Appendix

Intended for online publication only.

Survey

The survey intro was:

- "Consider the party incumbency effect in elections for statewide offices in the U.S. in recent decades.

  When using a regression discontinuity design that exploits variation from close elections we find that party incumbency increases the two-party vote share in the next election by around 8-9 percentage points on average.

  As is well known, this effect estimate only refers to very close elections that are decided within a narrow window around the 50% vote share threshold of winning, e.g. elections in which the party barley won with a vote of 50.5%.

  Here we are interested in your expectation of what the party incumbency effect might be in districts where the winner received a vote share that was substantially higher than the 50% threshold."

The first question was:

- "Consider the party incumbency effect in districts where the winner received between 50% and 60% of the vote.

  Do you expect the party incumbency effect in these districts to be smaller or larger than in districts right at the 50% threshold?"

- Answer options:
  - incumbency effect is smaller than at the 50% threshold
  - incumbency effect is about the same as at the 50% threshold
  - incumbency effect is smaller than at the 50% threshold

The next question was:

- "What magnitude do you expect for the party incumbency effect in districts where the winner received between 50% and 60% of the vote?

  Please move the slider to your expected effect size (e.g. 1 means you expect a 1 percentage points increase in incumbent party vote share). As a reminder: the effect at the 50% threshold is around 8-9 percentage points."

Data

Here we provide more information about the statewide elections dataset we employ in the paper. Table A.1 shows the number of data points used in the analysis with Control Set 3, the most parsimonious of the control sets. Specifically, each cell is the total number of data points entering the sample for a particular state and office, across the full range of values of the RD bandwidth or CIA window. The table does *not* count data points that have missing values for the outcome variable or for any of the control variables, so as to correspond precisely to the regression results reported. Note that some states have 0's in some columns reflecting the fact that those states do not hold elections for those offices (e.g., Alaska does not elect its attorney general and New Jersey does not elect any state executive office other than governor).

Table A.1 – Observations in Data Set, by State and Office. Each cell provides the total number of data points in the dataset used for analysis, subset to observations with no missing values for Control Set 3.

State	# Att Genl	# Auditor	# Gov	# LT Gov	# Senate	# Sec State	# Treasurer	Min Year	Max Year
AK	0	0	9	0	15	0	0	1960	2006
AL	12	9	14	11	16	9	10	1950	2006
AR	7	1	25	10	14	7	3	1948	2006
AZ	20	10	21	0	19	21	20	1948	2006
CA	15	0	14	15	19	15	14	1950	2006
CO	18	6	18	7	19	18	18	1948	2006
CT	15	0	15	8	19	16	16	1948	2006
DE	15	24	15	15	21	0	24	1948	2008
FL	9	0	15	0	20	12	11	1950	2006
GA	14	0	13	15	18	14	4	1950	2006
HI	0	0	10	0	15	0	0	1962	2006
IA	22	21	22	16	19	22	22	1948	2006
ID	15	10	15	15	18	14	12	1950	2006
IL	15	4	14	4	19	14	18	1948	2006
IN	15	21	15	5	19	21	21	1948	2008
KS	22	11	22	12	18	22	23	1948	2006
KY	15	15	15	10	19	15	15	1950	2007
LA	5	5	5	5	7	5	5	1950	1968
MA	20	18	20	10	19	19	19	1948	2006
MD	15	0	15	0	20	0	0	1950	2006
ME	0	0	16	0	20	0	0	1948	2006
MI	20	7	20	7	20	20	7	1948	2006
MN	19	15	18	9	20	18	16	1948	2006
MO	15	15	15	15	20	15	15	1950	2008
MS	15	14	13	13	18	14	15	1948	2007
MT	15	14	15	5	20	14	5	1948	2008
NC	15	15	15	15	19	15	15	1950	2008
ND	20	20	20	10	19	20	20	1948	2008
NE	19	20	20	10	20	20	18	1948	2006
NH	0	0	30	0	19	0	0	1948	2010
NJ	0	0	15	0	20	0	0	1948	2006
NM	21	20	21	7	20	21	20	1948	2006
NV	15	0	15	15	20	14	14	1950	2006
NY	15	0	15	0	19	0	0	1950	2006
OH	18	15	18	9	20	18	18	1948	2006
OK	11	13	15	15	19	6	12	1950	2006
OR	12	0	16	0	19	16	16	1948	2010
PA	7	15	15	4	20	4	15	1950	2008
RI	27	0	27	27	21	27	27	1948	2006
SC	8	0	15	9	18	8	6	1950	2006
SD	22	22	22	12	19	22	22	1948	2006
TN	0	0	13	0	20	0	0	1948	2006
TX	20	0	22	20	20	0	14	1948	2006
UT	16	15	14	0	20	6	14	1950	2008
VA	13	0	13	14	12	0	0	1948	2006
VT	24	26	31	32	19	28	21	1948	2010
WA	14	15	15	15	20	15	15	1950	2008
WI	20	0	21	10	20	21	21	1948	2006
WV	15	15	15	0	20	14	15	1948	2008
WY	0	15	15	0	21	13	15	1948	2008

Balance Checks for Table 4

Here we check the overlap in the covariate distributions. Tables A.2 and A.3 below summarize for the 5% and the 10% window the covariate balance in the raw and adjusted data for the conditioning set 1, the most extensive set of control variables. Overall we find that there is sufficient overlap in both of these windows for which we found the conditional independence assumption to be plausible. While there are significant imbalances in the raw data, these imbalances largely disappear in the unmatched or reweighed data; the means are close together, the p-values from the difference in means tests are all insignificant at conventional levels, and the variance ratios are close to one. Taken together these results suggests that there is enough covariate overlap in these windows to allow for a robust identification; a fact that is consistent with the finding that the incumbency effects estimates presented above do not vary much across the different adjustment methods.

