# Statistical Methods III: Spring 2013

Jonathan Wand

Stanford University

RDD + applications

# Outline

# Incumbency advantage

**Question:** Does incumbency provide an electoral advantage?

Regression approach (e.g., Gelman–King)

$$\mathbb{E}[V_{t+1}] = \beta_0 + \beta_1 P_t + \beta_2 (P_t \times R_{t+1}) + \beta_3 V_t,$$

- $V_t \in [0, 1]$ is the *Dem Share* in election $t$
- $P_t \in \{-1, 1\}$ is the *Winning Party* in election $t$
- $R_{t+1} \in \{0, 1\}$ indicates whether *Incumbent Runs* in election $t + 1$

Q: which parameter is "incumbency advantage"?
What are the possible threats to inference?

- strategic exit (decision of Incumbents)
- strategic entry (decision of Challengers)
- selection effect of elections (better candidates win)

Q: which of these are "corrected" by regression?

# Incumbency advantage

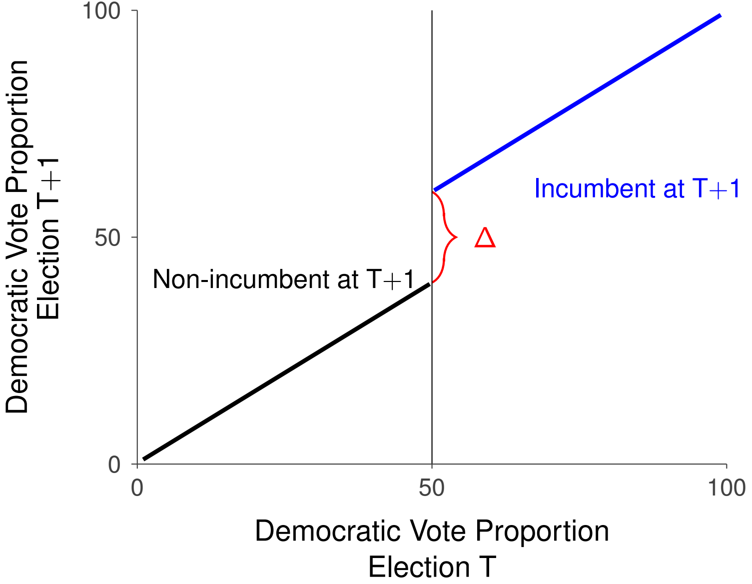
**Question:** Does incumbency provide an electoral advantage?

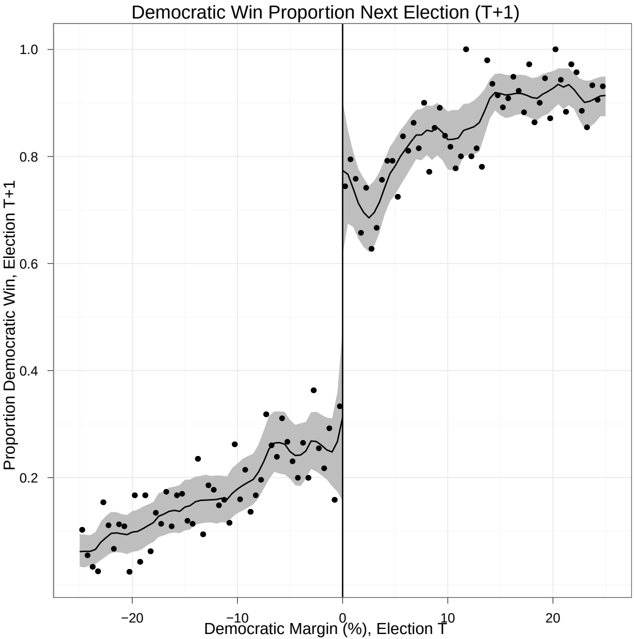
Quasi experimental approach (e.g., RDD, Lee 2008)

$$\Delta = E[Y_{t+1} \mid \text{won with 50 perc}] = E[Y_{t+1} \mid \text{lost with 50 perc.}]$$

At $t + 1$, compare vote share of *party* which just won at time $t$, with those that just lost at $t$.

# RDD – election discontinuity




Democratic Win Proportion Next Election (T+1)
Proportion Democratic Win, Election T+1
1.0
0.8
0.6
0.4
0.2
−20
−10
0
10
20
Democratic Margin (%), Election T

# RDD v regression

Which problem are same, which are different from regression?

