title

Jason Poulos, Kellie Ottoboni

Motivation

-

Overview of experiment

- Oregon Health Insurance Experiment

Estimating treatment effects

- Neyman-Rubin framework: each $i = \{1, ..., N\}$ participants have four potential outcomes, Y_{ist} for $s = 0, 1$ and $t = 0, 1$
- Define W, S, T, C, D, Y

Estimating treatment effects

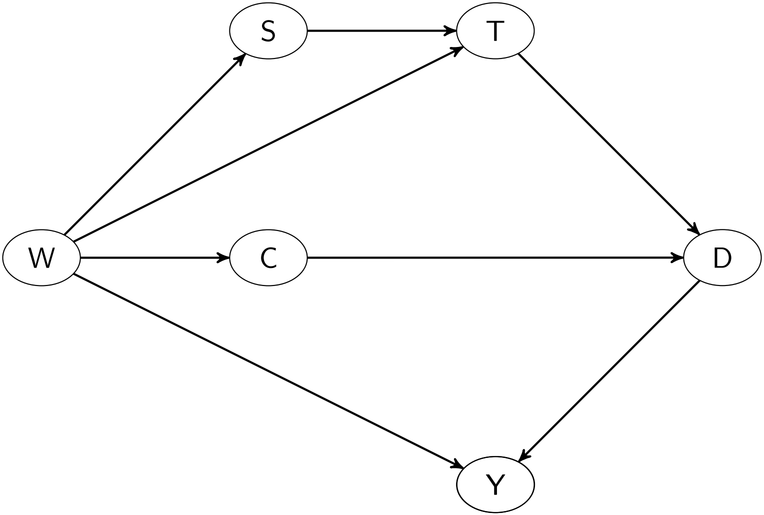


Figure: Causal diagram indicating the conditional independence assumptions needed to estimate the PATT.

Estimating treatment effects (cont.)

Assumption 1

Consistency under parallel studies: for all i and for $t = 0, 1$,

$$Y_{i0t} = Y_{i1t}$$

Estimating treatment effects (cont.)

Assumption 2

Strong ignorability of sample assignment for treated:

$$(Y_{01}, Y_{11}) \perp\!\!\!\perp S \mid (W, T = 1, C = 1), 0 < \mathbb{P}(S = 1 \mid W, T = 1, C = 1) < 1$$

Potential outcomes for treatment are independent of sample assignment for individuals with the same covariates W and assignment to treatment.

Assumption 3

Strong ignorability of sample assignment for controls:

$$(Y_{00}, Y_{10}) \perp\!\!\!\perp S \mid (W, T = 1, C = 1), 0 < \mathbb{P}(S = 1 \mid W, T = 1, C = 1) < 1$$

Potential outcomes for control are independent of sample assignment for individuals with the same covariates W and assignment to treatment.

Estimating treatment effects (cont.)

Assumption 4

Stable unit treatment value assumption (SUTVA):

$$Y_{ist}^{L_i} = Y_{ist}^{L_j}, \forall i \neq j$$

where L_j is the treatment and sample assignment vector for unit j.

Assumption 5

Conditional independence of compliance and assignment:

$$C \perp\!\!\!\perp T = 1 \mid W, 0 < \mathbb{P}(C = 1 \mid W) < 1$$

Estimating treatment effects (cont.)

Assumption 6

Monotonicity:

$$T_i \geq D_i, \forall i$$

This assumption implies that there are no defiers and that crossover is only possible from treatment to control.

Assumption 7

Exclusion restriction: For non-compliers

$$Y_{11} = Y_{10}$$

The treatment assignment affects the response only through the treatment received. In particular, the treatment effect may only be non-zero for compliers.

Estimating treatment effects (cont.)

Theorem

Under assumptions (1) - (7),

$$\tau_{PATT} = \mathbb{E}_{01}\left[\mathbb{E}\left(Y_{11} \mid S=1, D=1, W\right)\right] - \mathbb{E}_{01}\left[\mathbb{E}\left(Y_{10} \mid S=1, T=0, C=1,\right.\right.$$

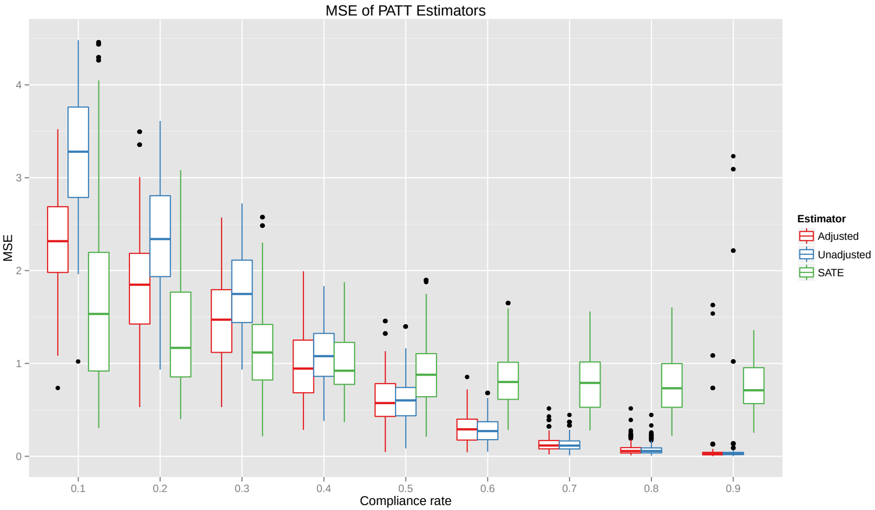
where $\mathbb{E}_{01}\left[\mathbb{E}(\cdot \mid \ldots, W)\right]$ *denotes the expectation with respect to the distribution of* W *in the treated individuals in the target population.*

