Review

Andrew Bertoli

4 December 2013

Roadmap

1. Review
2. Midterm

Review

Basic Definition of Causality (for this class):

The causal effect of X on Y is the difference between Y when X is present and Y when X is absent.

Review

Key Points

1. Need a counterfactual
2. Causal inference is a missing data problem

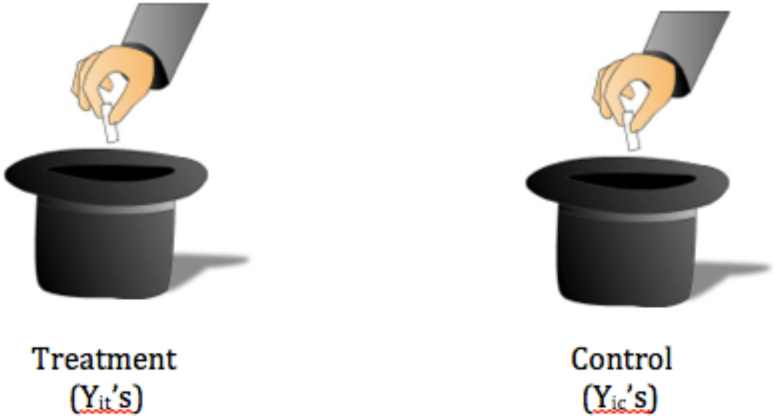
Review

Key Points

1. Need a counterfactual
2. Causal inference is a missing data problem

Review

Running an experiment is like randomly sampling from all the Y_{it}'s and Y_{ic}'s.



Review

Assumptions for an Experiment

1. $\{Y_{it}, Y_{ic}\} \perp\!\!\!\perp T_i$
2. $\{Y_{it}^{l}, Y_{ic}^{l}\} = \{Y_{it}^{L}, Y_{ic}^{L}\}$ (SUTVA or non-interference)

Review

Assumptions for Regression

1. $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \epsilon$

2. All independent and control variables are fixed (no measurement error)

3. There is no deterministic linear relationship between the X variables (no collinearity)

4. $E[\epsilon_i] = 0$ for all i

5. $\epsilon_i \sim N(0, \sigma^2)$ for all i

Assumptions 1-4 are necessary for $\hat{\beta}$ to be unbiased. Assumption 5 is required for the standard errors, p-values, and confidence intervals to be correct.

Review

A researcher wants to determine if there is discrimination against women in the workforce. He uses the model

$$\text{Salary} = a + b \cdot \text{Education} + c \cdot \text{Experience} + d \cdot \text{Man} + \epsilon$$

where man is a dummy variable indicating whether the person is a man.

Review

For our conclusions to be right, the model needs to be right. This means

a) The effect of every extra year of education should be constant, regardless of whether it's from 2nd to 3rd grade or from 11th to 12th grade.

b) Likewise, the effect of every extra year of experience should be constant.

c) The ϵ_i must be drawn i.i.d. $N(0, \sigma^2)$

d) We can think about what it means to manipulate someone's gender

e) Our treatment cannot affect the control variables

Review

Assumptions for Matching

1. $\{Y_{it}, Y_{ic}\} \perp\!\!\!\perp T_i | X$
2. $\{Y_{it}^{I}, Y_{ic}^{I}\} = \{Y_{it}^{L}, Y_{ic}^{L}\}$ (SUTVA or non-interference)

Review

Types of Matching

1. Mahalonobis Distance
2. Propensity Score
3. Genetic
4. Synthetic

Matching

Potential Problems with Matching

1. There might not be support in the data.

2. There might be support, but you chose the wrong X's.

3. You might have support and the right X's, but your formula for the propensity score is wrong (if you are doing propensity score matching) or your controls do not each follow an elliptic distribution (if you are doing Mahalanobis distance matching).

4. Everything else worked, but there was noise in your controls.

5. Everything worked perfectly, but people will still be skeptical or think that you p-hacked.

Review

Achieving better balance solves one of these problems.

1. There might not be support in the data.

2. There might be support, but you chose the wrong X's.

3. You might have support and the right X's, but your formula for the propensity score is wrong (if you are doing propensity score matching) or your controls do not each follow an elliptic distribution (if you are doing Mahalanobis distance matching).

