Detecting Heterogeneous Treatment Effect with Instrumental Variables

Michael Johnson, Hyunseung Kang

University of Wisconsin-Madison
WNAR
June 24, 2019

Table of contents

Introduction

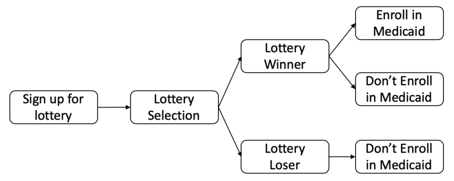
Oregon Health Insurance Experiment

- In 2008, investigators interested in understanding causal effect of **Medicaid** on various financial and health outcomes
- In an ideal world: RCT and randomly force/deny Medicaid on eligible participants
- A (popular) alternative: An encouragement (IV) design that by **random chance** nudges people to enroll in Medicaid.

OHIE website, https://www.nber.org/oregon/1.home.html

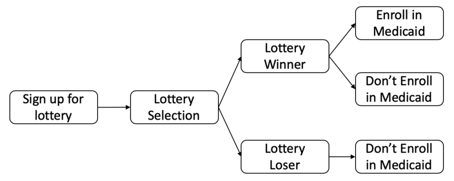
Oregon Health Insurance Experiment

IV design for studying effect of Medicaid:

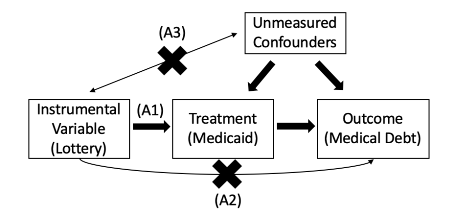


Oregon Health Insurance Experiment

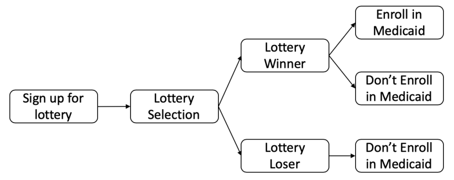
IV design for studying effect of Medicaid:



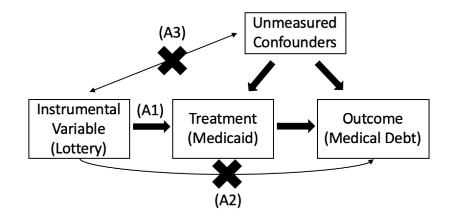
If the core IV assumptions are satisfied, then the complier average treatment effect (CATE) can be estimated as the effect of Medicaid (Angrist et al. 1996)



IV design for studying effect of Medicaid:



If the core IV assumptions are satisfied, then the complier average treatment effect (CATE) can be estimated as the effect of Medicaid (Angrist et al. 1996)



Intent-to-treat measures effect of lottery on the outcome

From CATE to Heterogeneous CATE

- An investigator may be curious how the effect of Medicaid on money owed for medical expenses varies over differing levels of education, sex, or race

From CATE to Heterogeneous CATE

- An investigator may be curious how the effect of Medicaid on money owed for medical expenses varies over differing levels of education, sex, or race
  - The **effect of Medicaid** on money owed for medical expenses varying over levels of covariate X

$$\lambda_0 = E\left[R_1 - R_0 \mid X = x, C = \textit{complier}\right]$$

From CATE to Heterogeneous CATE

- An investigator may be curious how the effect of Medicaid on money owed for medical expenses varies over differing levels of education, sex, or race
    - The **effect of Medicaid** on money owed for medical expenses varying over levels of covariate X

$$\lambda_0 = E\left[R_1 - R_0 \mid X = x, C = \mathit{complier}\right]$$

- Groups may be known prior to analysis, but what if they're not?

From CATE to Heterogeneous CATE

- An investigator may be curious how the effect of Medicaid on money owed for medical expenses varies over differing levels of education, sex, or race
    - The **effect of Medicaid** on money owed for medical expenses varying over levels of covariate X

$$\lambda_0 = E\left[R_1 - R_0 \mid X = x, C = complier\right]$$

- Groups may be known prior to analysis, but what if they're not?
- **Goal**: Estimation & honest inference of discovered heterogeneous CATE

Inferring Heterogeneity from Exploratory Methods with IV

Proposed Two-Staged Solution

Data pre-processing: Pair match lottery winners and losers on pre-lottery covariates

Proposed Two-Staged Solution

Data pre-processing: Pair match lottery winners and losers on pre-lottery covariates

