Using Regression Discontinuity to Uncover the Personal Incumbency Advantage*

Robert S. Erikson†
Professor of Political Science
Columbia University

Rocío Titiunik‡
Assistant Professor of Political Science
University of Michigan

July 5th, 2013

* We thank John Bullock, Matias Cattaneo, Sanford Gordon, Justin Grimmer, Shigeo Hirano, Gregory Huber, Gary King, Kelly Rader, Eric Schickler, Jasjeet Sekhon, Robert Van Houweling, Jonathan Wand, seminar participants at Harvard University, New York University, Princeton University, Stanford University, the University of California at Berkeley, and Yale University, and conference participants at the 2011 MIT American Politics Conference for helpful comments and suggestions.

† Corresponding author. rse14@columbia.edu, http://www.columbia.edu/~rse14, Political Science Department, Columbia University, 420 W. 118th St, New York, NY 10027.

‡ titiunik@umich.edu, http://www.umich.edu/~titiunik, Center for Political Studies, Institute for Social Research, University of Michigan, P.O. Box 1248, Ann Arbor, MI 48106-1248.

Abstract

We study the conditions under which estimating the incumbency advantage using a regression discontinuity (RD) design recovers the personal incumbency advantage in a two-party system. Expanding on the interpretation proposed by Lee (2008), who introduced RD as a method for estimating the party incumbency advantage, we develop a simple model that leads to unbiased estimates of the personal incumbency advantage and yields the surprising result that the RD effect double counts the personal incumbency advantage. Obtaining this result is possible under certain conditions, which we argue hold in the U.S. House when the analysis is restricted to open seats. We apply our model to analyze non-southern open-seat U.S. House elections between 1968 and 2008, where we estimate a personal incumbency advantage of about 7 percentage points. We also explore the estimation of the incumbency advantage beyond the limited RD conditions where knife-edge electoral shifts create the leverage for causal inference.

1 Introduction

The incumbency advantage, the added votes that a candidate receives due to his or her incumbency status, is one of the most studied topics in American politics. Throughout fifty years, scholars have proposed different methods to obtain a valid measure of this quantity. This has proven difficult, as there are several challenges. A first consideration is that candidates who win elections tend to be systematically different from those who do not—more experienced, better funded, more knowledgeable about public policy or more adept at public speaking. In other words, a skillful politician is both more likely to become an incumbent and continuously obtain high vote shares. Second, elections are usually contested by two major party candidates, and each can affect the vote margin. Incumbents may gain or lose votes due to the quality of their opponents. Third, incumbents are strategic in deciding when to retire, and tend to quit whenever they are expected to do badly. All these complications present challenges to finding an unbiased measure of the incumbency advantage.

The literature has dealt with these challenges in different ways. The "sophomore surge" estimates the within-individual electoral gain obtained by newly elected candidates in their reelection as freshmen. The "retirement slump" estimates the party's vote loss when its incumbent retires (Erikson 1971, 1972; Cover and Mayhew, 1977).[1] The "slurge" is the average of the surge and the slump (Alford and Brady 1989). The Gelman and King (1990) method

[1] Cover and Mayhew can claim credit for dubbing the terms "sophomore surge and retirement slump."

regresses the vote on the lagged vote and appropriate dummies for the incumbent's party and whether the incumbent runs.[2] As discussed below, these methods require strong assumptions.

A recent and popular strategy to estimate the incumbency advantage was developed by Lee (2008), who proposed to use close elections to estimate the incumbency advantage with a regression discontinuity (RD) design. The idea behind the RD design is simple: in very close elections, the winning candidate may be determined essentially by chance. When this happens, for example, U.S. House districts where the Democratic candidate barely wins are virtually identical to districts where the Republican candidate barely wins, approximating random assignment of the election outcome, and making it theoretically straightforward to estimate the effect of the party of the winner—incumbency—upon the partisan vote in the next election.

Political scientists have been quick to add the RD design to the toolbox of incumbency advantage methods, but have not given much attention to how the quantity that is estimated by this strategy is related to the concept of the incumbency advantage as traditionally defined. In the literature, the incumbency advantage is sometimes conceptualized as a party advantage but more often as a personal advantage to the incumbent candidate. The personal incumbency advantage is the votes gained by a candidate once she becomes an incumbent--from constituency service, name recognition, and the like. This is the advantage that accrues between the first race as a non-incumbent and the second running as a freshman congressmember (Erikson 1971). The party

[2] Further variations include estimating the sophomore surge by restricting the analysis to races where the same candidates run in multiple elections (Levitt and Wolfram 1997), and exploiting natural experiments such as redistricting (Ansolabehere, Snyder, and Stewart 2000) or unexpected deaths (Cox and Katz 2002).

advantage is usually conceptualized as the advantage to a party from having an incumbent run (Gelman and King 1990). This incorporates the personal advantage just discussed and the quality advantage that incumbents hold from the simple fact that, everything else being equal, candidates with more personal appeal tend not only to win elections but to win repeatedly (see Levitt and Wolfram 1997 and Zaller 1998, on the distinction.) If candidate quality did not matter, the personal and the party advantages would be identical. Still a third source is the scare-off advantage (Cox and Katz 1996), whereby strong incumbents add to their vote by scaring off potentially strong opponents. The scare-off effect (which we abbreviate simply as "the scareoff") can be decomposed into the direct scareoff arising from the incumbent's incumbency status itself and the indirect scareoff arising from the incumbent's higher-than-average quality. Incumbency-induced scareoff is an additional component of the personal incumbency advantage.

The typical RD estimate of the incumbency advantage in the U.S. House (or any other two-party setting) compares the Democratic party's vote share at election $t+1$ in districts were the Democratic party barely lost at t to the Democratic party's vote share at $t+1$ in districts were the party barely won at t.[3] This implementation of RD purposely abandons the focus on the individual incumbent's advantage pervasive in the incumbency advantage literature, and focuses instead in the overall advantage to a party from being the incumbent party, *regardless of whether the winning candidate already is the incumbent or wins an open seat and regardless of whether*

[3] Since U.S. House elections are contested by only two parties (approximately), defining the RD estimand in terms of the Democratic Party rather than in terms of the Republican Party is done without loss of generality. This is true of any two-party setting; in multi-party systems, however, estimates for different parties will not be mirror images.

the winner runs again in the next election. Because of these features, the RD estimand has been commonly interpreted as a measure of the *party* incumbency advantage.

