in[0,1]} \{\tilde{q}_1(v) - \tilde{q}_0(v)\} > \delta \quad \text{and} \quad H_0^-: \max_{v\in[0,1]} \{\tilde{q}_1(v) - \tilde{q}_0(v)\} < -\delta$$

- Construct one-sided point-wise CIs: $\widehat{U}_{1-\alpha}(v)$ and $\widehat{L}_{1-\alpha}(v)$

$$\text{reject}: H_0^+ \text{ at } \alpha \text{ level} \iff \max_{v\in[0,1]} \widehat{U}_{1-\alpha}(v) < \delta$$

$\rightsquigarrow$ We reject H_0 if we reject both H_0^+ and H_0^-

CI construction | Choose delta

Using Three-wave Panel to Assess the Assumption

- Some respondents of CCES 2010–12 panel are reinterviewed in 2014

Using Three-wave Panel to Assess the Assumption

- Some respondents of CCES 2010–12 panel are reinterviewed in 2014
- Focus on 2817 respondents who
  - Did not have the “prior exposure” as of 2010
  - Were not treated between 2010 and 2012

Using Three-wave Panel to Assess the Assumption

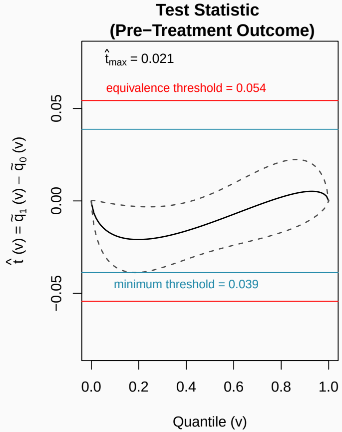
- Some respondents of CCES 2010–12 panel are reinterviewed in 2014
- Focus on 2817 respondents who
  - Did not have the "prior exposure" as of 2010
  - Were not treated between 2010 and 2012
- Approx. 25% of them (667) are newly treated between 2012 and 2014
  - 28 shootings

Using Three-wave Panel to Assess the Assumption

- Some respondents of CCES 2010–12 panel are reinterviewed in 2014
- Focus on 2817 respondents who
    - Did not have the "prior exposure" as of 2010
    - Were not treated between 2010 and 2012
- Approx. 25% of them (667) are newly treated between 2012 and 2014
    - 28 shootings
- Assess the distributional parallel trends assumption using 2010–12

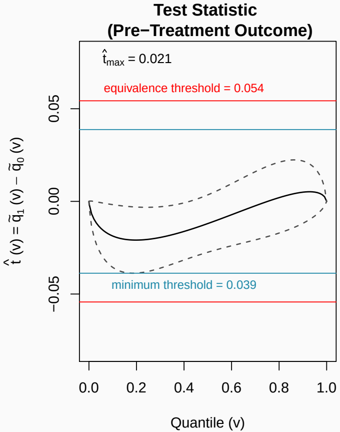
Using Three-wave Panel to Assess the Assumption

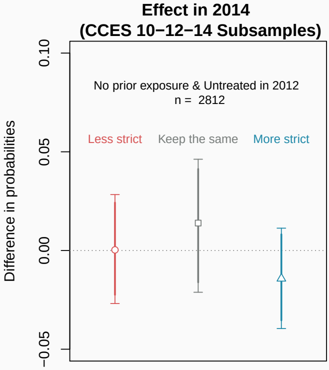
- Some respondents of CCES 2010–12 panel are reinterviewed in 2014
- Focus on 2817 respondents who
  - Did not have the "prior exposure" as of 2010
  - Were not treated between 2010 and 2012
- Approx. 25% of them (667) are newly treated between 2012 and 2014
  - 28 shootings
- Assess the distributional parallel trends assumption using 2010–12



Using Three-wave Panel to Assess the Assumption

- Some respondents of CCES 2010–12 panel are reinterviewed in 2014
- Focus on 2817 respondents who
  - Did not have the “prior exposure” as of 2010
  - Were not treated between 2010 and 2012
- Approx. 25% of them (667) are newly treated between 2012 and 2014
  - 28 shootings
- Assess the distributional parallel trends assumption using 2010–12





Concluding Remarks

- Difference-in-differences is widely used in social science research

Concluding Remarks

- Difference-in-differences is widely used in social science research
- Linearity assumption in DID is inappropriate for ordinal outcomes

Concluding Remarks

- Difference-in-differences is widely used in social science research
- Linearity assumption in DID is inappropriate for ordinal outcomes
- Propose a latent variable framework to address the issue