Table A.2 – Balance Before and After Covariate Adjustment (Window 5%)

Adjustment	Covariate	Mean Tr	Mean Co	S.Diff	T.pval	Var.Ratio
Unmatched	Dem Share t-1	5.89	−0.84	0.45	0.00	1.19
Unmatched	Dem Share t-2	3.72	1.40	0.14	0.17	1.18
Unmatched	Normal Votet t-1	52.05	50.80	0.21	0.05	0.98
Unmatched	Normal Vote t-2	51.69	51.15	0.08	0.43	0.98
Unmatched	Midterm Slump t	−0.17	−0.17	−0.01	0.93	1.02
After Genetic Matching	Dem Share t-1	5.89	5.38	0.02	0.16	1.38
After Genetic Matching	Dem Share t-2	3.72	4.20	−0.02	0.16	1.11
After Genetic Matching	Normal Votet t-1	52.05	52.40	−0.04	0.17	1.01
After Genetic Matching	Normal Vote t-2	51.69	51.57	0.01	0.74	1.13
After Genetic Matching	Midterm Slump t	−0.17	−0.18	0.00	0.65	1.02
After Entropy Balancing	Dem Share t-1	5.89	5.89	0.00	1.00	1.16
After Entropy Balancing	Dem Share t-2	3.72	3.72	0.00	1.00	1.03
After Entropy Balancing	Normal Votet t-1	52.05	52.05	0.00	1.00	0.80
After Entropy Balancing	Normal Vote t-2	51.69	51.69	0.00	1.00	0.86
After Entropy Balancing	Midterm Slump t	−0.17	−0.17	0.00	1.00	0.99

S.Diff=Standardized difference in means; T-pval=p-value from difference in means test; Var.Ratio: Ratio of variances

Table A.3 – Balance Before and After Covariate Adjustment (Window 10%)

Adjustment	Covariate	Mean Tr	Mean Co	S.Diff	T.pval	Var.Ratio
Unmatched	Dem Share t-1	6.41	−1.06	0.50	0.00	1.02
Unmatched	Dem Share t-2	5.09	0.22	0.30	0.00	1.31
Unmatched	Normal Votet t-1	52.54	50.14	0.39	0.00	1.01
Unmatched	Normal Vote t-2	52.35	50.46	0.29	0.00	1.11
Unmatched	Midterm Slump t	−0.19	−0.10	−0.16	0.04	0.96
After Genetic Matching	Dem Share t-1	6.41	6.11	0.01	0.35	1.12
After Genetic Matching	Dem Share t-2	5.09	5.22	−0.01	0.36	1.03
After Genetic Matching	Normal Votet t-1	52.54	52.53	0.00	0.95	1.15
After Genetic Matching	Normal Vote t-2	52.35	52.22	0.01	0.35	1.14
After Genetic Matching	Midterm Slump t	−0.19	−0.18	−0.01	0.32	1.01
After Entropy Balancing	Dem Share t-1	6.41	6.41	0.00	1.00	0.69
After Entropy Balancing	Dem Share t-2	5.09	5.09	0.00	1.00	0.94
After Entropy Balancing	Normal Votet t-1	52.54	52.54	0.00	1.00	0.71
After Entropy Balancing	Normal Vote t-2	52.35	52.35	0.00	1.00	0.85
After Entropy Balancing	Midterm Slump t	−0.19	−0.19	0.00	1.00	0.99

S.Diff=Standardized difference in means; T-pval=p-value from difference in means test; Var.Ratio: Ratio of variances

Sensitivity Analysis for Table 4

Here we provide the results from Rosenbaum sensitivity analysis for the matching based incumbency effect estimates presented in Table 4 (Rosenbaum 2002). The goal of the Rosenbaum sensitivity tests is to examine the degree of hidden bias due to an unobserved confounder that would be needed to explain away the incumbency effect estimates. The degree of hidden bias is determined with the Rosenbaum Gamma parameter, Γ, which measures the departure from a study that is free of bias. More precisely, it is defined as the upper bound on the degree to which two matched units that are similar on the observed covariates may nonetheless differ in their a priori odds of receiving the treatment (i.e. incumbency) due to differences in an unobserved confounder. This unobserved confounder is assumed to be a near-perfect predictor of the outcome (i.e. the vote share in the next election at $t + 1$). For example, if $\Gamma = 1$ the study is free of hidden bias because the odds of treatment assignment is the same for both units (as in a randomized experiment). If $\Gamma = 2$, we allow for substantial hidden bias since one of the two units that are matched on the covariates might still be up to twice as likely to receive the treatment due to differences on the powerful omitted variable.

Table **??** reports the lowest Γ values at which the incumbency effect estimates turn insignificant. The results turn out to be highly robust to hidden bias with Γ values ranging between 4 and 6 across the windows and conditioning sets. This implies that only a very strong hidden bias could explain away the incumbency effects. Net of the observed covariates, an unobserved confounder would need to be a near-perfect predictor of vote shares and produce a 4- to 6-fold increase in the odds of treatment assignment. This level of insensitivity to hidden bias far exceeds those typically found for social science studies where Γ values are commonly in the range of 1-2 (Keele 2010; Rosenbaum 2002, 2005).