Our paper offers an expanded interpretation of the RD design whereby one can back out unbiased estimates of the desired personal incumbency advantage. We present a simple model where a party's vote share in a district is determined by common year trends, the district's overall partisanship, the quality of candidates, and the personal incumbency advantage. Following the political science tradition, in our model the personal incumbency advantage in a given district stems solely from the seat holder’s vote gain from being an incumbent We embed this model in an RD design, and define the estimand of interest accordingly, as the difference in the party's vote share between districts where the party barely won and barely lost the previous election.

We show that, under our model, the RD design allows the estimation of the personal incumbency advantage without bias. This is possible under certain conditions, which we argue hold in the U.S. House for open seats rather than those contested by incumbents. The main result of our model is that in open seats the RD design double-counts the personal incumbency advantage or, in other words, the RD effect is twice the personal incumbency advantage. This occurs because each of the two groups that are compared in this restricted RD design, districts where a party barely won the open seat election and those where it barely lost, has a freshman incumbent candidate running for reelection. Since these freshman incumbents are, by construction, from different parties, the incumbency advantage increases the party's vote share in the barely-winner group and decreases the party's vote share in the barely-loser group, leading to a double-counting phenomenon. Moreover, although our model focuses on open seats where our assumptions are most plausible, in the Supplemental Online Appendix we show that this phenomenon affects all seats. Indeed, we believe this double-counting feature of the RD design, which has not been noticed before, is one of the main reasons why empirical RD estimates of the incumbency advantage are generally much larger than those obtained with other methods.

The paper is organized as follows. Section 2 discusses the RD estimate of the incumbency advantage, and the most commonly used measures of incumbency advantage in the political science literature. Section 3 presents our model and discusses assumptions under which the personal incumbency advantage can be obtained from RD estimates. Section 4 applies our model to estimate the personal incumbency advantage in non-southern U.S. House elections in the period 1968-2008. Section 5 explores the applicability of our model to incumbent-held seats. Section 6 compares the estimates from our model to other common incumbency advantage estimates, and Section 7 discusses whether our local RD-based estimates provide information about the global incumbency advantage. Section 8 concludes. Additional analyses are provided in the Supplemental Online Appendix, including a generalization of our model, a recast of our model in terms of potential outcomes, and validity checks of the RD assumptions in our sample.

2 Using RD to estimate the incumbency advantage

Since Lee (2008) first introduced the idea of using the RD design to estimate the incumbency advantage, scholars have adopted the design to study the electoral advantage of incumbents both in the U.S. and elsewhere (Broockman 2009; Butler 2009; Caughey and Sekhon 2011; Golden and Picci 2012; Grimmer, Hersh, Feinstein and Carpenter 2011; Hainmueller and Kern 2008; Trounstine 2011; Uppal 2009; Uppal 2010). The original setup assumes a two-party system, with the analysis usually focusing on the Democratic share of the two-party vote—the "score" or "running variable" that determines whether a district's Democratic candidate wins a given election, which we assume occurs at time t. When the Democratic vote share exceeds 50%, the Democratic candidate becomes an incumbent; and when this vote share falls below

50%, the Republican candidate becomes the incumbent. This discontinuity at the 50% cutoff, under appropriate smoothness assumptions (see Lee 2008; Imbens and Lemieux 2008), can be used to identify the incumbency advantage at this point. Informally, these assumptions require that at the 50% threshold winner and loser districts have continuous “pre-treatment” (and possibly unobserved) characteristics. The intuitive[4] interpretation is that districts where the Democratic party barely loses the election at t, say obtain a vote share of 49.9%, are on average similar to districts where Democratic candidates barely win the election, say obtain a vote share of 50.1%, in terms of important characteristics such as partisanship, candidate quality, and previous vote shares. Thus, near the 50% cutoff, winning or losing at t can be regarded *as if* randomly assigned, and we can compare the vote share of barely winner districts and barely loser districts in the following election, which we assume occurs at $t+1$, to recover the incumbency advantage *at the 50% cutoff*. The parameter of interest is the local average treatment effect of a Democratic win at the cutoff—measured as the difference between the average vote shares in districts where the Democratic party barely won the previous election and districts where the Democratic party barely lost the previous election.

This basic setup, used by all previous scholars who have estimated the incumbency advantage with an RD design, is our starting point. This framework emphasizes a *party level* analysis: the treatment group is defined as those districts where the Democratic party barely won the election at t, and the control group as those districts where the Democratic party barely lost

[4] Despite the fact that this is a common and intuitive interpretation of the RD estimand, stronger assumptions than continuity are required to ensure that such an as-if random interpretation is valid. We ignore this distinction here, but see Cattaneo, Frandsen and Titiunik (2013) for details.

the election at t. This leads naturally to the definition of the RD estimand that we stated above: the average vote share at $t+1$ in treatment districts minus the average vote share at $t+1$ in control districts.

This RD estimand is markedly different from most predecessors in the incumbency advantage literature. Sophomore surge and retirement slump focused, respectively, on the electoral gain (adjusted for year effects) enjoyed by freshmen U.S. Congress members in their first reelection attempt and the electoral loss (adjusted for year effects) suffered by a party once their veteran incumbents retires. The focus of these measures was on the *personal* incumbency advantage, as they attempted to measure the added vote increment enjoyed by an individual candidate for being an elected incumbent. Gelman and King (1990) offered another, frequently employed, estimate that models the vote as a function of candidate incumbency status, controlling for the lagged vote and the party holding the seat (plus year effects).

As estimates of the personal incumbency advantage, each of these traditional measures have obvious biases. The sophomore slump represents both the gain from incumbency and the change in opponent quality from one election to the next. The retirement slump and Gelman King measure both represent the combination of the lost personal incumbency advantage but also the loss of the incumbent's (presumably positive) personal quality. Moreover, all these measures—particular the latter set—suffer from potential selection bias, since incumbents are more likely to retire when electoral circumstances are adverse.