Stage 1: Find potential effect modifiers of Medicaid through exploratory machine learning methods (CART)

- Compute $|Y|$ to prevent effects of double-dipping data
- Use CART with $|Y|$ as outcome and X as predictors to find potential effect modifiers

Proposed Two-Staged Solution

Data pre-processing: Pair match lottery winners and losers on pre-lottery covariates

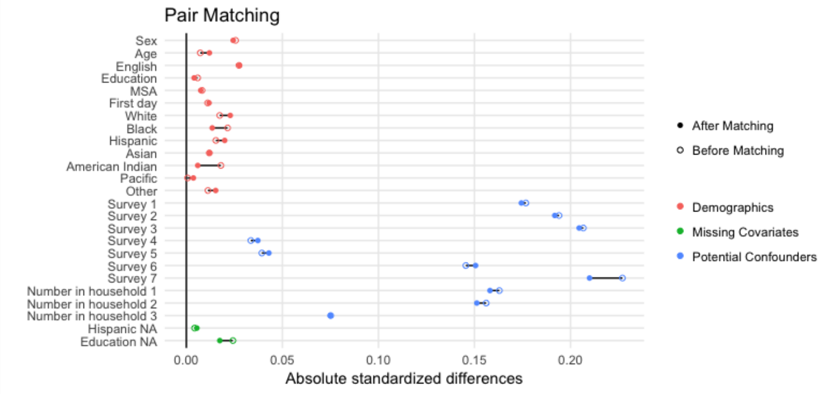
Stage 1: Find potential effect modifiers of Medicaid through exploratory machine learning methods (CART)

- Compute |Y| to prevent effects of double-dipping data
- Use CART with |Y| as outcome and X as predictors to find potential effect modifiers

Stage 2: Simultaneous testing of effect modifiers of Medicaid

- Use closed testing on CART tree

Data Pre-Processing: Matching



Stage 1: Absolute Value of Pairwise Differences $|\mathbf{Y}|$

- **Potential treatment**: d_{zij}; **Potential outcome**: r_{zij}, for $z \in \{0, 1\}$
- Similar to the usual Fisher's sharp null,
  assume $r_{1ij} - r_{0ij} = \lambda(d_{1ij} - d_{0ij})$ for all i, j
  - $r_{1ij} - \lambda d_{1ij} = r_{0ij} - \lambda d_{0ij}$ for all i, j

Stage 1: Absolute Value of Pairwise Differences $|\mathbf{Y}|$

- **Potential treatment**: d_{zij}; **Potential outcome**: r_{zij}, for $z \in \{0, 1\}$
- Similar to the usual Fisher's sharp null, assume $r_{1ij} - r_{0ij} = \lambda(d_{1ij} - d_{0ij})$ for all i, j
    - $r_{1ij} - \lambda d_{1ij} = r_{0ij} - \lambda d_{0ij}$ for all i, j

Matched Pair i	Individual 1	Individual 2
Possibility 1	Lottery Winner	Lottery Loser
Observe	$r_1 - \lambda d_1$	$r_0 - \lambda d_0$
Under null	$r_0 - \lambda d_0$	$r_0 - \lambda d_0$
Possibility 2	Lottery Loser	Lottery Winner
Observe	$r_0 - \lambda d_0$	$r_1 - \lambda d_1$
Under null	$r_0 - \lambda d_0$	$r_0 - \lambda d_0$

Stage 1: Absolute Value of Pairwise Differences $|\mathbf{Y}|$

- **Potential treatment**: d_{zij}; **Potential outcome**: r_{zij}, for $z \in \{0, 1\}$
- Similar to the usual Fisher's sharp null, assume $r_{1ij} - r_{0ij} = \lambda(d_{1ij} - d_{0ij})$ for all i, j
  - $r_{1ij} - \lambda d_{1ij} = r_{0ij} - \lambda d_{0ij}$ for all i, j

Matched Pair i	Individual 1	Individual 2
Possibility 1	Lottery Winner	Lottery Loser
Observe	$r_1 - \lambda d_1$	$r_0 - \lambda d_0$
Under null	$r_0 - \lambda d_0$	$r_0 - \lambda d_0$
Possibility 2	Lottery Loser	Lottery Winner
Observe	$r_0 - \lambda d_0$	$r_1 - \lambda d_1$
Under null	$r_0 - \lambda d_0$	$r_0 - \lambda d_0$

