

Are Close Elections Randomly Determined?*

Justin Grimmer [†] Eitan Hersh [‡] Brian Feinstein [§] Daniel Carpenter [¶]

October 22, 2010

Abstract

Elections with small margins of victory represent an important form of electoral competition and, increasingly, an opportunity for causal inference. Scholars using regression discontinuity designs (RDD) have interpreted the winners of close elections as randomly separated from the losers, using marginal election results as an experimental assignment of office-holding to one candidate versus the other. In this paper we suggest that marginal elections may not be as random as RDD analysts suggest. We draw upon the simple intuition that elections that are expected to be close will attract greater campaign expenditures before the election and invite legal challenges and even fraud after the election. We present theoretical models that predict systematic differences between winners and losers, even in elections with the thinnest victory margins. We test predictions of our models on a dataset of all House elections from 1946 to 1990. We demonstrate that candidates whose parties hold structural advantages in their district are systematically more likely to win close elections. Our findings call into question the use of close elections for causal inference and demonstrate that marginal elections mask structural advantages that are troubling normatively.

*We thank Dan Lee for helpful discussant comments and participants at the Midwest Political Science Association Annual Conference. For helpful discussions we thank Daniel Butler, Gary Cox, Andy Eggers, Jens Hainmueller, Daniel Hopkins, David Lee, Holger Kern, Gary King, and Clayton Nall. All remaining errors, omissions, and interpretations remain ours.

[†]Assistant Professor, Department of Political Science, Stanford University; Encina Hall West 616 Serra St., Palo Alto, CA, 94305

[‡]Ph.D. Candidate, Department of Government, Harvard University. 1737 Cambridge St., Cambridge, MA 02138

[§]J.D. candidate, Harvard Law School.

[¶]Allie S. Freed Professor of Government. Department of Government, Harvard University. 1737 Cambridge St., Cambridge, MA 02138.

Competitive majoritarian elections comprise perhaps the defining feature of democratic republics. The question of whether these elections are truly competitive has become a central criterion in the assessment of democracy, whether qualitative (Bensel, 2004) or quantitative (Gasiorowski, 1996; Vanhanen, 2000; Przeworski et al., 2000). The idea is rather simple and compelling; if those who hold power have little chance of becoming unseated, whether through elections or other means, then the political system tends toward autocracy in fact, whatever its formal institutions may suggest.

Not even the world’s mature democracies can take for granted the prevalence of electoral competition. The existence of competitive elections depends not merely upon institutions such as universal adult suffrage, open candidate qualification, reduced barriers to entry, and free press and speech protections, but also on how elections unfold behaviorally. In many cases, formally democratic systems persist with surprising rarity. In the United States, scholars have puzzled over the disappearance of “marginal elections” (Fiorina, 1977), or close contests in which each candidate or party would have plausible incentives to show responsiveness to voter preferences and concerns. The vast literature on the “incumbency advantage” in American congressional elections is, in part, a reflection on this reduced electoral competition (Ansolabehere, Snyder Jr. and Stewart III, 2000). Some critics have gone so far as to suggest that the lack of electoral competition makes the concept of democracy problematic itself. Elections for political office may not, in and of themselves, suffice for representative government; indeed, elections without competition may create fictions of popular sovereignty (McCormick, 2001).

Despite their historically increasing scarcity, marginal elections have become important in another way. In recent years economists, political scientists, statisticians and other scholars have begun to exploit the properties of marginal elections for purposes of causal inference (e.g., Lee 2008 and Eggers and Hainmueller 2009). Using a sophisticated technology of statistical inference and the intuition that close elections are near-randomly determined, these scholars have essentially treated the winners and losers of marginal elections as randomly assigned to “election winner” (treatment) and “election loser” (control) groups. As the margin gets close, in other words, the winner of the election is determined as if it were the result of a fair coin toss. In quite powerful analyses, these

scholars have shown theoretically that only very simple and easy-to-satisfy assumptions are needed to identify causal effects of interest (Hahn, Todd and van der Klaauw, 2001; Lee, 2008). Drawing upon these methods, causal inference designs from marginal elections have been skillfully used to demonstrate incumbency advantage (Lee, 2008), policy responsiveness (Lee, Moretti and Butler, 2004) and rents from office holding (Eggers and Hainmueller, 2009).

In this paper we consider properties of marginal elections that cast some doubt on whether they are truly randomly determined. Our initial purpose is to question the utility of close elections for causal inference designs. In addition, one of our larger purposes is to demonstrate that marginal elections may mask structural advantages for certain candidates and parties, therefore calling into question many of the normative appeals for marginal elections. In doing so, we draw upon a basic intuition of strategic electoral politics: in single non-transferable vote systems where the “winner takes all” – where the value from votes garnered in a close but losing effort is zero – the effort and advantages to be deployed by a candidate or party will be much more effective in a close election than in a rout. In other words, close elections are those where differences of campaign resources, structural advantages, and even fraud should most show themselves. As a result, marginal elections are the ones that will attract the greatest campaign effort and resources, and close contests will also attract the deployment of structural advantages.

If our hypotheses are correct about the effects of this resource flood, then close elections may fall disproportionately to the candidate with certain structural advantages. And if this is true, then the near-randomness of these contests – and their utility for causal inference – must be called into question. So too might the conclusions of regression discontinuity designs be revisited. If, for instance, it is shown that the winners of close elections are more likely than the losers to go onto richer earnings (Eggers and Hainmueller, 2009; Snyder and Querubin, 2008), one might ask whether the effect is due to winning office, or whether some property of the candidate that correlates with winning elections is the same property that leads to higher post-career earnings. For example, winning candidates may have better class-position, higher skill levels, or better access to the party elite. The idea that winning marginal elections reflects resource and structural advantages may also help explain why these individuals are reelected at higher rates in subsequent contests (Lee,

2008). Candidates better able to exploit their party’s structural advantages may also be better able to exploit the tools of incumbency once they arrive in Washington or have increased access to fundraising opportunities before the next election.

To formalize our hypotheses, we begin with two types of models of electoral “manipulation”, one model of campaigning before Election Day, one model of legal challenges and fraud after. Our first model makes the intuitive prediction that campaign expenditure will depend upon the predicted margin of the race. The model formalizes the intuition that equilibrium campaigning decreases as the expected margin of a race increases. For marginal elections, then, any asymmetries in campaign resources, skills, structural advantages and other candidate properties will become *magnified* in marginal elections. This implies that there will be systematic differences within narrow bandwidths of the discontinuity. Our second model examines manipulation of electoral results after an election, making the prediction that systematically manipulated elections will give the appearance of the razor-thin differences necessary for valid RDDs. Our models predict that candidates with structural advantages are better able to manipulate votes after the election, leading to the prediction that the winners of close elections differ systematically from the losers, confounding the estimates from RDDs.

We test the predictions of our theoretical models using a data set of Congressional elections after World War II. We aggregate data that are indicative of structural advantages in a district. Specifically, we employ data on the party controlling the Governor’s office at the time of the election, as well as data on the party controlling the election administration—such as the Secretary of State’s office. Our analyses indicate that candidates with structural advantages in a district—sharing the same party with the Governor or the Secretary of State—translates into a systematic advantage in extremely close elections. In some instances, these candidates are over *ten percentage points* more likely to win the election. This is indicative of the systematic determination of extremely close elections.

Before proceeding, we offer two qualifications. First, our analyses do not suggest that regression discontinuity designs are necessarily invalid. In cases where the distribution of election outcomes does not satisfy the properties we attribute theoretically and empirically to marginal elections, RDD

designs may stand as robust designs for causal inference. So too, one interpretation of our findings is that analysts simply need to take into account these structural advantages in a matching design where scholars match on partisan advantages. Still, the theoretical basis of our paper suggests that there may be unobservable differences in candidates in close elections, differential advantages for which statistical analysts cannot fully measure or account. Second, our analyses do not by themselves form the basis for any sort of general critique of elections and competitive democracy. More research would be needed to follow upon the inquiries here, yet the idea that close elections may be less stochastic than commonly presumed opens both normative and positive questions, to which we return in our conclusions.

1 Marginal Elections and Their Properties

1.1 Regression Discontinuity Designs

The idea that close elections embed a random component that pushes a winner “over the top” is made as a useful statistical assumption. But underlying this statistical assumption are several assumptions about the politics of close elections. In a world of two candidates and one office, a really competitive race is one that both candidates have a shot at winning. Taken to the extreme, this assumption about competition presumes that as the race gets close to equal vote shares, the outcome is determined as if a fair coin were tossed. This randomness creates opportunities for what is commonly called a “natural experiment.” If winning a marginal election is determined by the flip of a coin, then the background characteristics of candidates, parties, and districts that normally confound analyses are rendered orthogonal. This enables a study of a wide-range of consequences from winning office—rents, subsequent election advantages, a portfolio of policy choices, and policy outcomes—that are otherwise deeply confounded. Recent scholars describe the resulting exercise as “exploiting the quasirandom assignment of office in very close races” (Eggers and Hainmueller 2009).