Lee's RD design defines the incumbency advantage differently, as the "the overall causal impact of being the current incumbent party in a district on the votes obtained in the district's election" (Lee 2008, p.682), or, in other words, as the electoral gain to the incumbent party in a district relative to the vote share that the party would have obtained if it had not won the previous election and consequently had not been the incumbent. Once the incumbency advantage is defined as the subsequent electoral advantage to a party from winning rather than losing an election, RD arguably becomes the best available research design to estimate this parameter. But this is different from the personal incumbency advantage of elected candidates, which is the parameter we are interested in estimating. Lee's model, with its emphasis on a party-level analysis, *explicitly ignores whether the winning incumbent candidate at t runs again at $t+1$*, thus sidestepping the consequences of selective attrition, that is, the fact that candidates elected at t may decide to retire at $t+1$. Lee's formulation also ignores the distinction between incumbent races and open seats at election t. As will see, these two sets of circumstances provide different expectations for the vote shift at the 50-50 break point.

Congressional elections come in two types: open seats, with no incumbent in the race, and incumbent races, where the incumbent candidate seeks reelection. Incumbent races can be further subdivided into seats contested by Democratic and Republican incumbents. Figure 1 displays separate graphs for open seats, seats where the Democratic incumbent seeks reelection, and seats where the Republican seeks reelection. In each panel, one can observe an apparent incumbency advantage to the winning party of about 15 points. It is also evident that incumbent races where the incumbent obtains in the vicinity of 50 % of the vote are fairly rare events.

[Figure 1 Goes Here]

Below we offer assumptions under which a simple model of incumbency advantage plus a RD design allow us to recover an unbiased estimate of the personal incumbency advantage. The plausibility of these assumptions, we argue, are plausible for open seats. Under these assumptions, we show that the vote gain from t to t+1 represents twice the gain from a personal incumbency advantage.

3 Recovering the personal incumbency advantage from open seats

Consider a set of open seats at time *t*, defined as those with no incumbent running. As in the classic RD design, our unit of observation is the congressional district, which we index by i, and V_{it} and V_{it+1} are the Democratic share of the two-party U.S. House vote at elections *t* and $t+1$, respectively, in district i. The Republican share of the vote is $1-V_{it+1}$ so it suffices to focus on V_{it+1}.

We model V_{it+1} as

$$V_{it+1} = \quad Par_{it+1} + \theta I_{it+1} + (D_{it+1} - R_{it+1}) + e_{it+1} \tag{1}$$

where we introduce the concept of *Par*, a measure of the baseline vote for the Democratic party in district i, given the district's partisanship, the election year's partisan trend, no incumbent candidate, and Democratic and Republican candidates of average quality. In equation 1, Par_{it+1} is the district's *Par* at $t+1$; D_{it+1} and R_{it+1} are the *added* quality of the Democratic and Republican candidates, respectively, running at $t+1$ in district i above the quality of the average open seat candidate in their respective parties. Candidate quality is normed to be conditional on *Par*. For any value of *Par*, each party's expected candidate quality (conditional on it being an open seat) is zero. Thus, $D_{it+1}=0$ and $R_{it+1}=0$ when both candidates are of average quality (conditional on an open seat and *Par*) and $(D_{it+1}-R_{it+1})$ is the "quality differential" between the Democratic and Republican candidates. Finally, I_{it+1} equals 1 if the newly elected time *t* winner is a Democrat who seeks reelection at $t+1$, -1 if the newly elected time *t* winner is a Republican seeking reelection at *t*+1, and 0 if there is no incumbent running at $t+1$; e_{it+1} is a residual.

We only include cases where the time t winner seeks reelection at $t+1$ -- implying that I_{it+1} is never zero. By excluding districts that are open seats at $t+1$, we observe (in the language of experimental design) only "compliers" with the intended treatment (running as an incumbent). Including only compliers can produce bias from selective mortality, a possibility here if anticipation of the $t+1$ outcome affects retirement decisions. However, as we show in the Appendix, in the U.S. House almost all freshman incumbents seek reelection and the ones that do retire tend to do so for reasons unrelated to their expected vote share.

Our model also assumes that the RD design is valid and its continuity assumptions hold or, in other words, that the outcome of very close races at *t* is decided as-if randomly by factors that are not systematically related to future electoral success. The validity of these RD assumptions for U.S. House elections has recently become a matter of contention. Caughey and Sekhon (2011) show that for races in the period 1942-2008 decided by 0.5 percentage points or less, the incumbent party wins far more seats than it loses, a discrepancy greater than can be attributed to chance. Inferring that incumbent parties have the capability to sort into the winning column when the election is known to be extremely close, Caughey and Sekhon cast doubt on using RD for studying the incumbency advantage in U.S. House elections. Eggers et al. (2013), however, argue that Caughey and Sekhon's findings likely arose by chance, since they are specific to the sample analyzed and do not hold for other time periods. Regardless of the resolution of this debate, we show in the next section and in the Supplemental Online Appendix that there is no evidence of sorting in the sample of U.S. House races that we analyze here, which includes only non-southern open seats between 1968 and 2008.

Henceforth, we use V_{it+1}^{w} and V_{it+1}^{l} to denote, informally, the **average** vote share in Democrats' barely-winner and barely-loser districts, respectively,[5] a notation we also use to denote analogous averages for the other variables in our model. Taking averages for barely-winner and barely-loser districts in Equation (2) yields

$$V_{t+1}^{w} = Par_{t+1}^{w} + \theta \cdot I_{t+1}^{w} + \left(D_{t+1}^{w} - R_{t+1}^{w}\right) \tag{3}$$

$$V_{t+1}^{l} = Par_{t+1}^{l} + \theta \cdot I_{t+1}^{l} + \left(D_{t+1}^{l} - R_{t+1}^{l}\right),$$

where for open-seat districts that at t have very close elections, we treat incumbents' decisions to retire at $t+1$ as non-strategic (conditional on *Par*), so that $e_{t+1}^{w} = e_{t+1}^{l} = 0$.[6]

