Using Regression Discontinuity to investigate the Incumbency Advantage in US House of Representatives elections

Slootmaekers Bram

December 2025

Abstract

This paper estimates the causal effect of incumbency on electoral success in U.S. House of Representatives elections using a regression discontinuity design. Incumbent representatives enjoy remarkably high reelection rates, but whether this reflects genuine advantages obtained by holding office or whether this reflects the selection of higher quality candidates remains an open question with important implications for democratic accountability. Following Lee (2008), we exploit the discontinuity in the assignment of incumbency status in close elections to identify the causal incumbency advantage, comparing candidates who barely won to those who barely lost. Using data on U.S. House elections, we find that barely winning an election increases the probability of winning the subsequent election by approximately 35 percentage points. This is a substantial effect but notably smaller than the 60 percentage point coefficient observed in naive OLS comparisons. This difference is caused by omitted variable bias: Stronger or more popular candidates win and continue to win, with or without an incumbency advantage. The strength of a candidate is hard to fully capture with observed variables and is correlated with incumbency status, causing endogeneity of this covariate. We perform sensitivity analysis across multiple bandwidths and kernel specifications. Furthermore, we perform validity tests to assess the critiques around the local randomization assumption raised by Caughey and Sekhon (2011) through density tests and covariate balance checks. We found no evidence of manipulation or systematic sorting at the electoral threshold.

1 Introduction

The reelection rates of members of the incumbent party in the U.S. House of Representatives have consistently been reaching levels around 90% over the last decades, Lee (2008). Representatives of the incumbent party are extremely likely to win, the same goes for individual candidates that hold a seat as a representative. This pattern raises questions about the health and fairness of American democracy: Is this overwhelming success of elected representatives reflecting voter satisfaction or is it caused by structural advantages that help incumbents win elections regardless of their performance in their elected time periods? The latter case would be a sign that intended accountability mechanisms that protect democracy are not working as intended. In a healthy democracy, elections should allow voters to hold their representatives accountable for their performance. If incumbents win elections primarily because of an advantage inherent to holding office, rather than because of their merits and policies, the mechanism of accountability is not working properly. In this way, a large incumbency advantage could reduce competition, decrease turnover and impede election of more competent leaders, potentially protecting incompetent officials from accountability. Therefore, understanding the magnitude and establishing whether there is a causal relationship between incumbency and the probability of winning an election is of significant importance.

The estimation of this effect, however, is non-trivial. Simple regression models struggle distinguishing the incumbency advantage from differences in candidate strength and popularity. Winners may continue to win not because holding office offers significant advantages, but because they possess superior political skills that made them likely to win initially and also in following elections. The quality of a candidate influences their probability of winning, yet this factor is hard to completely capture using observed control variables. This results in biased estimates that overestimate the importance of the incumbency status. Consistent estimates require slightly more sophisticated approaches. In this paper we tackle this identification with regression discontinuity (RD) similar to Lee (2008). The intuition is as follows: candidates that win an election by the slightest of margins are supposed to be similar to the candidates that lost said election by just a small difference. By comparing the electoral performance of bare winners and bare losers, candidates with similar strength, in the next election, we can isolate the causal effect of incumbency from the influence of candidate quality.

In Lee (2008) RD is used to estimate the incumbency advantage of U.S. House of Representatives candidates. His paper indicates that barely winning an election significantly increases the party's vote share in the next contest and his paper opened the door to many papers using RD to estimate causal relationships in similar contexts. However, Lee (2008)'s approach was not free of scrutiny. Subsequent work questioned whether close elections truly approximate random assignment. Caughey and Sekhon (2011) provides a critique, showing that winners and losers of close U.S. House elections may not be as similar as initially assumed. They argue that covariates such as campaign spending, experience, and pre-election expectations differ among winners and losers of close elections. This critique suggests that candidates may exert control over close elections through superior resources or organizational advantages, implying that close winners are not valid counterfactuals for close losers. They argue that if true, RD estimates in this setting may be biased. In response, Eggers et al. (2015) reassess the validity of electoral RD designs using data from over 40,000 close races across multiple countries and electoral contexts. They argue that even when covariate imbalances are present, RD designs are still preferable to global regression approaches, given their transparent assumptions and local nature. Further clarification is provided in de la Cuesta and Imai (2016). They clarify a common misunderstanding in the literature. The original conditions for valid RD designs as described in Hahn et al. (2001) entail a continuity assumption, not strict local randomization. Under continuity, covariates may differ on either side of the cutoff, as long as potential outcomes evolve smoothly.

