PA

Much Ado About Nothing: RDD and the Incumbency Advantage

Robert S. Erikson[1] and Kelly Rader[2]

[1] Political Science, Columbia University, New York, NY 10027, USA
[2] Political Science, Yale University, New Haven, CT 06511, USA. Email: kelly.rader@yale.edu

Abstract

An influential paper by Caughey and Sekhon (2011a) suggests that the outcomes of very close US House elections in the postwar era may not be as-if random, thus calling into question this application of regression discontinuity for causal inference. We show that while incumbent party candidates are more likely to win close House elections, those who win are no different on observable characteristics from those who lose. Further, all differences in observable characteristics between barely winning Democrats and barely winning Republicans vanish conditional on which party is the incumbent. Any source of a special incumbent party advantage in close elections must be due to variables that cannot be observed. This finding supports the conclusion of Eggers et al. (2015) that Caughey and Sekhon's discovery of lopsided wins by incumbents in close races is a mere statistical fluke.

Regression discontinuity (RD) is a useful tool for causal inference in natural settings. The idea is simple. When a treatment is assigned based on a threshold, cases barely below the threshold are approximately identical in all respects to those barely on the other side of the threshold. Thus, untreated cases near the threshold can serve as counterfactuals for treated cases near the threshold, and researchers can estimate a local average treatment effect with minimal assumptions. RD has an important place in the analysis of two-candidate elections since very close elections on each side of the 50 percent vote threshold are similar in all aspects except for a virtual coin flip for who wins. If its assumptions are met, RD allows causal inference regarding the consequences of a party winning versus losing on its possible incumbency advantage in later elections (e.g., Lee 2008; Erikson and Titiunik 2015) or outcomes of elections for other offices (Folke and Snyder 2012; Erikson, Folke, and Snyder 2015).

Yet the validity of RD for causal inference has come under challenge. An influential and award-winning article,[1] Caughey and Sekhon (2011a), sounds a warning about the validity of the RD design when applied to US House elections in the postwar era. This article (hereafter CS) notes an interesting oddity in the distribution of outcomes for very close elections—elections within a margin of 0.5 percentage points of vote for the incumbent party.

The purpose of this letter is to show that, contrary to the impression one might get from a first reading of CS, there are *no* observable differences in covariates between barely winning and barely losing incumbent party candidates that could account for the lopsided incumbent party win rate. Additionally, any differences in observable variables between elections barely won by Democrats and those barely lost by Democrats vanish conditional on which party is the incumbent party.[2] This fact rules out incumbent party prowess at exploiting variables that are measured here as the source of the CS incumbent party win rate oddity.

Authors' note: The authors would like to thank three anonymous reviewers for their helpful comments. All replication materials are available at the *Political Analysis* Dataverse (Erikson and Rader 2016, doi:10.7910/DVN/567RS6).

1 Caughey and Sekhon (2011a) was the recipient of the 2012 Warren Miller prize for the best paper in *Political Analysis* published in the previous year.

2 Elections barely won by Democrats by a two-party vote are, by definition, elections that are barely lost by Republicans, and vice versa.

1 Lopsided Incumbent Party Win Rate

CS shows that in very close House elections from the postwar era, those in which the incumbent party received a two-party vote share between 49.25 and 50.25 percent, the incumbent party candidate won more often than would be expected by chance.[3] In the 85 elections in this narrow range, the incumbent party won 59 times and the challenger party only 26.[4] CS speculates that incumbent parties might have some unidentified source of leverage over the opposition that allows them to boost their candidate over the finish line when (and only when) the contest is known to be very close.

The CS result has been unsettling, in part because of its seeming implausibility,[5] and in part because of its potential impact on RD analysis of elections. Eggers *et al.* (2015) shows that the same anomaly is not found elsewhere in legislative elections around the world, nor in other US elections, including US House elections before 1945. Eggers *et al.* (2015) suggests that the CS finding is a "fluke." But if a fluke, what kind of fluke is it? Did CS identify an unexpected causal dynamic that is in operation only for modern US House elections? Or is it just a statistical anomaly, a "statistically significant" pattern that could readily be found by chance when searching over many legislatures?