Combining the expressions in equation (3), the RD estimand becomes

$$\begin{aligned}\tau^{RD} \quad &= V_{t+1}^{w} - V_{t+1}^{l} \\ &= \left(Par_{t+1}^{w} - Par_{t+1}^{l}\right) + \theta \cdot \left(I_{t+1}^{w} - I_{t+1}^{l}\right) + \left(D_{t+1}^{w} - R_{t+1}^{w}\right) - \left(D_{t+1}^{l} - R_{t+1}^{l}\right)\end{aligned} \tag{4}$$

We are interested in learning about θ but, as shown in equation (4), this parameter is not immediately available: τ^{RD} recovers the personal incumbency advantage plus the difference in

---

[5] More formally, for any variable y, the w and l subscripts indicate, respectively, the right and left limits of $E\left(y \mid V_{it} = v\right)$ when v approaches 50 from above and below, i.e. $y^{w} \equiv \lim_{v \to \frac{1}{2}^{+}} E(y \mid V_{it} = v)$ and $y^{l} \equiv \lim_{v \to \frac{1}{2}^{-}} E(y \mid V_{it} = v)$. We adopt the more intuitive interpretation of averages for simplicity of exposition.

[6] Strictly speaking, we do not need mean-zero residuals, just $e_{t+1}^{w} = e_{t+1}^{l}$. In other words, the RD allows us to relax the assumption of non-strategic retirement to an assumption of continuity of retirement decisions. See Section 3 in the Supplemental Appendix for details.

Par ($Par^w_{t+1} - Par^l_{t+1}$) and the candidate quality differentials $(D^w_{t+1} - R^w_{t+1}) - (D^l_{t+1} - R^l_{t+1})$. Note that, given our assumption of non-retirement at $t+1$, $I^w_{t+1} = 1$ and $I^l_{t+1} = -1$.

The difference in district Par does not pose a major obstacle. By virtue of the RD assumptions, barely-winner and barely-loser districts' Par at t are on average equal to each other, $Par^w_t = Par^l_t$. Any average change in Par between t and $t+1$ affects barely winner and barely loser districts similarly, so that the equality still holds in the following election, $Par^w_{t+1} = Par^l_{t+1}$. Under this simplifying assumption, it follows that

$$\tau^{RD} = \theta \cdot \left(I^w_{t+1} - I^l_{t+1} \right) + QD^w_{t+1} + QD^l_{t+1} \qquad (5)$$

where we have defined the *quality differential* terms $QD^w_{t+1} \equiv (D^w_{t+1} - R^w_{t+1})$ and $QD^l_{t+1} \equiv (R^l_{t+1} - D^l_{t+1})$.

Equation (5) shows the RD estimand as the sum of three terms. The first term is the direct personal incumbency advantage, θ, multiplied by $(I^w_{t+1} - I^l_{t+1}) = 2$. The second term, QD^w_{t+1}, is the average difference in quality between the Democratic and Republican candidates at $t+1$ in districts where the Democrat barely wins. The third term, QD^l_{t+1}, is the average difference in quality between the Republican and Democratic candidates at $t+1$ in districts where the Democrat barely loses.

To study the quality differential terms, we first consider what happens at election t. The as-if randomness of close races in the RD design guarantees that, at t, Democratic candidates in barely-winner districts are of equal average quality to Democratic candidates in barely-loser districts ($D^w_t = D^l_t$), and Republican candidates in barely-winner districts are of equal average quality than Republican candidates in barely-loser districts ($R^w_t = R^l_t$). Moreover, although generally in any given race the winner candidate will tend to be of higher quality than the loser candidate, we assume that in very close open seat races winner and loser will tend to be of the same average quality. Since we define quality relative to *Par*, this implies that in open seat races near the 50% cutoff, $D_{it}^{w} = R_{it}^{w} = D_{it}^{l} = R_{it}^{l} = 0$.

Next we turn to $t+1$, when the open seat winners become first-term incumbents. In order to characterize the quality differential in this election, we assume that the open seat winners at time t maintain their original quality from time t, so that $D_{t+1}^{w} = D_{t}^{w} = 0$ and $R_{t+1}^{l} = R_{t}^{l} = 0$.[7] This leaves only D_{t+1}^{l} and R_{t+1}^{w} from equation (5), which we assume are negative under a scare off argument. At time $t+1$, the barely-losing parties (at time t) must select new candidates from their candidate pools. But instead of average quality candidates as at time t, they may be forced to pick worse candidates due to a scare off effect: potential high-quality challengers, knowing that incumbents have access to resources they can exploit for electoral advantage, may be discouraged from entering the race (Cox and Katz 1996). In other words, as the freshman incumbents gain their new incumbency advantage, they may adversely affect the quality of their opponents. This strategic behavior results in $t+1$ challengers of lower average quality than the typical open seat candidate, thus leading to $D_{t+1}^{l} < 0$ and $R_{t+1}^{w} < 0$ -- recall that D_{t+1}^{l} (R_{t+1}^{w})

---

[7] The assumption that the $t+1$ incumbent's quality at $t+1$ will be the quality at time t is plausible as an expectation. Suppose instead the model were to incorporate some regression to the mean for incumbent quality. That model could be restated so that candidates revert to their personal mean, which then becomes our quality variable. New shocks at t+1 would enter the error term. The only difference would be that with regression to the mean there would be some autocorrelation of the error term.

represents the quality of the Democratic (Republican) challenger in districts where there is a Republican (Democratic)_incumbent at $t+1$.

Importantly, the scareoff may itself be decomposed into two sources. First, challengers may be deterred from entering the race simply because their opponent is an incumbent who has access to perquisites of office which are believed to translate into an electoral advantage. We call this the *incumbency scareoff*, as it arises directly from the incumbency status of the incumbent candidate. Second, challengers may also be deterred from entering the race when the incumbent is of very high quality, which decreases the chances of a challenger's victory even in the absence of a direct incumbency advantage. We call this the *quality scareoff*. Importantly, in competitive open seat races, our assumptions guarantee that the quality scareoff is zero, since the winner and the loser are assumed to be of equal (and average) quality. In these races, the only component of the scareoff is the incumbency scareoff. Hereafter, we use the symbol σ to refer to the *total scareoff*, that is, to the sum of the incumbency and quality scareoff. In close open seats, the quality scareoff component of σ will be zero, but in incumbent-held seats both terms will be nonzero.