This paper aims to contribute to existing literature by revisiting the original research question as proposed in Lee (2008), but doing this while adhering to modern best practices and using new recent data. We employ bias corrected RD estimation methods as developed in Calonico et al. (2014), which can be found in the "rdrobust" package in R. These modern methods provide more reliable inference than the procedures available when Lee's original paper was published. Additionally, we also perform sensitivity analysis, investigating the effect of using different kernels and bandwidths for our RD design and we contrast naive regression estimates with our RD estimates. We also address the methodological critiques raised by Caughey and Sekhon (2011) through multiple validity tests. We conduct McCrary (2008) density tests to check for manipulation of the running variable and examine whether pre-treatment covariates are balanced at the threshold.

2 Primary Hypothesis

The primary hypothesis of this paper is that winning a U.S. House election causally increases a candidate's probability of winning the subsequent election, independent of underlying candidate quality or popularity. This hypothesis is motivated by the persistent trend of high reelection rates observed among U.S. House incumbents. While incumbents win reelection at high rates, it is unclear whether this reflects advantages conferred by holding office, or by selection: stronger candidates are more likely to win, both initially and in subsequent elections.

There are several possible mechanisms through which incumbency may increase future electoral success. First, incumbents gain enhanced name recognition and visibility through their official duties and media coverage when they are holding office. Incumbents may also benefit from informational advantages that they obtained through resources that are only available to elected candidates. Another possibility is that voters might interpret previous electoral success as a precedent for competency. Finally, holding office may deter strong challengers who prefer to wait for open seat races.

At the same time, these mechanisms are difficult to separate from candidate quality. The hypothesis is therefore evaluated using a regression discontinuity design that exploits the discontinuity that stems from the binary allocation of incumbency status.

3 Empirical Methodology

3.1 Data

The dataset we use is MIT Election Data and Science Lab (2024). Originally, it contains the amount of votes each candidate received in elections for a certain seat. A certain seat is allocated to a combination of a state and district. It holds this information ranging from 1976 to 2024. Before we estimated our RD model, we first had to process the data. First we restricted our analysis to "regular" elections. This means that we delete observations belonging to write-in candidates, run-off elections and other special elections. Afterwards, for each election, we focus on the top two candidates, as these represent the meaningful electoral competition and allow us to define the vote margin (our running variable) in an unambiguous way. Once we have only these relevant observations we define the running variable for our RD design as follows:

$$\text{margin}_i = \begin{cases} \frac{V_{\max}-V_{\min}}{V_{\text{total}}} & \text{if candidate } i \text{ is the winner} \\ \frac{V_{\min}-V_{\max}}{V_{\text{total}}} & \text{if candidate } i \text{ is the loser} \end{cases}$$

Where $V_{\max}, V_{\min}, V_{\text{total}}$ are the vote count of the winner, runner up, and sum of those two respectively. Next, we create a seat ID, which tracks the combination of a state and district. The state and district together allow us to track the same office position over multiple elections, which we later use to get cluster-corrected standard errors. Finally, we create the outcome variable of our RD: win_{t+2} , which is one if the candidate wins the next election (two years later) and zero otherwise. Constructing this variable required addressing an important data quality issue: The names of candidates have not always been consistently recorded. For example, in 1978 the election for the first district in Alaska was won by "Don Young", and in 1980 he was reelected. However, for 1980 the name was recorded as "Donald E Young". Unfortunately, this was not an exception, as there were multiple cases where the same candidate would appear under a different name format. To deal with this to create win_{t+2}, we opted to say that a candidate was reelected if in the next election, for the same seat, a person with the same last name was elected. This approach has a potential limitation:

if a politician's child or other relative with the same surname wins the same seat two years later, our data would incorrectly code this as the original candidate's reelection. However, such succession in consecutive elections is sufficiently rare that we judge this risk to be less severe than the alternative of losing true matches due to name formatting inconsistencies.