CS presents evidence that appears to tilt the verdict strongly toward a newly discovered but unexpected electoral dynamic rather than a statistical fluke. Their Figure 2 (p. 394) shows that when the vote is measured *as percent Democratic*, instead of percent for the incumbent party, there is a clear imbalance in familiar covariates, such as previous experience in office and proportion of campaign spending by the Democrat. That is, variables that are associated with Democratic wins are more associated with Democratic wins than losses at the threshold. Since the incumbency advantage for US House elections is usually estimated with Democratic voting (not incumbent party voting) as the running variable, it would seem that the key assumptions of RD are seriously violated for estimating the incumbency advantage with RD. The control and treatment groups are not equivalent.

2 A Covariate Imbalance?

We show that what appears to be imbalance across many different things is actually only imbalance across one thing—incumbent party status. Because of the lopsided incumbent party win rate (the oddity to be explained), bare Democratic winners are more likely to be from districts previously held by Democrats (and so they are generally incumbents themselves), and bare Democratic losers are more likely to be from districts previously held by Republicans (and so they are generally challengers). Therefore, to compare bare Democratic winners to bare Democratic losers is to compare bare incumbent party winners to bare challenger losers. Because incumbents are different from challengers in many ways related to strength in the previous and present elections, this one imbalance in incumbent party status causes all of the others to appear. That is, the distributional imbalance in incumbent party vote share accounts for all of the covariate imbalance across bare Democratic winners and losers in Caughey and Sekhon's Figure 2!

Indeed, conditioning on which party is the incumbent party makes all of the observed imbalances between bare Democratic winners and bare Democratic losers disappear. Using the CS data (Caughey and Sekhon 2011b), our Figures 1 and 2 replicate the tests in CS Figure 2 separately for elections in which the Democrats are the incumbent party and Republicans are

---

3 The imbalance in incumbent win rate was first noted in Snyder (2005).

4 The gap is sufficient to reject the null of no discontinuity at the winning threshold in the distribution of incumbent party vote margin, with $p = 0.004$, using the test for sorting proposed by McCrary (2008), with bin size $= 0.5$ and bandwidth $= 5$.

5 If incumbent parties can manipulate the result of close elections, this process must overcome the following hurdles. Many incumbents (or their campaign staff) must have the power to observe or intuit tiny differentials in expected election outcomes, far smaller than can be detected with confidence from any poll. Second, when confronted by the likelihood of a very close loss (and only then), the specially endowed incumbents must have the means to overcome the slim electoral hurdle. Opposition parties would have to lack these powers, at least to the extent that incumbent parties do, even though by definition the two parties are at virtually equal political strength, at least for the present election.

Variable Name	Valid Cases	Treated Mean	Control Mean
Dem % t − 1	33	61	61
Dem % Margin t − 1	33	24	24
Inc's D1 NOMINATE	33	−0.29	−0.27
Dem Inc in Race	33	0.84	0.75
Dem's # Prev Terms	33	2.9	3.1
Rep's # Prev Terms	33	0.20	0.62
Rep Experience Adv	33	0.04	0.00
Dem Experience Adv	33	0.72	0.75
Partisan Swing	33	−6.1	−0.001
CQ Rating {−1, 0, 1}	29	0.50	0.57
Dem Spending %	22	57	55
Dem Donation %	14	60	63
Dem Sec of State	33	0.56	0.25
Dem Governor	33	0.48	0.50
Dem Pres % Margin	32	−0.06	−0.09
Dem−held Open Seat	33	0.16	0.25
Voter Turnout %	33	33	33
Pct Gov't Worker	29	5.2	4.7
Pct Urban	29	73	70
Pct Black	29	5.0	6.5
Pct Foreign Born	29	4.8	6.3

Figure 1. Balance plot for Democratic winners and losers in elections in which the Democratic candidate's relative margin was within 0.50 percentage points (led or lost by less than 0.5 percent) and the Democrats were the incumbent party.

the incumbent party, respectively.[6,7] With the possible exception of partisan swing in the case of elections with Democratic party incumbents, there are no statistically distinguishable differences between bare Democratic winners and losers when the Democrats are the incumbent party nor between bare Democratic winners and losers when the Republicans are the incumbent party. Of course, conditioning on which party is the incumbent does not eliminate the possibility that some unobserved confounder is still imbalanced within these groups. But any such confounder is not only unmeasured but also does not result in imbalance in the variables we have measured.