Estimation Procedure

1. Using the group assigned to treatment in the RCT ($S = 1$, $T = 1$), train a model to predict complier status C using the covariates W.
2. Predict who in the RCT assigned to control *would have* complied to treatment had they been assigned to the treatment group.
3. For the group of observed compliers to treatment and predicted compliers in the control group, train a model to predict the response using as features the covariates W and the treatment T (assigned and observed are the same, for these individuals). This model gives $\mathbb{E}(Y_{1t} \mid S = 1, C = 1, T = t, W)$ for $t = 0, 1$.
4. For all individuals who received treatment in the population ($S = 0, D = 1$), estimate their potential outcomes Y_{10} and Y_{11} using the model from step 3. The mean counterfactual Y_{11} minus the mean counterfactual Y_{10} is the estimate of τ_{PATT}.

Simulation

- describe simulation design?



MSE of PATT Estimators
MSE
0
1
2
3
4
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
Compliance rate
Estimator
Adjusted
Unadjusted
SATE

Data

- List of 1805 lottery participants compiled by **?**
- Fortunate drawer records for 1805 and 1807 lotteries [**??**]
- Roster of officeholders published by Georgia Archives (Trustee period – 1847)
- Roll call votes extracted from Journals of the House and Senate of the State of Georgia
  - 15 votes: emancipate certain slaves; facilitate introduction of slaves into state and prevent slaves being carried out of the state; and punish slaves and free blacks
- Individual property tax records (1790 – 1865) [**??**]

Linking participants with officeholders

1. Manually deduplicate 1807 records matched with officeholders based on exact match of surname and Soundex codes of first name
2. Randomly split matched records into training (60%) and test (40%) sets
3. Fit an algorithmic model using random forests [Breiman, 2001] on training set with features common to both datasets (test set error rate of 35%)
4. Use model to deduplicate 1805 lottery records matched with officeholders

Results

Table: Officeholding by treatment assignment.

Panel A: 1805 sample							
Response		**Control**	$\%_m$	**Treated**	$\%_n$	**All**	$\%_N$
Officeholder	0	16747	93.1	3058	93.2	19805	93.2
	1	1234	6.9	222	6.8	1456	6.8
	all	17981	100.0	3280	100.0	21261	100.0
Panel B: Combined sample							
Officeholder	0	16747	93.1	10813	93.9	27560	93.4
	1	1234	6.9	705	6.1	1939	6.6
	all	17981	100.0	11518	100.0	29499	100.0

Notes: Distribution of officeholders by treatment assignment for sample of 1805 lottery participants (Panel A) and combined sample of 1805 participants and 1807 fortunate drawers (Panel B). Both samples exclude women orphans, and pretreatment officeholders.

Results (cont.)

Table: Support for slavery by treatment assignment.

Panel A: 1805 sample							
Variable	**Treatment**	**N**	**Min.**	**Mean**	**Max.**	**S.d.**	**#NA**
Support for slavery	0	255	0	0.737	1	0.366	963
	1	219	0	0.722	1	0.382	632
	all	474	0	0.730	1	0.373	1595
Panel B: Combined sample							
Support for slavery	0	401	0	0.735	1	0.371	1353
	1	216	0	0.757	1	0.369	763
	all	617	0	0.743	1	0.370	2116

Distribution of the outcome variable, by treatment assignment, for 1805 participants (Panel A) or combined sample of 1805 participants and 1807 fortunate drawers (Panel B) who held office in the General Assembly before 1848 . 'Support for slavery' is the mean of votes in favor of slavery.

Results (cont.)

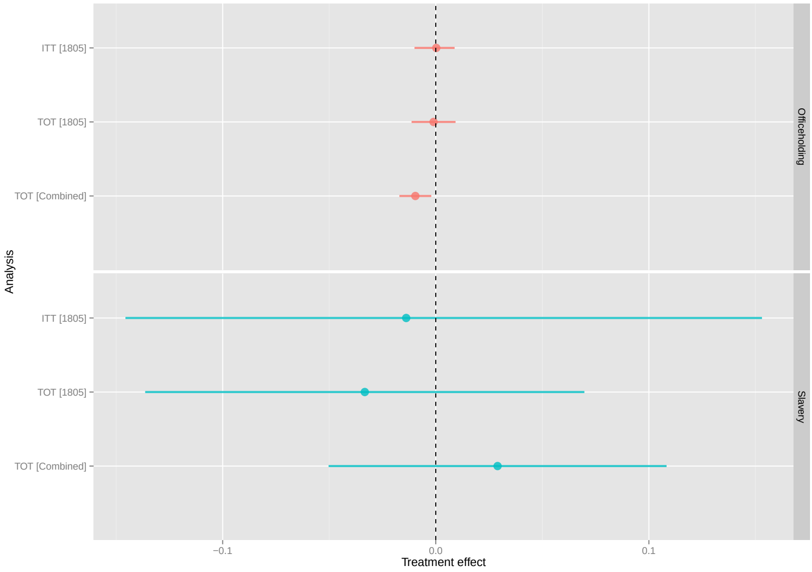


Figure: Heterogenous treatment effects for 1805 lottery participants.

