Quantitative Research Methods IV - 17.806

Recitation, Week 11.

Topic: Sensitivity Analysis.

Benjamín Muñoz

April 28, 2023

MIT

Massachusetts Institute of Technology

Table of contents

1/ Introduction

Identification Assumptions in Observational Studies

- **Absence** of a randomized treatment.
- In **expectation**, the randomization of the treatment (D) balances observed and unobserved characteristics between the treatment and control groups ($Y^0, Y^1 \perp\!\!\!\perp D$).

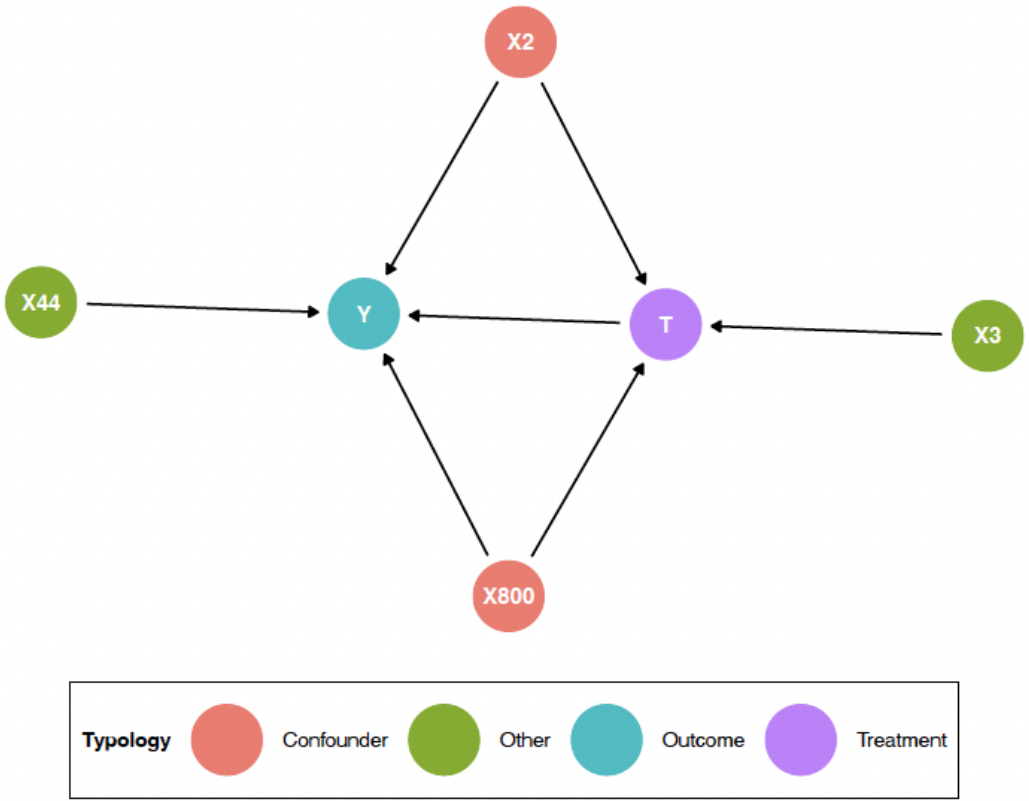
Identification Assumptions of the ATE

- **SUTVA:** No interference and Consistency $(Y(D_1, D_2, \ldots, D_N)_i = Y(D_1', D_2', \ldots, D_N')_i$ if $D_i = D_i')$.
- **Conditional Ignorability:** $Y_i^0, Y_i^1 \perp\!\!\!\perp D_i | X_i = x$, for any $x \in \mathcal{X}$. (For the ATT $\rightsquigarrow Y_i^0 \perp\!\!\!\perp D_i | X_i = x$; For the ATU $\rightsquigarrow Y_i^1 \perp\!\!\!\perp D_i | X_i = x$)
- **Common Support:** $0 < P(D_i = 1 | X_i = x) < 1$ for any $x \in \mathcal{X}$. (For the ATT $\rightsquigarrow P(D_i = 1 | X_i = x) < 1$; For the ATU $\rightsquigarrow 0 < P(D_i = 1 | X_i = x))$

Confounding

- Threat to internal validity of causal inference = **Confounding**.
- **Conditional Ignorability** = **No Unmeasured Confounders**.
- **Typical Definiton:** a confounder as a pre-treatment variable that was associated with the treatment and the outcome (conditional on the exposure) (see VanderWeele, 2013 for a discussion).