We can also rearrange the data to estimate the imbalance at the threshold for the incumbent party vote for all cases, combining Republican- and Democrat-held seats. Figure 3 shows that

6 These plots exclude covariates shown in CS Figure 2 that lack variation in measurement when conditioning on which party is the incumbent. Following CS, we test for differences using Fisher's exact tests for dichotomous variables (denoted by circles) and exact Wilcoxon rank sum tests for continuous and ordinal variables (denoted by diamonds) in Figures 1–3.

7 Eggers *et al.* (2015, p. 263) do a similar analysis using ordinary least squares and controlling for the party of the incumbent.

Variable Name	Valid Cases	Treated Mean	Control Mean
Dem % t − 1	52	37	42
Dem % Margin t − 1	51	−22	−16
Inc's D1 NOMINATE	52	0.30	0.33
Rep Inc in Race	52	0.67	0.76
Dem's # Prev Terms	52	0.28	0.47
Rep's # Prev Terms	52	1.8	3.2
Rep Experience Adv	48	0.62	0.78
Dem Experience Adv	48	0.12	0.06
Partisan Swing	52	4.5	4.9
CQ Rating {−1, 0, 1}	40	−0.23	−0.52
Dem Spending %	25	47	41
Dem Donation %	20	49	40
Dem Sec of State	52	0.33	0.32
Dem Governor	52	0.28	0.47
Dem Pres % Margin	47	−0.13	−0.11
Rep−held Open Seat	52	0.33	0.24
Voter Turnout %	52	42	35
Pct Gov't Worker	44	5.0	4.3
Pct Urban	44	65	64
Pct Black	44	4.8	4.7
Pct Foreign Born	44	2.8	3.7

Figure 2. Balance plot for Democratic winners and losers in elections in which the Democratic candidate's relative margin was within 0.50 percentage points (led or lost by less than 0.5 percent) and the Republicans were the incumbent party.