Concluding Remarks

- Difference-in-differences is widely used in social science research
- Linearity assumption in DID is inappropriate for ordinal outcomes
- Propose a latent variable framework to address the issue
- Revisit the recent debate on the relationship between the mass shootings and the attitudes toward gun control:

Concluding Remarks

- Difference-in-differences is widely used in social science research
- Linearity assumption in DID is inappropriate for ordinal outcomes
- Propose a latent variable framework to address the issue
- Revisit the recent debate on the relationship between the mass shootings and the attitudes toward gun control:
  - Find that effects are concentrated among Democrats who do not have "prior exposure" to shootings and among Independents

References

- Yamauchi, Soichiro. (2020). "Difference-in-Differences for Ordinal Outcomes: Application to the Effect of Mass Shootings on Attitudes towards Gun Control" *Working Paper*.
- `orddid`: R package for implementing the difference-in-difference for the ordinal outcomes. Available at github.com/soichiroy/orddid

Send comments and suggestions to
syamauchi@g.harvard.edu

For more information
soichiroy.github.io

Additional Results

Treating as Continuous Outcome

- Consider a cross-sectional setting:

$$\zeta_j = \Pr(Y_i(1) = j) - \Pr(Y_i(0) = j)$$

- Rescale the outcome: $\widetilde{Y}_i = Y_i/(J-1)$
- The difference-in-means estiamtor on $\widetilde{Y}_i$ can be written as

$$\widehat{\tau}_{\text{DiM}} = \sum_{j=1}^{J} (J-j)^{-1} \widehat{\zeta}_j$$

  where

$$\widehat{\tau}_{\text{DiM}} = \frac{1}{n_1} \sum_{i=1}^{n} D_i Y_i - \frac{1}{n_0} \sum_{i=1}^{n} (1 - D_i) Y_i$$

  $\rightsquigarrow$ Weighted average of $\widehat{\zeta}_j$ with weights are $1/(J-j)$
- This can potentially cancel out the effects: E.g., $\widehat{\zeta}_1 > 0$ and $\widehat{\zeta}_2 < 0$

Invariance of Causal Effect to Choice of Cutoffs

- **Proposition**: Suppose $Y_{it} \in \mathcal{J} \equiv \{0, 1, 2\}$ and $U \sim \mathcal{N}(0, 1)$. Let $\boldsymbol{\kappa}$ and $\boldsymbol{\kappa}'$ be the different sets of cutoffs. Then, for all $j \in \mathcal{J}$.

$$\widehat{\zeta}_j(\boldsymbol{\kappa}) = \widehat{\zeta}_j(\boldsymbol{\kappa}')$$

- **Intuition**:

(1) Assumption is imposed on the quantile scale (i.e., distributional PT)

$\rightsquigarrow$ counterfactual distribution is identified as long as quantile info. is preserved

(2) Changing cutoffs affect mean & scale $\rightsquigarrow$ transform the latent variables

(3) But quantile information is preserved, $\Pr(Y^* \leq \kappa_1) = \Pr(\tilde{Y}^* \leq \kappa_1')$

$$\int_{\kappa_1}^{\kappa_2} \phi((y^* - \mu_{00})/\sigma_{00}) dy^* = \underbrace{\Pr(Y_{00} = 1)}_{\text{observed prob.}} = \int_{\kappa_1'}^{\kappa_2'} \phi((y^* - \mu_{00}')/\sigma_{00}') dy^*$$

$\rightsquigarrow$ uniquely recovers the counterfactual distribution Y_{11}^*

Back

Identification of Latent Variables

- Suppose that the cutoffs are fixed at κ_1 and κ_2 for $Y_{dt} = j \in \{0, 1, 2\}$. Then, μ_{dt} and σ_{dt} in $Y^*_{dt} \sim \mu_{dt} + \sigma_{dt} U$ are uniquely identified from the observed probability distribution.
- Proof: Suppose that U has the density $f_U(u)$. Then, we can form a non-linear system of equations

$$\Pr(Y_{dt} = 0) = \int_{-\infty}^{\kappa_1} f_U((y^* - \mu_{dt})/\sigma_{dt}) dy^*$$
$$\Pr(Y_{dt} = 2) = \int_{\kappa_2}^{\infty} f_U((y^* - \mu_{dt})/\sigma_{dt}) dy^*$$

which are sufficient for estimating μ and σ.

