Conceptual Issues in Causal Inference

Jasjeet S. Sekhon

May 19, 2013

Unification and Single Interventions

- So many possible interventions, so little time
- Spend most of my time:
  - evaluating a given empirical design/identification strategy with specific data and a specific intervention
- SWIGs are easier to interpret, but are also closer to how some of us think
- If there is no experiment can be performed to verify a particular assumption, what are we doing?

Dorn's Question

- If there is no experiment can be performed to verify a particular assumption, what are we doing?
- New theory: the absence of an edge in a graph in terms of the absence of a population level direct effect
  - Absence of an edge no longer implies: $Y(x) = Y(x')$
- New theory: interventions are only defined on a strict subset of variables
  - not all the counterfactual implied by NPSEM exist
- Question: SWIGS versus structural equations?

Backdoor Criterion

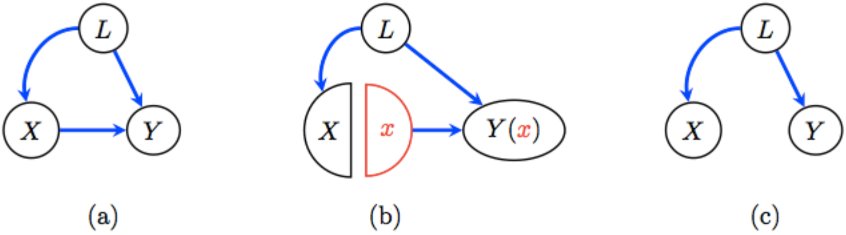


Figure 5: Adjusting for confounding. (a) The original causal graph. (b) The template $\mathcal{G}(x)$, which shows that $Y(x) \perp\!\!\!\perp X \mid L$. (c) The DAG $\mathcal{G}_{\underline{X}}$ obtained by removing edges from X advocated in Pearl (1995, 2000, 2009) to check his 'backdoor condition'.

Backdoor Criterion and Descendants

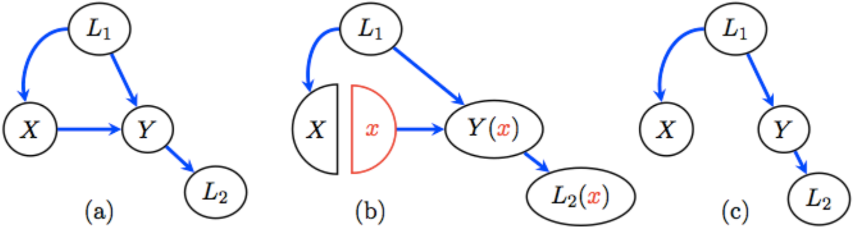
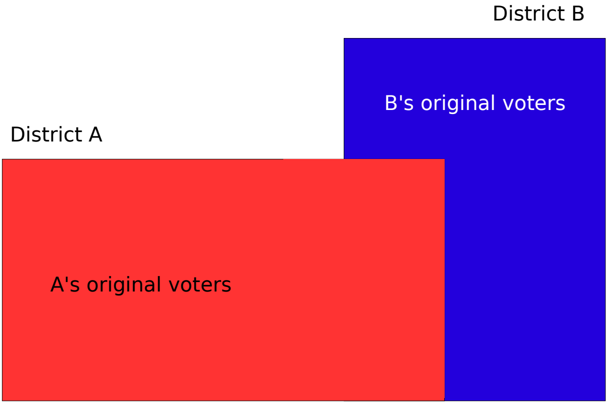
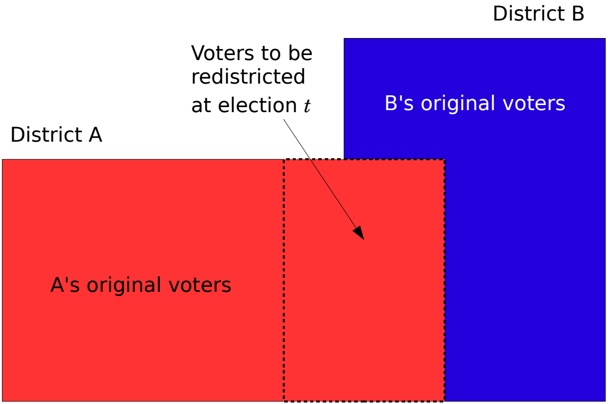


Figure 7: Simplification of the backdoor criterion. (a) The original causal graph $\mathcal{G}$. (b) The template $\mathcal{G}(x)$, which shows that $Y(x) \perp\!\!\!\perp X \mid L_1$, but does not imply $Y(x) \perp\!\!\!\perp X \mid \{L_1, L_2\}$ when there exists an arrow from X to Y, i.e. the null hypothesis is false. (c) The DAG $\mathcal{G}_{\underline{X}}$ obtained by removing edges from X advocated in Pearl (2000, 2009).