Table A.4 – Sensitivity Analysis for Incumbency Effects in Less Competitive Districts. Table presents Gamma values from Rosenbaum sensitivity tests for the matching based estimates of the incumbency effects presented in Table 4. The reported Gamma values measure the degree of hidden bias from an unobserved confounder at which the effect estimates would turn insignificant.

Sensitivity of Incumbency Effect Estimates in Less Competitive Districts

	Control Set 1: Dem Share$_{t-1}$ Dem Share$_{t-2}$ Normal Vote$_{t-1}$ Normal Vote$_{t-2}$ Midterm Slump$_t$		Control Set 2: Dem Share$_{t-1}$ Dem Share$_{t-2}$ Normal Vote$_{t-1}$ Midterm Slump$_t$		Control Set 3: Dem Share$_{t-1}$ Normal Vote$_{t-1}$	
Window	Gamma HL	Gamma p-val	Gamma HL	Gamma p-val	Gamma HL	Gamma p-val
5	1.00	1.60	1.00	1.30	1.00	1.30
10	1.00	1.60	1.00	1.60	1.00	1.50
15	1.00	1.60	1.00	1.50	1.00	1.50

Gamma HL: The lowest Rosenbaum Gamma at which the lower bound of the Hodges-Lehman point estimate of the incumbency effect remains above zero. Gamma p-val: The lowest Rosenbaum Gamma at which the upper bound of the p-value from a Wilcoxon Sign Rank Test turns insignificant. Window: Sample used to estimate the effect by comparing winners and losers.

CIA Tests without Covariates

In this section, we replicate the CIA tests without using any control variables. As we see, at some windows the tests continue to look solid. However, at the 10% window the estimate coefficients are roughly three times as large as with controls.

Table A.5 – Conditional Independence Tests. Presents CIA tests from equation 4 without any covariates to the left of the discontinuity (D=0) and to the right (D=1). The CIA appears to be questionable at windows as small as size 10, and fails at 15.

| | **No Controls:** | |
Window	D=0	D=1
5	0.09	0.25
	(0.30)	(0.28)
	N=480	N=487
10	0.17	0.14
	(0.11)	(0.12)
	N=923	N=904
15	0.31	0.13
	(0.07)	(0.07)
	N=1309	N=1263
20	0.32	0.18
	(0.05)	(0.05)
	N=1560	N=1550
25	0.31	0.28
	(0.04)	(0.04)
	N=1748	N=1793
30	0.29	0.30
	(0.03)	(0.03)
	N=1864	N=1966
35	0.31	0.32
	(0.03)	(0.03)
	N=1954	N=2095
40	0.30	0.29
	(0.03)	(0.03)
	N=2012	N=2186

Robust standard errors in parentheses. $V_{i,t}$ and $Y_{i,t+1}$ measured in percentage points.

Balance Checks for Table 5

The Tables A.6-A.8 below report the balance statistics for the samples used for the effect estimation in Table 5, where we exclude the observations that are right above the threshold. Again, we find that there is sufficient overlap in both of these windows for which we found the conditional independence assumption to be plausible.

Table A.6 – Balance Before and After Covariate Adjustment (Margins of 5-10%)

Adjustment	Covariate	Mean Tr	Mean Co	S.Diff	T.pval	Var.Ratio
Unmatched	Dem Share t-1	7.29	−1.57	0.60	0.00	1.06
Unmatched	Dem Share t-2	6.79	−0.42	0.46	0.00	1.44
Unmatched	Normal Votet t-1	53.30	49.74	0.58	0.00	1.11
Unmatched	Normal Vote t-2	52.98	50.01	0.45	0.00	1.21
Unmatched	Midterm Slump t	−0.19	−0.11	−0.16	0.07	0.93
After Genetic Matching	Dem Share t-1	7.29	7.19	0.01	0.70	1.01
After Genetic Matching	Dem Share t-2	6.79	6.76	0.00	0.94	1.08
After Genetic Matching	Normal Votet t-1	53.30	53.30	0.00	0.97	1.14
After Genetic Matching	Normal Vote t-2	52.98	52.90	0.01	0.71	1.18
After Genetic Matching	Midterm Slump t	−0.19	−0.19	0.00	1.00	0.99
After Entropy Balancing	Dem Share t-1	7.29	7.29	0.00	1.00	0.65
After Entropy Balancing	Dem Share t-2	6.79	6.79	0.00	1.00	0.87
After Entropy Balancing	Normal Votet t-1	53.30	53.30	0.00	1.00	0.66
After Entropy Balancing	Normal Vote t-2	52.98	52.98	0.00	1.00	0.79
After Entropy Balancing	Midterm Slump t	−0.19	−0.19	0.00	1.00	0.96

S.Diff=Standardized difference in means; T-pval=p-value from difference in means test; Var.Ratio:
Ratio of variances

Table A.7 – Balance Before and After Covariate Adjustment (Margins 5-15%)