they are not statistically distinguishable on any observables in the CS dataset. Incumbent party candidates who barely win did not have substantially larger wins in the previous election than did incumbent party losers. Bare incumbent party winners do not differ ideologically from their losing counterparts. They are not more or less likely to be in a race with an actual incumbent from either party (as opposed to an open seat). They have not served more terms, nor do they have a substantial experience advantage. They are not more or less affected by national partisan swings. Their races are not more or less predictable by *CQ*. They do not spend more money or raise more money relative to their opponents than do their barely losing counterparts relative to their opponents. They are not more likely to share a party with their governor or secretary of state. They are not more likely to run in open seats previously held by either party. Their districts' presidential voting is no different. They do not run in higher or lower turnout elections, and their constituents are not demographically different. In short, while incumbent party candidates
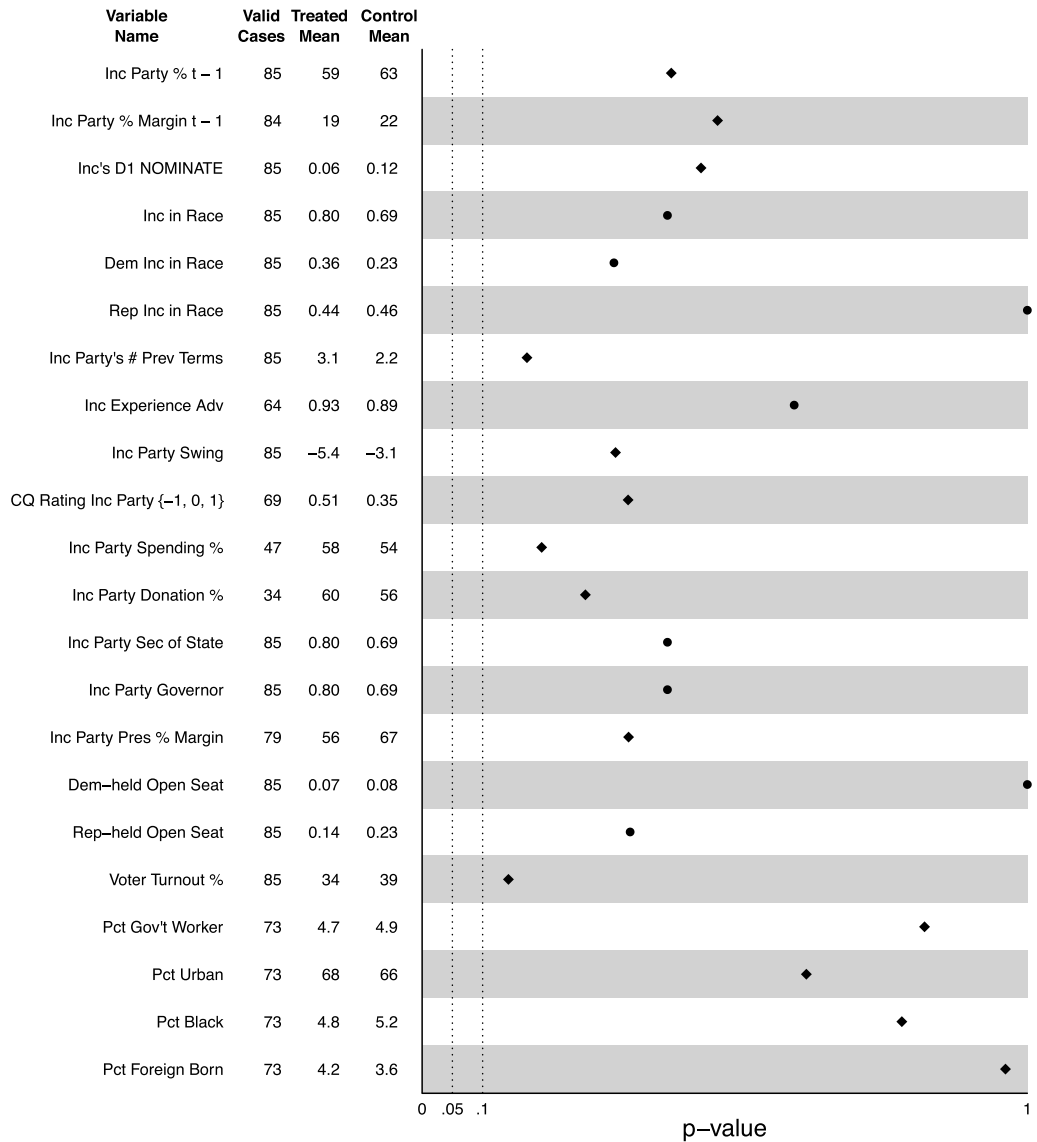
Variable Name	Valid Cases	Treated Mean	Control Mean
Inc Party % t − 1	85	59	63
Inc Party % Margin t − 1	84	19	22
Inc's D1 NOMINATE	85	0.06	0.12
Inc in Race	85	0.80	0.69
Dem Inc in Race	85	0.36	0.23
Rep Inc in Race	85	0.44	0.46
Inc Party's # Prev Terms	85	3.1	2.2
Inc Experience Adv	64	0.93	0.89
Inc Party Swing	85	−5.4	−3.1
CQ Rating Inc Party {−1, 0, 1}	69	0.51	0.35
Inc Party Spending %	47	58	54
Inc Party Donation %	34	60	56
Inc Party Sec of State	85	0.80	0.69
Inc Party Governor	85	0.80	0.69
Inc Party Pres % Margin	79	56	67
Dem−held Open Seat	85	0.07	0.08
Rep−held Open Seat	85	0.14	0.23
Voter Turnout %	85	34	39
Pct Gov't Worker	73	4.7	4.9
Pct Urban	73	68	66
Pct Black	73	4.8	5.2
Pct Foreign Born	73	4.2	3.6

p−value

Figure 3. Balance plot for incumbent party winners and losers in elections in which the incumbent party candidate's relative margin was within 0.50 percentage points (led or lost by less than 0.5 percent).

have won more close elections in the postwar era than they have lost, those who win are not statistically distinguishable from those who lose. If incumbent party candidates do indeed have special leverage to push themselves over the edge in very close elections, that leverage does not manifest itself in terms of any observed imbalance herein across those who use it successfully and those who fail. To be clear, we are not arguing that we have shown that such special leverage does not exist. But our findings do suggest that the source of such leverage, if it exists, is probably outside the realm of measurable variables.