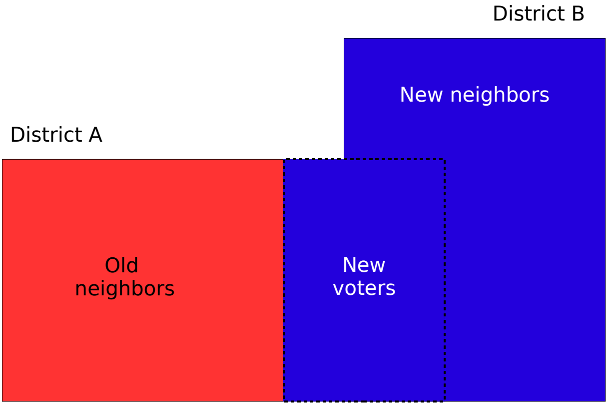
Before one-time redistricting



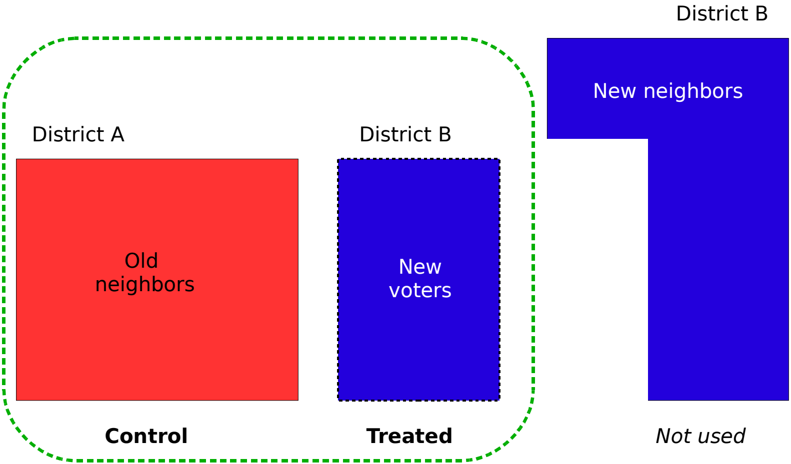
One-time redistricting



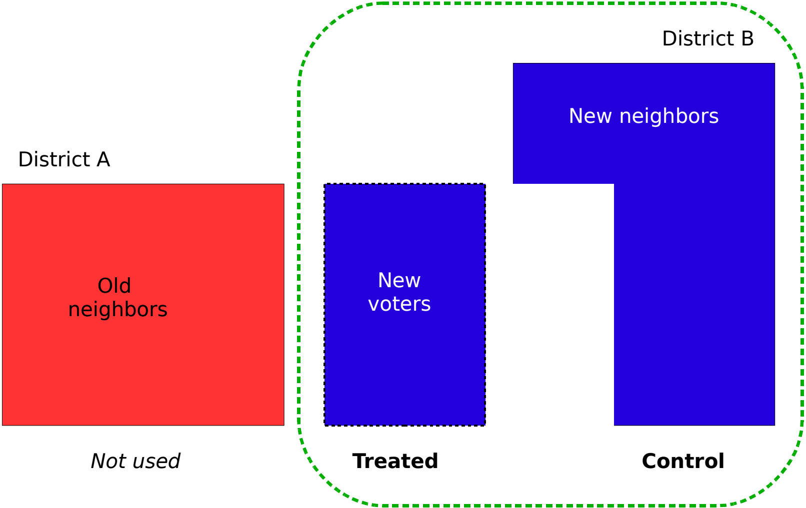
After one-time redistricting



First identification strategy: old-neighbors design



Second identification strategy: new neighbors design



Formally

- Let T_i be equal to 1 if precinct i is moved from one district to another before election t and equal to 0 if it is not moved
- Let D_i be equal to 1 if precinct i has new voters in its district at t and equal to 0 otherwise
- Let $Y_{00}(i,t)$ be precinct i's outcome $T_i = 0$ and $D_i = 0$ it is not moved and does not have new neighbors
- Let $Y_{01}(i,t)$ be precinct i's outcome if $T_i = 0$ and $D_i = 1$ the precinct is not moved and has new neighbors
- Let $Y_{11}(i,t)$ be precinct i's outcome if $T_i = 1$ and $D_i = 1$ the precinct is moved and has new neighbors

Fundamental Problem of Causal Inference

For each precinct, we observe only one of its three potential outcomes:

$$
\begin{aligned}
Y(i,t) \quad = \quad & Y_{00}(i,t) \cdot (1 - T_i) \cdot (1 - D_i) + \\
& Y_{01}(i,t) \cdot (1 - T_i) \cdot D_i + \\
& Y_{11}(i,t) \cdot T_i \cdot D_i
\end{aligned}
$$