Figure 1: Confounding Paths in DAG



Robustness and Credibility

- **General Idea of Validity:** method to determine the robustness of an assessment by examining the extent to which results are affected by changes in methods, models, values of unmeasured variables, or assumptions.
    1. Plausibility of Design:
        - Balance Tests.
        - Placebo Tests.
        - Sensitivity Analysis
    2. Alternative Modelling Strategies.
    3. Measurement Error.
    4. Missing Data.

$\rightsquigarrow$ **Sensitivity Analysis:** assessing the potential influence of **unmeasured** confounding on causal conclusions (Bias Analysis).

2/ The E-Value

Risk Ratio as QoI

- **Risk Ratio:** or Relative Risk is a measure of the risk of a certain event happening in one group (Exposed or Treated) compared to the risk of the same event happening in another group (Unexposed or Control).
- Formula for CI: $\ln(\widehat{RR}) \pm z\sqrt{\frac{B/A}{A+B} + \frac{D/C}{C+D}}$.
- It's not the same as the Odds Ratio ($OR = \frac{A+D}{B+C}$).

Predictor	Outcome: Yes	Outcome: No
Yes	A	B
No	C	D

$$RR = \frac{(A/(A+B))}{(C/(C+D))}$$

Estimation of RR

1. Unconditional Strategy: 2×2 Cross-Tabulation.
2. Conditional Strategy: Cochran-Mantel-Haenszel Method.
3. Conditional Strategy: Log-Binomial Regression:
   - Logistic Regression: $\text{logit}(p_i) = \frac{p_i}{1-p_i} = X^T\beta \rightsquigarrow exp(\beta_k)$ is the odds ratio corresponding to a 1-unit change in X_k (ceteris paribus).
   - Log-Binomial Regression $\log(p_i) = X^T\beta \rightsquigarrow exp(\beta_k)$ is the risk ratio corresponding to a 1-unit change in X_k (ceteris paribus).

Bounding Factor

- Ding and VanderWeele (2016). Let's start with the notation:
  - Treatment (or exposure) **D**.
  - Outcome **Y**.
  - Measured confounders **X**
  - Unmeasured confounders **U**.

- Now, the Quantities:
  - **Observed RR:** $RR^{Obs}_{DY|X} = \frac{P(Y=1|D=1,X=x)}{P(Y=1|D=0,X=x)}$.
  - **RR of Treatment on Confounder:** $RR_{DU,k|x} = \frac{P(U=k|D=1,X=1)}{P(U=k|D=0,X=1)}$
  - **Maximum of the Effect of U on D among the Unexposed:** $RR_{UY|D=0,X} = \frac{max_k P(Y=1|D=0,X=x,U=k)}{min_k P(Y=1|D=0,X=x,U=k)}$
  - **Maximum of the Effect of U on D among the Exposed:** $RR_{UY|D=1,X} = \frac{max_k P(Y=1|D=1,X=x,U=k)}{min_k P(Y=1|D=1,X=x,U=k)}$
  - **True RR:** $RR^{True}_{DY|X} = \frac{\sum_{k=0}^{K-1} P(Y=1|D=1,X=x,U=k) \times P(U=k|X=x)}{\sum_{k=0}^{K-1} P(Y=1|D=0,X=x,U=k) \times P(U=k|X=x)}$

Bounding Factor

Key Result: Even in presence of unmeasured confounding the true RR must be at least as large

$$RR^{True}_{DY|X} \geq \frac{RR^{Obs}_{DY|X}}{\frac{RR_{DU} \times RR_{UY}}{RR_{DU} + RR_{UY} - 1}}$$

$$\frac{RR_{DU} \times RR_{UY}}{RR_{DU} + RR_{UY} - 1} \geq \frac{RR^{Obs}_{DY}}{RR^{True}_{DY}}$$

- The Joint Bounding Factor is always smaller than both of the RR_{DU} and RR_{UY}.

Bounding Factor

TABLE 1. Magnitudes of the Joint Bounding Factor for Different Combinations of the Exposure–Confounder Association RR_{EU} and the Confounder–Outcome Association RR_{UD}