The argument for a regression discontinuity design is powerful, particularly when one considers how hard it is to exactly identify causal effects from observational data. In causal inference, we are

primarily interested comparing two counterfactual states of the world. For a running example in this paper, we are interested in measuring the *incumbency advantage* or the effect of incumbency status on electoral support (for example, Erikson 1971 ; Gelman and King 1990). We follow Lee’s (2008) example and consider the effect of incumbency on support for Democrats in Congressional districts. To measure the incumbency advantage, we need to compare the percent of the vote for Democrats in district i under “treatment” $Z_i(1)$, with a Democrat incumbent in district i , and the percent of the vote for Democrats in district i under control $Z_i(0)$, or without a Democrat incumbent in the district. The fundamental problem of causal inference ensures that for each district i we observe only response under treatment or response under control (Holland, 1986),

$$Z_i = D_i Z_i(1) - (1 - D_i) Z_i(0)$$

where D_i is equal to 1 if the Democrat candidate wins the election and 0 otherwise. Given the impossibility of identifying individual level treatment effects, the goal of many causal studies is to identify the *Average Treatment Effect* (ATE), or the average response to treatment for a population of Congressional districts,¹

$$\text{ATE} = E[Z(1) - Z(0)].$$

In general, the systematic selection that plagues observational data will make identifying the ATE difficult, if not impossible. For example, we might consider contrasting the average electoral support for the observed incumbents $E[Z(1)|D = 1]$ to the average electoral support for the observed opponents $E[Z(0)|D = 0]$. But, there are systematic differences between incumbents and challengers in any electoral cycle that are unrelated to the advantages of holding office. For example, incumbents had to win an election to obtain the incumbency status, and therefore they may have systematically better candidate quality than the non-incumbent candidates. This difference between incumbents and their challengers, along with a host of other potential confounders, implies that $E[Z(1)|D_i = 1] \neq E[Z(1)]$ and the same confounding ensures that $E[Z(0)|D_i = 0] \neq E[Z(0)]$.

¹Throughout we will suppose that the expectation is over the relevant districts.

Together this implies the well known fact about observational data: the naive difference in means will fail to identify the ATE (Morgan and Winship, 2007).

Political scientists regularly employ regression models or use matching procedures in an attempt to remove confounding. But both methods rely upon *selection on observables*: the assumption that we have the exact set of covariates that remove all systematic differences between incumbents and challengers (Morgan and Winship, 2007). Further, unless exact stratification on the covariates is possible, we also must assume that we have identified the proper functional form for a regression, the correct specification of a propensity score (Rosenbaum and Rubin, 1983), or a combination of other matching algorithms that lead to comparable treatment and control groups (Sekhon, 2010; Hainmueller, 2010). Certainly the careful application of regression, matching, and their combination can reduce the confounding, but exact identification of any causal effect remains unlikely (Ho et al., 2007).

The insight of the regression discontinuity design is that identification of a *local* average treatment effect is possible, even from observational data that are otherwise deeply confounded. RDDs focus on identification of a treatment effect at a covariate level that constitutes a threshold for treatment assignment: below the threshold level of the covariate the subjects are assigned to control, above the threshold they are assigned to treatment. In electoral studies that employ RDDs, it is common to focus on vote share in the previous election, x , with studies attempting to identify the causal effect of incumbency *at the discontinuity*, or at the level of voter support that determines the election winner, $x = \frac{1}{2}$. We will denote the causal effect at the threshold of $\frac{1}{2}$ of vote share by,

$$\text{ATE}_{1/2} = \text{E}[Z(1) - Z(0)|x = 1/2],$$

or the average difference between electoral support for Democrats in districts with a Democrat incumbent, less the electoral support for Democrats in districts without a Democrat incumbent, given that the vote share in the previous election was $x = 1/2$.

Identification of $\text{ATE}_{1/2}$ from observational data requires two continuity assumptions. Specifically, we assume that $\text{E}[Z(0)|x]$, expected support for non-incumbent Democrats, given previous vote share x , and $\text{E}[Z(1)|x]$, expected support for incumbent Democrats given previous vote share

x , are continuous in x (Hahn, Todd and van der Klaauw, 2001; Lee, 2008; Imbens and Lemieux, 2008).² The continuity assumptions identify the causal effect of interest by overcoming of the fundamental problem of causal inference, but only at the threshold. As we approach 0.5 from either side, the continuity of the functions ensures that $E[Z(0)|X = 0.5] = \lim_{x \uparrow 0.5} E[Z(0)|X = x]$ and that $E[Z(1)|X = 0.5] = \lim_{x \downarrow 0.5} E[Z(1)|X = x]$. And therefore,

$$\begin{aligned} E[Z(1) - Z(0)|X = 0.5] &= \lim_{x \downarrow 0.5} E[Z(1)|X = x] - \lim_{x \uparrow 0.5} E[Z(0)|X = x] \\ &= ATE_{1/2} \end{aligned}$$

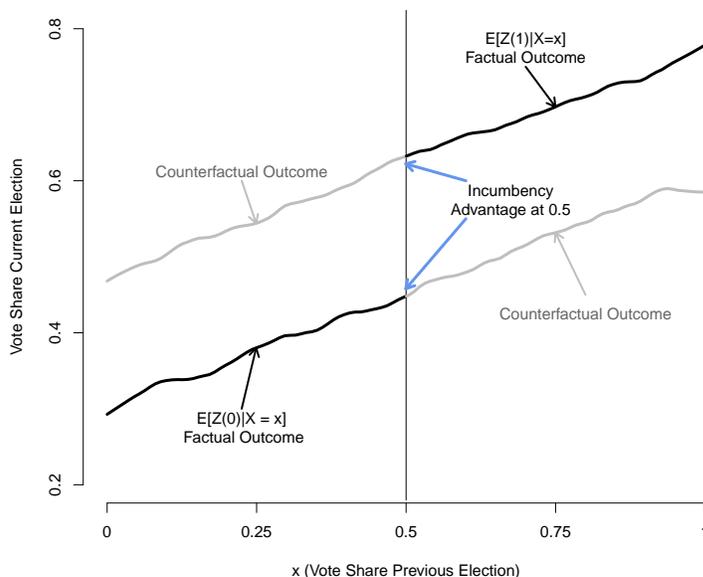
In other words, the continuity assumptions allow us to simultaneously observe $E[Z(1)|X = 0.5]$ and $E[Z(0)|X = 0.5]$.

To better understand this assumption, Figure 1 provides a graphical depiction. In Figure 1 the black lines represent the observed conditional expectations and the gray lines are the counterfactual conditional expectations, those that are not observed. Notice that the black and gray lines are connected continuously at 0.5. This continuity implies that there are no systematic differences between the treatment and control groups, immediately around the discontinuity. This then implies that, as we approach 0.5 from below in the limit, the expected value of the control observations provide the correct counterfactual value for the treated observations. Likewise, in the limit as we approach the discontinuity from above, the treated observations provide the correct counterfactual responses for the control units. The result is that the difference, $E[Z(1)|X = 0.5] - E[Z(0)|X = 0.5]$ identifies $ATE_{1/2}$.

The continuity assumptions at the marginal elections is the key to RDDs identifying $ATE_{1/2}$. These assumptions, and their more general variants, are regularly trumpeted as weak assumptions that provide robust identification in many different contexts. In the next section, we suggest that

²This is stronger than actually needed to identify the causal effect of interest, as both Imbens and Lemieux (2008) and Lee (2008) observe. However, the more general assumptions preserve the basic intuition that we motivate here and suffer from similar vulnerabilities. In general, we can restrict the continuity assumption to the discontinuity (Imbens and Lemieux, 2008). Even more generally, we might suppose that we observe vote share x , but fail to observe some effort level W . Then, it need only be the case that the cdf of x conditional on w , $F(x|W)$ is continuously differentiable in x at $x = 1/2$. As we will see all the assumptions rely on the critical assumption that, at the discontinuity, observations are just as likely to be above the threshold as they are to be below the threshold (which is why the continuity assumptions are so critical).

Figure 1: Graphical Presentation of Assumptions to Identify $ATE_{1/2}$



This figure provides a graphical demonstration of the assumptions used to identify $ATE_{1/2}$ in regression discontinuity designs. The black lines represent the observed relationship between electoral support as a non-incumbent ($E[Z(0)|X = x]$) and electoral support as an incumbent ($E[Z(1)|X = x]$). The gray lines are the counterfactual, or unobserved functions. The critical assumption is that both conditional-expectation functions are continuous. In the limit as we approach the discontinuity, there are no systematic differences between challengers and incumbents, otherwise, there would be a discontinuity in the conditional-regression functions. The absence of these discontinuities implies the identification of $ATE_{1/2}$.

the continuity assumptions may be more restrictive than previously suspected.