The combination of these conditions (restriction to open seats at t, definition of quality relative to *Par*, winner and loser of average quality, scare off effect) leads to $QD^w_{t+1} = (D^w_{t+1} - R^w_{t+1}) = -R^w_{t+1} > 0$ and $QD^l_{t+1} = (R^l_{t+1} - D^l_{t+1}) = -D^l_{t+1} > 0$, that is, to a quality differential that is positive and equal to the incumbency scare off effect in both barely-winning and barely-losing districts. For simplicity, we assume that the (incumbency) scareoff is equal for both parties: $-D^l_{t+1} = -R^w_{t+1} \equiv \sigma$.

Under these conditions, the RD estimand in equation (5) simplifies to

$$\tau^{RD} \qquad = 2\theta + (-R^{w}_{t+1} - D^{l}_{t+1})$$

$$= 2\theta + 2\sigma\,,$$

which shows that *the usual RD estimate is twice the personal incumbency advantage*. Thus, we can recover the personal incumbency advantage from the RD design simply by dividing by two:

$$\frac{\tau^{RD}}{2} = \theta + \sigma \tag{6}$$

Why does this double-counting phenomenon arise? In *barely-winner* districts, the Democratic vote share *increases* by the personal incumbency advantage because the Democratic candidate is an incumbent --who has both direct access to the benefits of office and the ability to scare off strong Republican opponents. Conversely, in *barely-loser* districts, the Democratic vote share *decreases* by the personal incumbency because the Democratic candidate is running against a *Republican* incumbent --who is the one with access to the benefits of office and the ability to deter strong Democratic challengers from entering the race. Thus, for the subset of districts where incumbents are running for reelection at $t+1$, when we subtract the $t+1$ Democratic party's vote share in barely-loser districts from the Democratic vote share in barely-winner districts, we count the personal incumbency twice.

To review, we claim that for open seats, the effect of the time *t* regression discontinuity at the 50-50 breakpoint on the subsequent *t*+1 vote is twice the personal incumbency advantage. Under the assumption that winners of open seats retire rarely and only for non-strategic reasons, it is safe to ignore cases where the freshman open seat winner does not seek reelection. We have argued that at the 50-50 discontinuity, on average the time *t* Republican and Democratic open seat candidates are of equal quality, allowing the time *t*+1 differential between Democratic winning cases and Democratic losing cases to represent twice the personal incumbency advantage (direct plus scareoff) free of consideration of time *t* candidate quality. Conveniently, the canceling out (on average) of candidate quality not only helps us to estimate the personal incumbency advantage without bias, it also is an estimate of the local party effect at 50-50. With (at the limit) no differential in Republican and Democratic quality when comparing winners and losers of very close open seats, the personal incumbency advantage is also the party incumbency advantage.

4 RD Estimates of the Personal Incumbency Advantage in the U.S. House

We now apply our theoretical findings to first-year incumbents following their open seat victories. Our data source is the CQ Voting and Elections Collection for the 1968-2008 period, excluding South and Border states -- below we refer to our sample of states as "non-southern" states for the sake of brevity. We start in 1968 because of the well-documented growth of the incumbency advantage in the 1960s. We exclude South and Border state districts because—at least early in the period—it is difficult to estimate their partisan loyalties from presidential voting, a matter that takes on importance in the analysis below.[8]

The unit of observation is a congressional district, the running variable or score is the Democratic margin of victory at t and the outcome variable is the Democratic share of the total vote at $t+1$. The incumbency treatment variable, I_{t+1}, takes on values -1 if the time t winner (and thus freshman incumbent) is a Republican and +1 if a Democrat. There are 399 cases. Excluded are the 14 instances when an open seat winner did not contest the next election. Also

[8] The states excluded are Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia and West Virginia.

excluded are instances where redistricting occurs between the open-seat election and the freshman election.

For part of our results, we measure the vote as the deviation from *Par*. Recall that *Par* represents the expected open-seat House vote given district partisanship and year effects. We measure *Par* as follows. First, for non-southern open seat districts, we regress the House vote on the current or most recent presidential vote in the district, plus year dummies. The predictions from this equation comprise *Par.* For open seats, deviations from *Par* represent the net effects of the Democratic and Republican candidates' quality, plus random error.

The results are shown several ways. Table 1 shows results of parametric OLS estimation of the RD effect where the t+1 vote is predicted as a function of the vote at t and the time t winner discontinuity. Table 2 shows the results of non-parametric local linear regression with triangular kernel and bandwidth chosen optimally following Imbens and Kalyanaraman (2012). Table 3 offers linear regression results within 2 percentage points of the 50-50 cutoff for the residual vote measure.

Parametric analysis

Table 1 displays the OLS equations. The first column' equation is the bare-bones specification. The dependent variable is the Democratic vote at $t+1$. The independent variables are the Democratic vote at t and the incumbency treatment. There are no controls. The estimated incumbency advantage is 6.80 percentage points. The second column's equation adds year dummy variables, which add precision as seen by the reduction in the RMSE. Now the estimated incumbency effect rises almost a point beyond the column 1 estimate, with a slightly tighter standard error. The third column adds the presidential vote as a further control, giving the tightest prediction of all and a slightly higher estimated incumbency advantage of 7.85.

[Table 1 Goes Here]

Column 4 shows the result when the dependent variable is reframed as the residual from *Par*. This is of interest because the running variable (time t vote) now is virtually unrelated to the dependent variable. The incumbency estimate is 8.00. In the fifth column the dependent variable is the residual surge or the change in the residual vote from t to $t+1$. Here, the running variable has a slightly negative coefficient, in contrast with the positive coefficient for the discontinuity, in this case 7.70. Finally, the last column shows the results of a global polynomial that regresses the vote share at t+1 on a fourth-order polynomial of the running variable and the same polynomial interacted with the treatment, a common parametric approach to estimate the effect at the discontinuity (see Lee 2008). The coefficient of 7.29 percentage points reported is the difference in the global polynomial approximation at each side of the cutoff, roughly similar to those reported in the previous columns.