Bounding Factor	RR_{UD} 1.3	1.5	1.8	2	2.5	3	3.5	4	5	6	8	10
RR_{EU}												
1.3	1.06	1.08	1.11	1.13	1.16	1.18	1.20	1.21	1.23	1.24	1.25	1.26
1.5	1.08	1.12	1.17	1.20	1.25	1.29	1.31	1.33	1.36	1.38	1.41	1.43
1.8	1.11	1.17	1.25	1.29	1.36	1.42	1.47	1.50	1.55	1.59	1.64	1.67
2	1.13	1.20	1.29	1.33	1.43	1.50	1.56	1.60	1.67	1.71	1.78	1.82
2.5	1.16	1.25	1.36	1.43	1.56	1.67	1.75	1.82	1.92	2.00	2.11	2.17
3	1.18	1.29	1.42	1.50	1.67	1.80	1.91	2.00	2.14	2.25	2.40	2.50
3.5	1.20	1.31	1.47	1.56	1.75	1.91	2.04	2.15	2.33	2.47	2.67	2.80
4	1.21	1.33	1.50	1.60	1.82	2.00	2.15	2.29	2.50	2.67	2.91	3.08
5	1.23	1.36	1.55	1.67	1.92	2.14	2.33	2.50	2.78	3.00	3.33	3.57
6	1.24	1.38	1.59	1.71	2.00	2.25	2.47	2.67	3.00	3.27	3.69	4.00
8	1.25	1.41	1.64	1.78	2.11	2.40	2.67	2.91	3.33	3.69	4.27	4.71
10	1.26	1.43	1.67	1.82	2.17	2.50	2.80	3.08	3.57	4.00	4.71	5.26

The E-Value (VanderWeele and Ding, 2017).

1. The minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association, conditional on the measured confounders.
2. Based on two components: the strength of the association between U and Y, strength of the association between U and D.
3. Calculation:
   - If $RR > 1 \rightsquigarrow$ E-Value $RR + \sqrt{RR \times (RR - 1)}$.
   - If $RR < 1 \rightsquigarrow RR^* = \frac{1}{RR}$ E-Value $RR^* + \sqrt{RR^* \times (RR^* - 1)}$

R Code

```
### Load packages
library(EValue)          # E-Value

### Calculate E-Value
eval <- evalues.RR(est =  RR,
                   lo = lwr,
                   hi = 0.upr)

### Bias Plot
bias_plot(RR, xmax = max)
```

3/ The R-Value

OVB and Partial R^2 Parameterization

1. Restatement of typical sensitivity analysis:

$$\begin{aligned}\widehat{Bias} &= \hat{\gamma} \times \hat{\delta} \\ &= \text{Impact} \times \text{Imbalance}\end{aligned}$$

2. Based on OLS Mechanics: OVB does not require a causal (Potential Outcomes) interpretation.

3. **Impact** ($\hat{\gamma}$): how looking at different subgroups of the confounder impacts (changes) our Best Linear Prediction of the Outcome.

4. **Imbalance** ($\hat{\delta}$): the difference in the Linear Expectation of the confounder, when comparing individuals with the same values for the covariates, but differing by one unit on the treatment.

OVB and Partial R^2 Parameterization

5. Description in terms of Partial R^2: proportion of unexplained variance that we will attribute to confounding for purposes of sensitivity analysis.

6. The Partial R^2 used measures the relationship between the Confounder and the Outcome ($R^2_{Y\sim U|D,X}$) and between the Confounder and the Treatment ($R^2_{D\sim U|X}$).

Sensitivity Statistics

1. Simple workflow.

R Code

```
### Load packages
library(sensemakr)

### Report basic results
sensitivity_stats(model, treatment = "D")

### Run sensitivity analysis
sensitivity_01 <-  sensemakr(model = model,
treatment = "D",
benchmark_covariates = "X",
q = 0.5, kd = 1:3)


### Create plots
plot(sensitivity_01)
plot(sensitivity_01, sensitivity.of="t-value")
ovb_extreme_plot(model = model,
treatment = "D",
r2yz.dx = c(1, 0.75, 0.5))
```

```
Null hypothesis: q = 0.5 and reduce = TRUE
-- This means we are considering biases that reduce the absolute value of
   the current estimate.
-- The null hypothesis deemed problematic is H0:tau = -0.379

Unadjusted Estimates of 'abd':
  Coef. estimate: -0.7581
  Standard Error: 0.2169
  t-value (H0:tau = -0.379): -1.7475

Sensitivity Statistics:
  Partial R2 of treatment with outcome: 0.0165
  Robustness Value, q = 0.5: 0.0627
  Robustness Value, q = 0.5, alpha = 0.05: 0

Verbal interpretation of sensitivity statistics:

-- Partial R2 of the treatment with the outcome: an extreme confounder
   (orthogonal to the covariates) that explains 100% of the residual
   variance of the outcome, would need to explain at least 1.65% of the
   residual variance of the treatment to fully account for the observed
   estimated effect.

-- Robustness Value, q = 0.5: unobserved confounders (orthogonal to the
   covariates) that explain more than 6.27% of the residual variance of
   both the treatment and the outcome are strong enough to bring the
   point estimate to -0.379 (a bias of 50% of the original estimate).
   Conversely, unobserved confounders that do not explain more than 6.27%
   of the residual variance of both the treatment and the outcome are not
   strong enough to bring the point estimate to -0.379.

-- Robustness Value, q = 0.5, alpha = 0.05: unobserved confounders
   (orthogonal to the covariates) that explain more than 0% of the
   residual variance of both the treatment and the outcome are strong
   enough to bring the estimate to a range where it is no longer
   'statistically different' from -0.379 (a bias of 50% of the original
   estimate), at the significance level of alpha = 0.05. Conversely,
   unobserved confounders that do not explain more than 0% of the
   residual variance of both the treatment and the outcome are not strong
   enough to bring the estimate to a range where it is no longer
   'statistically different' from -0.379, at the significance level of
   alpha = 0.05.
```