3.2 OLS regression baseline

The most straightforward, and naive, way of investigating the relationship between incumbency status and the probability of winning the next election would be to use ordinary least squares regression. We first explore this approach to establish a benchmark for comparison, while acknowledging its limitations. The first model one could try to estimate the incumbency advantage could be to linearly regress the probability of winning the next election on the covariate which indicates whether the candidate won the current election (is the incumbent). The linear probability model would look as follows:

$$\text{win}_{i,t+2} = \beta_0 + \beta_1 * \text{win}_{i,t} + \epsilon_{i,t} \tag{1}$$

Inference using model 1 raises a multiple of concerns. First, in this setting the assumption of homoskedastic error terms looks pretty unreasonable. It would not be hard to imagine that the variance of the error term would differ across different districts, states and time. The same worries apply concerning autocorrelation of the error terms: It seems entirely plausible that error terms concerning the same office seat would be correlated over time. To account for these properties we use cluster-corrected standard errors, where we cluster at the seat level (state-district combinations). A more serious problem with this approach is the omitted variable bias. There are bundles of factors that make a candidate more likely to win an election, political talent, strategy, funds, etc. Some of those qualities are specific to a candidate and make them likely to win, both initially and in the subsequent elections. This causes the covariate ($\text{win}_{i,t}$) to be correlated with the error term, leading to inconsistent estimates for β_1. Using the observed data available, we can partially address the observable differences, by including more covariates and estimate the following model.

$$\text{win}_{i,t+2} = \beta_0 + \beta_1 * \text{win}_{i,t} + \beta_2 * margin_{i,t} + \beta_3 * \text{year}_t + \epsilon_{i,t} \tag{2}$$

In equation 2, we include the vote margin from the current election ($margin_{i,t}$) to capture a candidates strength, candidates who win by larger margins may possess superior political skills that persist over time. We also include the election year to control for secular trends in incumbency advantage or electoral competitiveness. We can extend on this by including year fixed effects and state fixed effects to absorb time dependent national factors and state differences in political culture and electoral competitiveness. The resulting model is as follows:

$$win_{i,t+2} = \beta_0 + \beta_1 * win_{i,t} + \beta_2 * margin_{i,t} + \alpha_t + \gamma_s + \epsilon_{i,t} \tag{3}$$

Where α_t and γ_s are the fixed effects reflecting the year and state respectively. Even these richer specifications, however, cannot fully account for the unobserved candidate quality. Vote margin, while informative, is itself an outcome of the same unobserved qualities we are trying to control for. Therefore, $\hat{\beta}_1$ remains likely biased upward in all three specifications, strong candidates win and continue to win, with or without an incumbency advantage. These OLS estimates provide a useful benchmark, illustrating what naive comparisons would suggest about the incumbency advantage, but they should not be interpreted causally.

3.3 Regression discontinuity

To perform valid causal inferences, we employ a sharp regression discontinuity (RD) design to identify the causal effect of incumbency on electoral success in subsequent elections. The key intuition behind this technique is that extremely close elections are decided by quasi-random events such as voter turnout and others, implying that the differences between candidates that barely win an election compared to those that barely lose are minimal. Therefore, we can look at elections with very narrow differences as a quasi-random assignment of the treatment variable (incumbency status). We exploit this to bypass the problem of unobserved heterogeneity to make causal inference. The identification of the RD rests on two main assumptions Hahn et al. (2001):

- Continuity of potential outcomes: In the absence of treatment, the conditional expectation of the outcome would be continuous at the threshold. Meaning that in the absence of treatment, individuals just above and just below the cutoff would have had similar outcomes. This would be violated if, for example, candidates who barely win differ systematically from those who barely lose on unobserved characteristics that also affect reelection probability.

- No precise manipulation of the running variable: Candidates cannot precisely control whether they finish just above or just below the threshold. This would be violated if, for example, candidates could strategically manipulate vote counts in close elections.

Under these assumptions comparing outcomes just above and below the threshold identifies the causal effect of incumbency. We estimate a local linear regression with robust bias-corrected inference as developed in Calonico et al. (2014) and implemented in the "rdrobust" package in R. This approach fits separate linear regressions on either side of the threshold using observations within a certain bandwidth of the cutoff. We also assess the robustness of our estimates by conducting a sensitivity analysis across multiple bandwidths ($h \in \{0.01, 0.025, 0.05, 0.10, 0.25\}$) and kernel functions (triangular, uniform, Epanechnikov). This addresses the concerns about the potentially arbitrary nature of bandwidth and kernel selection.