Adjustment	Covariate	Mean Tr	Mean Co	S.Diff	T.pval	Var.Ratio
Unmatched	Dem Share t-1	7.96	−1.57	0.64	0.00	1.10
Unmatched	Dem Share t-2	8.57	−0.42	0.57	0.00	1.34
Unmatched	Normal Votet t-1	53.47	49.74	0.60	0.00	1.06
Unmatched	Normal Vote t-2	53.41	50.01	0.51	0.00	1.19
Unmatched	Midterm Slump t	−0.16	−0.11	−0.09	0.20	0.94
After Genetic Matching	Dem Share t-1	7.96	7.85	0.01	0.49	1.02
After Genetic Matching	Dem Share t-2	8.57	8.46	0.00	0.49	1.03
After Genetic Matching	Normal Votet t-1	53.47	53.39	0.01	0.50	1.11
After Genetic Matching	Normal Vote t-2	53.41	53.34	0.01	0.70	1.22
After Genetic Matching	Midterm Slump t	−0.16	−0.16	0.01	0.53	1.03
After Entropy Balancing	Dem Share t-1	7.96	7.96	0.00	1.00	0.65
After Entropy Balancing	Dem Share t-2	8.57	8.57	0.00	1.00	0.73
After Entropy Balancing	Normal Votet t-1	53.47	53.47	0.00	1.00	0.61
After Entropy Balancing	Normal Vote t-2	53.41	53.41	0.00	1.00	0.73
After Entropy Balancing	Midterm Slump t	−0.16	−0.16	0.00	1.00	0.96

S.Diff=Standardized difference in means; T-pval=p-value from difference in means test; Var.Ratio:
Ratio of variances

Table A.8 – Balance Before and After Covariate Adjustment (Margins 10-15%)

Adjustment	Covariate	Mean Tr	Mean Co	S.Diff	T.pval	Var.Ratio
Unmatched	Dem Share t-1	8.72	−1.57	0.69	0.00	1.15
Unmatched	Dem Share t-2	10.60	−0.42	0.72	0.00	1.23
Unmatched	Normal Votet t-1	53.66	49.74	0.65	0.00	1.00
Unmatched	Normal Vote t-2	53.90	50.01	0.59	0.00	1.16
Unmatched	Midterm Slump t	−0.12	−0.11	−0.02	0.84	0.95
After Genetic Matching	Dem Share t-1	8.72	8.62	0.00	0.59	1.02
After Genetic Matching	Dem Share t-2	10.60	10.53	0.00	0.80	1.02
After Genetic Matching	Normal Votet t-1	53.66	53.62	0.01	0.75	1.03
After Genetic Matching	Normal Vote t-2	53.90	53.86	0.00	0.86	1.28
After Genetic Matching	Midterm Slump t	−0.12	−0.12	0.01	0.68	1.05
After Entropy Balancing	Dem Share t-1	8.72	8.72	0.00	1.00	0.65
After Entropy Balancing	Dem Share t-2	10.60	10.60	0.00	1.00	0.59
After Entropy Balancing	Normal Votet t-1	53.66	53.66	0.00	1.00	0.56
After Entropy Balancing	Normal Vote t-2	53.90	53.90	0.00	1.00	0.67
After Entropy Balancing	Midterm Slump t	−0.12	−0.12	0.00	1.00	0.96

S.Diff=Standardized difference in means; T-pval=p-value from difference in means test; Var.Ratio: Ratio of variances

Estimates for Republicans

In this section, we replicate the analysis focusing instead on Republican rather than Democratic incumbency. Because the RD estimate is the difference in vote share across the party of interest in treated and control districts, redefining the treatment from Democratic to Republican incumbency, itself, has no effect on the estimated results at the discontinuity. However, away from the discontinuity, focusing on the Republicans is akin to calculating the the average treatment effect for the control units (ATC). As the plot shows, when we re-focus on this analysis we again find the same pattern of results.

Table A.9 – **Conditional Independence Tests for Republicans.** Presents CIA tests from equation 4 to the left of the discontinuity (D=0) and to the right (D=1). The CIA appears to be satisfied at windows as large as size 10, and partially satisfied at 15.

	Control Set 1: Rep $Share_{t-1}$ Rep $Share_{t-2}$ Normal $Vote_{t-1}$ Normal $Vote_{t-2}$ Midterm $Slump_t$		Control Set 2: Rep $Share_{t-1}$ Rep $Share_{t-2}$ Normal $Vote_{t-1}$ Midterm $Slump_t$		Control Set 3: Rep $Share_{t-1}$ Normal $Vote_{t-1}$	
Window	D=0	D=1	D=0	D=1	D=0	D=1
5	0.33	-0.05	0.27	-0.05	0.21	0.02
	(0.29)	(0.34)	(0.30)	(0.33)	(0.28)	(0.31)
	N=446	N=441	N=446	N=441	N=474	N=471
10	0.06	0.08	0.06	0.08	0.09	0.12
	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)	(0.11)
	N=810	N=837	N=810	N=837	N=866	N=899
15	0.05	0.29	0.05	0.29	0.11	0.31
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
	N=1131	N=1170	N=1131	N=1170	N=1201	N=1255
20	0.10	0.32	0.11	0.32	0.15	0.33
	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)
	N=1386	N=1389	N=1386	N=1389	N=1471	N=1485
25	0.18	0.32	0.18	0.32	0.23	0.32
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
	N=1614	N=1553	N=1614	N=1553	N=1709	N=1655
30	0.20	0.30	0.21	0.29	0.24	0.30
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
	N=1782	N=1655	N=1782	N=1655	N=1879	N=1761
35	0.22	0.30	0.22	0.30	0.26	0.30
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
	N=1897	N=1736	N=1897	N=1736	N=2003	N=1844
40	0.18	0.29	0.18	0.29	0.22	0.30
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
	N=1979	N=1783	N=1979	N=1783	N=2093	N=1894

Robust standard errors in parentheses. $V_{i,t}$ and $Y_{i,t+1}$ measured in percentage points.