4. Everything else worked, but there was noise in your controls.

5. Everything worked perfectly, but people will still be skeptical or think that you p-hacked.

Review

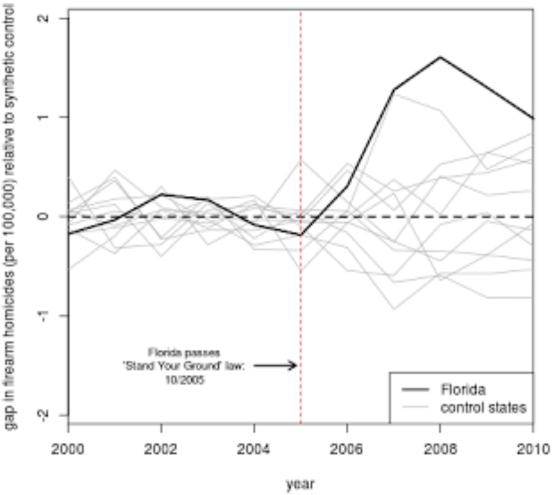
Suggestions

1. As a placebo test, redo the matching without the previous outcome and test the previous outcome. The two groups should be balanced on the previous outcome.

2. If possible, plot your outcome variable as a function of time before and after treatment. The treated and control units should look similar before treatment, but diverge afterwards (remember to also do this without matching on the previous outcome).

3. In general, matching studies that just show better balance and post-treatment tests should not be trusted.

Review



gap in firearm homicides (per 100,000) relative to synthetic control
2
1
0
-1
-2
2000
2002
2004
2006
2008
2010
year
Florida passes
'Stand Your Ground' law:
10/2005
Florida
control states

Review

Suggestions

1. As a placebo test, redo the matching without the previous outcome and test the previous outcome. The two groups should be balanced on the previous outcome.

2. If possible, plot your outcome variable as a function of time before and after treatment. The treated and control units should look similar before treatment, but diverge afterwards (remember to also do this without matching on the previous outcome).

3. In general, matching studies that just show better balance and post-treatment tests should not be trusted.

Review

Two Approaches for Regression Discontinuities

1. As-if randomness in a window around the cut-point
2. Continuity in potential outcomes at the cut-point

Review

Suggestions

1. Make sure that your results hold for both approaches if possible

2. Check balance across a number of covariates to ensure that there is no evidence of sorting.

3. In general, whenever you face a difficult decision like the size of RD window or the method of bootstrapping standard errors, justify the choice you made and report the results for other reasonable choices in the online appendix.

4. You can often use a difference-in-differences estimator to help eliminate bias. This will also probably decrease the standard errors of your tests.

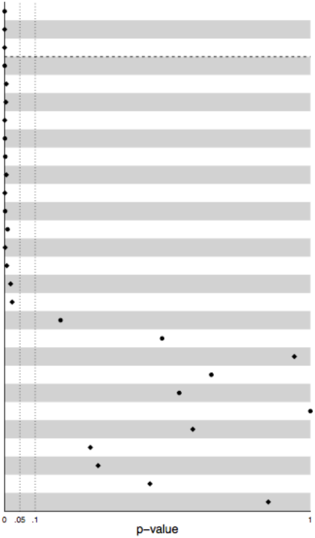
5. As a robustness check for the as-if randomness approach, use regression to control for a number of covariates.

Review

Warning: RDs do not always work.