- $\left|Y_i\right| = \left| \pm \left(r_{0i1} - \lambda d_{0i1} - (r_{0i2} - \lambda d_{0i2})\right)\right|$

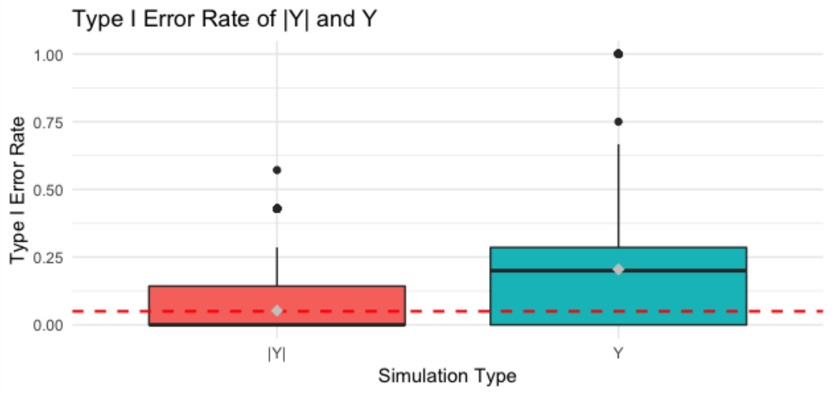
Stage 1: Absolute Value of Pairwise Differences $|\mathbf{Y}|$

- **Potential treatment**: d_{zij}; **Potential outcome**: r_{zij}, for $z \in \{0, 1\}$
- Similar to the usual Fisher's sharp null, assume $r_{1ij} - r_{0ij} = \lambda(d_{1ij} - d_{0ij})$ for all i, j
  - $r_{1ij} - \lambda d_{1ij} = r_{0ij} - \lambda d_{0ij}$ for all i, j

Matched Pair i	Individual 1	Individual 2
Possibility 1	Lottery Winner	Lottery Loser
Observe	$r_1 - \lambda d_1$	$r_0 - \lambda d_0$
Under null	$r_0 - \lambda d_0$	$r_0 - \lambda d_0$
Possibility 2	Lottery Loser	Lottery Winner
Observe	$r_0 - \lambda d_0$	$r_1 - \lambda d_1$
Under null	$r_0 - \lambda d_0$	$r_0 - \lambda d_0$

- $\left|Y_i\right| = \left|\left(r_{0i1} - \lambda d_{0i1} - (r_{0i2} - \lambda d_{0i2})\right)\right|$

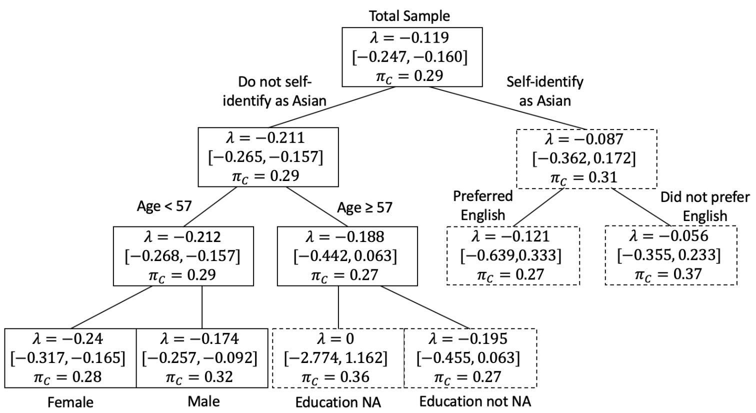
Stage 1: Why $|Y|$, not Y?



- **tl;dr** Without $|Y|$, the data is double-dipped and the type I error is inflated

Stage 1: Using |Y| and CART to Discover Potential Effect Modifiers

- Y: Whether or not an individual currently owes money for medical expenses
- $H_{0_s} : \lambda = \lambda_s = 0$



Stage 2: Closed Testing for Strong Familywise Error Control

- How do you test to whether a *potential* effect modifier is a *true* effect modifier?