Back

Alternative Formula of Identification

- Suppose $Y_{dt} = j \in \{0, 1, 2\}$. Let $v_1 = F_{01}(\kappa_1)$ and $v_2 = F_{01}(\kappa_2)$ where $\boldsymbol{\kappa}$ is a set of fixed cutoffs.
- Under the assumptions, we identify μ_{11} and σ_{11} by the following system of non-linear equations:

$$q_0(v_1) = \int_{-\infty}^{F_{10}^{-1}(v_1)} f_U((y^* - \mu_{11})/\sigma_{11}) dy^*$$
$$q_0(v_2) = \int_{-\infty}^{F_{10}^{-1}(v_2)} f_U((y^* - \mu_{11})/\sigma_{11}) dy^*.$$

Constructing Confidence Intervals for Testing

- Let $t(v) = \tilde{q}_1(v) - \tilde{q}_0(v)$
- Point-wise upper and lower $(1-\alpha)$ level confidence intervals:

$$\widehat{U}_{1-\alpha}(v) = \hat{t}(v) + \Phi^{-1}(1-\alpha)\sqrt{\mathrm{Var}(\hat{t}(v))/n}$$
$$\widehat{L}_{1-\alpha}(v) = \hat{t}(v) - \Phi^{-1}(1-\alpha)\sqrt{\mathrm{Var}(\hat{t}(v))/n}$$

- **Proposition**

$$\Pr\left(\max_{v\in[0,1]} t(v) \leq \max_{v'\in[0,1]} \widehat{U}_{1-\alpha}(v')\right) \geq 1-\alpha$$
$$\Pr\left(\min_{v\in[0,1]} t(v) \geq \min_{v'\in[0,1]} \widehat{L}_{1-\alpha}(v')\right) \geq 1-\alpha$$

Back

Choosing Delta

- Value of δ reflect the admissible level of “non equivalence”

  $\rightsquigarrow$ Larger values of δ correspond to lenient thresholds
- Calibrate δ based on the rejection threshold for the KS test

$$\delta_n = \min\left\{\sqrt{-\log(\alpha)/2}\sqrt{\frac{n_1+n_0}{n_1 n_0}}, 1\right\}$$

  and take $\alpha = 0.05$
- Can report the equivalence CI: minimum possible value of δ at α level

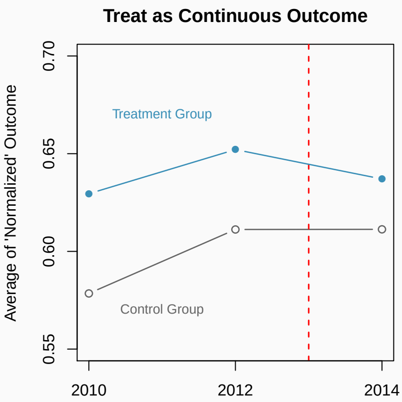
$$\delta_{\min,n} = \max_{v\in[0,1]}\left\{\left|\widehat{U}_{1-\alpha}(v)\right|, \left|\widehat{L}_{1-\alpha}(v)\right|\right\}$$

  $\rightsquigarrow$ Equivalence CI is given by $[-\delta_{\min,n}, \delta_{\min,n}]$