These alternative specifications yield somewhat different estimates of the incumbency advantage, but all in the range from just under 7 points to 8 points and all are highly significant statistically, as would be expected. This estimate is of the personal incumbency advantage—the extra edge an incumbent gains once becoming an incumbent. As discussed above, if one is interested in the outcome of a close open-seat election in terms of the differential between a Republican and a Democratic victory, the effect is twice the coefficient—at 14 to 16 points.

Non-parametric analysis

Table 2 turns to the local linear regression estimates. Informally, this estimation method involves choosing a neighborhood around the 50 percent cutoff, and estimating two weighted regressions of the outcome on a constant and the running variable, one for barely-winner districts and another for barely-loser districts, where the weights are a function of the distance between the observation's value of the running variable and the 50 percent threshold. The RD effect is obtained by subtracting the estimated constant in the barely-loser regression from the estimated constant in the barely-winner regression. For very small neighborhoods around the discontinuity cutoff, this local linear estimation approach is similar to performing a simple difference in means between barely-winner and barely-loser districts.[9]

[Table 2 Goes Here]

As shown in column (1) of Table 2, the standard RD effect in our non-southern open-seat sample is 15.12%: when the Democratic party barely wins election t, its vote share in election $t+1$ is about 15 percentage points higher than it would have been if it had barely lost. The vote share in districts where the Democratic party barely lost at t is about 40.55%, which means that winning a very close race increases the Democratic party's average vote share to about 55.67%. In order to recover the personal incumbency effect from this estimate, we divide this number by two. Our estimate of the personal incumbency advantage is therefore

[9] We also estimated the effects using a difference in means within a 5% margin of victory, a difference in means within a 2% margin of victory, and also 4th order global polynomial that uses the entire data as shown in Table 1. All methods yield similar results to those reported in Table 2 and leave the overall conclusions of our analysis unchanged.

7.56%, with a 95% confidence interval ranging from 4.98 to 10.14. This is very similar to the parametric point estimates reported above. We report conventional standard errors to follow common practice, although a mean-squared-error optimal bandwidth is technically too large for these standard errors to be valid. Using the robust confidence intervals developed by Calonico, Cattaneo and Titiunik (2013) yields very similar results (not reported).

Residual Vote in the 48-52 Band

Next, we use the residual vote measure of the vote at t and $t+1$. Table 3 presents the details, where we consider all results for open seat winners where the time t vote is within two points of 50-50. Whereas the time t average vote margin for the time t winner within this band is 51.0 percent (not shown), of greater importance is the time t residual vote (relative to *Par*), which on average is very close to zero. This means that on average, the quality differential between the winning and loosing candidate in the close election (which approximates a coin flip) is close to what would be expected for an open seat, given the district's partisanship and year effects (*Par*). It follows that since the quality advantage is nil and the outcome is a coin flip, the winners of open seats in the 50-50 range are of average quality, given the competitive district.

[Table 3 Goes Here]

At $t+1$, however, our open seat winners run for reelection. Now as incumbents, their average vote zooms over 7 percent. Measured as the residual vote shift from t to $t+1$, we see a sophomore surge of 7.10 points.[10] Whereas in general the sophomore surge may be biased downward because the initial victory is partially due to good luck and a worse than average opponent, this bias is absent when the contest is close. Moreover, the problem of strategic retirement vanishes completely in very competitive open seat races: in the period we study, only one open seat winner within the 48-52% vote window retired at election $t+1$.

A Cautionary Pause: Checking the RD Assumptions

As mentioned above, Caughey and Sekhon (2011) sounded a warning about the validity of the RD design when applied to U.S. House elections in the period 1942-2008, arising from a greater frequency of barely-winning incumbent party cases than barely-losing incumbent party cases. The window where they find the asymmetry is quite small, within a lead of less than 0.5 percentage points (i.e., a window of 49.75-50.25 in two-party elections). Do the Caughey-Sekhon findings contaminate our RD estimates for non-southern open seat candidates in the period 1968-2008? We do not believe so, for several reasons.

First, even if it is plausible that incumbent candidates can boost their vote at the margin given a tight race, it is implausible that the retiring incumbent's party holds the same power and the same will to protect the successor candidate in an open seat. In our sample of non-southern 1968-2008 open seats, the incumbent party holds no special edge in races decided within the narrow band of 49.75-50.25 of the two-party vote. In this window, the incumbent party wins

---

[10] If the running variable (time *t* vote) is included as a regressor in the regression equation shown in Table 3, this variable has a negative but non-significant impact and yields a 9.45 point incumbency advantage estimate.

barely more often than the challenger party (8 vs. 6), a mild asymmetry that is quite consistent with a 50% probability of treatment assignment.[11]

In our analysis of this section, the treatment is a [close] Democratic victory and the control is a [close] Republican victory in the open seat race. Thus, we also investigated whether there exists observable covariance imbalances between tight open seat races won by Democrats and those won by Republicans. In our sample of non-southern open seats in 1968-2008, we find no evidence of statistically significant covariate imbalance between the treatment and control group in races decided by 1.0 percentage point or less. In Section 4 of the Supplemental Online Appendix, we analyze several pre-treatment covariates for races in this range, including previous Democratic vote and campaign spending. We report both balance tests (using simple *t*-tests as well as randomization inference tests to account for the small sample size) and placebo tests in which we estimate the RD effect using a local linear regression with optimal bandwidth. In all cases, there are no statistically significant imbalances in pre-treatment covariates.

Finally, suppose that all our objections are wrong, so that within the tiny band where the expected outcome is between 49.75 and 50.25 percent of the two-party vote, incumbent parties can sometimes exert an extra force to push the successor candidate of the incumbent's party over the top. Would the contamination of the RD design be severe? Suppose that among observed

[11]The probability of observing 8 or more successes out of 14 trials of a Bernoulli experiment with hypothesized probability of success equal to 1/2 is 0.395. Like Caughey-Sekhon, we also conducted a Fisher's exact test of significance where the two binary variables are party of the incumbent and time *t* winner. Restricted to our sample, the results are not significant for any bandwidth within 49-51 percent of the two-party vote.

incumbent-party winners within this window, about half owed their victories to special powers that only the local congressional incumbent party possesses. At t+1, the freshman incumbent is not likely to be in the narrow circumstance where they exert this special force once again. The only violation is a possible distortion of the running variable by a fraction of a percentage point.