3 If No Fluke

We have argued that the disproportionate winning record of incumbent candidates when the vote is very close to the 50–50 threshold is implausible as more than just a statistical "fluke." To call it a fluke is to assert the absence of sorting at the threshold—to deny that when close to the threshold, some candidates have a special ability to manipulate themselves onto the winning side. But what if our assertion is incorrect? In particular, what exactly would be the consequences for using RD to estimate the incumbency advantage?
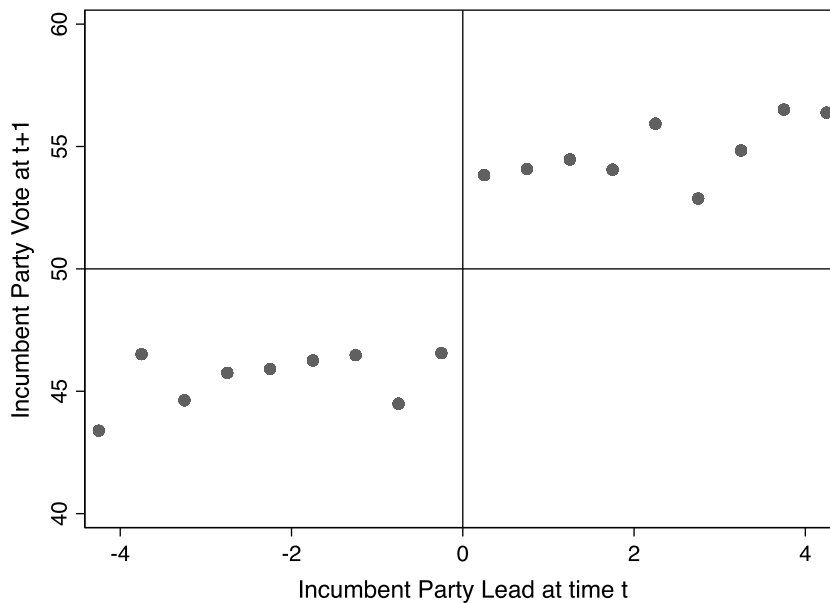
Figure 4. Mean incumbent party vote margin within 0.50 bins.

The confounding condition would be that some candidates, mainly incumbents, have a unique ability to exert the necessary effort for a final spurt to win very close elections. This same unmeasured endowment could also operate to distort the vote in the next election, at $t + 1$. Consider then an RD analysis to estimate the $t + 1$ incumbency advantage by comparing districts where the incumbent barely wins and where the incumbent barely loses in election t in terms of their party's performance in election $t + 1$. It would be easy to imagine that if barely winning incumbents at time t are exceptionally strong candidates, the estimate of the $t + 1$ incumbency advantage would be biased upwards.

What would be the remedy? A simple solution known as "donut RD" would be to exclude the potentially offending cases closest to the threshold and conduct an RD predicting the $t + 1$ vote from the incumbent vote running variable plus passing the threshold (Almond and Doyle 2011; Barreca, Lindo, and Waddell 2016). As Barreca, Lindo, and Waddell (2016) states, "dropping observations in the immediate vicinity of the treatment threshold should be thought of as a useful robustness check that has the potential to highlight misspecification in any RD design." The statistical justification would be the assumption of the continuation of the functional form in the vicinity of the threshold.

To visualize this setup, Figure 4 uses the CS data to predict the $t + 1$ vote from the time t vote differential (incumbent party minus the opposition), averaged within bins with the narrow width of half a percentage point. The key observations are the two bins nearest to the threshold. We could estimate the incumbency advantage using only the difference in the $t + 1$ vote for these two bins. Or we could summarize the data with a regression equation predicting the $t + 1$ vote from the time t vote plus an indicator variable for the incumbent party winning at time t. Or, as suggested above, we could perform the regression while omitting the two bins closest to the threshold.[8]

It is evident that when applied to the Caughey–Sekhon data, any of these three methods would yield similar estimates of the incumbency advantage.[9] This similarity should help to assuage

---

8 Figure 4 includes incumbent-contested seats and open seats in the same RD model. Unopposed $t + 1$ outcomes are excluded.