- strategic exit (decision of Incumbents)?
- strategic entry (decision of Challengers)?
- selection effect of elections (better candidates win)?

# RDD: notation

Let

- $Z$ is a continuous variable,
  termed "forcing" variable: along which discontinuity occurs
  e.g., vote share, GRE score
- $c$ deterministic and exogenously given threshold for discontinutity
  e.g., 50%+1 (majority, plurality can also work); 790=scholarship.

# RDD: Assignment mechanism

- Sharp design: assignment is determistic function of known, common cutpoint $c$ and value of $Z_i$

$$D_i = \begin{cases} 1 & \text{if } Z_i \geq c \\ 0 & \text{if } Z_i < c \end{cases}$$

- Fuzzy design: assignment probabilistic function of known, common cutpoint $c$ and value of $Z_i$

$$\lim_{z \downarrow c} P(D_i = 1) \neq \lim_{z \uparrow c} P(D_i = 1)$$

- Potential outcome is function of both $Z_i$ and potentially $D_i$

$$Y_i(0) = \mu_0(Z_i) + \epsilon_{0i}$$

$$Y_i(1) = \mu_1(Z_i) + \epsilon_{1i}$$

# RDD: Parameter

Treatment at discontinuity (TAD)

$$TAD = E[Y_i(1) \mid D_i = 1, Z_i = c] - E[Y_i(0) \mid D_i = 1, Z_i = c]$$

NOTE:

- again, we do not observe $Y_i(0)$ and $Y_i(1)$,
- because at $c$ we do observe $Y_i(0)$ (sharp RDD)

Study limits approaching each side of the threshold

$$\Delta_{RD} = \lim_{z \downarrow c} E[Y_i \mid Z_i = z] - \lim_{z \uparrow c} E[Y_i \mid Z_i = z]$$

Under what conditions is TAD identified at limit?

# RDD: Sources of bias

$$
\begin{aligned}
\Delta_{RD} = & \lim_{z \downarrow c} E[Y_i \mid Z_i = z] - \lim_{z \uparrow c} E[Y_i \mid Z_i = z] \\
& + E[Y_i(1) \mid D_i = 1, Z_i = c] - E[Y_i(1) \mid D_i = 1, Z_i = c] \\
& + E[Y_i(0) \mid D_i = 1, Z_i = c] - E[Y_i(0) \mid D_i = 1, Z_i = c] \\
& + E[Y_i(0) \mid D_i = 0, Z_i = c] - E[Y_i(0) \mid D_i = 0, Z_i = c]
\end{aligned}
$$

$$
\begin{aligned}
= & E[Y_i(1) \mid D_i = 1, Z_i = c] - E[Y_i(0) \mid D_i = 1, Z_i = c] \\
& + E[Y_i(0) \mid D_i = 1, Z_i = c] - E[Y_i(0) \mid D_i = 0, Z_i = c] \\
& + \lim_{z \downarrow c} E[Y_i \mid Z_i = z] - E[Y_i(1) \mid D_i = 1, Z_i = c] \\
& + E[Y_i(0) \mid D_i = 0, Z_i = c] - \lim_{z \uparrow c} E[Y_i \mid Z_i = z]
\end{aligned}
$$

# RDD: Sources of bias

1. $E[Y_i(0) \mid D_i = 1, Z_i = c] - E[Y_i(0) \mid D_i = 0, Z_i = c]$

   selection bias: can indviduals control which side they are on? randomization would makes this zero

2. $\lim_{z \downarrow c} E[Y_i \mid Z_i = z] - E[Y_i(1) \mid D_i = 1, Z_i = c]$

   gap approaching limit, from above: continuity in $\mu_1(Z_i)$ would make this zero

3. $E[Y_i(0) \mid D_i = 0, Z_i = c] - \lim_{z \uparrow c} E[Y_i \mid Z_i = z]$

   gap approaching limit, from below: continuity in $\mu_0(Z_i)$ would make this zero

# RDD: Assumptions

Let there be a known, fixed $z = c$ such that

(1) discontinuity in treatment assignment

$$\lim_{z \downarrow c^+} Pr(D = 1 \mid Z = z) \neq \lim_{z \uparrow c^-} Pr(D = 1 \mid Z = z)$$

(2) continuity in potential outcomes

$$\lim_{z \downarrow c^+} Pr(Y_j \leq r \mid Z = z) = \lim_{z \uparrow c^-} Pr(Y_j \leq r \mid Z = z) \quad (j = 0, 1)$$