Back

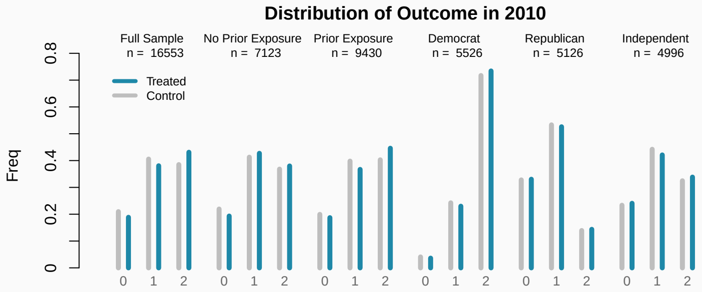
Additional Empirical Analysis

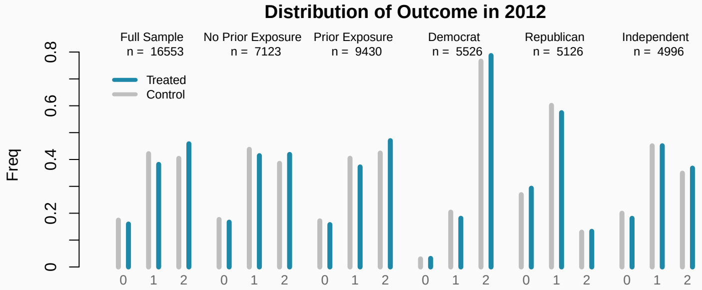
Treating as a Continuous Outcome



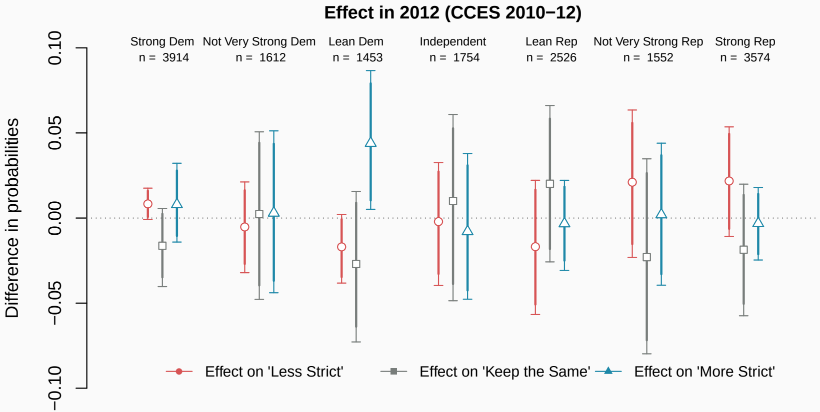
Back

Outcome Distributions by Sub-Group

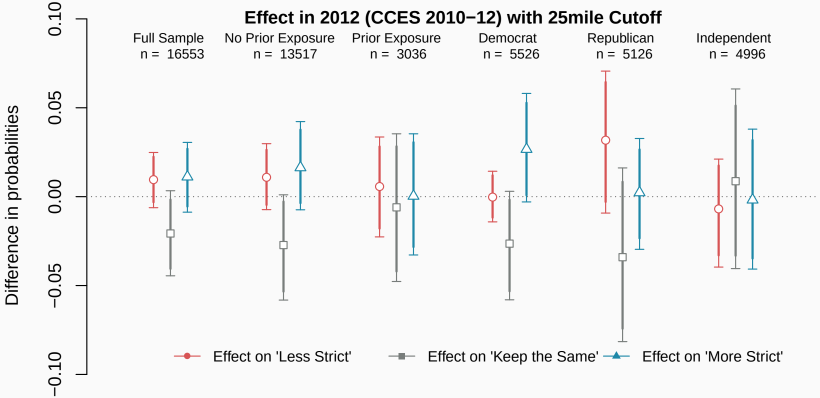




Party ID based on 7-point Scale Measure

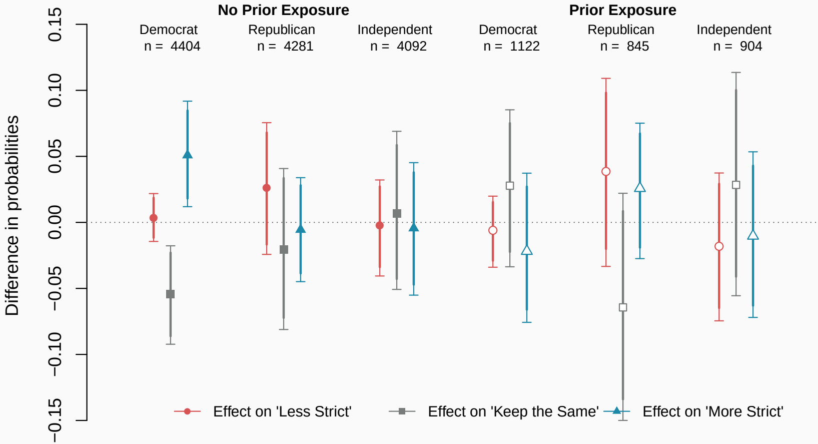


Different Distance Threshold: 25 Miles



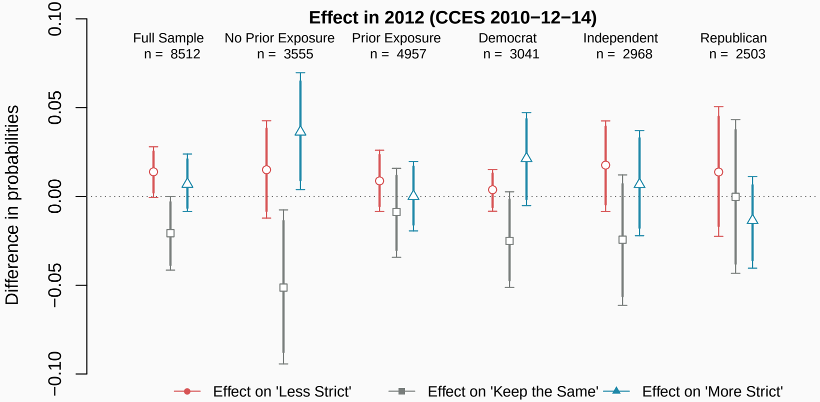
Back

Different Distance Threshold: 25 Miles



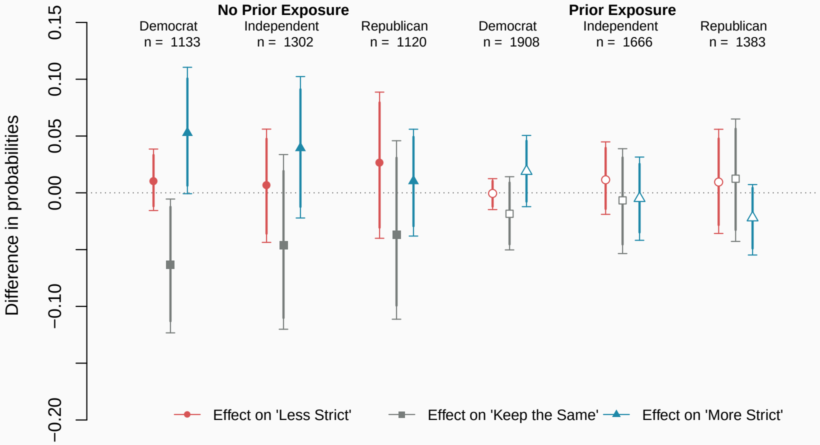
Back

Three-wave Sub-sample: Effect in 2012



Back

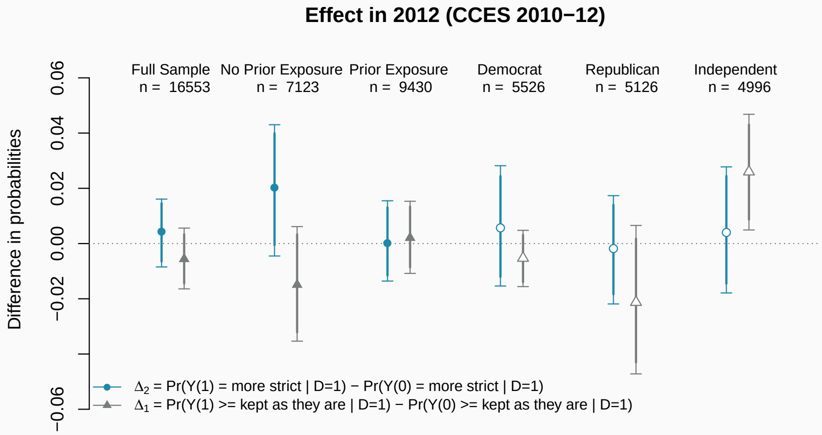