1.2 What Can Go Wrong? Practical and Theoretical Concerns

Recent applications of RDD designs draw heavily upon the continuity logic developed in Hahn, Todd and van der Klaauw (2001) and Lee (2008). As Eggers and Hainmueller (2009: 11) remark of their study of British parliamentary elections (where emphasis is added),

Following pioneering work by Lee (2008), we note that in very close elections, the assignment to political office is largely based on random factors. Although winning candidates may generally be different from losing candidates at the time of the election (e.g., better looks, more money, greater speaking ability), *there is no reason to expect the winners and losers of elections decided by razor-thin margins to systematically differ in any way.* The RD design therefore attempts to estimate the difference in wealth

precisely at the threshold where winners and losers are decided (i.e., where the margin of victory approaches zero). If local random assignment holds at the threshold, the RD estimate can thus be as credible as an estimate from a randomized experiment.

We highlight two potential problems with the continuity logic: one practical, one theoretical. In practice, the key problem in application of RDD designs to election is that data constraints and statistical power requirements means that too few elections with razor-thin margins are available for most analyses. Hence the analyst must choose a bandwidth – for purposes of election analyses, a margin of victory into which the sample cases fall, thus specifying a sample from which races with margins larger than the bandwidth are excluded. Eggers and Hainmueller (2009) examine parliamentary elections in Great Britain, comparing winners and losers of marginal races. They use a statistical criterion to choose a bandwidth of 15 percentage points in vote share; hence races decided by a 57-43 margin lie in their sample. Eggers and Hainmueller (2009) then compare winners and losers within this bandwidth, and also implement a regression model and matching analysis on this sample, including their “full set of covariates (including schooling, university education, occupation, gender, year of birth, and year of death).”

The selection of bandwidths represents a disconnect between the theoretical results that justify the use of regression discontinuity designs and their actual application. Regression discontinuity proofs are based on an assumption of an infinite (or extremely large) sample that allows for no extrapolation at the discontinuity. In any application, however, there will be insufficient data at the margin to perform the described limit and still retain enough statistical power to reject any null hypotheses. This forces the selection of a bandwidth and the borrowing of information across the bandwidth to extrapolate to the discontinuity. If factors are balanced at the discontinuity, but imbalanced in areas very close to the discontinuity and within the bandwidth, then the result could be a badly biased estimate of the $ATE_{1/2}$. And if regression or matching estimators are used within the bandwidth, but fail to include characteristics that are imbalanced a short distance from the discontinuity, the analyses will be unable to identify $ATE_{1/2}$, or any other unconfounded causal effect. In short, extrapolation matters in real applications, even if it is ignored in the econometric proofs.

Table 1: Summary of Assumptions and Potential Issues with RDD Designs in Marginal Elections

- 1) Treatment is essentially randomized to winners and losers only in the limit, yet researchers must choose a bandwidth. Nothing about optimal bandwidth choice gets around this problem.
- 2) There are no post-assignment (post-voting) discontinuities such as legal challenges or fraud that may affect assignment to winners and losers.

A second problem is the possibility of sorting around a discontinuity, which renders RDD estimates no better than the estimates from observational studies. Once an initial ballot count is announced in a close race all sides know, with certainty, how many votes they will need to legally challenge or how many ballots they will need to stuff in order to win the election. This enables stealing of elections with extremely small margins. Building on this intuition, below, we present a game of post-election manipulation that predicts candidates will use their resources to systematically secure office. The manipulation will result in candidates doing just enough to “steal” an election from their opponent—creating the impression of marginal elections that are actually systematically determined. If candidates can deterministically sort around the boarder, RDDs no longer provide valid estimates of $ATE_{1/2}$ or another causal effect of interest. Intuitively, sorting represents a type of selection, breaking the protocol of an experiment. More technically, sorting creates a discontinuity in $E[Z(1)|X = x]$ and $E[Z(0)|X = x]$ functions.³ The result is that $E[Z(0)|X = 1/2]$ no longer provides a valid estimate of the counterfactual losing response for candidates tha just happen to win. The result is bias in an unknown direction and of unknown size.

In the following sections we provide theoretical and empirical evidence that both problems discussed here are likely to manifest in Congressional election data. Any bandwidth choice – even if done under conditions of algorithmic optimality – will leave a sample of elections that are, by definition, marginal. These marginal elections will attract greater campaign investments, such as advertising, deployment of structural advantages, and mobilization efforts. Indeed, as the margins

³In the more general proof in Lee (2008) we can think of the discontinuity occurring in the measure on the unobserved (effort) variable W . If $g(w)$ is continuous, then each observation is just as likely to be in the treated arm or the control arm at the discontinuity. If there is a discontinuity, however, some observations are systematically more likely to be in treatment than control. This breaks the weighted average conditions in Lee’s (2008) Proposition 2b and 3b.

get smaller, our models suggest that candidates will invest more of these resources in the race. Any systematic differences in candidate resources, quality, advantages and other variables – and one can (and we do) think of candidate equality on these dimensions as a measure-zero event – within narrow bandwidths will confound the causal effects of interest, unless there is sufficient data to focus only on outcomes at the margin. Unless these variables are controlled for (or matched upon), or unless they are uncorrelated with the outcome variable, the *applied discontinuity analysis* will not achieve randomization across the 50-percent threshold. And even if the conditions are met for randomization at the discontinuity, close elections are the most likely to be subjected to legal challenges and most at risk for electoral fraud. Post-election manipulations of vote results are deterministic, resulting in sorting around the discontinuity. This renders RDDs no better equipped to estimate causal effects than other methods for observational data.

2 How Do Campaigns Purposefully Sort Around the Discontinuity?

Politicians do not participate in elections only as candidates; they also have a hand in managing nearly every decision of the electoral process, from deciding the boundaries of electoral jurisdictions to the system of voter registration, from the format of the ballot to the mobilization of supporters. Moreover, some politicians, namely those associated with the dominant political party in their respective states and districts, play a far greater role in the process than their competitors. Consequently, we consider the potential for the origin of structural advantages in districts and the potential for purposeful sorting around the discontinuity.

Dominant parties may have a very good sense of how close a given election is going to be ahead of time. These parties may understand the pulse of the voters and the landscape of the district. If the election does not look close, they need not waste their resources. If it looks very close, they may employ massive resources to put themselves over the 50% mark. And immediately after Election Day, but before the results are certified, parties know with certainty the number of votes necessary to win an election. Dominant parties are able to use their influence on legal proceedings, the ability

to certify electoral results, or even their opportunity to commit fraud to tip electoral results.

We consider two possible pathways for manipulation by dominant parties in close elections, one before the election, one after.⁴ Long before Election Day, dominant parties are able to craft Congressional districts to accomplish their electoral goals. If one political party dominates a state's political offices, it can reap significant advantages by creating favorable legislative districts. Strategic redistricting was one of the first causes hypothesized for the decades-long trend of fewer and fewer close elections in the U.S. Congress (Tuft, 1973). But a growing consensus has emerged that redistricting is not the cause of the vanishing marginals (e.g., Ferejohn 1977; Abramowitz, Alexander and Gunning 2006, Ferejohn 1977), because political parties rarely construct safe districts. Rather, the optimal strategy for a dominant party is to create districts in which its candidates can all win by slight margins, allowing the party to gain more seats overall (Gopoian and West, 1984; Campagna and Grofman, 1990; Desposato and Petrocik, 2003). The result are systematic differences in narrow bands around a discontinuity, although there will still be balance at the discontinuity.

After Election Day, but before the certification of electoral results there is the opportunity for electoral fraud and legal challenges. The dominant party or candidate, likely in control of key functions of election administration, clearly has more opportunities to perpetuate fraud than out-partisans. Caro (1990) recounts how Lyndon Johnson exploited his connections in Texas to steal a Senate primary election from Coke Stevenson, producing just enough fraudulent ballots to defeat his opponent (this is also recounted in Snyder (2005)). Similarly, structural partisan advantages shaped the outcome of the Florida recount during the 2000 presidential election. The Republican Secretary of State, Katherine Harris, certified candidate George W. Bush as the winner under a cloud of partisan favoring. The Florida Supreme Court, filled with Democrat appointees, extended recounts, raising the suspicion that the Court was aiding Gore's effort. And, of course, the United States Supreme Court's 5-4 decision that ended the post-election dispute was vilified as partisan.