(3) Continuity in other covariates, possible confounders

$$\lim_{z \downarrow c^+} Pr(X_j \leq r \mid Z = z) = \lim_{z \uparrow c^-} Pr(X_j \leq r \mid Z = z) \quad (j = 0, 1)$$

there are no jumps at the treatment threshold $c$

# RDD: Assumptions

Let there be a known, fixed $z = c$ such that

(1.b) discontinuity in treatment assignment, sharp design

$$\lim_{z \downarrow c^+} Pr(D = 1 \mid Z = z) = 1$$

$$\lim_{z \uparrow c^-} Pr(D = 1 \mid Z = z) = 0$$

Meaning: individual receives treatment iff observed covariate $Z_i$ crosses known threshold $c$

(2.b) continuity in expectation of potential outcomes

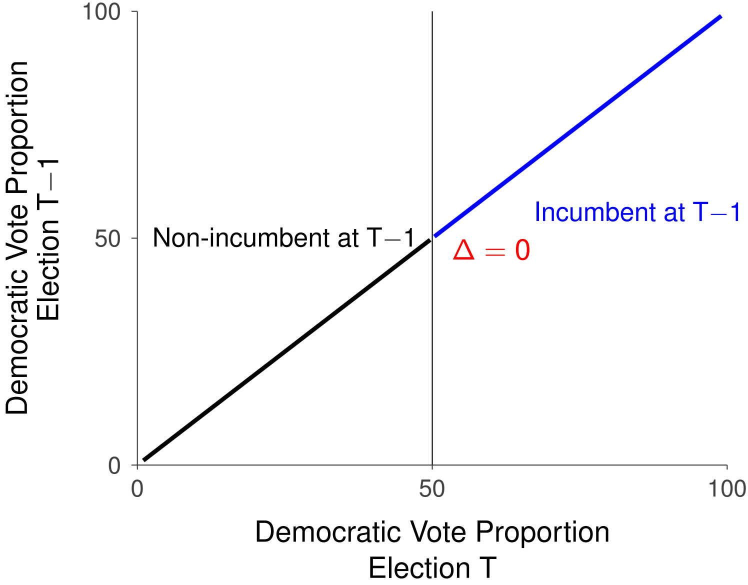
$$\lim_{z \downarrow c^+} E(Y_j \mid Z = z) = \lim_{z \uparrow c^-} E(Y_j \mid Z = z) \quad (j = 0, 1)$$

(3.b) Continuity in expectation of covariates

$$\lim_{z \downarrow c^+} E(X_j \mid Z = z) = \lim_{z \uparrow c^-} E(X_j \mid Z = z) \quad (j = 0, 1)$$

Q: what does this imply about comparison of $X$ on either side of $c$?

# RDD – placebo test, balance test



# RDD: homogeneous treatment effects, sharp RDD

Assume (1.b) and (2.b), and $\tau = Y_1 - Y_0$, such that,

$$Y_i = \tau D_i + Y_{0i}$$

Using assumptions (1.a)-(1.c) and homogenous effect model, let's prove that

$$\begin{aligned}\Delta_{RD} &= \lim_{z \downarrow c^+} E[Y|Z=z] - \lim_{z \uparrow c^-} E[Y|Z=z] \\ &= \tau \\ &= E[Y(1) - Y(0)|Z=c]\end{aligned}$$

# RDD: homogeneous treatment effects, sharp RDD

Let $\epsilon > 0$, and consider, by (2.b, continuity) and homogeneity,

$$\begin{aligned} Y^+ = \lim_{z \downarrow c^+} E[Y|Z=z] &= \lim_{\epsilon \downarrow 0} E[Y_i(1)|Z_i = c+\epsilon] \\ &= \lim_{\epsilon \downarrow 0} E[Y_i(0)+\tau|Z_i = c+\epsilon] \\ &= \tau + \lim_{\epsilon \downarrow 0} E[Y_i(0)|Z_i = c+\epsilon] \\ &= \tau + E[Y_i(0)|Z_i = c] \end{aligned}$$

and

$$\begin{aligned} Y^- = \lim_{z \uparrow c^-} E[Y|Z=z] &= \lim_{\epsilon \uparrow 0} E[Y_i(0)|Z_i = c-\epsilon] \\ &= E[Y_i(0)|Z_i = c] \end{aligned}$$