9 Using a linear RD regression with incumbent party vote differential as the running variable and cases within the 7-point bandwidth yields a 7.59 estimate of the incumbency effect (1.01 SE). Deleting the barely winning and barely losing cases (donut hole RD) adjusts the coefficient upward slightly to 7.86 (1.18 SE). Comparing incumbent-winning and incumbent-losing cases for the cases within 0.5 of the threshold only yields a 7.27-point differential (SE = 1.92) in their $t + 1$ vote means.

---

concern that the observational imbalance at the threshold might present a problem for estimating the incumbency advantage. If we were still worried, we could use the estimate that omits the questionable observations near the threshold.

4 The Bottom Line

Let us return to the central question about the lopsided incumbent win rate. Is it evidence of a real causal process governing postwar US congressional elections or is it something to be ignored as a statistical fluke? Our reanalysis shows that one cannot separate incumbent winners and losers based on district or candidate attributes. Thus it is highly unlikely that incumbent parties, disproportionately successful in close elections, use measurable covariates to push themselves over the finish line. By elimination, any special resources that give a true advantage for the incumbent party in close elections must be intangible incumbent party assets that are not measured here. Although we cannot be certain that what cannot be observed does not exist, the absence of a covariate imbalance in terms of observable variables pushes the verdict in the direction of fluke rather than causal process.

If it is a fluke, the tendency for the incumbent party to win very tight elections would not matter for using RD to estimate the incumbency advantage. And if it is no fluke, and Caughey and Sekhon identify a real phenomenon with close US House elections, then what of using RD to estimate the incumbency advantage? Although it is not ideal to discard cases close to the threshold when performing RD, it is permissible to do so if necessary to eliminate suspected sorting around the threshold. Our analysis suggests that for modern US House elections, subtracting observations near the threshold makes little difference in the estimation of the incumbency advantage.

References

Almond, Douglas, and Joseph J. Doyle. 2011. After midnight: A regression discontinuity design in length of postpartum hospital stays. *American Economic Journal: Economic Policy* 3:1–34.

Barreca, Alan I., Jason M. Lindo, and Glen R. Waddell. 2016. Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry* 54:268–293.

Caughey, Devin M., and Jasjeet S. Sekhon. 2011a. Elections and the regression discontinuity design: Lessons from close U.S. house races, 1942–2008. *Political Analysis* 19:385–408.

Caughey, Devin M., and Jasjeet S. Sekhon. 2011b. Replication data for: Elections and the regression-discontinuity design: lessons from close U.S. house Races, 1942–2008, hdl:1902.1/16357, Harvard Dataverse, V2.

Eggers, Andrew, Olle Folke, Anthony Fowler, Jens Hainmueller, Andrew B. Hall, and James M. Snyder. 2015. On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races. *American Journal of Political Science* 59(1):259–274.

Erikson, Robert S., Olle Folke, and James M. Snyder. 2015. A gubernatorial helping hand? How governors affect presidential elections. *Journal of Politics* 77(2):491–504.

Erikson, Robert S., and Kelly Rader. 2016. Replication data for: Much Ado about nothing: RDD and the incumbency advantage, doi:10.7910/DVN/567RS6, Harvard Dataverse, V1, [UNF:6:SGAYMsO3MwwA9UFLrvFscQ==].

Erikson, Robert S., and Rocío Titiunik. 2015. Using regression discontinuity to uncover the personal incumbency advantage. *Quarterly Journal of Political Science* 10(1):101–119.

Folke, Olle, and James M. Snyder. 2012. Gubernatorial midterm slumps. *American Journal of Political Science* 56(4):931–948.

Lee, David S. 2008. Randomized experiments from non-random selection in U.S. house elections. *Journal of Econometrics* 142(2):675–697.

McCrary, Justin. 2008. Manipulation of the running variable in regression discontinuity design: A density test. *Journal of Econometrics* 142(2):698–714.

Snyder, Jason. 2005. Detecting manipulation in U.S. house elections. Working paper, Haas School of Business, UC Berkeley.