Stage 2: Closed Testing for Strong Familywise Error Control

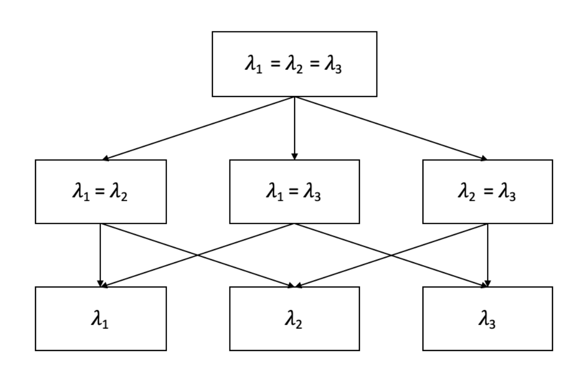
- How do you test to whether a *potential* effect modifier is a *true* effect modifier?
- How do you do this without inflating the familywise type I error rate?

Stage 2: Closed Testing for Strong Familywise Error Control

- How do you test to whether a *potential* effect modifier is a *true* effect modifier?
- How do you do this without inflating the familywise type I error rate?

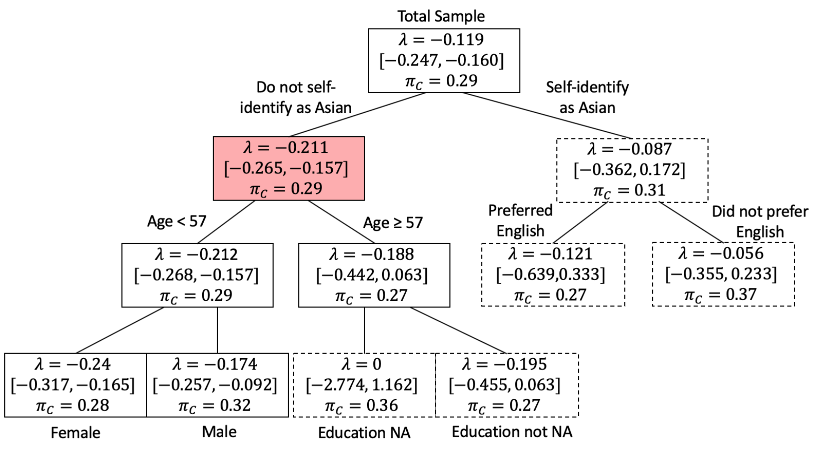
- Solution: Closed testing (Marcus et al. 1976) for nested or tree-structured hypotheses

Stage 2: Closed Testing for Strong Familywise Error Control

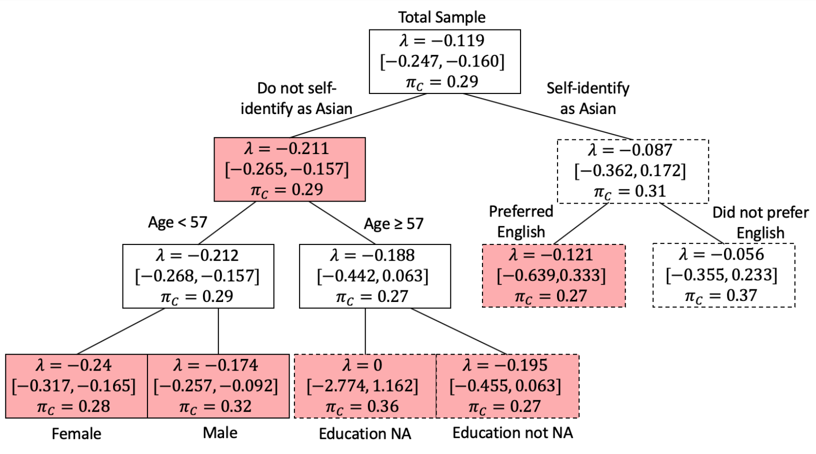


- A family of hypotheses is closed under intersection
- Each hypothesis is level α
    - Kang et al. (2016) test for matching and IV setting

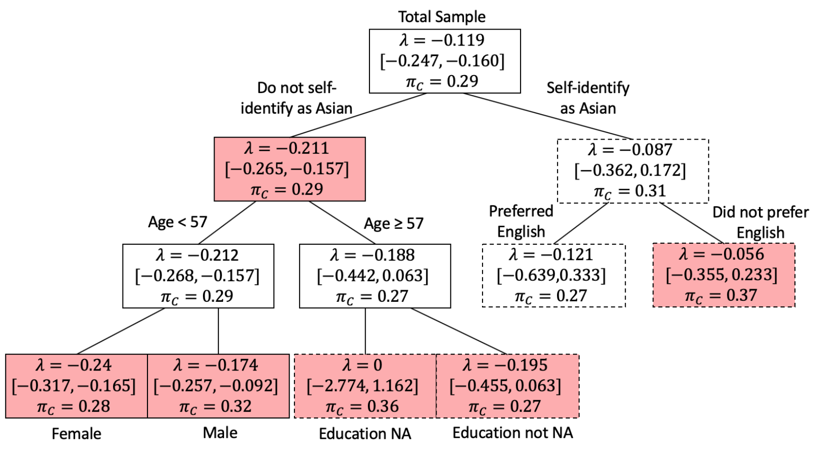
Stage 2: Closed Testing for Strong Familywise Error Control



