Estimating population average treatment effects from experiments with noncompliance

Kellie Ottoboni Jason Poulos

Oct. 15, 2015

Motivation

- RCTs are the "gold standard" for estimating the causal effect of a treatment
  - External validity is an issue when RCT participants don't reflect the target population
  - Non-compliance to treatment assignment biases estimates of the sample average treatment effect (SATE) towards 0
- Idea: reweight responses in the treatment group of RCT compliers to estimate population average treatment effect on the treated (PATT)
  - Hartman et al. [2015] develop a nonparametric reweighting method to extend SATE to PATT
  - We extend this method to the case of one-way crossover

Estimating treatment effects

- Neyman-Rubin framework: each $i = \{1, ..., N\}$ participants have four potential outcomes, Y_{ist} for $s = 0, 1$ and $t = 0, 1$
  - S = study assignment: S=1 for RCT, S=0 for population/observational study
  - T = treatment assignment: T = 1 for treatment, T = 0 for control
  - D = treatment received
- Other variables
  - W = observed covariates
  - C = compliance to treatment
  - Y = response

Estimating treatment effects (cont.)

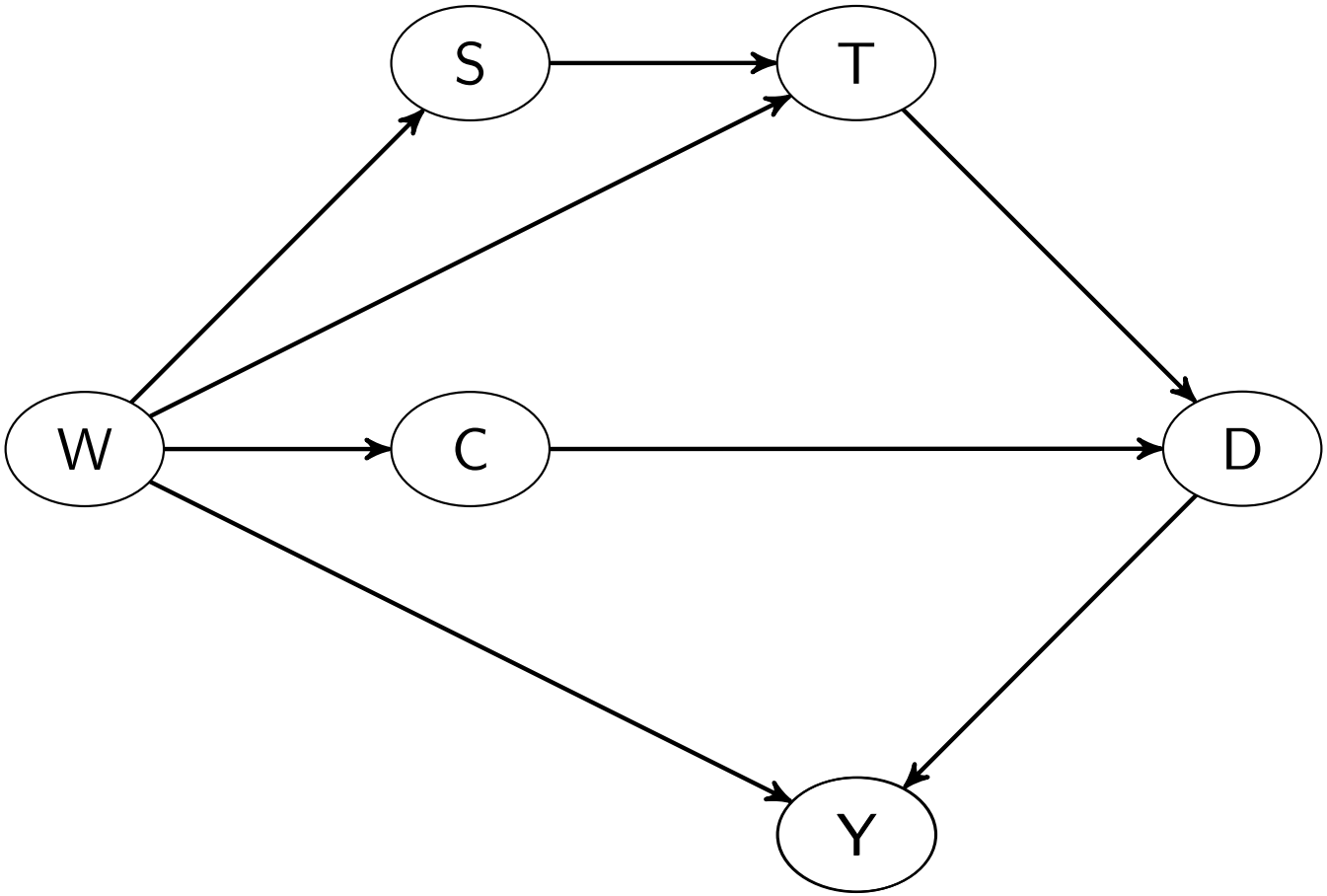


Figure: Causal diagram indicating the conditional independence assumptions needed to estimate the PATT.