Outcome: *educ*

Treatment:	Est.	S.E.	t-value	$R^2_{Y\sim D\|\mathbf{X}}$	$RV_{q=0.5}$	$RV_{q=0.5,\alpha=0.05}$
abd	-0.758	0.217	-1.747	1.6%	6.3%	0%
df = 729						

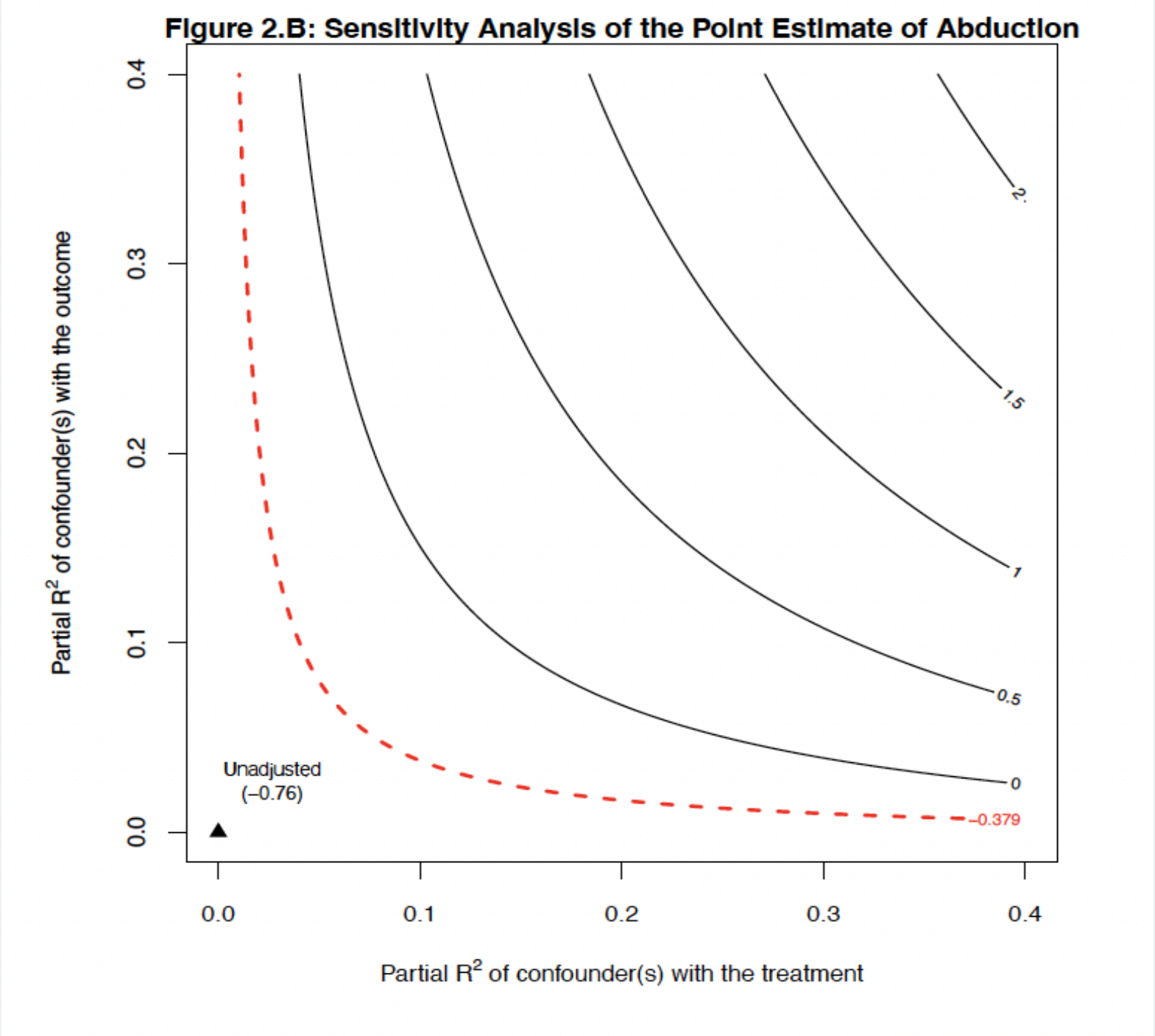
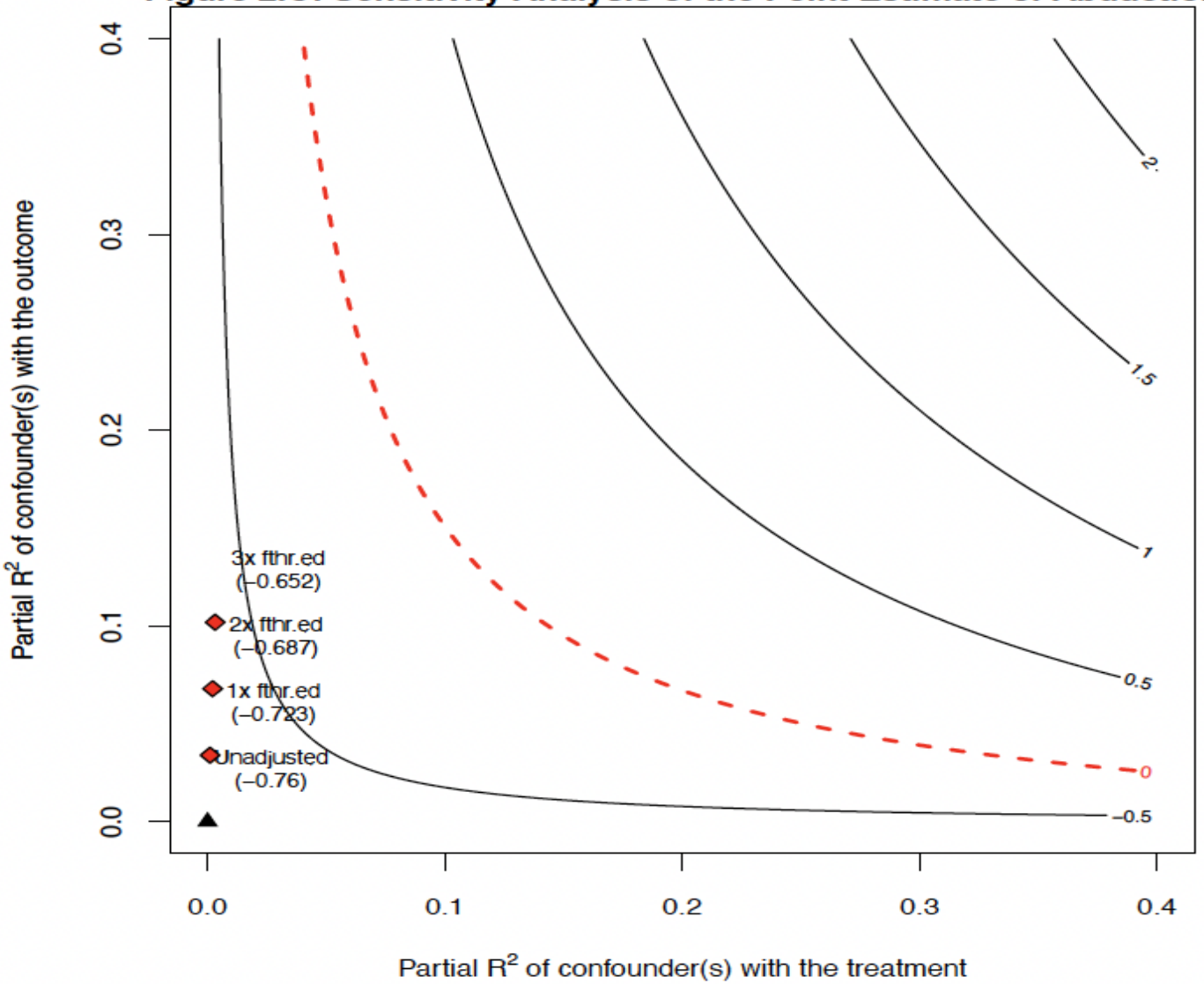