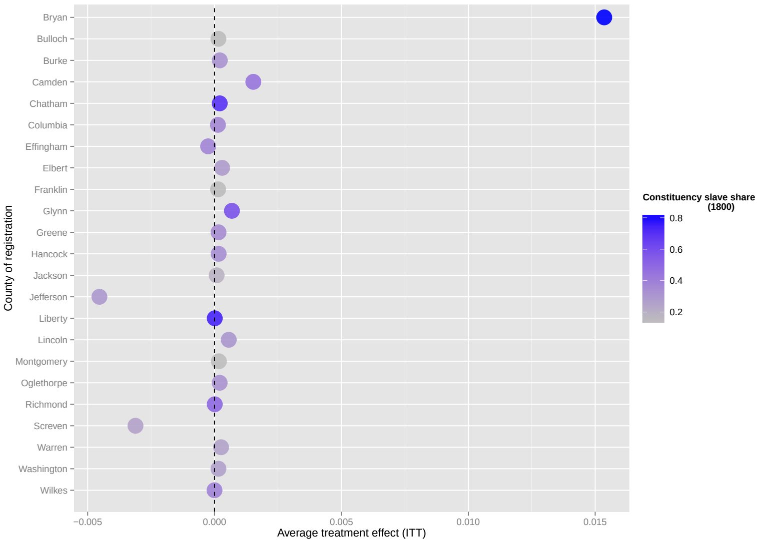


Figure: Pre– and posttreatment wealth densities for legislator–participants who voted on roll calls.

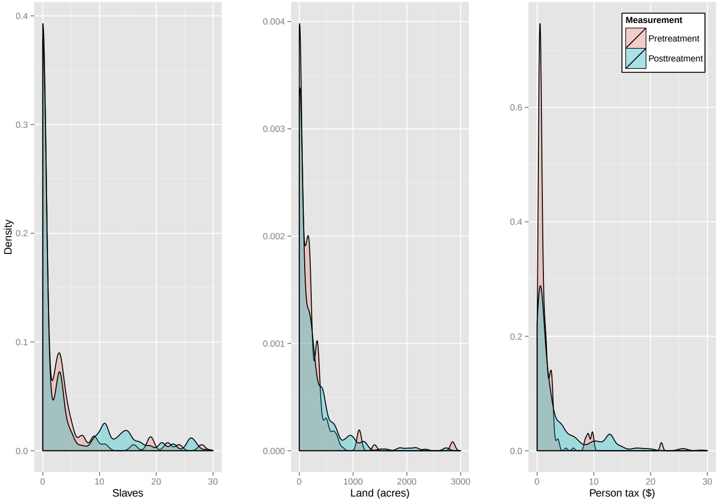
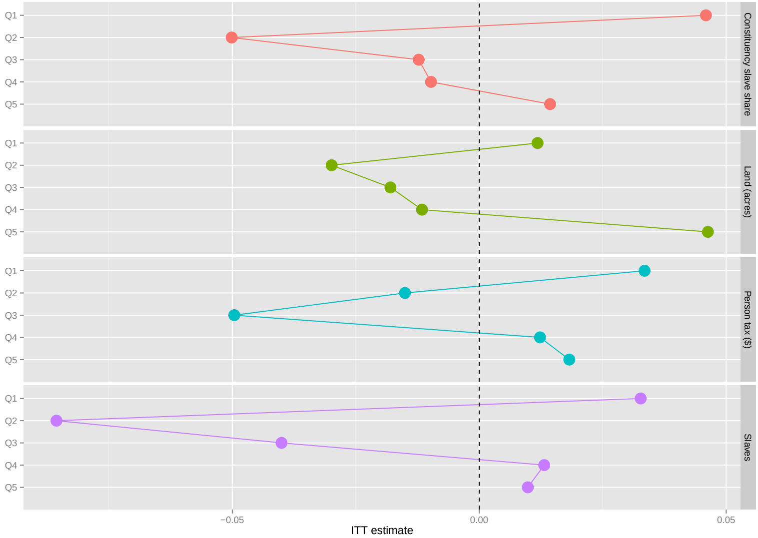


Figure: Heterogenous treatment effects for legislators who voted on roll calls.



Discussion

- If property wealth influences political power, we should be able to find evidence in antebellum South
- Officeholding: tight CIs on zero effect implies evidence "in favor of" null
- Elite ideology: too much uncertainty to detect significant treatment effect
  - Substantial heterogeneity in treatment effect according to pretreatment wealth

Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.