So,

$$Y^+ - Y^- = \tau = E[Y(1) - Y(0)|Z = c]$$

# RDD: heterogeneous treatment effects

Assume (1.b) and (2.b), suppose $\tau_i = Y_{1i} - Y_{0i}$

**Theorem** $E[\tau_i | Z_i = c] = Y^+ - Y^-$

**Proof**

$$\lim_{\epsilon \downarrow 0} E[Y_i | Z_i = c + \epsilon_i] = \lim_{\epsilon \downarrow 0} E[Y_i(0) + \tau_i | Z_i = c + \epsilon_i]$$
$$= E[\tau_i | Z_i = c] + E[Y_i(0) | Z_i = c]$$

$$\lim_{\epsilon \uparrow 0} E[Y_i | Z_i = c + \epsilon_i] = \lim_{\epsilon \uparrow 0} E[Y_i(0) | Z_i = c + \epsilon_i]$$
$$= E[Y_i(0) | Z_i = c]$$

$$Y^+ - Y^- = E[\tau_i | Z_i = c] + E[Y_i(0) | Z_i = c] - E[Y_i(0) | Z_i = c] = E[\tau_i | Z_i = c]$$

# RDD: homogeneous treatment effects, fuzzy

Under fuzzy, can solve for

$$\tau = \frac{\lim_{z \downarrow c^+} E(Y|Z=z) - \lim_{z \uparrow c^-} E(Y|Z=z)}{\lim_{z \downarrow c^+} E(D_i|Z=z) - \lim_{z \uparrow c^-} E(D_i|Z=z)}$$

Q: What do you call this estimator?
Q: Where have you seen it before?

Simplifies under sharp

$$\tau = \lim_{z \downarrow c^+} E(Y|Z=z) - \lim_{z \uparrow c^-} E(Y|Z=z)$$

Q: How/why?

# RDD: Estimation

Consider

$$Y^{+} = \lim_{z \downarrow c} E[Y_i \mid Z_i = z]$$

$$Y^{-} = \lim_{z \uparrow c} E[Y_i \mid Z_i = z]$$

Can we estimate these quantities?

What would we need?

- Do we have enough (or any) data at the limit for $c$ ?
- At the limit can we observe both $D = 1$ and $D = 0$?

# RDD: Estimation

Estimator, for some $\delta$ that *you* choose

$$Y^{+} = \frac{\sum_{i=1}^{n} Y_i \cdot 1(Z_i \in [c, c+\delta])}{\sum_{i=1}^{n} 1(Z_i \in [c, c+\delta])}$$

$$Y^{+} = \frac{\sum_{i=1}^{n} Y_i \cdot 1(Z_i \in [c, c+\delta])}{\sum_{i=1}^{n} 1(Z_i \in [c, c+\delta])}$$

# RDD

Hahn, Todd, deKlaauw (2001)

- treatment assignment insufficient to identify TAD
- proved need for continuity restrictions
- but hard to adjudicate / unclear behavioral implications/assumptions

Lee (2007)

- key is that location of $Z$ around threshold smoothly probabilistic
- no one near threshold can "control" precisely which side of $c$ they are on

# RDD: Lee

Let $V = Z - c$ be the observable forcing variable

- Condition (2b) $V_i$ must be drawn from distribution that is sufficiently smooth
  specifically, $F(v|w)$ is continuously differentiable at $v = 0$
- $V$ can be a function of (unobserved) type/effort $W$
- $V$ can be correlated with potential outcomes
- assignment

$$D_i = 1(V_i \geq 0)$$

What does continuity of $V$ give?

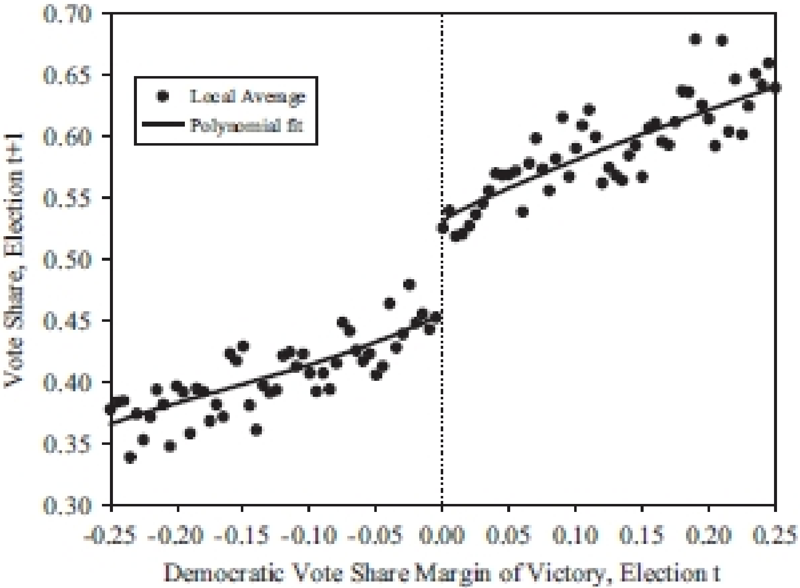
- probability of being just above/below $c$ the same
- people can't pile up on one side of $c$
- same distribution of characteristics on either side of $c$