Figure 2.B: Sensitivity Analysis of the Point Estimate of Abduction
Partial R2 of confounder(s) with the outcome
0.0
0.1
0.2
0.3
0.4
2
1.5
1
0.5
0
−0.379
Unadjusted
(−0.76)
0.0
0.1
0.2
0.3
0.4
Partial R2 of confounder(s) with the treatment

Figure 2.C: Sensitivity Analysis of the Point Estimate of Abduction



Outcome: *educ*

Treatment:	Est.	S.E.	t-value	$R^2_{Y\sim D\|\mathbf{X}}$	$RV_{q=1}$	$RV_{q=1,\alpha=0.05}$
abd	-0.758	0.217	-3.495	1.6%	12.1%	5.5%
df = 729		*Bound (1x fthr.ed)*: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 3.4%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0.1%				

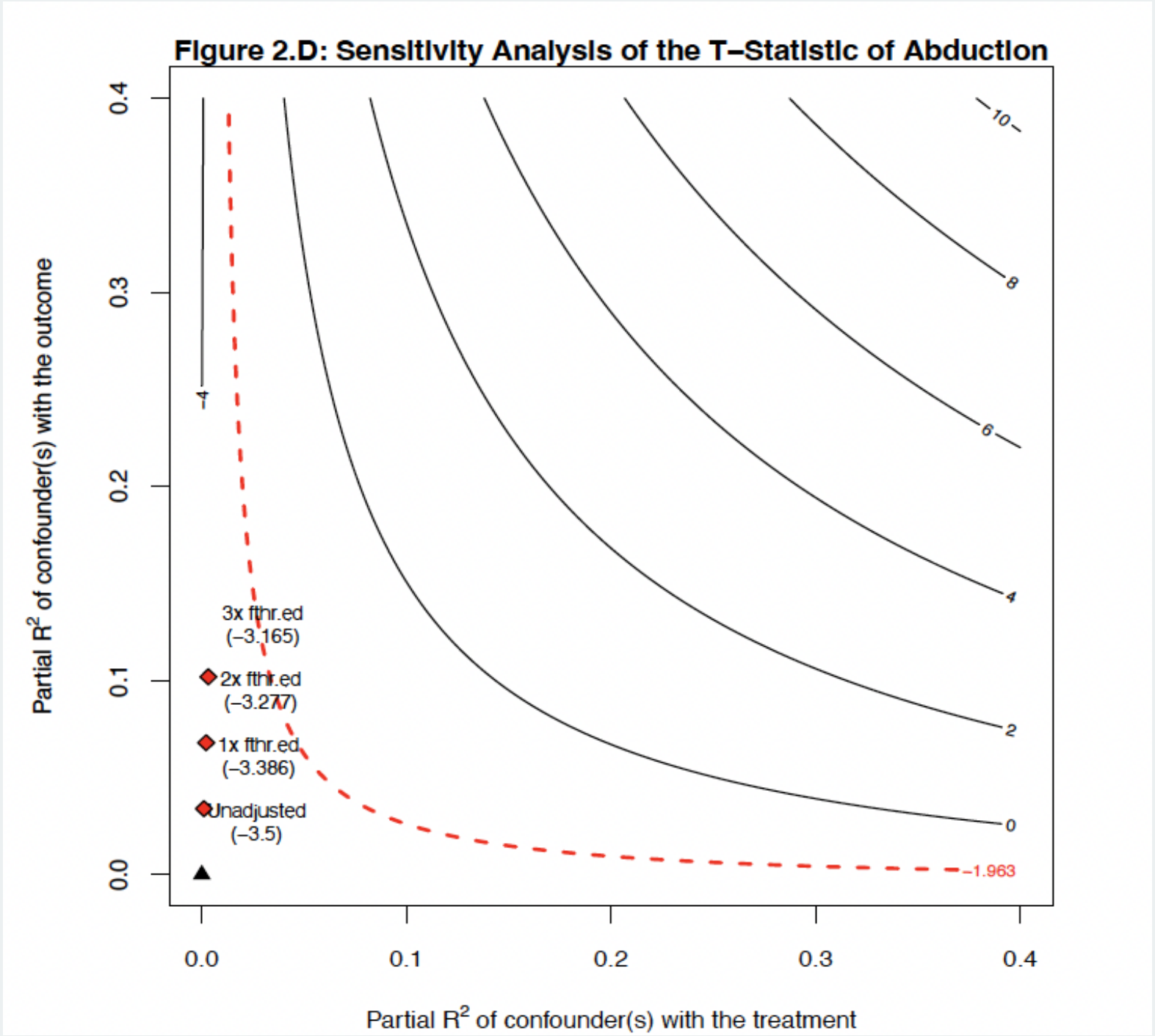
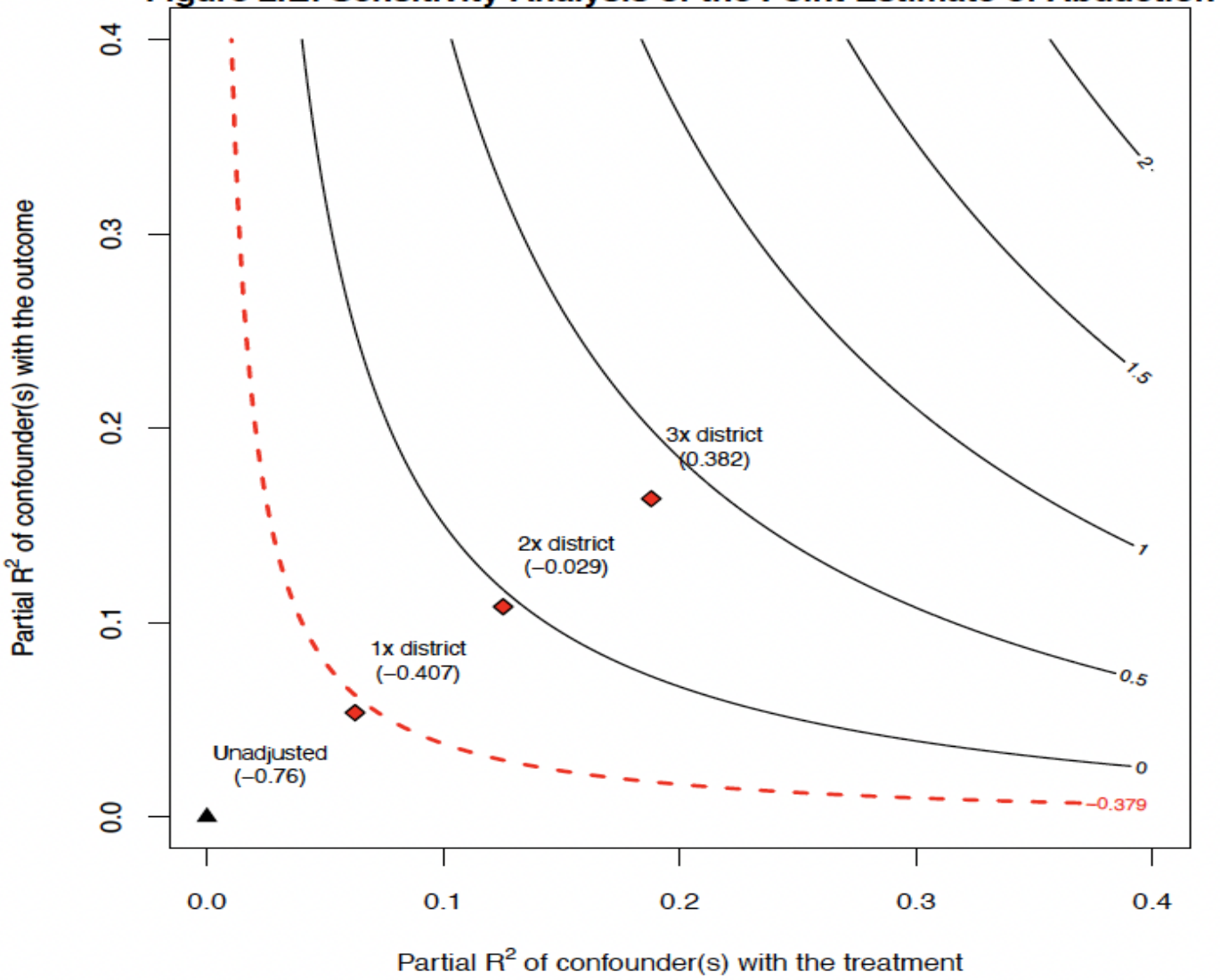