3.4 Validity Tests

In the context of U.S House of Representative elections, the two foundational identifying assumptions of RD have been scrutinized in the literature. More specifically, Caughey and Sekhon (2011) argues that the first assumption in 3.3 is violated. We perform the McCrary density test, McCrary (2008), to test whether the density of the running variable is continuous at the threshold and whether there are signs of significant manipulation at the cut-off. The McCrary density test estimates the density function separately on each side of the cutoff and tests whether these densities are equal at the threshold. Under the null hypothesis of no manipulation, the density of the running variable should be continuous at the cutoff. If candidates with superior resources or organization can systematically push themselves just over the winning threshold in close races, then winners and losers at the margin are not comparable, violating our identifying assumption. To further analyze the first assumption, we also perform additional covariate balance tests. We test whether predetermined covariates are balanced at the threshold. If candidates who barely win differ systematically from those who barely lose on the observed variables, this could indicate sorting around the cutoff, which threatens our first identifying assumption. We examine two key covariates: election year and party. For those two covariates, we estimate a standard RD specification treating the covariate as the outcome variable and test whether there is a discontinuity at margin = 0.

4 Empirical Results

4.1 OLS regression baseline

Before presenting our regression discontinuity estimates, we first establish a baseline using ordinary least squares regression to illustrate the extent of the omitted variable bias. Table 1, contains the regression results for the three models discussed in 3.2. All three specifications use cluster-robust standard errors at the seat level to account for potential correlation of error terms for observations regarding the same seat. With 505 unique seats in our sample, clustering is both feasible and appropriate given concerns about heteroskedasticity and autocorrelation in the error structure.

Table 1: Naive OLS Estimates of Incumbency Advantage

	Model 1	Model 2	Model 3
Coefficient Win_t	0.6347911	0.60077453	0.600809
Standard Error	0.0058942	0.01136732	0.011394
P-value $P(\lvert T\rvert > t)$	2.2e-16	2.2e-16	2.2e-16

Notes: Cluster-robust standard errors (clustered at seat level)

Estimation of model 1 shows a coefficient of 0.635, which would mean that the expected probability of a candidate winning the next election increases by around 63.5 % when the candidate won the last election (is the incumbent). The estimate is highly significant relative to all common significance levels. This coefficient of the treatment variable in this model is the highest among all the simple OLS estimates. This reflects the fact that model 1 is our most simple and naive approach. There are no other variables capturing the overall candidate strength in the model, therefore this OLS regression fails the most in disentangling the candidate quality from the incumbency advantage. Hence the biggest upward bias of them all.

Model 2 adds control variables for the vote margin in the current election and the election year. The inclusion of these covariates reduces the estimated effect slightly to 0.601. The vote margin serves as a proxy for candidate strength. The year variable controls for secular trends in incumbency advantage or changes in electoral competitiveness over the nearly five decades covered by our data. Despite these additional controls, the estimated incumbency advantage remains rather large and highly significant. We see that these control variables have already eliminated a small part of the upward bias.

Model 3 represents our richest OLS specification, including both year and state fixed effects alongside the vote margin. The year fixed effects absorb time-varying national factors that might affect all House races simultaneously such as national economic conditions or party strength. The state fixed effects control for time invariant differences across states such as political culture and institutional features. This specification produces a coefficient of 0.601, nearly identical to Model 2, indicating that state-level heterogeneity does not substantially alter the estimated relationship once we control for vote margin and temporal trends.

Importantly, these OLS estimates should not be interpreted causally. As has been repeated several times already, the fundamental problem is omitted variable bias stemming from unobserved candidate quality. Candidates possess varying levels of political talent, charisma, fundraising ability, etc. These characteristics are difficult to measure comprehensively but strongly influence both initial and subsequent electoral success. Strong candidates win initially and continue to win in future elections, not necessarily because incumbency offers advantages, but because they possess enduring qualities that make them formidable competitors regardless of incumbency status.

4.2 Regression discontinuity

Before estimating our RD model, we first explore how the running variable and outcome variable relate to each other.

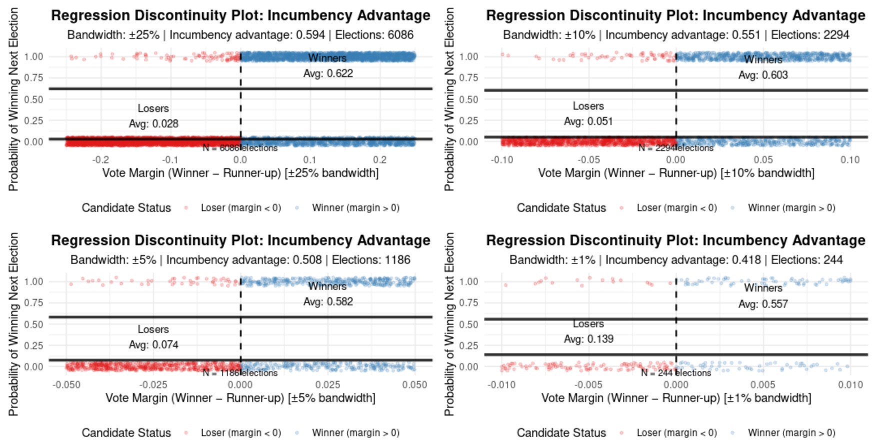

Figure 1: Average chance of winning of incumbents (blue) vs non-incumbents (red) for different bandwidths