We can estimate two different ATT's:

$$ATT_0 \equiv E\left[Y_{11}(i,t) - Y_{00}(i,t) \mid T_i = 1, D_i = 1\right]$$

$$ATT_1 \equiv E\left[Y_{11}(i,t) - Y_{01}(i,t) \mid T_i = 1, D_i = 1\right]$$

Identification of ATT_0

$$ATT_0 \equiv E\left[Y_{11}(i,t) - Y_{00}(i,t) \mid T_i = 1, D_i = 1\right]$$

Is identified if:

$$E\left[Y_{00}(i,t) \mid T_i = 1, D_i = 1\right] = E\left[Y_{00}(i,t) \mid T_i = 0, D_i = 0\right]$$

ATT_0 requires that voters who stay in A and voters who are moved from A to B would have the same average outcomes if they hadn't been moved.

Identification of ATT_1

$$ATT_1 \equiv E\left[Y_{11}(i,t) - Y_{01}(i,t) \mid T_i = 1, D_i = 1\right]$$

Is identified if:

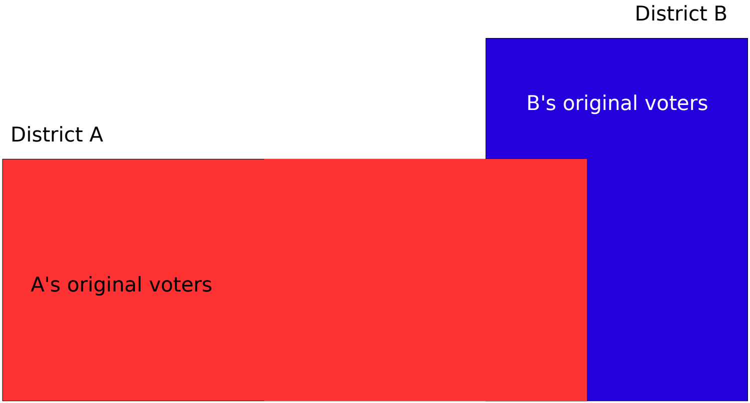
$$E\left[Y_{01}(i,t) \mid T_i = 1, D_i = 1\right] = E\left[Y_{01}(i,t) \mid T_i = 0, D_i = 1\right]$$

ATT_1 requires that voters who are originally in B and voters who are moved from A to B would have the same average outcomes if A's voters would not have been moved even though they would be in different districts.

Randomization does not imply that B's old voters are a valid counterfactual for B's new voters

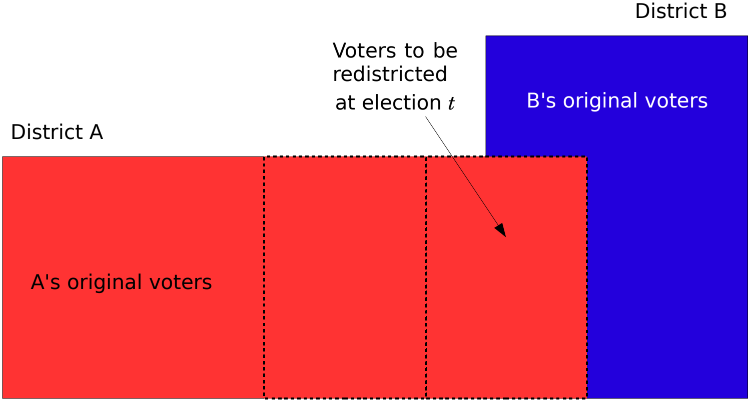
Before two-time redistricting

(election *t-1*)



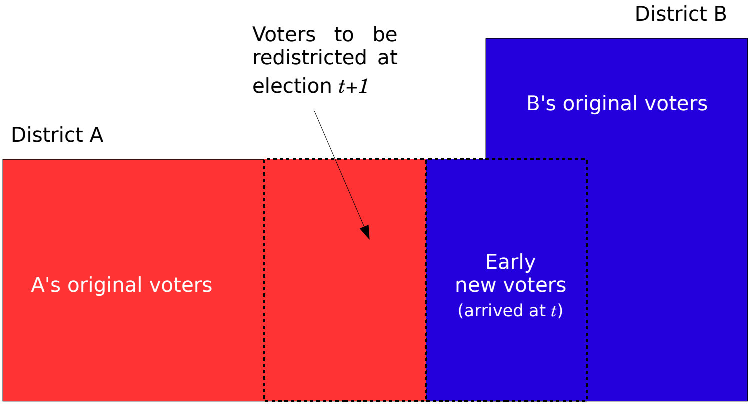
Before two-time redistricting

(election *t-1*)