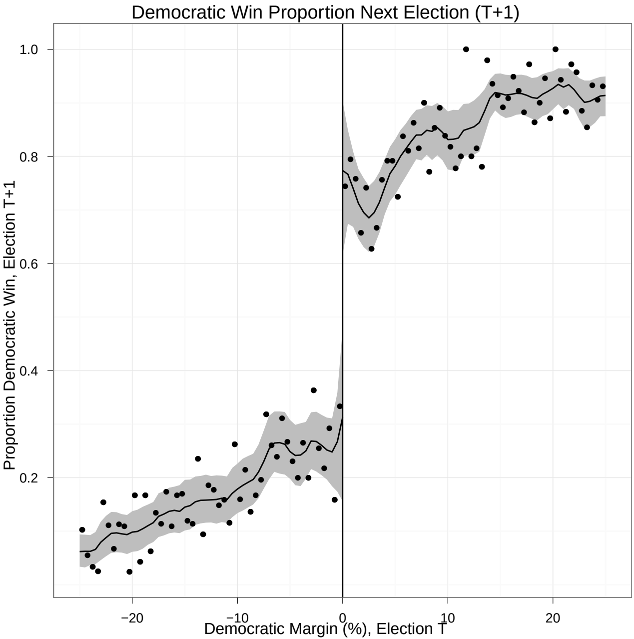
# RDD: difficulties

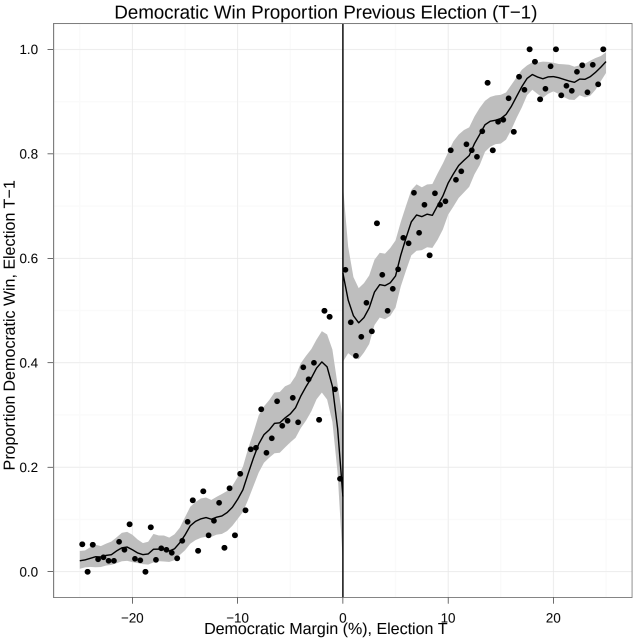
- selection / manipulation of forcing vairable
  - Can someone stop/force which side they are on
  - Do we change who is around the cutpoint individuals fighting to get over cutpoint?

  May be able to test for both of these; only the latter is problematic
- estimation / functional form
  - Is there enough data near cutpoint?
  - To what extent do results change as a function of model choice
- specificity: context / localness
  - Is TAD really of interest?
  - To what extent does TAD describe potential effects elsewhere on $Z$