Three-wave Sub-sample: Effect in 2012



Back

Cumulative Effect: Two-wave Panel

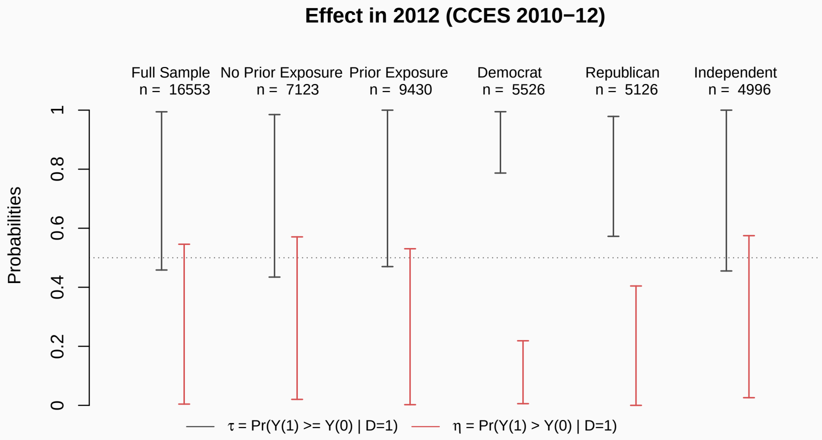
$$\Delta_j = \Pr(Y_{i1}(1) \geq j \mid D_i = 1) - \Pr(Y_{i1}(0) \geq j \mid D_i = 1)$$



Back

Bound Results: Two-wave Panel

$$\tau = \Pr(Y_{i1}(1) \geq Y_{i1}(0) \mid D_i = 1), \quad \text{and} \quad \eta = \Pr(Y_{i1}(1) > Y_{i1}(0) \mid D_i = 1)$$



Back

Definition of Mass Shootings

Cases involving the following:

1. Firearms as the primary weapon used,
2. Attacks on non-family members of the general public
3. Attacks in which at least three or more individuals were injured or killed

Back