Caughey and Sekhon provide a valuable note of caution about using RD blindly, without examining the applicability of the usual RD assumptions and the consequences for their failure. Setting aside the broader implications of their findings, we argue that their paper does not pose a threat to our particular analysis of the incumbency advantage, estimated for open seat winners in non-southern races in the period 1968-2008.

Summary

With several variations, we have estimated the personal incumbency advantage for close open seats to be about 7 percentage points. Double this value and one estimates the downstream value to a party from wining or losing a very close open-seat election is roughly 14 percentage points. We have argued that these estimates are biased neither upward or downward. This claim of unbiasedness does not extend to close elections when one of the candidate is the incumbent seeking reelection. Still, incumbent seats provide further clues to understanding the incumbency advantage. We turn to incumbent-contested seats next.

5 Exploring Incumbent-Contested Seats

Inference for incumbent-held seats poses a more serious challenge than for open seats. When discussing open seats with close races, it was safe to assume that winning candidates' performance at $t+1$ is due to the direct and indirect effects of incumbency. This is because the determination of the time t winner is a virtual coin flip between two candidates who in expectation are of average quality, given Par. For incumbent races, we can make no such assumption. We have no reason to expect either incumbents or challengers to be average candidates in close races. This leaves no reason to suspect that the winner's quality at $t+1$ will be average. The math for open seats no longer applies, as the incumbent-challenger quality differential can be either positive or negative. See Section 2 in the Supplemental Online Appendix for a derivation of this general case where the sign of the quality differentials is no longer known.

Consider the case where the incumbent is a Democrat in a tight reelection contest in year t who either survives to seek further reelection at t+1 or is defeated by a successful Republican challenger. The RD equation then is

$$\begin{aligned}\tau^{RD} &= (D_{t+1}^{w}+\theta)(1+\rho)+(R_{t+1}^{l}+\theta)(1+\rho)\\ &= 2(1+\rho)(D_{t+1}^{w}+R_{t+1}^{l}+\theta)\end{aligned}$$

where D_{t+1} and R_{t+1} refer to the expected quality of the winning incumbent and winning challenger, respectively. The difference from the open seat case is that for incumbent races, $D_{t+1}^{w}\neq 0$ and $R_{t+1}^{l}\neq 0$. Embattled incumbents at the cusp of losing tend to be below average candidates. Their challengers who take them to the 50-50 level might be better than average candidates, except that they may be scared off by their opponent's incumbency. A separate problem is strategic retirements of incumbents who barely survive reelection. Retirees tend to be the least popular incumbents.[12]

---

[12] Further, the Caughey-Sekhon objections arise once again since incumbents do win a disproportionate high share when thrust into a very close race. However, observable covariate imbalances between incumbent races closely won by Democrats and those closely won by

Clearly, regression discontinuity is not a safe method for estimating the direct incumbency advantage (θ) or the combined direct and indirect (scareoff) incumbency advantage ($\theta+\sigma$). Yet RD can be used to achieve a more modest goal. Armed with our estimate of $\theta+\sigma$ from our open seat analysis, we can perform a sensitivity analysis to estimate the plausible range of the average quality of incumbents and challengers in close incumbent races. More importantly, we can estimate the plausible range of σ, the indirect scareoff effect, which allows us to back out an estimate of θ, the direct incumbency effect.

Table 4 shows basic data for close incumbent races for U.S. House seats where the incumbent's time t vote is within the 48-52 percent band. The data are shown from the perspective of the time t incumbent candidate, whether Republican or Democrat. That is, the residual vote is measured as the percent for the time t incumbent party, relative to *Par*. Comparing with the results in Table 3, where we present a similar analysis for open seats, we observe that in their subsequent race at $t+1$, the winners of incumbent-held seats perform more poorly than their open seat counterparts. Surviving incumbents win only 5.14 percentage points more than *Par* and successful challengers only 6.34 more than *Par*, while the analogous gain in open seats is 7.10 points. These comparisons suggest that both candidates in close incumbent races tend to be of worse than the average quality.

[Table 4 Goes Here]

Republicans are due to mixing races with a Democrat incumbent with those with a Republican incumbent. When the vote is measured as the incumbent share of the two-party vote, as we do in this section, this control for incumbency eliminates these covariance imbalances. See Eggers et al. (2013) and Erikson and Rader (2013).

Table 4 contains only three numbers relevant for estimating the dynamics of close incumbent races: the time $t+1$ residual vote for surviving time t incumbents seeking reelection (+5.14), the $t+1$ residual vote given a time t incumbent loss (-6.34), and the time t residual vote. We have two estimates for the time t residual vote, one for incumbent winners and one for incumbent losers. In theory these should be identical, but sampling error generates separate estimates. We use the weighted (by sample size) mean, which equals +1.45. This estimate includes all tight time *t* races, including those by incumbents who retire at *t*+1. Twenty-two percent of the surviving incumbents from close races quit rather than run at *t*+1 and their average residual vote was a full 2.00 points lower than the 78 percent who sought reelection. We take this differential into account in our discussion below.[13]

We divide our analysis in three steps. First, we can account for the $t+1$ residual vote for surviving incumbents as follows:

$$+5.14 = \theta_I + \sigma_I + Q_I \tag{7}$$

where θ_I is the incumbency advantage specific to incumbents in close races who seek reelection, σ_I is the total scareoff specific to incumbents after winning a close race, and Q_I is the vote increment due to the incumbent's inherent quality.

---

[13] Selective retirement is a problem only for veteran incumbent survivors, not for successful challengers. All 62 successful challenger from our sample of close contests sought reelection at *t*+1. Incumbent retirements appear specifically a function of their own poor quality and not anticipating adverse partisan swings or the *t*+1 challenger's quality. Retirees do not face worse outcomes at t+1 in terms of *Par.*

Second, we can analogously account for the $t+1$ vote for the time t incumbent party when it falls to a successful challenger:

$$-6.34 = -\theta_C - \sigma_C - Q_C \tag{8}$$

where θ_C is the incumbency advantage specific to successful challengers in close races, σ_C is the total scareoff specific to successful challengers in their first reelection attempt, and Q_C is a quality vote term analogous to Q_I.