As a more detailed example of systematic manipulation, consider the recent post-election dispute between 2008 Minnesota U.S. Senate candidates Norm Coleman and Al Franken. The first stage

⁴There are, of course, many more pathways for manipulation possible.

of the post-election dispute involved a recount of contested paper ballots that were submitted at the polls on Election Day. These were basically the equivalent of “hanging-chad” issues: ballots that were marked, but not marked exactly right. A non-partisan panel went through the ballots in question. The result was a success for Franken. The Election Day count was Coleman up by 215; after these ballots were sorted through, Franken took a lead by 49 votes.⁵

But the second stage of the recount reveals how structural advantages can determine the outcomes of very close elections. Several absentee ballots had been submitted but not counted. The campaigns agreed to open up the envelopes of 953 absentee ballots and count the votes inside. And this agreement had the appearance of unbiasedness: each campaign had the power to veto absentee ballots that they thought were invalid, but had to raise the objection *before* the envelopes were opened. The recount went exceedingly well for the Franken campaign, whose lead jumped to 225 votes after this stage of the recount, essentially ensuring that Franken would win the election.

How did the Franken campaign gain such a huge lead from a set of votes that both parties could have rejected? The key is that the absentee ballots came in the mail, revealing the names of the voters and their address information on the envelopes. This enabled the campaigns to perpetrate two forms of cherry-picking. First, the campaigns could selectively contact voters who had submitted absentee ballots but whose votes were not counted and encourage them to complain or provide them with legal aid. The two campaigns demanded lists from the election office of people who requested absentee ballots but who were not marked as having voted. They then could merge these records with their statistical model predicting each person’s level of support in the Senate race (presumably based on voter demographics, campaign contacts, and other micro-targeted information) and selectively call citizens favoring their respective candidates. If the Democrats had access to a superior voter file than the Republicans, this could have helped them gain votes.

The second form of cherry-picking is that the competing campaigns sorted through the absentee ballots together, and each campaign could veto the inclusion of disputed ballots they thought should not be counted. Nate Silver, then of the website *fvethirtyeight.com*, suggested that the Franken campaign may have been seriously advantaged in this veto process. The Coleman campaign vetoed

⁵Rachel E. Stassen-Berger, “Franken Leads by 50,” *St. Paul Pioneer Press*, December 29, 2008.

ballots based on the partisan composition of the precinct or county where the ballots were cast. The Franken campaign vetoed ballots based on the individual characteristics of the actual voter whose ballot was in dispute. Based on the counties that the 953 absentee ballots came from, observers predicted that Franken would receive 52% of the recounted absentee ballots. In fact, he received 61% of them.⁶

The Minnesota recount demonstrates how structural advantages determine close elections. Franken likely won the recount because Democrats had a better voter list, better access to the list, or a better model to identify likely supporters than Republicans. The recount also demonstrates the possibility for post-election manipulation, even in elections with national implications. This recount was extremely high profile, receiving attention from both the liberal and conservative leaning media; yet the Franken campaign was able to deploy its advantages to win the election.

3 Theoretical Model

We now formalize this intuition about campaigning and post-election manipulation and how they affect identification of causal effects in RDDs with two formal models. The idea that the expected margin of an election can draw greater effort from its contestants and their allies can be usefully formalized; the formalization not only ratifies the intuition but also draws attention and lends clarity to the underlying variables that matter most in examining these elections. There are, of course, many models of elections – such as spatial models of vote choice – but the essential properties of the models we seek are not those that examine voter choice or aggregation, nor the production of information (as in models of negative advertising). Instead, we seek simple but generalizable models that describe campaign dynamics, both before and after an election.

To that end, we consider a model of two candidates who observe a pre-election poll. In response to this information, the candidate (and/or the parties) spend costly resources in an attempt to

⁶Details about the Minnesota senate recount, cherry-picking, and selective vetoing of absentee ballots are described in news articles and web blogs, such as: Nate Silver, “Franken Jumps Out to 225-Vote Lead on Strength of Absentee Ballots,” *fivethirtyeight.com*, January 3, 2009; Nate Silver, “In Minnesota, End of Beginning Starts Today,” *fivethirtyeight.com*, January 3, 2009; Bob Collins, “Recount Q & A,” *Minnesota Public Radio News*, January 3, 2009; Eric Kleefeld, “Friendly Coleman Witness: They Cherry-Picked Me,” *Talkingpointsmemo.com*, January 29, 2009; “Senate Contest Day 4: ‘Cherry-picking,’” *The Uptake*, January 30, 2009.

increase their vote shares. These attempts meet with stochastic success, a random component still partially determines the outcome of the election. Under equilibrium campaigning in this model, resources are directed into districts that pre-election polls reveal to be competitive. This magnifies structural advantages and subsequently causes systematic differences between winners and losers within narrow bandwidths around the discontinuity. In our supplemental appendix we generalize this model using a differential game and demonstrate that our same predictions hold in this much more general model.

Our second model formalizes post-election challenges that are an important element of marginal elections. In this model, candidates observe the post-election, but pre-certification, vote totals. Then both candidates employ a set of tools to modify the final electoral total. Under equilibrium in this model, we show that resource advantaged candidates are able to “steal” elections from their disadvantaged opponents. This causes systematic sorting around the discontinuity, which confounds the causal effects estimated using RDD.

Both models sacrifice a focus upon information production (the equilibria and dynamics are not Bayesian), but they are useful for describing the dynamics of campaigns and the behavior of contestants as margins get smaller or larger both before and after the election. Both models preserve the rational choice properties of campaigns while permitting fully dynamic modeling that embeds candidates’ valuations of the future.⁷

3.1 A Simple Model of Campaigning

We begin our analysis with a simple model of resource investment during campaigns. Our model demonstrates that resources from both parties will converge upon close elections and that institutional advantages for one party will make them systematically more likely to win close elections. The result is that the parties that hold an institutional advantage in a state will be systematically more likely to win close elections.

We suppose that there are two candidates, 1 and 2, who are competing in an election. Our game proceeds in two stages. First, a poll that reveals to the candidates the current vote share

⁷Models with greater behavioral realism are possible and desirable, but are beyond the scope of analysis here.

in the election x_0 . After observing this poll the candidates make a decision about how much to invest in the campaign. Let c_1 denote the resources for candidate 1 and c_2 denote the resources for candidate 2. After the candidates make their investment decision, the final vote share is revealed, with the vote share for candidate 1 given by

$$x_1 = \gamma_1 c_1 - \gamma_2 c_2 + w \tag{3.1}$$

where γ_1 and γ_2 represent a multiplier on the campaign's investments and w is a draw from a $\text{Normal}(x_0, \sigma_0^2)$. The vote share for candidate 2 is given by $x_2 = 1 - x_1$. γ_1 and γ_2 capture one manifestation of candidates' institutional capacity during an election. Candidates with stronger party backing may be able to receive more return for their investments than their opponent.

Candidates' utilities are a combination of the cost of the campaign and their probability of obtaining the returns from office. Let k_1 and k_2 be multipliers that capture how efficiently candidates are able to invest their money during an election. Then, the candidates' utility functions are given by,

$$\begin{aligned} U_{\text{cand1}}(c_1, c_2) &= \text{Prob}(x_1 \geq 0.5) - k_1 \exp(c_1) \\ U_{\text{cand2}}(c_1, c_2) &= \text{Prob}(x_2 \geq 0.5) - k_2 \exp(c_2) \end{aligned}$$

To summarize, our game proceeds in three stages:

- 1) A poll result x_0 is revealed to the candidates
- 2) Candidates make their campaign investments c_1 and c_2
- 3) Vote share is revealed and payoffs are realized

Proposition 1 in the appendix proves that there is a pure strategy symmetric Nash equilibrium. To provide comparative statistics on this equilibrium we employ two simulations to demonstrate two primary points of our analysis. First, we show that an equilibrium response from both candidates is to invest more in closer elections.⁸ For both simulations, we will analyze an election where

⁸A formal comparative static will likely reveal that the amount invested in any one election is non-decreasing,

Candidate 1 has a resource advantage over Candidate 2, $\gamma_1 > \gamma_2$. Our first simulation demonstrates that, in the equilibrium, candidates invest more in close elections. The left-hand plot in Figure 2 shows that closer preelection polls induce more investment from candidates. To demonstrate this, we varied the preelection poll from 0.5—indicative of a very close election—to 0.7 and 0.3—indicative of an uncompetitive election.

As Figure 2 illustrates, the closer election induces more investment from both candidates. The result of this increased investment is systematic differences in who wins elections. In the right-hand plot in Figure 2 shows that equilibrium strategies predict that candidates with resource advantages will be systematically more likely to win close elections, even within very small bandwidths. This figure varies the size of the bandwidth along the horizontal axis, from wider (a 25% bandwidth) to more narrow (using the predictions from a polynomial regression model at the discontinuity). The vertical axis presents the average difference in resources between candidates who win and those that lose.