Estimating treatment effects (cont.)

Theorem

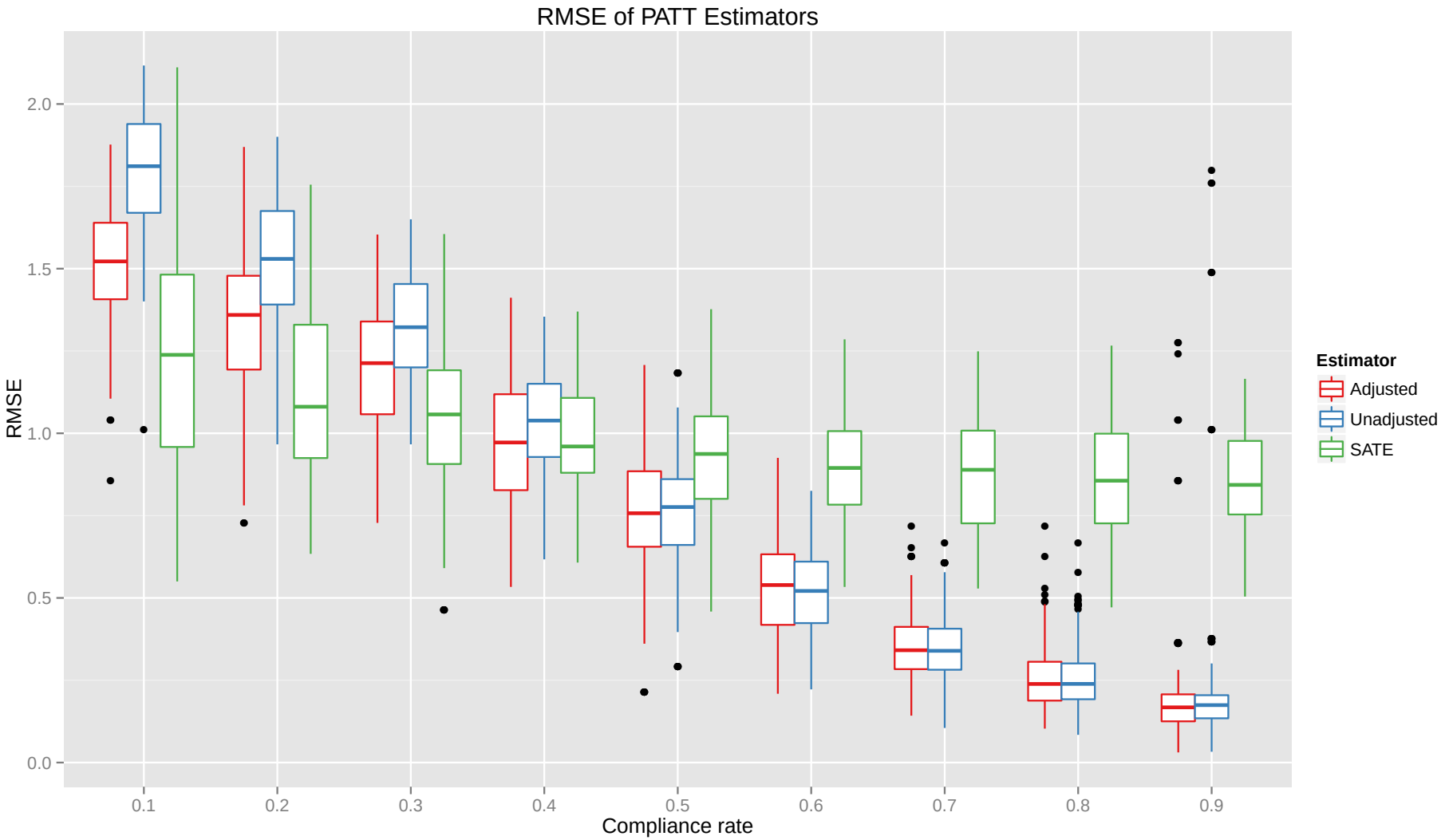
Under assumptions (1) - (7),

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

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.*

Simulation Design

- Generate a population of 30,000 with 3 observable covariates W
- Set S, T, C, Y to be linear functions of W, with some Gaussian noise
- Heterogeneous treatment effect: magnitude of effect depends on one of the covariates
- Sample 5,000 "randomizables" for RCT and 5,000 "observables" for observational study. Enroll individuals according to S
- Predict would-be compliers in the RCT control group using logistic regression
- Estimate response curve in RCT compliers using a random forest
- Use model to estimate potential outcomes in the observational study to estimate τ_{PATT}