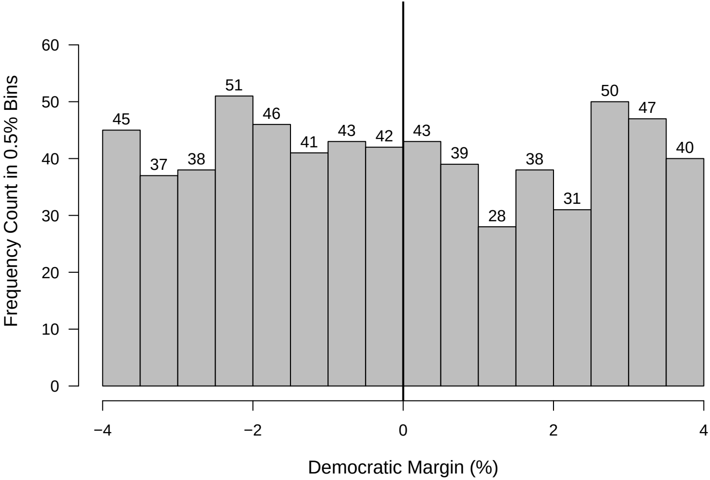
# RDD: Lee (2008) vote share, t+1



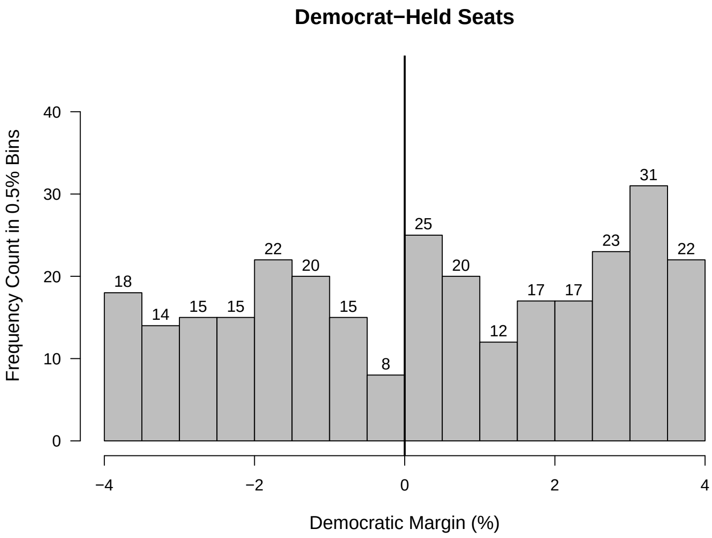

Democratic Win Proportion Next Election (T+1)
Proportion Democratic Win, Election T+1
1.0
0.8
0.6
0.4
0.2
−20
−10
0
10
20
Democratic Margin (%), Election T


Democratic Win Proportion Previous Election (T−1)
Proportion Democratic Win, Election T−1
1.0
0.8
0.6
0.4
0.2
0.0
−20
−10
0
10
20
Democratic Margin (%), Election T

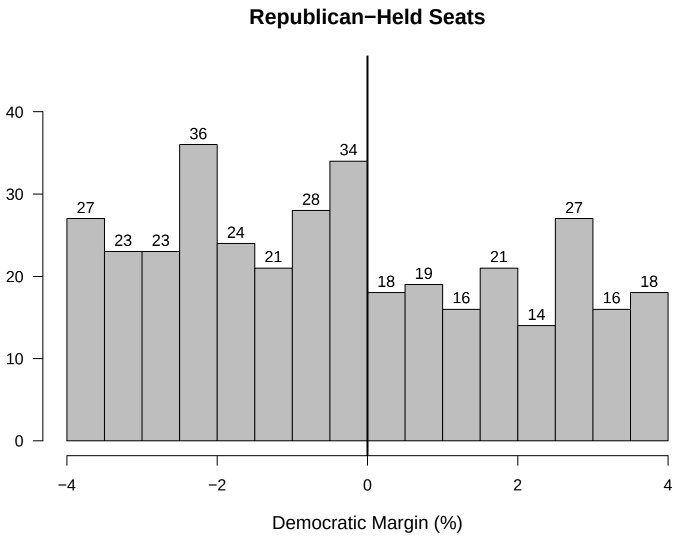
# Democratic Margin in Close Elections at *t*