Review

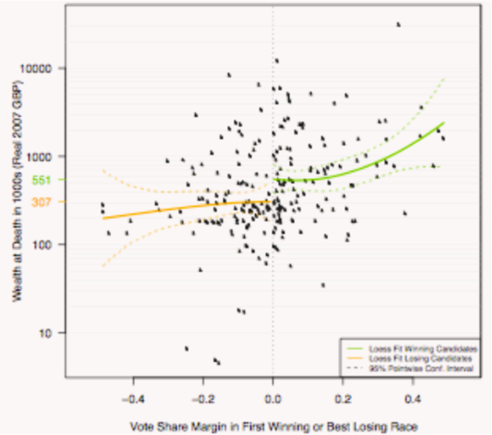
Variable Name	Valid Cases	Treated Mean	Control Mean
Dem Win t + 1	*85*	*0.74*	*0.33*
Dem % t + 1	*85*	*53*	*43*
Dem % Margin t + 1	*84*	*8.6*	*−8.0*
Dem Win t − 1	85	0.58	0.19
Dem % t − 1	85	51	45
Dem % Margin t − 1	84	5.2	−6.4
Inc's D1 NOMINATE	85	−0.05	0.21
Dem Inc in Race	85	0.49	0.14
Rep Inc in Race	85	0.28	0.62
Dem's # Prev Terms	85	1.8	0.98
Rep's # Prev Terms	85	0.86	2.7
Rep Experience Adv	81	0.27	0.62
Dem Experience Adv	81	0.49	0.20
Partisan Swing	85	−1.7	4.0
CQ Rating {−1, 0, 1}	69	0.23	−0.29
Dem Spending %	47	54	45
Dem Donation %	34	56	45
Dem Sec of State	85	0.47	0.31
Dem Governor	85	0.40	0.48
Dem Pres % Margin	79	−0.09	−0.10
Dem–held Open Seat	85	0.09	0.05
Rep–held Open Seat	85	0.14	0.19
Open Seat	85	0.23	0.24
Voter Turnout %	85	37	34
Pct Gov't Worker	73	5.1	4.4
Pct Urban	73	70	65
Pct Black	73	4.9	5.0
Pct Foreign Born	73	4.0	4.1



Review

Warning: Even when it is random for close pairs, sometimes you still can't answer the question of interest.

Review



Wealth at Death in 1000s (Real 2007 GBP)
10000
1000
551
307
100
10
-0.4
-0.2
0.0
0.2
0.4
Vote Share Margin in First Winning or Best Losing Race
Loess Fit Winning Candidates
Loess Fit Losing Candidates
95% Pointwise Conf. Interval

Review

Natural Experiments

1. Nature controls for observable and unobservable factors
2. Also more credible because it constrains the researcher

Review

Suggestions

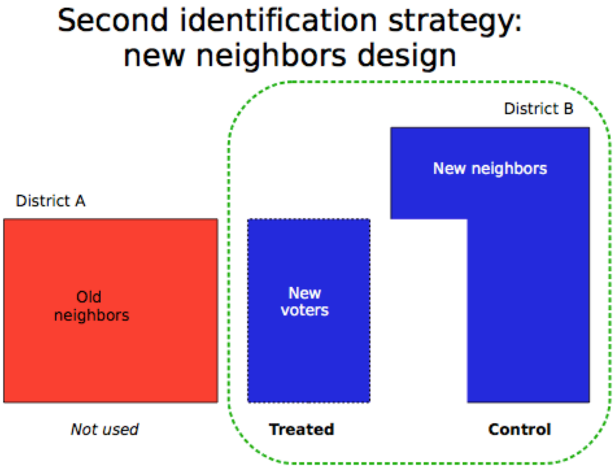
1. Check balance across a number of covariates.

2. Use a difference-in-differences estimator if possible.

3. Make sure that you define the correct treatment and control groups.

4. You will frequently run into the same problems that many other experiments have (non-compliance, non-respondents, external validity)

Review



Second identification strategy:
new neighbors design
District B
New neighbors
District A
Old
neighbors
New
voters
Not used
Treated
Control

Review

Instrumental Variables

We frequently have exogenous variation for treatments that we don't really care about, like weather or natural disasters.

IV designs try to exploit this variation to make inferences about treatments that we do care about.

Review

Assumptions

1. SUTVA
2. As-if randomness of Z
3. Exclusion Restriction (Z only affects Y through T)
4. Non-zero causal effect of Z on T
5. Z either increases or decreases T for all units (no defiers)

Review

Conclusions

1. The right methodological choices depend on the question and the data.

2. The are many questions in social science that statistics cannot answer.

3. Regression and matching studies often disguise the fact that there is no implied manipulation.

4. Regression discontinuities and natural experiments are usually more credible, but you always need to check for evidence that the design worked.

5. Even when the design works perfectly, flukes can happen. Replication and case studies greatly improve the credibility of the results.

Midterm