Table A.10 – Incumbency Effects for Republicans in Less Competitive Districts and at the Threshold. The top panel presents incumbency effect estimates in less competitive districts based on the conditional independence assumption for different windows and covariate adjustment methods. The bottom panel presents for comparison the incumbency effect estimates at the threshold based on a regression discontinuity design for different bandwidths.

Incumbency Effect in Less Competitive Districts

	Control Set 1: $Rep\ Share_{t-1}$ $Rep\ Share_{t-2}$ $Normal\ Vote_{t-1}$ $Normal\ Vote_{t-2}$ $Midterm\ Slump_t$			Control Set 2: $Rep\ Share_{t-1}$ $Rep\ Share_{t-2}$ $Normal\ Vote_{t-1}$ $Midterm\ Slump_t$			Control Set 3: $Rep\ Share_{t-1}$ $Normal\ Vote_{t-1}$		
Window	OLS	Match	Weight	OLS	Match	Weight	OLS	Match	Weight
5	8.03	7.90	8.02	8.03	7.79	8.02	7.59	7.66	7.57
	(0.63)	(0.76)	(0.67)	(0.63)	(0.79)	(0.67)	(0.60)	(0.79)	(0.63)
	N=887	N=887	N=887	N=887	N=887	N=887	N=945	N=945	N=945
10	8.31	8.93	8.23	8.33	8.59	8.30	8.14	8.03	8.10
	(0.47)	(0.63)	(0.47)	(0.47)	(0.58)	(0.47)	(0.45)	(0.59)	(0.46)
	N=1647	N=1647	N=1647	N=1647	N=1647	N=1647	N=1765	N=1765	N=1765
15	9.33	9.48	9.52	9.35	9.59	9.61	9.31	9.32	9.50
	(0.43)	(0.51)	(0.51)	(0.43)	(0.53)	(0.51)	(0.41)	(0.55)	(0.49)
	N=2301	N=2301	N=2301	N=2301	N=2301	N=2301	N=2456	N=2456	N=2456

Incumbency Effect at the Threshold (RD estimates)

Bandwidth	Local Linear	Local Linear	Local Linear
1	9.99	9.99	9.36
	(3.44)	(3.44)	(3.22)
	N=178	N=178	N=191
2	8.73	8.73	8.41
	(2.23)	(2.23)	(2.08)
	N=361	N=361	N=384
5	7.52	7.52	7.22
	(1.30)	(1.30)	(1.22)
	N=887	N=887	N=945

Covariate adjustments are: OLS - Linear regression; Match: One-to-one nearest neighbor matching with replacement and bias adjustment; Weight: Entropy balancing; Local linear: Local linear RD regression. Robust standard errors in parentheses. Window: Sample used to estimate the effect by comparing winners and losers. Bandwidth: Sample used to estimate the RD effect at the threshold. $Y_{i,t+1}$ measured in percentage points, 0–100.

Figure A.1 – Incumbency Effects for Republicans in Less Competitive Districts and at the Threshold. Figure shows the incumbency effect estimates in less competitive districts based on the conditional independence assumption for windows between 1% and 20% (based on the regression adjustment with conditioning set 1). For comparison the Figure at the very left also shows the RD based estimate of the incumbency effect at the threshold (based on the local linear regression with a 5% bandwidth).
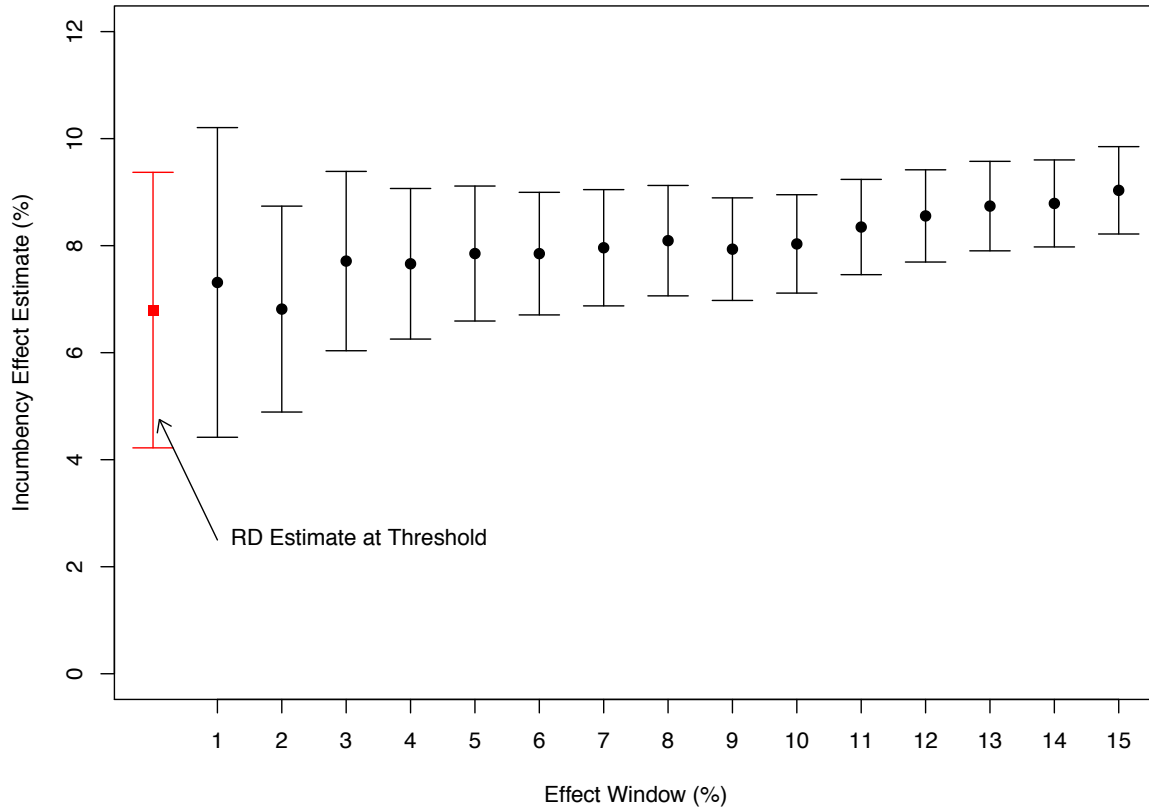
Results for Scare-off Effects