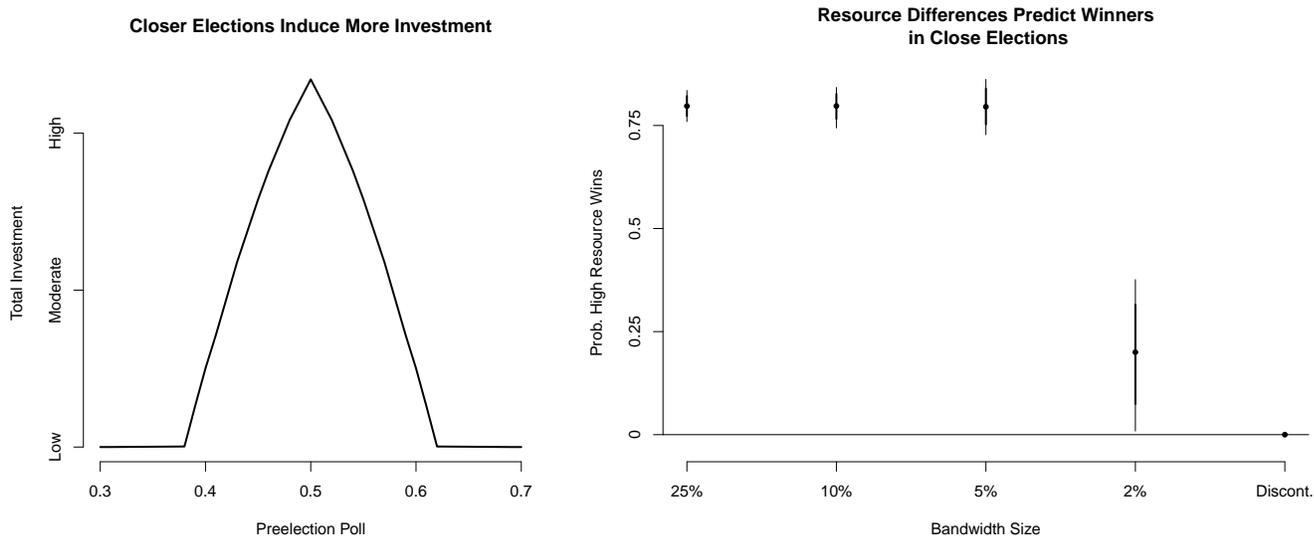
The right-hand plot in Figure 2 shows that our model predicts systematic differences exist between winners and losers, even in very close elections. Even elections decided by less than 2% points, we expect that winners will have systematically greater resources than losers. This model predicts, therefore, that empirical analyses that rely upon wide bandwidths will provide poor estimates of $ATE_{1/2}$. But, because of the randomization after the candidates invest their resources, the model predicts that the resources will be balanced at 0.5, which is demonstrated with the zero estimate at the far right.

3.2 Systematic Differences at the Discontinuity

Our models of campaigning predict that candidates with an institutional advantage in a district are systematically more likely to win close elections, even within very narrow bandwidths. But the models do not predict that RDDs will provide invalid estimates in the limit. The randomness inherent in each model predicts that the estimate at the discontinuity will be an unbiased estimate of the treatment effect at the discontinuity, so long as there are sufficient observations to estimate

because some elections an equilibrium response is to not campaign.

Figure 2: Close Elections Induce Greater Campaigning



This figure demonstrates two predictions from the simple campaigning model. The left-hand plot shows that the game predicts more resources invested in close elections. The right-hand plot presents the prediction of systematic differences in winners and losers in even close elections.

the effect exactly at the threshold for winning the elections. The important implication for the study of close elections is that commonly used bandwidths are unable to identify the desired treatment effect. In principle, however, enough data could be collected to identify the desired causal effect if sufficiently narrow bandwidths are employed.

Campaigns represent only one method candidates and parties can employ to manipulate vote totals. After an election, they are able to employ legal and illegal means to alter the official tally. This manipulation represents a type of sorting, a violation of the assumptions necessary for RDD to identify valid causal effects. In extremely close elections, both parties will file legal complaints, demand recounts, challenge ballots and use their resources to obtain a desired certified vote total. Parties and candidates are able to use more nefarious methods to obtain their desired results. Candidates can stuff ballot boxes, use the votes of citizens long deceased, or commit a variety of other components of fraud that will systematically alter the outcome of the close election. For example, Caro (1990) details how the leading candidate in Texas elections would “hold out” their

fraudulent ballots to ensure that they remain ahead of their opponent (Caro, 1990)[310].

In this section we discuss a simple game that captures this post-election manipulation. We model a sequence of “legal” challenges and show that candidates with a resource advantage are able to systematically claim elections using legal challenges that their opponent would have won in the absence of such challenges.⁹

Suppose that a campaign has occurred and both candidates have observed the vote share x_c . After observing this electoral result, the game proceeds in three stages. In the first stage of the game, the candidate ahead after the campaign (if $x_c > 0.5$, Candidate 1, if $x_c < 0.5$ Candidate 2) makes a decision about how much to invest in post-election manipulation. In the second stage of the game the other campaign decides on how much to invest in their legal challenges. We will denote both campaigns investment by l_1 and l_2 . The final stage of the game is the realization of election results, which we assume are a consequence of the following process,

$$x_l = \eta_1 l_1 - \eta_2 l_2 + x_c$$

where η_1 and η_2 represent Candidate 1 and 2’s institutional capacity to manipulate post-election results, respectively. If $\eta_1 > \eta_2$, a candidate is more effectively able to manipulate election results. After deciding on the amount to invest, payoffs are realized. Crucially, notice that there is no random component in this process, as both parties now know with certainty the number of votes they will need to tilt the election in their favor.

To finish specifying the game, we assume the utility function for the two candidates are given by,

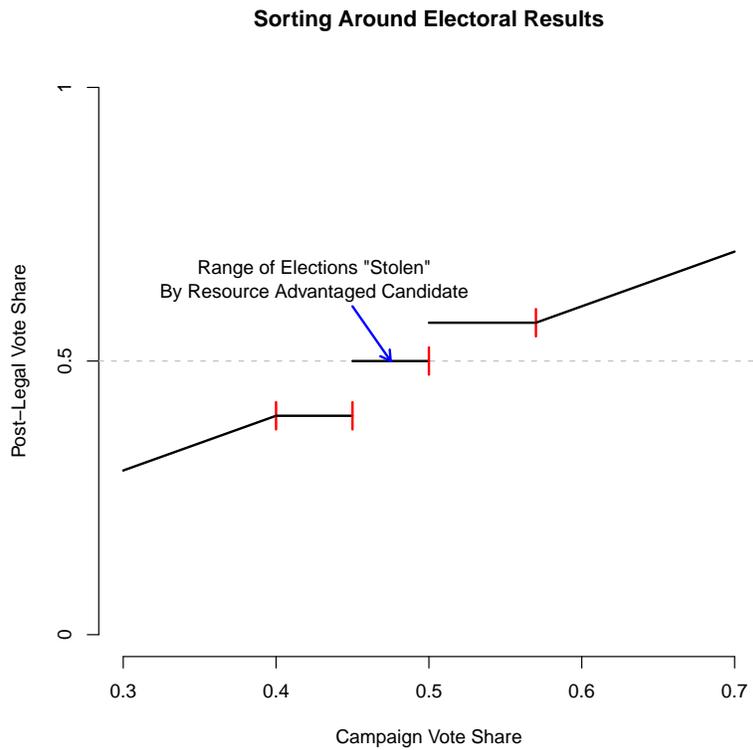
$$U_1(l_1, l_2) = \begin{cases} -k_1 l_1^2 & \text{if } x_l \leq 0.5 \\ 1 - k_1 l_1^2 & \text{if } x_l > 0.5, \end{cases} \quad U_2(l_1, l_2) = \begin{cases} 1 - k_2 l_2^2 & \text{if } x_l \leq 0.5 \\ -k_2 l_2^2 & \text{if } x_l > 0.5, \end{cases}$$

where k_1 and k_2 encode the cost multiplier to both candidates.

Proposition 2 in the Appendix describes a pure-strategy sub-game perfect Nash Equilibrium to this game. It predicts that a candidate with a resource advantage will be able to manipulate

⁹We use legal challenges to avoid appropriating fraudulent motivations or deeds to party officials. But certainly, our model is intended to include both legal and illegal methods of post-election vote manipulation.

Figure 3: Resource Advantages Allow Candidates to Steal Election



This figure presents the equilibrium predictions from the simple post-election manipulation game, predicting that candidates can employ their resource advantages to systematically win extremely close elections.

election results after the fact, ensuring her victory. In this way the candidate is able to “steal” the election: even the public voted for Candidate 2 in the campaign, Candidate 1 emerges victorious through the manipulation. Figure 3 displays this dynamic demonstrating the area of vote stealing. The horizontal axis presents the pre-election vote share, the vertical axis is the vote share after the legal manipulation. The thick line through the plot presents the equilibrium election results, with the vertical red-lines denoting changes in the equilibrium strategy.