Figure 1 compares how for different bandwidths the average probability of winning differs among incumbents (blue) versus non-incumbents (red). We see that the average "incumbency advantage", which is calculated as the differences between the averages of the two groups, reduces uniformly as we look at closer elections (smaller bandwidths). This pattern reaffirms our hypothesis of omitted variable bias. Increasing the bandwidth introduces bigger differences in candidate quality between winners and losers. The fact that increasing the bandwidth results in higher average incumbency advantage shows that part of what appears to be an incumbency effect in broader samples actually reflects the selection of stronger candidates into incumbency status. When we focus on increasingly close elections, where candidates are most similar, the estimated advantage shrinks.

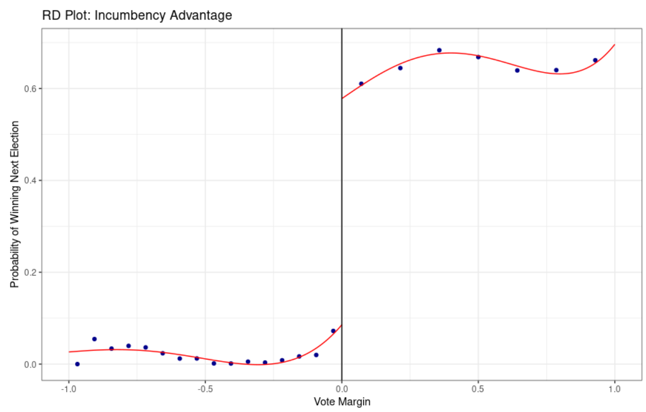

Figure 2: RD Plot: reelection probability by vote margin

Figure 2 presents binned scatterplot, which is standard in the context of RD designs, to provide a visualization of the relationship between vote margin (running variable) and reelection probability (outcome variable). The plot divides the running variable into evenly spaced bins on either side of the threshold and displays the average outcome within each bin. A clear discontinuity is visible at the cutoff: candidates who barely win (margin just above zero) have substantially higher reelection probabilities than candidates who barely lose (margin just below zero). This visual evidence strongly supports the existence of a causal incumbency advantage.

Table 2 presents our regression discontinuity estimates using bias-corrected robust inference procedures from Calonico et al. (2014). All specifications include election year as a covariate, and we examine sensitivity across five different bandwidths (0.25, 0.10, 0.05, 0.025, 0.01) and three kernel functions (triangular, uniform, Epanechnikov). Figure 3 visualizes the results.

Table 2: Regression Discontinuity Estimates of Incumbency Advantage

		Kernel Type		
Bandwidth	N	Triangular	Uniform	Epanechnikov
0.25	6,086	0.496***	0.526***	0.505***
		(0.023)	(0.020)	(0.022)
0.10	2,294	0.423***	0.446***	0.429***
		(0.039)	(0.034)	(0.037)
0.05	1,186	0.383***	0.398***	0.386***
		(0.055)	(0.050)	(0.053)
0.025	558	0.378***	0.380***	0.385***
		(0.076)	(0.070)	(0.075)
0.01	244	0.359***	0.302***	0.353***
		(0.108)	(0.101)	(0.107)

Notes: This table reports bias-corrected robust regression discontinuity estimates using the rdrobust package Calonico et al. (2014). Standard errors in parentheses. All specifications include election year as a covariate. N refers to the effective sample size within the bandwidth. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Several patterns are visible from Table 2 and figure 3. First, across all bandwidths and kernel choices, the estimated incumbency effect is positive and statistically significant, indicating a robust causal effect of winning an election on the probability of winning the subsequent election. This finding is consistent with the visual evidence from Figure 2, which shows a clear discontinuity in reelection probability at the cut-off.