In this section, we apply the technique to statewide races not to estimate the overall incumbency advantage away from the 50-50 threshold, but instead to investigate the scare-off effect in these races. The analysis thus parallels the "mechanisms" analyses of Section 5 in the paper.

First, in Table A.11 we present the CIA tests for the scare-off outcome variable, the net quality differential between the Democratic and Republican candidates at $t + 1$. This quality differential variable naturally takes the values 1, when the Democratic at $t + 1$ is experienced and the Republican is not, 0 when neither candidate at $t + 1$ is experienced, and -1 when the Republican is and the Democrat is not. To make the coefficients more legible, in this table we estimate the effects after multiplying the variable by 100 so that it runs from -100 to 100.

As the table shows, we tend to find relatively small coefficients, especially at the 15% window, and we cannot reject the null of no slope. When reading the table, bear in mind that the outcome variable is scaled to be in some sense twice as large as in the analysis on vote share, since vote share runs 0-100 and this net quality differential variable runs -100 to 100.

Next, in Table A.12 and Figure A.2, we estimate the scare-off effects away from the threshold and compare them to the RD estimates at the threshold. Regardless of the estimation technique, control set, or window, we find a flat scare-off effect quite comparable to the RD estimates. We focus our comparison on the 5% bandwidth RD estimate (last row of table); estimates at smaller bandwidths are far less stable due to the smaller sample sizes and the coarseness of the outcome variable. As a result, we conclude that scare-off does not appear to vary much if at all in the same 15% window for which we found no change in the incumbency advantage away from 50-50. One explanation for the lack of change in the incumbency advantage away from the threshold may therefore be the fact that incumbents in less competitive districts or no more or less able to induce experienced candidates from challenging them.

Table A.11 – Conditional Independence Tests When Outcome Variable is Net Candidate Quality Differential. Presents CIA tests from equation 4 to the left of the discontinuity (D=0) and to the right (D=1).

Window	Control Set 1: Dem $Share_{t-1}$ Dem $Share_{t-2}$ Normal $Vote_{t-1}$ Normal $Vote_{t-2}$ Midterm $Slump_t$		Control Set 2: Dem $Share_{t-1}$ Dem $Share_{t-2}$ Normal $Vote_{t-1}$ Midterm $Slump_t$		Control Set 3: Dem $Share_{t-1}$ Normal $Vote_{t-1}$	
	D=0	D=1	D=0	D=1	D=0	D=1
5	-2.17	2.54	-2.02	2.49	-2.23	2.09
	(1.97)	(2.08)	(1.94)	(2.08)	(1.85)	(2.05)
	N=441	N=446	N=441	N=446	N=471	N=474
10	-0.89	0.00	-0.86	-0.04	-0.85	0.30
	(0.72)	(0.76)	(0.72)	(0.76)	(0.69)	(0.75)
	N=837	N=811	N=837	N=811	N=899	N=866
15	0.55	-0.33	0.56	-0.35	0.64	-0.29
	(0.41)	(0.45)	(0.41)	(0.45)	(0.39)	(0.45)
	N=1170	N=1132	N=1170	N=1132	N=1255	N=1201
20	0.17	-0.02	0.18	-0.01	0.28	0.07
	(0.30)	(0.31)	(0.30)	(0.31)	(0.29)	(0.30)
	N=1389	N=1387	N=1389	N=1387	N=1485	N=1471
25	0.34	0.23	0.35	0.24	0.42	0.29
	(0.23)	(0.23)	(0.23)	(0.23)	(0.22)	(0.22)
	N=1553	N=1615	N=1553	N=1615	N=1655	N=1709
30	0.31	0.34	0.33	0.36	0.39	0.40
	(0.19)	(0.18)	(0.19)	(0.19)	(0.18)	(0.18)
	N=1655	N=1783	N=1655	N=1783	N=1761	N=1879
35	0.20	0.30	0.20	0.32	0.28	0.40
	(0.17)	(0.16)	(0.17)	(0.16)	(0.16)	(0.16)
	N=1736	N=1898	N=1736	N=1898	N=1844	N=2003
40	0.25	0.24	0.26	0.26	0.32	0.35
	(0.16)	(0.14)	(0.16)	(0.14)	(0.15)	(0.14)
	N=1783	N=1980	N=1783	N=1980	N=1894	N=2093

Robust standard errors in parentheses. $V_{i,t}$ and $Y_{i,t+1}$ measured in percentage points.