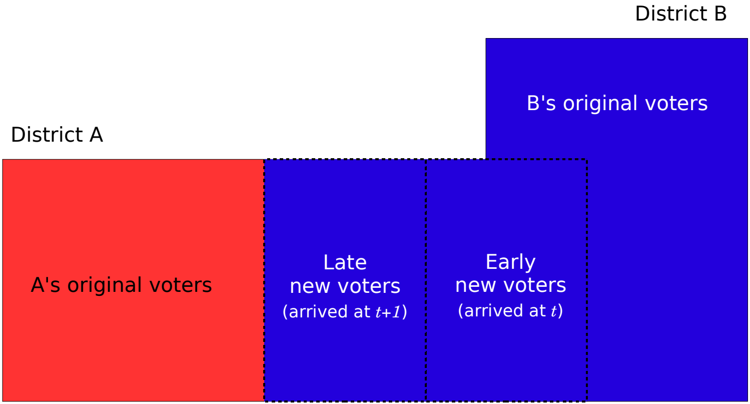
Two-time redistricting

(election t)



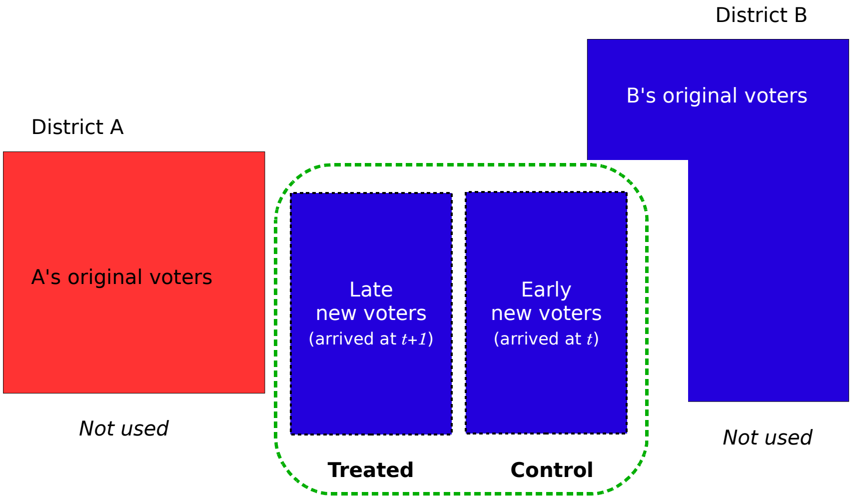
Two-time redistricting

(election $t+1$)



After two-time redistricting

(election $t+1$)



The Best Design: Multiple Redistrictings

Let $W_{i,t+1} = 1$ if precinct i is moved from district A to district B at election $t + 1$.

Let $W_{i,t+1} = 0$ if precinct i is moved from A to B at election t.

Let $Y_0\left(i, t+1\right)$ denote the outcome of i at election $t + 1$ if $W_{i,t+1} = 0$

Let $Y_1\left(i, t+1\right)$ denote the outcome of i at election $t + 1$ if $W_{i,t+1} = 1$

The parameter of interest ATT_B is

$$ATT_B \equiv E\left[Y_1\left(i, t+1\right) - Y_0\left(i, t+1\right) \mid W_{i,t+1} = 1\right]$$

The parameter of interest ATT_B is

$$ATT_B \equiv E\left[Y_1(i,t+1) - Y_0(i,t+1) \mid W_{i,t+1} = 1\right]$$

which is identified by

$$E\left[Y_0(i,t+1) \mid W_{i,t+1} = 1\right] = E\left[Y_0(i,t+1) \mid W_{i,t+1} = 0\right]$$

By randomization we have

$$E\left[Y_0(i,t-1) \mid W_{i,t+1} = 1\right] = E\left[Y_0(i,t-1) \mid W_{i,t+1} = 0\right]$$

Randomization along with the following stability assumption provides identification:

$$\begin{aligned} E\left[Y_0(i,t+1) - Y_0(i,t-1) \mid W_{i,t+1} = 1\right] &= \\ E\left[Y_0(i,t+1) - Y_0(i,t-1) \mid W_{i,t+1} = 0\right] \end{aligned}$$

One-Time Redistricting Designs

There are two obvious designs under the randomization where everyone is in a certain district at $t-1$ and some precincts are randomly moved to another district at t.

- Compare new voters with old neighbors
- Compare new voters with new neighbors
- Randomization ensures exchangeability for the first design, but not the second
- This occurs because the history of new voters with new neighbors is not balanced by randomization