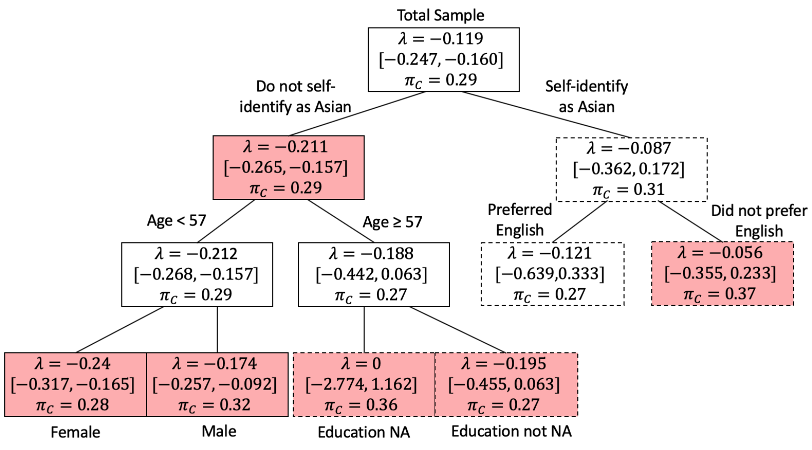
Stage 2: Closed Testing for Strong Familywise Error Control



Stage 2: Closed Testing for Strong Familywise Error Control



Stage 2: Closed Testing for Strong Familywise Error Control



Theorem

Above procedure **strongly** *controls FWER.*

Conclusion

Summary and Future Work

Summary

- Valid simultaneous discovery and inference of heterogeneous treatment effect using instrumental variables
- Use $|Y|$ with CART to discover potential effect modifiers
- Use closed testing on CART to test for true potential effect modifiers and strongly control for family wise error rate
- Medicaid had a larger impact on complying young men and women who do not self-identify as Asian in reducing medical debt

Future Work

- Determine how CART splits based on treatment effect and compliance rate
- Investigate method with other machine learning exploratory methods
  - LASSO, Random Forests

References

- Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434), 444-455.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. New York: Chapman and Hall/CRC.
- Finkelstein, A., Taubman, S., Wright, B., ... & Oregon Health Study Group. (2012). The Oregon health insurance experiment: evidence from the first year. *The Quarterly Journal of Economics, 127*(3), 1057-1106.
- Hsu, J. Y., Small, D. S., & Rosenbaum, P. R. (2013). Effect modification and design sensitivity in observational studies. *Journal of the American Statistical Association, 108*(501), 135-148.
- Hsu, J. Y., Zubizarreta, J. R., Small, D. S., & Rosenbaum, P. R. (2015). Strong control of the familywise error rate in observational studies that discover effect modification by exploratory methods. *Biometrika, 102*(4), 767-782.
- Kang, H., Kreuels, B., May, J., & Small, D. S. (2016). Full matching approach to instrumental variables estimation with application to the effect of malaria on stunting. *The Annals of Applied Statistics, 10*(1), 335-364.
- Marcus, R., Eric, P., & Gabriel, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika, 63*(3), 655-660.

Matching & IV

- **Potential treatment**: d_{zij} ; **Potential outcome**: $r_{zij}^{(d_{zij})}$ for $z \in \{0, 1\}$
- The estimand of interest is the effect ratio

$$\lambda = \frac{\sum_{i=1}^{I} \sum_{j=1}^{n_i} r_{1ij}^{(d_{1ij})} - r_{0ij}^{(d_{0ij})}}{\sum_{i=1}^{I} \sum_{j=1}^{n_i} d_{1ij} - d_{0ij}}$$

$$\stackrel{\text{IV}}{=} E[R_1 - R_0 \mid C = complier]$$