Table A.12 – Scare-off Effects in Less Competitive Districts and at the Threshold. The top panel presents scare-off effect estimates in less competitive districts based on the conditional independence assumption for different windows and covariate adjustment methods. The bottom panel presents for comparison the scare-off effect estimates at the threshold based on a regression discontinuity design for different bandwidths.

Scare-off Effect in Less Competitive Districts

	Control Set 1: Dem Share$_{t-1}$ Dem Share$_{t-2}$ Normal Vote$_{t-1}$ Normal Vote$_{t-2}$ Midterm Slump$_t$			Control Set 2: Dem Share$_{t-1}$ Dem Share$_{t-2}$ Normal Vote$_{t-1}$ Midterm Slump$_t$			Control Set 3: Dem Share$_{t-1}$ Normal Vote$_{t-1}$		
Window	OLS	Match	Weight	OLS	Match	Weight	OLS	Match	Weight
5	11.32	13.94	10.93	11.29	11.83	10.91	11.55	16.84	11.09
	(3.19)	(4.24)	(3.32)	(3.19)	(3.99)	(3.32)	(3.10)	(4.15)	(3.28)
	N=739	N=739	N=739	N=739	N=739	N=739	N=784	N=784	N=784
10	11.51	12.48	11.26	11.59	12.33	11.24	11.76	11.46	11.64
	(2.44)	(3.36)	(2.56)	(2.45)	(3.16)	(2.56)	(2.36)	(3.21)	(2.46)
	N=1356	N=1356	N=1356	N=1356	N=1356	N=1356	N=1451	N=1451	N=1451
15	11.72	11.23	11.28	11.77	11.59	11.29	11.87	10.63	11.67
	(2.09)	(3.33)	(2.32)	(2.09)	(2.75)	(2.32)	(2.02)	(2.49)	(2.18)
	N=1870	N=1870	N=1870	N=1870	N=1870	N=1870	N=1989	N=1989	N=1989

Scare-off Effect at the Threshold (RD estimates)

Bandwidth	Local Linear	Local Linear	Local Linear
1	0.04	0.04	0.31
	(16.55)	(16.55)	(15.99)
	N=145	N=145	N=153
2	4.74	4.74	5.63
	(9.94)	(9.94)	(9.69)
	N=295	N=295	N=310
5	12.16	12.16	13.34
	(6.19)	(6.19)	(6.03)
	N=739	N=739	N=784

Covariate adjustments are: OLS - Linear regression; Match: One-to-one nearest neighbor matching with replacement and bias adjustment; Weight: Entropy balancing; Local linear: Local linear RD regression. Robust standard errors in parentheses. Window: Sample used to estimate the effect by comparing winners and losers. Bandwidth: Sample used to estimate the RD effect at the threshold. $Y_{i,t+1}$ measured as $(-100, 0, 100)$.

Figure A.2 – Scare-off Effects in Less Competitive Districts and at the Threshold. Figure shows the scare-off effect estimates in less competitive districts based on the conditional independence assumption for windows between 1% and 20% (based on the regression adjustment with conditioning set 1). For comparison the Figure at the very left also shows the RD based estimate of the scare-off effect at the threshold (based on the local linear regression with a 5% bandwidth).
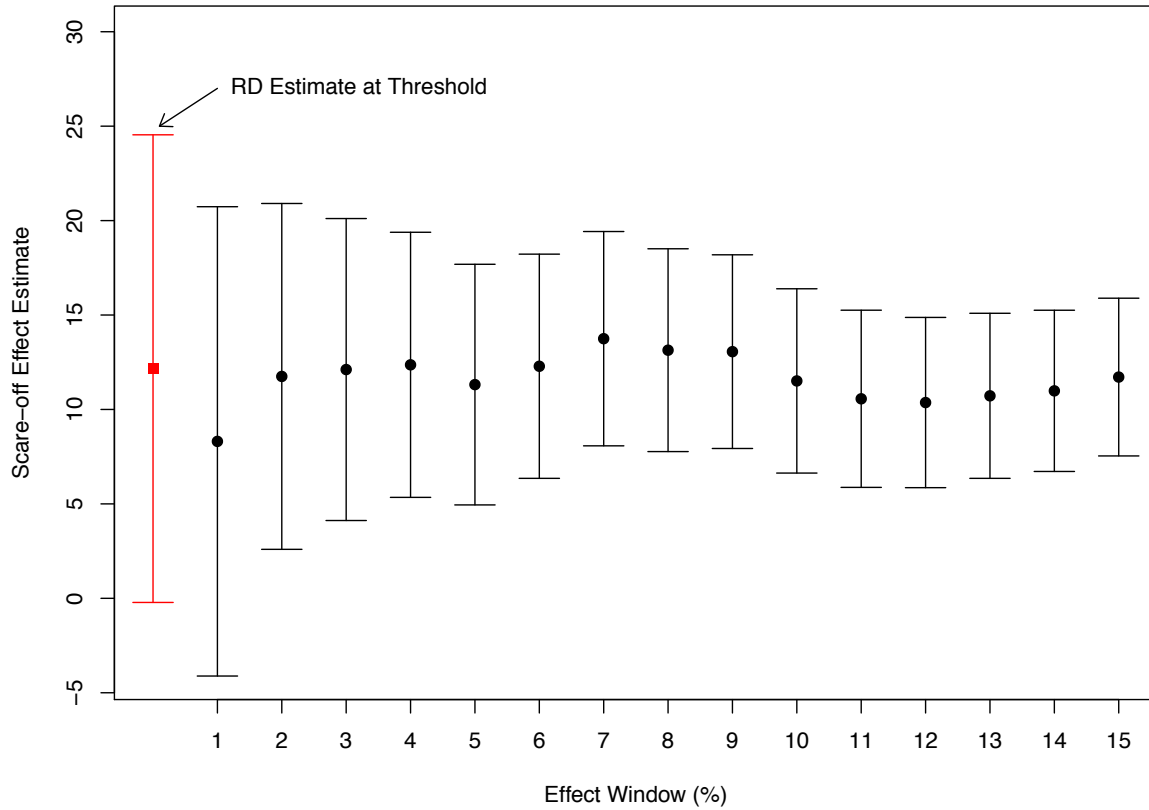
Applying the Technique to U.S. House Elections: An Example where the CIA Fails