Figure 3 shows clearly that the resource advantaged candidate is able to use legal challenges to secure victory in marginal election that originally favored their opponent. This represents *sorting* around the discontinuity, behavior that violates the assumptions necessary for RDD to identify valid causal effects. If candidate’s resource advantages help to determine whether they are able to steal marginal elections and subsequently affects their behavior in office, then the continuity assumption is violated. Specifically, candidates who just happen to win extremely close election will, on average, hold a resource advantage over the candidates that happen to just lose an election. This systematic difference then implies that $\lim_{x \uparrow 0.5} E[Z(1)|X = x] \neq \lim_{x \downarrow 0.5} E[Z(1)|X = x]$ and that $\lim_{x \uparrow 0.5} E[Z(0)|X = x] \neq \lim_{x \downarrow 0.5} E[Z(0)|X = x]$. The result is that RDD estimates are no better than estimates from other observational methods.

4 Empirical Analysis of Close Elections

Our theoretical models predict that there will be systematic differences in resources in very close elections and differences at the discontinuity in close elections if sorting occurs. If the differences in resources are correlated with the dependent variable, this will result in RDD failing to identify $ATE_{1/2}$. In this section we show that there are systematic differences in who wins very close House elections and these systematic differences are indicative of the importance of structural advantages in very close election. Analyzing House elections from 1945-1990, we show that candidates whose party controls the Governor’s mansion or the election administration are systematically more likely to win extremely close elections. This provides evidence that there are systematic determinants to very close elections, violating the practical and theoretical conditions necessary for RDDs to provide valid causal estimates.

Our empirical analysis requires data on election returns and measures of party control, which serve as indicators of a party's structural advantages in a state. For election data, we employ the ICPSR's "Candidate and Constituency Statistics of Elections in the United States, 1788-1990."¹⁰ Our analysis is restricted to elections for the U.S. House of Representatives from 1945 to 1990. The key variable is the *win margin*, defined as difference between the winning candidate and the second-place candidate in the race.

We consider two measures of partisan control. The first is the party of the state Governor during each election cycle. Gubernatorial partisan data was assembled from a database available through the National Governors Association (NGA) website. The second measure is the party of the Secretary of State or statewide election administrator. These data were assembled from several sources. The National Association of Secretary of States (NASS) publishes a list of all past Secretary of States (or equivalent state officers) and the dates they served. The party affiliations of these election administrators were found either on state-sponsored websites or on the "Political Graveyard."¹¹

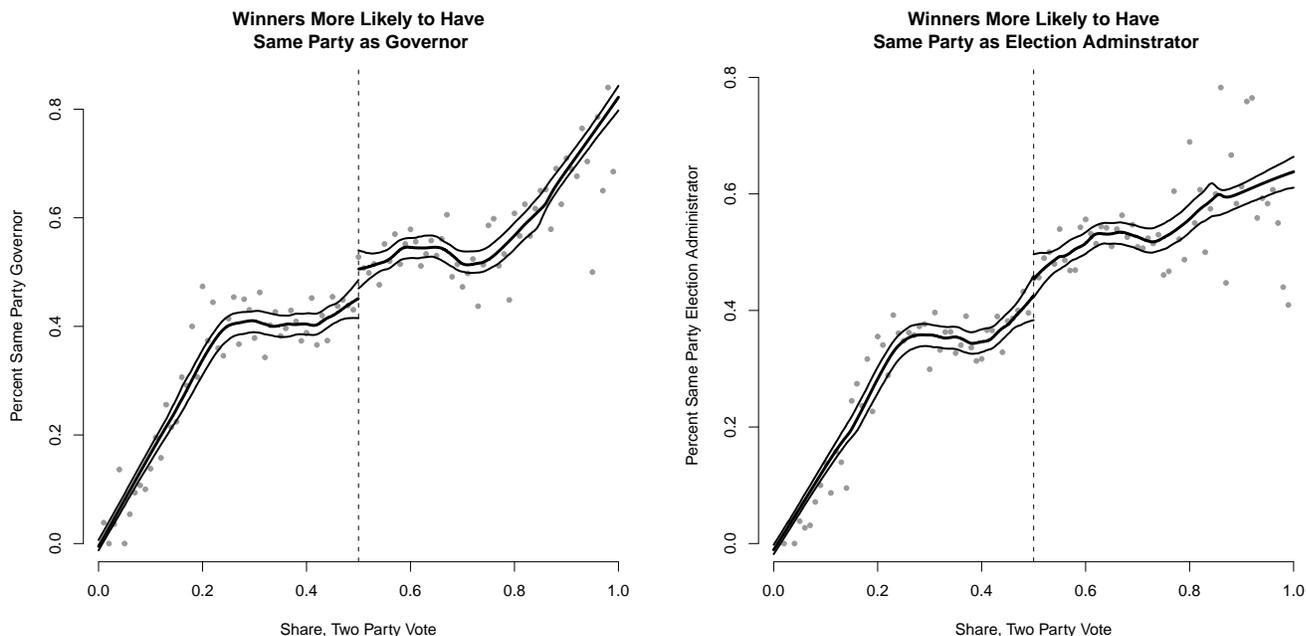
The left-hand plot in Figure 4 demonstrates that winners of extremely close elections also tend to hold structural political advantages. Along the horizontal axis is the share of the two-party vote. As we move towards the center, we observe increasingly close elections, until we reach the dotted-line which represents the discontinuity, or the 50% of votes necessary to win the election. The vertical axis measures the proportion of candidates from the same party as the Governor and the gray-dots create bins of legislators based on their vote share and measure the proportion of candidates with the same party as the Governor within each bin. The black-lines are nonparametric regressions of the proportion of candidates from the same party as their Governor against the two-party vote share. If marginal elections are essentially decided by a coin flip, we would expect the line to the left of 50% and the line to the right to meet exactly at the discontinuity.

But the large gap between the regression lines demonstrates that structural advantages are correlated with who wins extremely close elections. Candidates who barely won the election were

¹⁰Inter-University Consortium for Political and Social Research. Candidate and Constituency Statistics of Elections in the United States, 1788-1990 [Computer file]. ICPSR07757-v5. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1994.

¹¹Lawrence Kestenbaum, "The Political Graveyard, <http://politicalgraveyard.com>.

Figure 4: Gubernatorial and Electoral Administrative Control is Correlated with Winning Close Elections



This figure demonstrates that candidates who win very close elections are systematically more likely to belong to the same party as the Governor and the election administrator (Secretary of State). The large gaps as the discontinuity in agreement between candidates and Governors (left-hand plot) and candidates and electoral administrators (right-hand plot) show that structural political advantages predict who win close elections. This is evidence that pre-election campaigning and post-election legal challenges are inducing differences in marginal elections.

6 percentage points more likely to belong to the same party as the Governor than candidates who barely lost. Belonging to the same party as the election administrator also predicts who wins extremely close elections. The right-hand plot of Figure 4 shows that candidates who win extremely close elections are 5.5 percentage points more likely to belong to the same party than candidates who barely lose. Bootstapped confidence intervals for both differences reveal that we can confidently reject the null that the differences at the discontinuity are zero ($p_i 0.02$).