Second, the magnitude of the estimated effect declines systematically as the bandwidth narrows. At the largest bandwidth of 0.25, the estimated incumbency advantage ranges between approximately 0.50 and 0.53, depending on the kernel. As the bandwidth is reduced to 0.10 and 0.05, the estimates fall to the range of 0.38–0.45. At the smallest bandwidths, which focus on the closest elections, the estimates stabilize around 0.35–0.38, albeit with larger standard errors due to the reduced sample size. This pattern mirrors the descriptive evidence shown in Figure 1 and provides further support for the presence of omitted variable bias in broader comparisons. Wider bandwidths include elections in which winners and losers differ more substantially in candidate quality. Narrower bandwidths, where candidates are more comparable, deliver estimates closer to the true causal incumbency effect.

Third, the estimates are stable across kernel choices. For a given bandwidth, the triangular, uniform, and Epanechnikov kernels yield similar estimates. This robustness suggests that the results are not driven by arbitrary kernel choices, which strengthens confidence in the our findings.

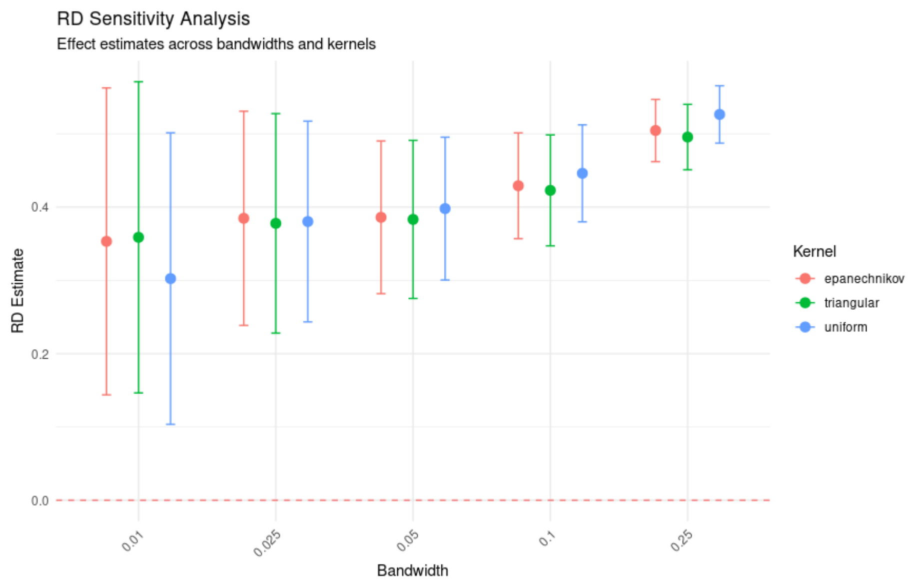

Figure 3: Results of the RD

Taken together, our preferred estimates suggest that barely winning a U.S. House election increases the probability of winning the subsequent election by approximately 35–40 percentage points. This effect is substantial but notably smaller than the 60–63 percentage point estimates obtained from naive OLS regressions in Section 4.1. The difference between the OLS and RD estimates is consistent with the idea that naive regressions overestimate the incumbency advantage by conflating it with candidate quality.

Finally, the bandwidth trade-off is evident in figure 3. Smaller bandwidths reduce bias by improving comparability between winners and losers, but at the cost of increased variance due to fewer observations. Larger bandwidths improve precision but risk reintroducing bias. The consistency of estimates across the different bandwidths suggests a large but not overwhelming incumbency advantage.

4.3 Validity tests

The validity of our regression discontinuity estimates depends on two identifying assumptions as mentioned in Section 3.3: continuity of potential outcomes and no precise manipulation of the running variable. These assumptions have been subject to scrutiny in the literature, particularly in Caughey and Sekhon (2011). The paper argued that winners and losers of close U.S. House elections may not be as comparable as initially assumed. We therefore conduct two sets of validity tests to assess whether these assumptions hold in our data.

4.3.1 Density Test

To test for potential manipulation of the running variable, we conduct the density test using the rddensity package as described in McCrary (2008). This test examines whether there is a discontinuity in the density of the running variable at the threshold. If candidates could precisely control election outcomes to barely win rather than barely lose having superior resources or even electoral fraud, for example, we would expect to observe an unusual concentration of observations just above the cutoff relative to just below it. Such bunching would violate the no manipulation assumption and potentially invalidate the results of our RD design. We restrict the sample to Democratic candidates only. This restriction is necessary to deal with forced symmetry: when we include both parties in the sample, every close election contributes two observations and due to the way we have defined our running variable, these two observations are mechanically symmetric around zero. This forced symmetry would mask genuine manipulation if it exists. By restricting to one party, we eliminate this artificial balance and allow the test to detect discontinuities in the density.