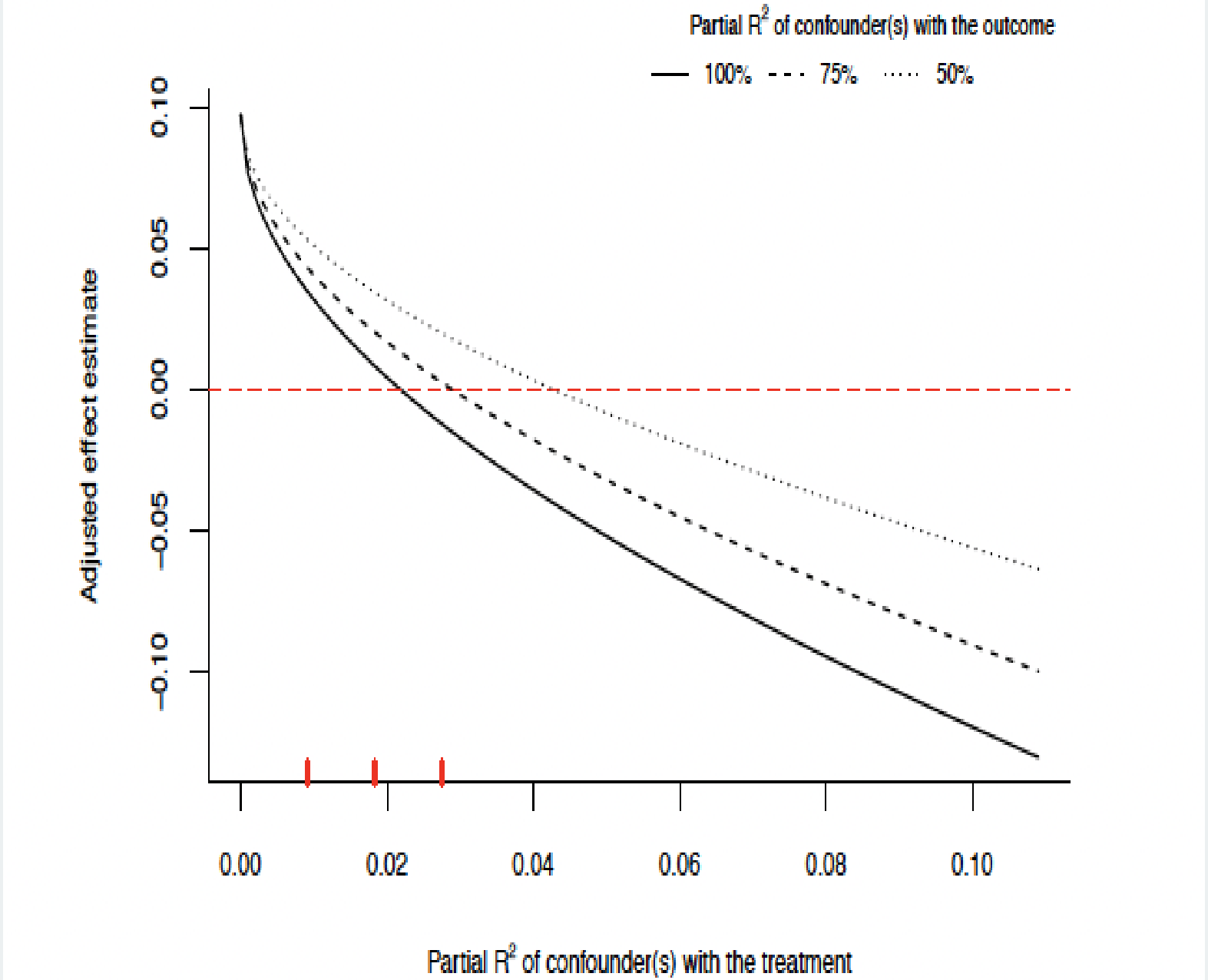
Figure 2.D: Sensitivity Analysis of the T–Statistic of Abduction
Partial R2 of confounder(s) with the outcome
Partial R2 of confounder(s) with the treatment
0.0
0.1
0.2
0.3
0.4
10
8
6
4
2
0
4
3x fthr.ed
(–3.165)
2x fthr.ed
(–3.277)
1x fthr.ed
(–3.386)
Unadjusted
(–3.5)
–1.963

Figure 2.E: Sensitivity Analysis of the Point Estimate of Abduction



Outcome: *educ*

Treatment:	Est.	S.E.	t-value	$R^2_{Y\sim D\|\mathbf{X}}$	$RV_{q=0.5}$	$RV_{q=0.5,\alpha=0.05}$
abd	-0.758	0.217	-1.747	1.6%	6.3%	0%
df = 729		*Bound (1x district)*: $R^2_{Y\sim Z\|\mathbf{X},D} = 5.4\%$, $R^2_{D\sim Z\|\mathbf{X}} = 6.3\%$				


Partial R2 of confounder(s) with the outcome
100%
75%
50%
Adjusted effect estimate
0.10
0.05
0.00
−0.05
−0.10
0.00
0.02
0.04
0.06
0.08
0.10
Partial R2 of confounder(s) with the treatment