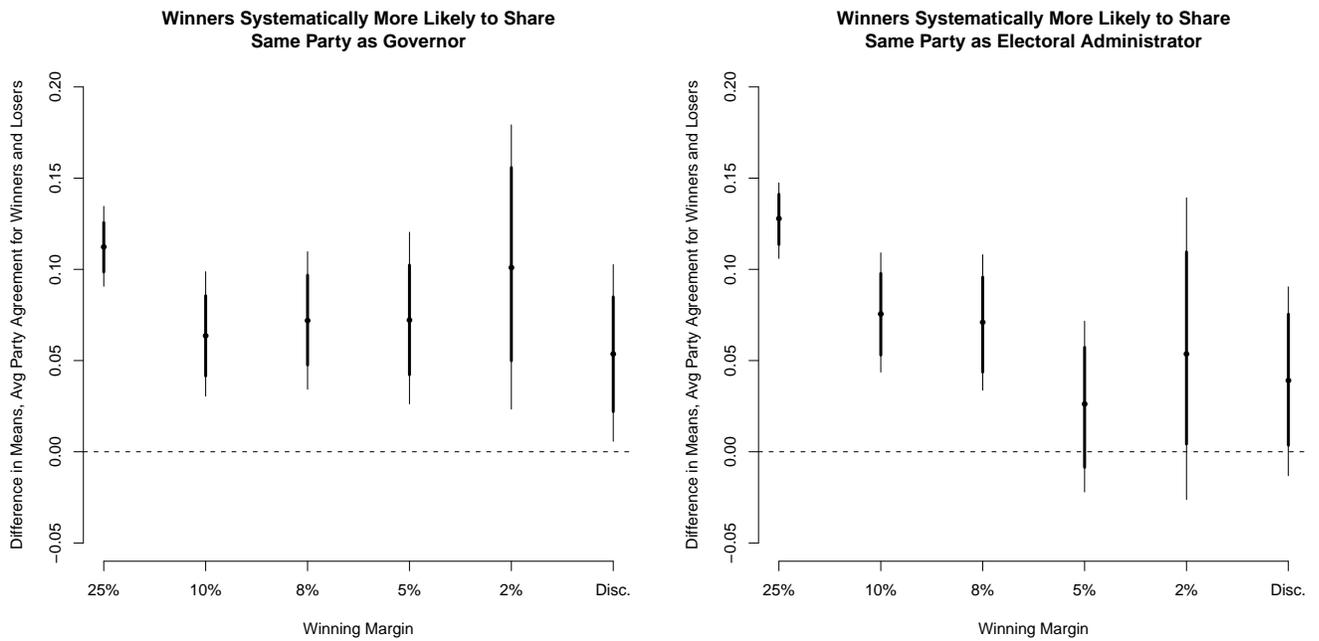
The systematic political differences between candidates who win and lose very close elections are robust to modeling choices and bandwidth selections. The left-hand plot in Figure 5 shows that candidates who belong to the same party as the Governor are systematically more likely to win marginal elections. We compare the proportion of candidates that win close elections in states with a Governor who belongs to the same party to the proportion of candidates who win close

elections in states with a Governor who belongs to a different party. The horizontal axis varies the election margin used to compute this difference. Moving left to right we consider increasingly closer elections—the estimate to the far-left captures elections decided by less than 25 percentage points, whereas the far-right estimate is a prediction of the difference at the discontinuity. For each difference, the thick and thin lines represent 80 and 95 percent confidence intervals for the difference, respectively.

Figure 4 shows that candidates from the same party as the Governor are systematically more likely to win than candidates from a different party. Even in extremely close election—elections decided within by less than 2% points, (51-49 or closer) candidates running in a state where the Governor is from the same party are *10 percentage points more likely to win the election*. The right-hand plot computes the difference for candidates and election administrators, demonstrating that candidates share the same higher probability of winning extremely close elections when they share a party label with the Secretary of State: an almost five percentage point higher probability of winning, although we fail to reject the null of no difference at standard levels of significance. The far right-hand estimates in both plots show that these comparisons within bandwidths align with our observations from Figure 4: systematic political differences distinguish winners and losers in close elections.

Readers might fairly ask why the party of the Governor or election administrator of a state might reflect advantages for a candidate for federal office in that state. We argue that the advantages are manifold, and that it is difficult if not impossible to break out different paths of causation. Some of the effect might have less to do with the officials themselves and more to do with what control of the offices represents: an indicator of underlying latent strength. The fact that a party controls the Governorship is surely a correlate of the strength of its state party organization. In close races, candidates and their party allies will seek to capitalize on superior (lower cost, higher leverage) mobilization, publicity and turnout operations that are available. So too, a Governor may be able to use her or his personal appeal and name recognition to campaign for candidates in close races. As an incumbent with superior access to fundraising (contributors often rationally invest in incumbents), the Governor may be able to redirect funds to a same-party candidate in a

Figure 5: Gubernatorial and Electoral Administrative Control is Correlated at a Wide Range of Bandwidths



This figure shows that candidates with structural advantages are systematically more likely to win close elections. The left-hand plot shows that candidates from the same party as the Governor win elections more often. The vertical axis compares the proportion of candidates who win elections and are from the same party as the Governor to the proportion of candidates who win close elections from a different party than the Governor. Moving left to right we consider increasingly close contests, ranging from 25 percentage point margins (62.5/37.5) at the far-left to elections decided by razor-thin margins at the discontinuity. Regardless of bandwidth, candidates from the Governor's party are more likely to win. The same basic pattern occurs in the right-hand plot, which demonstrates that candidates from the election administrator's party are systematically more likely to win close elections.

close election.

There are also more insidious possibilities. Governors and Secretaries of State may be able to influence election administration, may be able to deploy state officials and even government workers to dampen turnout in opposing constituencies while enhancing turnout in friendly constituencies. Or there may be partisan advantages in the counting and certification of ballots. It was a common joke among Democrats after the Bush-Gore election in 2000 that “it ain’t over until your brother counts the votes.” The presidential election that year was decided by a razor-thin margin in the state of Florida, where the state’s Governor, Jeb Bush, was the brother of Presidential candidate George W. Bush. The advantages of holding the Secretary of State’s office are more straightforward. Secretary’s of State design election rules that might dampen turnout for opponents. After the election, secretaries of state often conduct recounts and decide when to certify elections. Partisan control over these pivotal steps in the election process invites fraud and manipulation of post-election results.

The 2000 Florida dispute highlights most dramatically how these offices are used to advantage candidates. Through each step of the post-election dispute, key movers, like the Governor and Secretary of State, but also the state legislature, state Supreme Court, and U.S. Supreme Court, were all accused of using their authority to advance the chances of their preferred candidate’s success. The very fact that in the United States the election authorities are themselves elected as partisan actors or are appointed by partisan actors means that the closest elections may be anything but randomly decided. The differences we present in this section suggest that there are systematic components to elections that are assumed to be random. Even in elections decided by razor-thin margins, we observe important differences between winners and losers. And this result is extremely robust. Because these differences are likely to be correlated with many outcomes of interest, this represents an important objection to the use of RDDs to estimate causal effects.

5 Conclusion

Close elections may not reflect randomness and may be the precise venue where candidate and partisan advantages in organization most show themselves. In the single non-transferable vote

systems employed in the United States, where the plurality winner takes the full value of the seat, it is the marginal elections that will consume the focus of candidate and party resources during the campaign and even after the votes have been cast. Our analysis suggests that further research is in order before analysts confidently use close elections as quasi-experiments.

Our empirical analyses demonstrate that close elections are in fact systematically determined, in large measure because these contests attract disproportionate investment by the candidates. Our theoretical models provide two explanations for this result. First, RDD analysts rely upon a theoretical result about the causal assignment mechanism at the discontinuity point (one-half of the votes cast), but in practice they must choose a bandwidth which includes races with larger margins. These margins will include races where one candidate has systematic advantages over another candidate, and there are many reasons to believe that these advantages are correlated with downstream variables like later earnings, later voting patterns and policy outcomes. Second, elections do not end when people are finished casting their ballots. Ballots have to be counted and certified, and election results must be declared legal and legitimate. Candidates can also deploy advantages at this post-voting stage, breaking the continuity of the regression function at the 50 percent threshold.

Our theoretical arguments and empirical evidence imply that more care is needed before natural experiments are used to identify causal effects. Across applications, we believe that assumptions made to identify causal effects using a natural experiment carry with them important political assumptions. Therefore, to identify and exploit natural experiments, it is incumbent upon the researcher to understand and examine the underlying political process that determines the assignment mechanism. This includes more than quantitative demonstrations of balance in covariates across treatment and control groups. It may also require qualitative, ethnographic or historical study to show that how observations are assigned to treatment and control are not systematically related to the outcome of interest. To that end, we believe that future work should examine not only the practice of RDDs, but also the equally vital question of who wins close elections. Depending on the margin, what are predictors of winning these elections? How often do legal challenges occur and with what conditional probability of success? How can the various causal pathways of candidate

advantage be disentangled from one another?

The evidence we present also calls into question whether elections decided by razor-thin margins are truly marginal. Marginal elections, where either party has a chance to win, represent an important source of voter influence on national government (Mayhew, 1974). Certainly the marginals are vanishing, a large literature argues, but at least some “toss-up” seats remain. But our analyses present a bleaker portrait of party competition for seats. Even the closest elections are determined, at least in part, by systematic structural advantages of one party. This blunts the effectiveness of close elections as a tool to translate voter preferences into national government.

A Proofs Appendix

Proposition 1. *A pure strategy symmetric Nash-Equilibrium exists to the this game, with Candidate 1's equilibrium investment strategy given by*

$$\begin{aligned}
 c_1 = & \frac{1}{2(\gamma_1 - \gamma_2)^2} (\gamma_1 - \gamma_2 - 2\sigma_0^2 - 2\gamma_1 x_0 + 2\gamma_2 x_0 + 2(\gamma_1 - \gamma_2)\gamma_2 \log[\frac{k_1}{\gamma_1 t}] \\
 & + 2\gamma_1 \gamma_2 \log[\frac{k_1}{\gamma_1 t}] + 2\gamma_1 \gamma_2 \log[\frac{\gamma_2 t}{k_2}] - 2\gamma_2^2 \log[\frac{\gamma_2 t}{k_2}] \\
 & + 2(\sigma_2(-\gamma_1 + \gamma_2 + \sigma_0^2 + 2\gamma_1 x_0 - 2\gamma_2 x_0 - 2\gamma_1(\gamma_1 - \gamma_2) \log[\frac{k_1}{\gamma_1 t}] + 2\gamma_2(\gamma_2 - \gamma_1) \log[\frac{\gamma_2 t}{k_2}]))
 \end{aligned}$$

where $t = \frac{1}{2\pi\sigma_0^2}$.