Applying the technique employed in the paper to U.S. House elections is challenging due to redistricting, which is why we focused on statewide elections in the body of the paper. Because districts cannot be followed over long periods of time, it is more difficult to control for various lags of the normal vote like we do for statewide races. For example, if we use two periods of lags, we must throw out essentially all observations occurring in years ending with '2'—since they have no analog for the election occurring in the previous term—and all observations occurring in years ending with '4.'

As a result of this obstacle, it is more difficult to develop a set of proxy variables in the U.S. House. To illustrate this trouble, in Table A.13 we apply the technique to the U.S. House. Although we have tried many control sets, we focus on three illustrative ones in the table. In the first column, we control for two lags of the Democratic vote share, as well as for both the midterm slump (as defined in the paper) and presidential coattails, a variable defined to take on the value 1 for years in which the Democratic party won a presidential election, -1 for years in which the Democratic party lost a presidential election, and 0 otherwise. In the second column, we use the two lags and midterm slump, omitting the coattail variable. Finally, in the third column we simply use the two lags.

As the table shows, the CIA tests do not suggest the validity of the assumption regardless of window size or the control set used. Consider the 5% window (first row). While the standard errors do not always allow us to reject the null of no conditional relationship between the running variable and the outcome variable, the substantive size of the coefficients is large. This underscores the discussion in the paper: we must scrutinize not just the binary outcome of the statistical test (accept or reject), but the size of the coefficient. At larger windows, these coefficients remain relatively large, and we can reject the null of a zero slope.

We believe this provides a useful example of how a researcher might see a case in which he or she cannot generalize beyond the RD threshold. We should note for future work, however, that it may be possible to apply to the technique to U.S. House elections in other ways. For example, we have found that including district fixed effects appears to account for much of the remaining conditional relationship between the running variable and the outcome. While this result is encouraging, it should be clear that including such variables will make the estimation of effects much more difficult; units will have to be matched not just on the basis of the other control variables (like lagged vote shares and midterm slump) but also *within district*, i.e., matched only to other elections in the same district in a different time period. This might create problems with limited overlap and resulting model sensitivity.

Table A.13 – Conditional Independence Tests, U.S. House 1948–2012.
Presents CIA tests from equation 4 to the left of the discontinuity (D=0) and to the right (D=1). The CIA appears to fail in the U.S. House.

	Control Set 1: $Dem\ Share_{t-1}$ $Dem\ Share_{t-2}$ $Midterm\ Slump_t$ $Coattail_t$		Control Set 2: $Dem\ Share_{t-1}$ $Dem\ Share_{t-2}$ $Midterm\ Slump_t$		Control Set 3: $Dem\ Share_{t-1}$ $Dem\ Share_{t-2}$	
Window	D=0	D=1	D=0	D=1	D=0	D=1
5	0.65 (0.35) N=192	0.47 (0.41) N=226	0.67 (0.36) N=192	0.52 (0.41) N=226	0.72 (0.38) N=192	0.52 (0.42) N=227
10	0.42 (0.13) N=424	0.68 (0.19) N=415	0.46 (0.13) N=424	0.67 (0.18) N=415	0.40 (0.14) N=426	0.65 (0.18) N=419
15	0.39 (0.07) N= 665	0.32 (0.10) N= 576	0.41 (0.07) N= 665	0.32 (0.10) N= 576	0.37 (0.08) N= 669	0.30 (0.10) N= 581
20	0.40 (0.05) N= 903	0.26 (0.07) N= 733	0.42 (0.05) N= 903	0.25 (0.07) N= 733	0.41 (0.05) N= 912	0.24 (0.07) N= 738
25	0.34 (0.04) N=1153	0.25 (0.06) N= 898	0.35 (0.05) N=1153	0.25 (0.06) N= 898	0.34 (0.05) N=1165	0.24 (0.06) N= 905
30	0.34 (0.04) N=1381	0.29 (0.05) N=1080	0.36 (0.04) N=1381	0.29 (0.05) N=1080	0.35 (0.04) N=1394	0.29 (0.05) N=1089
35	0.31 (0.03) N=1581	0.30 (0.04) N=1254	0.33 (0.03) N=1581	0.30 (0.04) N=1254	0.32 (0.04) N=1595	0.29 (0.04) N=1266
40	0.34 (0.03) N=1733	0.32 (0.03) N=1431	0.36 (0.03) N=1733	0.33 (0.03) N=1431	0.35 (0.03) N=1747	0.31 (0.03) N=1443

Robust standard errors in parentheses. $V_{i,t}$ and $Y_{i,t+1}$ measured in percentage points.