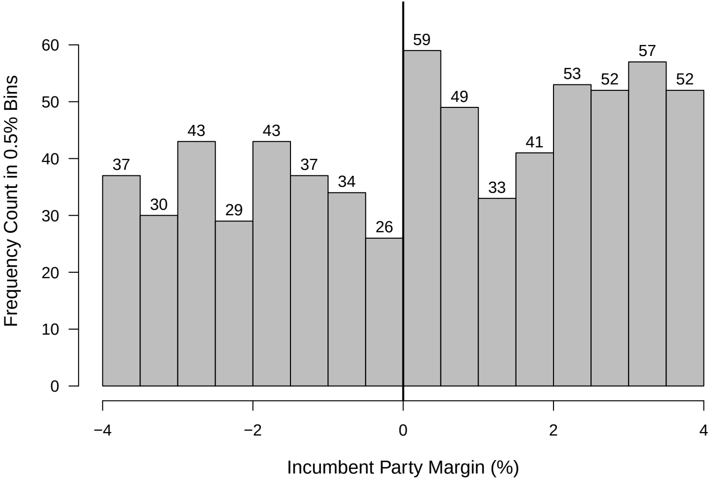
# Broken Down by Incumbent Party at $t$

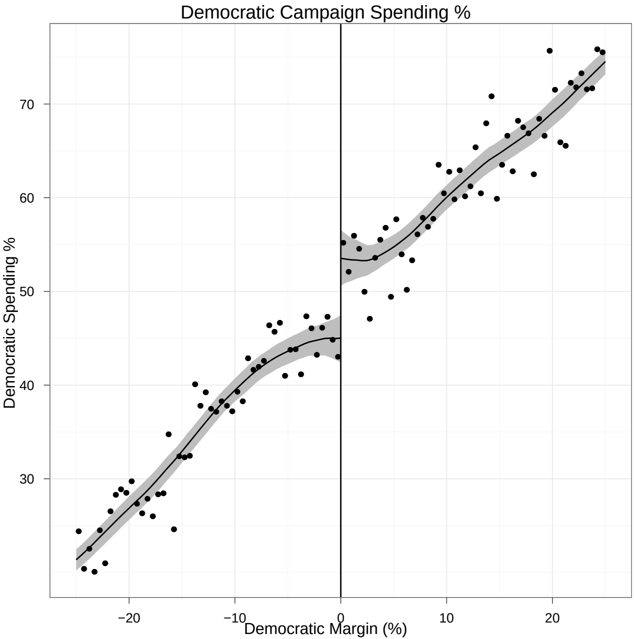


# Broken Down by Incumbent Party at $t$

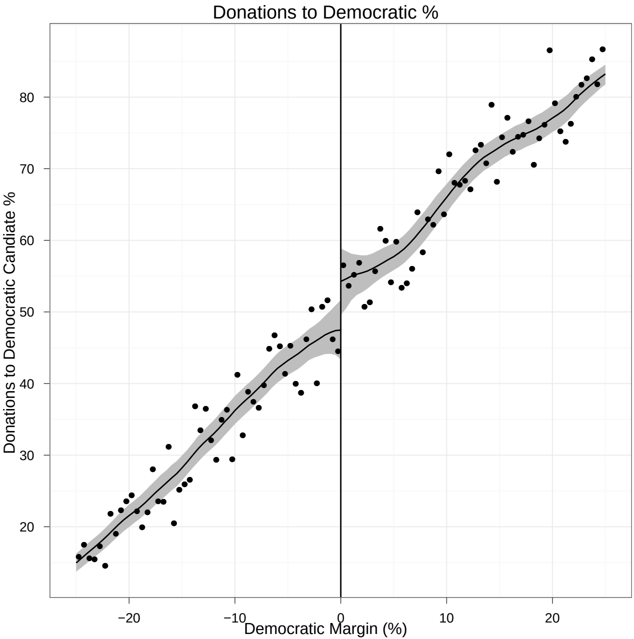


# Incumbent Party's Margin in Close Elections at *t*





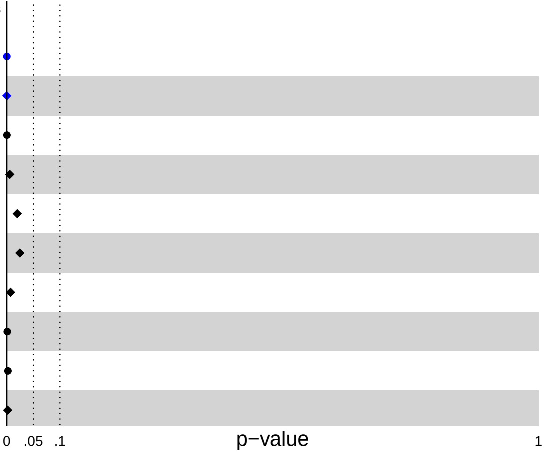
Democratic Campaign Spending %
Democratic Spending %
70
60
50
40
30
−20
−10
0
10
20
Democratic Margin (%)



Donations to Democratic %
Donations to Democratic Candiate %
80
70
60
50
40
30
20
−20
−10
0
10
20
Democratic Margin (%)