RMSE of PATT Estimators
RMSE
2.0
1.5
1.0
0.5
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
Compliance rate
Estimator
Adjusted
Unadjusted
SATE

Application: Oregon Health Insurance Experiment (OHIE)

- In 2008, $\approx$ 90,000 uninsured low-income adults participated in a lottery to receive Medicaid benefits [Finkelstein et al., 2012]
- Participants selected by the lottery won the opportunity for themselves and any household member to apply for Medicaid
- After sample exclusions, 29,834 participants were selected by the lottery; remaining 45,008 served as controls
- Two health care use responses from mail survey (N = 23,741): emergency room (ER) and primary care visits in past 12 months
- Compliance measure: indicator for whether participant was enrolled in Medicaid program during study period

Observational data

- Data on the target population from National Health Interview Study (NHIS) for Health Statistics for 2009–2013
- Restrict to respondents with income below 138% of FPL and on Medicaid ($N = 3,914$)
- Covariates and responses match OHIE

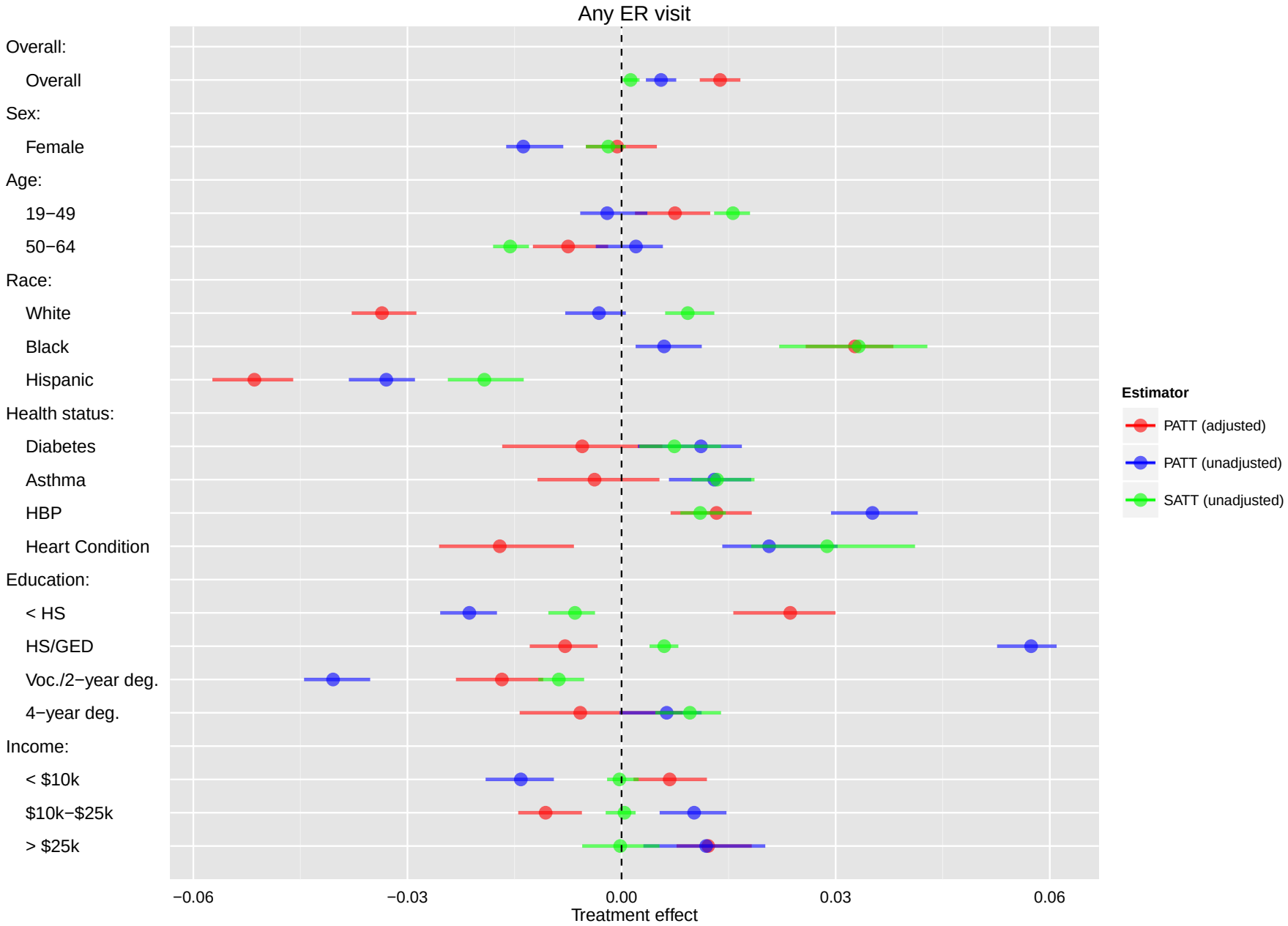
	OHIE control $n = 5476$		OHIE treated $n = 5193$		NHIS treated $n = 3382$	
Covariate	**n**	%	**n**	%	**n**	%
Sex:						
Female	3148	57.5	2920	56.2	2380	70.4
Age:						
19-49	1636	29.9	1367	26.3	2429	71.8
50-64	3840	70.1	3826	73.7	953	28.2
Race:						
White	4829	88.2	4393	84.6	1991	58.9
Black	243	4.4	197	3.8	1050	31.1
Hispanic	301	5.5	476	9.2	910	26.9
Health status:						
Diabetes	581	10.6	539	10.4	452	13.4
Asthma	1036	18.9	887	17.1	652	19.3
High blood pressure	1670	30.5	1418	27.3	1143	33.8
Heart condition	170	3.1	141	2.7	285	8.4
Education:						
Less than high school	1056	19.3	950	18.3	1183	35.0
High school diploma or GED	3081	56.3	2775	53.4	1076	31.8
Voc. training / 2-year degree	969	17.7	1031	19.9	934	27.6
4-year college degree or more	370	6.8	437	8.4	189	5.6
Income:						
$< \$10k$	5476	100.0	3204	61.7	1452	42.9
\$10k-\$25k	0	0.0	1616	31.1	1622	48.0
$> \$25k$	0	0.0	373	7.2	308	9.1
Response						
Any ER visit	1393	25.4	1301	25.1	881	26.1
Any outpatient visit	3299	60.2	3081	59.3	2116	62.6