Proof. We begin by calculating Candidate 1's probability of Winning.

$$\begin{aligned}
 \text{Prob}(x_1 \geq 0.5) &= \text{Prob}(\gamma_1 c_1 - \gamma_2 c_2 + w \geq 0.5) \\
 &= \text{Prob}(-\gamma_1 c_1 + \gamma_2 c_2 + 0.5 \leq w) \\
 &= \int_{-\gamma_1 c_1 + \gamma_2 c_2 + 0.5}^{\infty} f(w|x_0, \sigma_0^2) dw
 \end{aligned}$$

where $f(\cdot|\cdot)$ represents the normal density function. Of course, $\text{Prob}(x_2 \geq 0.5) = 1 - \text{Prob}(x_1 \geq 0.5)$.

To find the symmetric pure strategy Nash, we'll solve the first order conditions for both candidates, which sets up the following equations

$$\begin{aligned}
 \gamma_1 t \exp \left[- \left(\frac{(-\gamma_1 c_1 + \gamma_2 c_2 + 0.5 - x_0)^2}{2\sigma_0^2} \right) \right] &= k_1 \exp[c_1] \\
 \gamma_2 t \exp \left[- \left(\frac{(-\gamma_1 c_1 + \gamma_2 c_2 + 0.5 - x_0)^2}{2\sigma_0^2} \right) \right] &= k_2 \exp[c_2]
 \end{aligned}$$

And all that remains is to solve the simultaneous equations. □

Proposition 2. *Without loss of generality assume that $\eta_1/\sqrt{k_1} > \eta_2/\sqrt{k_2}$. A pure strategy sub-*

game perfect Nash-Equilibrium to the game is characterized by the following investments,

$$\begin{aligned}
l_1 = 0, l_2 = 0 & \quad \text{if} \quad x_c \leq 0.5 - \frac{\eta_1}{\sqrt{k_1}} \\
l_1 = 0, l_2 = x_c - \left(0.5 - \frac{\eta_1}{\sqrt{k_1}}\right) & \quad \text{if} \quad 0.5 - \frac{\eta_1}{\sqrt{k_1}} < x_c \leq 0.5 + \frac{\eta_2}{\sqrt{k_2}} - \frac{\eta_1}{\sqrt{k_1}} \\
l_1 = 0.5 - x_c, l_2 = 0 & \quad \text{if} \quad 0.5 + \frac{\eta_2}{\sqrt{k_2}} - \frac{\eta_1}{\sqrt{k_1}} < x_c \leq 0.5 \\
l_1 = 0.5 + \frac{\eta_2}{\sqrt{k_2}} - x_c, l_2 = 0 & \quad \text{if} \quad 0.5 < x_c \leq 0.5 + \frac{\eta_2}{\sqrt{k_2}} \\
l_1 = 0, l_2 = 0 & \quad \text{if} \quad 0.5 + \frac{\eta_2}{\sqrt{k_2}} < x_c
\end{aligned}$$

Proof. Candidates can always guarantee 0 by not investing. Therefore, the maximum possible investment amount is $l_1 = \frac{1}{\sqrt{k_1}}$, with a total possible effect of $\eta_1/\sqrt{k_1}$.

The challenger's subgame perfect strategy is to invest if a win is possible and to not invest if the legal challenges are too expensive. Therefore, in elections where the election is close, but x_c favors the low resource candidate, $0.5 - \frac{\eta_1}{\sqrt{k_1}} < x_c \leq 0.5 + \frac{\eta_2}{\sqrt{k_2}} - \frac{\eta_1}{\sqrt{k_1}}$ the winning candidate must invest to prevent the higher resource candidate from stealing the election.

Consider x_c such that $0.5 + \frac{\eta_2}{\sqrt{k_2}} - \frac{\eta_1}{\sqrt{k_1}} < x_c \leq 0.5 + \frac{\eta_2}{\sqrt{k_2}}$ and first suppose that the resource advantaged candidate is behind after the campaign, $x_c < 0.5$. In this case, the sub-game perfect response for the advantaged candidate is to invest enough to steal the election, $0.5 - x_c + l_2$. Since the resource advantaged candidate is able to outspend the first moving candidate, her equilibrium response is to not invest. If $x_c > 0.5$ then the resource advantaged candidate just needs to invest enough to deter the credible threat from her opponent.

This describes the complete sub-game perfect Nash equilibrium.

□

References

- Abramowitz, Alan I., Brad Alexander and Matthew Gunning. 2006. "Incumbency, Redistricting, and the Decline of Competition in U.S. House Elections." *Journal of Politics* 68(1):75–88.
- Ansola-behere, Stephen, James M. Snyder Jr. and Charles Stewart III. 2000. "Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage." *American Political Science Review* 44(1):17–34.
- Bensel, Richard. 2004. *The American Ballot Box in the Mid-Nineteenth Century*. Cambridge Univ Press.
- Campagna, Janet and Bernard Grofman. 1990. "Party Control and Partisan Bias in 1980s Congressional Redistricting." *Journal of Politics* 52(4):1242–1257.
- Caro, Robert. 1990. *The Years of Lyndon Johnson: Means of Ascent*. Vintage Press.
- Desposato, Scott W. and John R. Petrocik. 2003. "The Variable Incumbency Advantage: New Voters, Redistricting, and the Personal Vote." *American Journal of Political Science* 47(1):18–32.
- Eggers, Andy and Jens Hainmueller. 2009. "MPs for Sale? Returns to Office in Postwar British Politics." *American Political Science Review* 103(04):513–533.
- Erikson, Robert. 1971. "The Advantage of Incumbency in Congressional Elections." *Polity* 3(3):395–405.
- Ferejohn, John A. 1977. "On the Decline of Competition in Congressional Elections." *American Political Science Review* 71(1):166–176.
- Fiorina, Morris. 1977. *Congress: Keystone of the Washington Establishment*. Yale University Press.
- Gasiorowski, MJ. 1996. "An Overview of the Political Regime Change Dataset." *Comparative Political Studies* 29(4):469–483.

- Gelman, Andrew and Gary King. 1990. "Estimating Incumbency Advantage Without Bias." *American Journal of Political Science* 34(4):1142–1164.
- Gopoian, J. David and Darrell M. West. 1984. "Trading Secrets for Seats: Strategic Considerations in the Redistricting Process." *Journal of Politics* 46(4):1080–1096.
- Hahn, Jinyong, Petra Todd and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with Regression Discontinuity Design." *Econometrica* 69:201–209.
- Hainmueller, Jens. 2010. "Synthetic Matching for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." MIT Mimeo.
- Ho, Dan, Kosuke Imai, Gary King and Elizabeth Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* .
- Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396):945–960.
- Imbens, Guido and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2):615–635.
- Lee, David. 2008. "Randomized Experiments from Non-Random Selection in US House Elections." *Journal of Econometrics* 142(2):675–697.
- Lee, David, Enrico Moretti and Matthew Butler. 2004. "Do Voters Affect or Elect Policies? Evidence from the US House." *Quarterly Journal of Economics* 119(3):807–859.
- Mayhew, David R. 1974. "Congressional Elections: The Case of the Vanishing Marginals." *Polity* 6(3):295–317.
- McCormick, John. 2001. "Machiavellian Democracy: Controlling Elites with Ferocious Populism." *American Political Science Review* 95(2):297–313.
- Morgan, Stephen and Chris Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press.

- Przeworski, A, M Alvarez, J Cheibub and F Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. Cambridge University Press.
- Rosenbaum, Paul and Donald Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41–55.
- Sekhon, Jasjeet. 2010. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R." *Journal of Statistical Software* . Forthcoming.
- Snyder, J and P Querubin. 2008. "The Rents to Political Office in the US, 1840-1870." Massachusetts Institute of Technology, Mimeo.
- Snyder, Jason. 2005. "Detecting Manipulation in U.S. House Elections." Unpublished Manuscript.
- Tufte, Edward R. 1973. "The Relationship Between Seats and Votes in Two-Party Systems." *American Political Science Review* 67(2):540–554.
- Vanhanen, Tatu. 2000. "A New Dataset for Measuring Democracy 1810-1998." *Journal of Peace Research* 37(2):251–265.