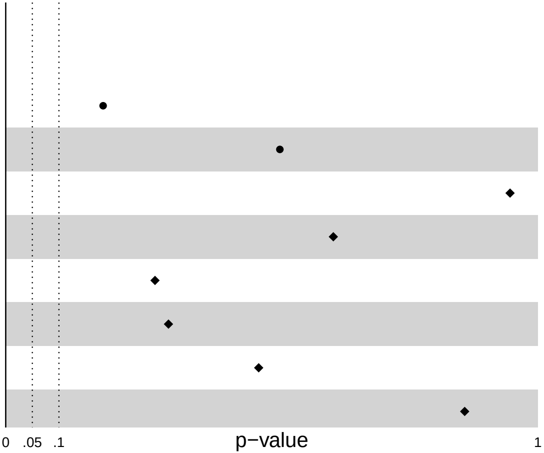
# Covariate Balance

| Variable Name | Valid Cases | D Win Mean | D Loss Mean |
|---|---|---|---|
| *Dem Win t + 1* | *85* | *0.74* | *0.33* |
| *Dem % t + 1* | *85* | *53* | *43* |
| Dem Win t − 1 | 85 | 0.58 | 0.19 |
| Dem % t − 1 | 85 | 51 | 45 |
| Dem Spending % | 47 | 54 | 45 |
| Dem Donation % | 34 | 56 | 45 |
| CQ Rating {−1, 0, 1} | 69 | 0.23 | −0.29 |
| Dem Inc in Race | 85 | 0.49 | 0.14 |
| Rep Inc in Race | 85 | 0.28 | 0.62 |
| Partisan Swing | 85 | −1.7 | 4.0 |



# Covariate Balance

| Variable Name | Valid Cases | D Win Mean | D Loss Mean |
|---|---|---|---|
| Dem Sec of State | 85 | 0.47 | 0.31 |
| Dem Governor | 85 | 0.40 | 0.48 |
| Dem Pres % Margin | 79 | −0.09 | −0.10 |
| 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 |



# National Partisan Swings

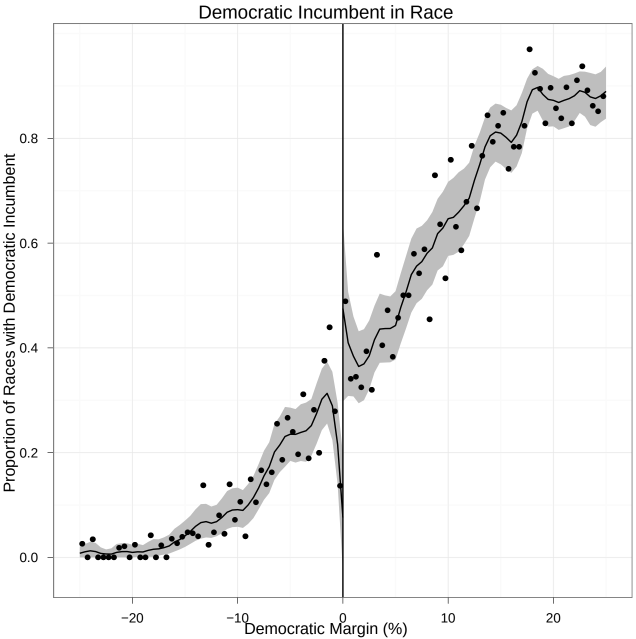
- Partisan swings are imbalanced:
  - 1958 (pro-Democratic tide): all 6 close elections occurred in Republican-held seats
  - 1994 (pro-Republican tide): all 5 close elections occurred in Democratic-held seats
  - Close elections do not generally occur in 50/50 districts

# There is Strategic Exit

Strategic exit among incumbents

- 20% who barely win retire prior to next election
- but *all* candidates who barely won first election ran for reelection

- Change over time:
  - 1994: less evidence of strategic exit, lots of strategic entry
  - 2006: 18 Republican open seats versus 9 Democratic open seats
  - 2010: In non-safe seats, 6 Republicans retired, but 15 Democrats retired


Democratic Incumbent in Race
Proportion of Races with Democratic Incumbent
0.0
0.2
0.4
0.6
0.8
−20
−10
0
10
20
Democratic Margin (%)

# RDD: practical suggestions

1. Graph data
2. Estimate using (flexible) linear regression in small bandwidth
3. check robustness of assumptions