Table 3 shows the results. Panel A shows the main density test. First we discuss the fact that the data-driven bandwidth is 0.150-0.156. This is substantially larger than the bandwidths of our RD design. However, this is perfectly reasonable. Estimation of density functions requires a larger bandwidth and it is common in RD literature for these bandwidths in the density tests to be significantly larger than those of the RD design. With this clarification out of the way we look at the results. The p-value of the test is 0.085, which is not small enough to reject the null hypothesis of no manipulation at the conventional 5% level. It is, however, "borderline insignificant". Therefore, to further strengthen our conclusions we would like some additional evidence to support our no manipulation assumption. That evidence can be found in Panel B. Panel B reports complementary binomial tests across multiple window lengths around the threshold. These tests compare the number of observations just below versus just above the cutoff under the null hypothesis of equal probability ($p = 0.5$). If manipulation were present, we would expect to see significantly more winners than losers in narrow windows around the threshold. However, across all ten window lengths examined, none of the binomial tests reject the null hypothesis. All p-values exceed 0.25, with many exceeding 0.50. The observations are remarkably balanced on both sides of the cutoff across all bandwidths, giving strong indication in favor of our assumption.

Table 3: McCrary Density Test for Manipulation

Panel A: Main Density Test		
Number of observations	7,434	
Observations left of cutoff	3,282	
Observations right of cutoff	4,152	
Kernel	Triangular	
Bandwidth method	Estimated	
	Left of c	Right of c
Effective number of obs	855	896
Bandwidth (h)	0.156	0.150
Robust test statistic: T = -1.72, p-value = 0.085		

Panel B: Binomial Tests at Various Window Lengths

Window Length / 2	Obs $<$ c	Obs $\geq$ c	P-value
0.003	20	24	0.652
0.006	37	38	1.000
0.009	51	58	0.566
0.013	70	69	1.000
0.016	90	88	0.940
0.019	103	103	1.000
0.022	129	112	0.303
0.025	149	129	0.254
0.028	168	154	0.469
0.032	182	171	0.595

Notes: Panel A reports results from the density test implemented using the `rddensity` package McCrary (2008). Panel B reports binomial tests at various window lengths, comparing the number of observations just below versus just above the cutoff.

4.3.2 Covariate Balance

As a second validity check, we examine whether pre-treatment covariates are balanced at the threshold. Under a valid RD design, potential outcomes must be continuous at the cutoff. Observable characteristics may differ as long as they evolve smoothly. If they do not evolve smoothly this could indicate potential sorting that could bias our estimates. We test for discontinuities in two key covariates: election year and party affiliation. These variables are chosen because they are clearly predetermined before the election results and capture potential sources of systematic differences between winners and losers.

Figure 4 presents the covariate balance test for election year. The plot shows no visible discontinuity at the cutoff. The average election year evolves smoothly through the threshold, with winners and losers in close elections drawn from similar time periods. Formal RD estimation confirms this visual impression: the estimated discontinuity in year at the cutoff is small in magnitude and statistically insignificant (p-value of 0.273). This result indicates that our sample composition does not change systematically around the threshold.

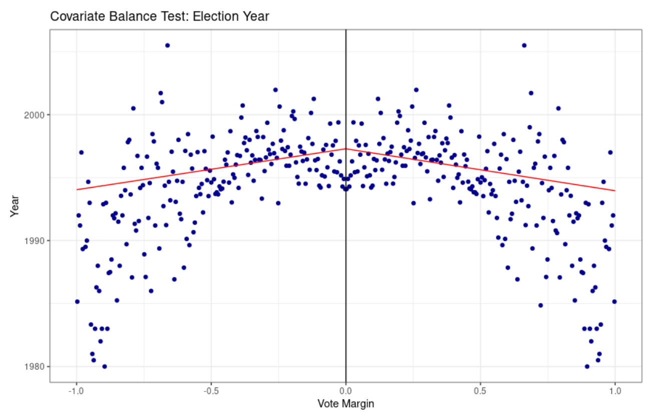

Figure 4: Covariate balance, year

Figure 5 presents the covariate balance test for party affiliation. We specifically test whether the probability of being a Democrat jumps discontinuously at the cutoff. If one party were systematically better at winning very close elections we would observe a discontinuity. However, the plot shows no evidence of such a jump. The probability of being a Democrat evolves smoothly through the threshold, and formal estimation yields a very small and statistically insignificant discontinuity (p-value of 0.995).