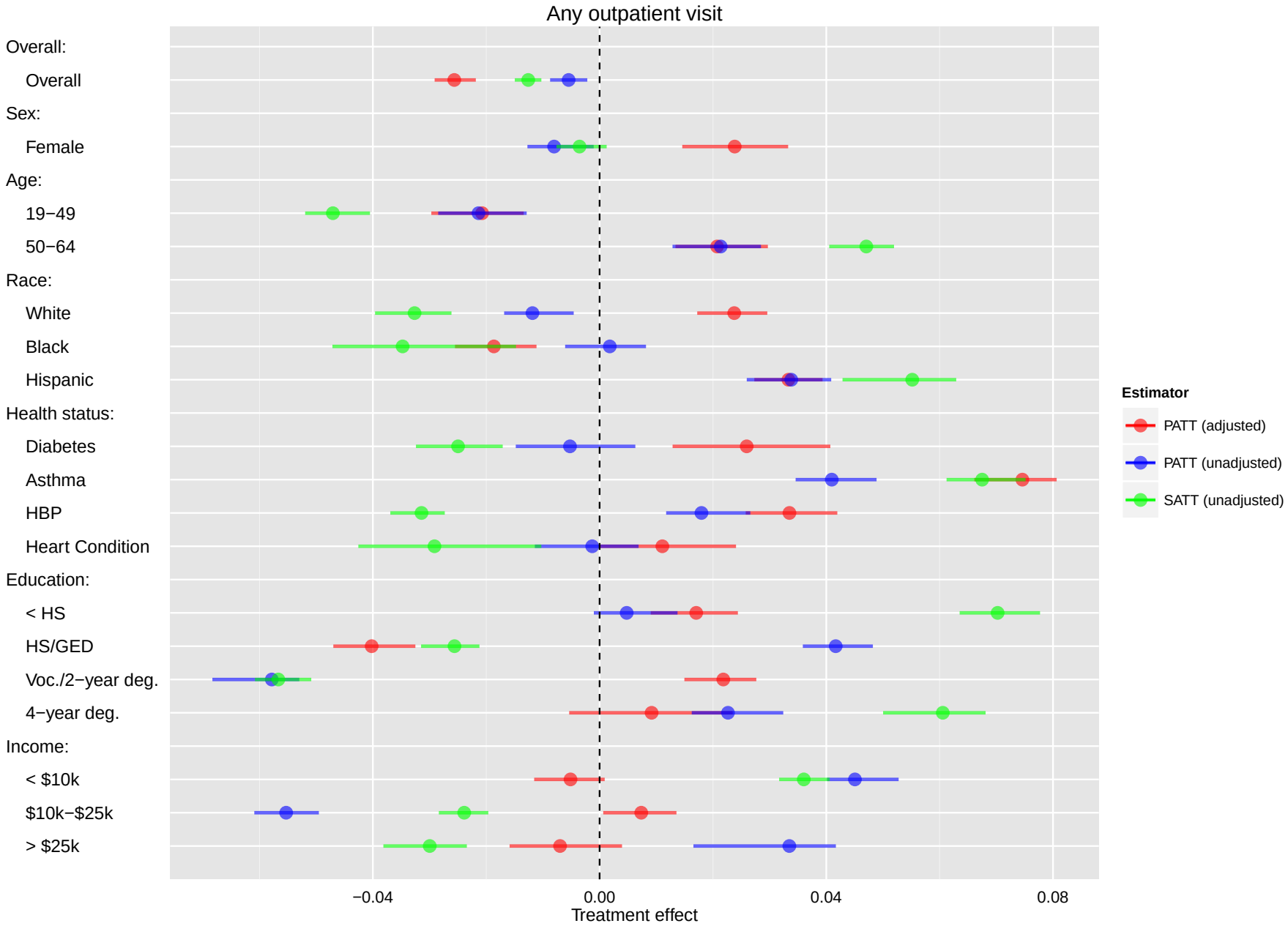
Checking Assumptions

- Monotonicity is violated: two-way crossover occurred in OHIE
  - 60% of treated did not enroll in Medicaid
  - 14% of controls enrolled in Medicaid during the study period
  - Cross-over from control to treatment is low relative to other direction
- Key assumption is strong ignorability: model of response given covariates is same for RCT and population
  - We have included all possible confounders we have data on

Estimation Procedure

1. Train random forests (RF) model to predict complier status, given covariates, for RCT treated
2. Use model to predict compliers among RCT controls
3. For observed and predicted compliers, train RF model to predict response using covariates and treatment as features
4. Using response model, estimate potential outcomes for population "compliers" on medicaid
5. τ_{PATT} is the difference in means between potential outcomes


Any ER visit
Overall:
Overall
Sex:
Female
Age:
19−49
50−64
Race:
White
Black
Hispanic
Health status:
Diabetes
Asthma
HBP
Heart Condition
Education:
< HS
HS/GED
Voc./2−year deg.
4−year deg.
Income:
< $10k
$10k−$25k
> $25k
−0.06
−0.03
0.00
0.03
0.06
Treatment effect
Estimator
PATT (adjusted)
PATT (unadjusted)
SATT (unadjusted)


Any outpatient visit
Overall:
Overall
Sex:
Female
Age:
19−49
50−64
Race:
White
Black
Hispanic
Health status:
Diabetes
Asthma
HBP
Heart Condition
Education:
< HS
HS/GED
Voc./2−year deg.
4−year deg.
Income:
< $10k
$10k−$25k
> $25k
−0.04
0.00
0.04
0.08
Treatment effect
Estimator
PATT (adjusted)
PATT (unadjusted)
SATT (unadjusted)

Conclusions

- Proposed estimator performs better than unadjusted estimator in simulations when compliance is low and can be predicted by observed covariates
- Adjusted and unadjusted τ_{PATT} estimates similar for primary care visits
- Negative treatment effect on population compliers for ER visits (unadjusted τ_{PATT} is positive)
- Substantial differences between sample and population estimates in terms of race and health status

Amy Finkelstein, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P Newhouse, Heidi Allen, Katherine Baicker, Oregon Health Study Group, et al. The oregon health insurance experiment: Evidence from the first year. *The quarterly journal of economics*, 127(3):1057, 2012.

National Center for Health Statistics. National health interview survey. `http://www.cdc.gov/nchs/nhis.htm`. April 2015.

Erin Hartman, Richard Grieve, Roland Ramsahai, and Jasjeet S Sekhon. From sate to patt: combining experimental with observational studies to estimate population treatment effects. *JR Stat. Soc. Ser. A*, 10:1111, 2015.

Appendix: estimator assumptions

Assumption 1

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

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

Appendix: estimator assumptions (cont.)

Assumption 2

Strong ignorability of sample assignment for treated:

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

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

Appendix: estimator assumptions (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:

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

Appendix: estimator assumptions (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.