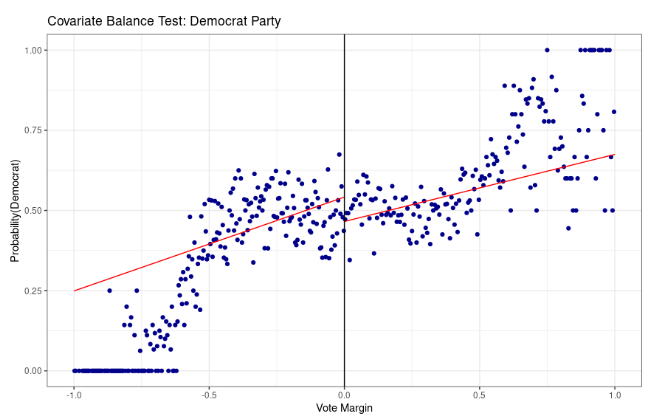

Figure 5: covariate balance, party

4.3.3 Interpretation

Taken together, these validity tests provide strong support for the identifying assumptions of our RD design. The density test shows no evidence of manipulation and the covariate balance tests show no systematic differences in election year or party composition between narrow winners and losers. These findings suggest that our identifying assumptions are valid and we can have a causal interpretation of the RD estimates.

Evidence of manipulation or covariate imbalance would cast doubt on the local randomization assumption and potentially bias our RD estimates. However, as clarified by de la Cuesta and Imai (2016), some covariate imbalance does not necessarily invalidate the RD design, as the formal requirement is continuity of potential outcomes, not strict local randomization.

5 Conclusion

This paper has revisited the question of whether incumbency causally increases electoral success in U.S. House of Representatives elections. Using a regression discontinuity design that exploits the discontinuous assignment of incumbency in close races, we aimed to separate the true incumbency advantage from differences in underlying candidate quality, similarly to Lee (2008).

The empirical results provide clear evidence of a substantial incumbency advantage. Candidates who barely win an election are around 35-40 percentage points more likely to win the subsequent election than candidates who barely lose, confirming that holding office offers real benefits in electoral races. At the same time, the comparison with naive OLS estimates shows the presence of omitted variable bias in standard regression approaches. Much of the apparent advantage of incumbents in simple models reflects the fact that stronger, more popular candidates are both more likely to win initially and in subsequent elections.

To assess the strength of our results we performed sensitivity analysis. We estimated the coefficient reflecting the incumbency advantage using RD designs with different bandwidths and kernel specifications. We observed that the point estimates are quite robust to the choice of different kernels. Furthermore, we saw that decreasing the bandwidths led to a stabilization of estimates to a value between 0.35 and 0.40. We also specifically addressed concerns around the validity of the identifying assumptions in electoral RD expressed by Caughey and Sekhon (2011). Both the McCrary density test McCrary (2008) and covariate balance test show no significant evidence against our identifying assumptions.

Overall, this study reinforces the conclusion that incumbency itself plays a causal role in shaping electoral outcomes. These results have implications for democratic accountability. The presence of a large incumbency advantage suggests that elections may not fully function as mechanisms for holding representatives accountable based on performance. Advantages associated with holding office can reduce electoral competition and protect incumbents, even when voter preferences are closely divided.

While we have confidence in the strength of our results, we are also aware that there is room for further research and extensions. Especially in terms of the quality and availability of data we faced some limitations. The inconsistent recording of candidate names, for example, required us to find some workaround. This workaround seems reasonable, but might introduce slight measurement errors. Even more valuable would be to extend the covariate balance tests to more comprehensive data. Access to information about candidate characteristics such as campaign spending would allow us to investigate the identifying assumptions in a more complete manner. Despite these limitations, our study seeks to contributes to the literature by providing updated empirical evidence on the causal effect of incumbency using modern regression discontinuity methods, performing sensitivity analyses, explicitly validating our identifying assumptions and using extended data.

References

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Caughey, D. and Sekhon, J. S. (2011). Elections and the regression discontinuity design: Lessons from close U.S. House races, 1942–2008. *Political Analysis*, 19(4):385–408.

de la Cuesta, B. and Imai, K. (2016). Misunderstandings about the regression discontinuity design in the study of close elections. *Annual Review of Political Science*, 19:375–396.

Eggers, A. C., Fowler, A., Hainmueller, J., Hall, A. B., and Snyder, J. M. (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.

Hahn, J., Todd, P., and van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

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

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

MIT Election Data and Science Lab (2024). U.S. House 1976–2024.