Third, we can account for the time t residual vote for the incumbent party of close races

$$+1.45 = \theta_I + Q_I - Q_C + \xi \tag{9}$$

where ξ is "luck" at time t. The luck factor is from momentary influences on the election not due to partisan and year effects or apart from the two candidates' personal votes as fixed effects. For incumbents caught in close races, the mean error term is not necessarily zero.

This makes seven unknowns for three equations. However, we can perform some sensitivity analysis. We make some plausible assumptions about some of the unknown terms and observe the consequences for other parameters. First, we assume that the direct incumbency advantage θ is identical for surviving incumbents and successful challengers who win close races, and that it equals the personal incumbency advantage for close winners of open seats. From the results in Table 3, we know this quantity to be

$$\theta_O = 7.10 - \sigma_O$$

where the subscript o stands for open seats (remember the quality terms are zero by assumption in open seats). In other words, we assume $\theta_I = \theta_C = \theta_O$.

As discussed in the previous section, for freshman winners of close *open seats*, the mean total scareoff is solely the challenger party's response to the freshman's new incumbency advantage. The mean scareoff is unaffected by the mean residual vote at time t because it is near zero for close open seats. For incumbent races, however, the mean total scareoff should include a response to the personal quality of the incumbent (quality scareoff) in addition to the mean incumbency advantage (incumbency scareoff).

Thus, we can conceptualize the scareoff as a proportion of the combined direct personal incumbency advantage plus vote due to candidate quality. We call this proportion ρ, as above. The residual vote of surviving incumbents (equation (7) above) is now

$$+5.14 = (1+\rho)(\theta + Q_I) \qquad (7a)$$

leading to $Q_I = \frac{-1.96}{1+\rho}$.[14] This estimate, however represents the quality only for incumbent "compliers" who seek reelection. Since the time *t* residual vote for all successful incumbents is 0.45 points less than the residual vote for compliers alone, we recalculate equation (7a) by subtracting 0.45 from the left-hand side.

$$+4.59 = (1+\rho)(\theta + Q_I) \qquad (7b)$$

[14] We get to this number after plugging in the expression $\theta(1+\rho) = 7.10$, which follows from the equation for θ_O and the assumption $\theta_I = \theta_C = \theta_O$.

Taking into account retirees, estimated time t incumbent quality in close contests is even more negative, $Q_I^* = \frac{-2.41}{1+\rho}$ where Q_I^* signifies the average quality of all incumbents, not just reelection seekers.

The $t+1$ residual vote when the time t incumbent loses is now

$$-6.34 = -(1+\rho)(\theta + Q_C) \qquad \text{(8a)}$$

leading to $Q_C = \frac{-0.76}{1+\rho}$. Note that either winner of a close incumbent contest is below the average quality if this were an open seat race. And the time t residual vote is

$$+1.45 = \theta + Q_I^* - Q_C + \xi \qquad \text{(9a)}$$

Summing (7b) and (8a), we get $-1.75 = (1+\rho)(Q_I^* - Q_C)$, thus $(Q_I^* - Q_C) = \frac{-1.75}{1+\rho}$. Substituting into (9a),

$$\begin{aligned} +1.45 \quad &= \theta + Q_I^* - Q_C + \xi \qquad \text{(9b)} \\ &= \frac{7.10 - 1.75}{1+\rho} + \xi \end{aligned}$$

Although this equation is still under identified, we have simplified to one equation and two unknowns. Clearly, some combination of scareoff (positive ρ) and bad luck (negative ξ) is necessary for equation (9b) to add up. Incumbents with close races lose their edge because their poor quality diminishes the usual quality scareoff advantage and also because of bad luck.

We can gain some clarity by drawing on our three equations to model the average gain of surviving incumbents in Table 4:

$$+3.58 \qquad = \rho(\theta + Q_I) + Q_C - \xi \qquad (10)$$

$$= \frac{5.14\rho - 0.76}{(1+\rho)} - \xi$$

Equation (10) has two components.[15] One is the mean incumbent's gain from replacing a $\frac{0.76}{(1+\rho)}$ total scareoff with an average scareoff, given an incumbent quality of $\frac{-1.96}{(1+\rho)}$. The second is the return to normal luck (zero) following the time t close call partially due to bad luck (ξ). We cannot separate these two effects. But we can set some plausible ceilings and floors for their values.

First, we assume that from time t to time $t+1$, the surviving incumbent gains from an increase in scareoff. This reasoning sets the floor for the $t+1$ scareoff at $\frac{0.76}{(1+\rho)}$, the vote the incumbent gains at time t from the challenger's quality. This lower bound is

[15] We obtain this equation by subtracting the residual vote for surviving incumbents at $t+1$ (equation 7a) from the residual vote of incumbents at t (equation 9a): $(1+\rho)(\theta + Q_I) - \{\theta + Q_I - Q_C + \xi\}$.

$\rho_{LB} = \frac{0.76}{5.14} = 0.15$.[16] Given this lower bound, the return to normal luck fully accounts for the incumbent gain represented by equation (10). Given the lower bound, we can estimate the upper bound of the direct personal incumbency advantage in close races, θ.

$$\theta_{UB} = \frac{7.10}{1+\rho_{LB}} = \frac{7.10}{1.15} = 6.17$$

In this extreme instance, the average gain for surviving incumbents is entirely due to the disappearance of their time t bad luck.

What is the upper bound for ρ? For incumbents to be threatened with a close contest, luck must either be neutral or bad. So we set a ceiling for luck at zero. Doing the math, it turns out that for luck to equal zero, ρ would need to be an implausibly high 2.78 as if almost all what appears to be the incumbency advantage is actually scareoff. We conclude that luck must indeed have been a contributor to these close contests being close.

So ρ could be anywhere between 0.15 and 2.78. A plausible guess is to set $\rho = 0.25$ so that mean total scareoff for open seat winners equals one fourth the size of θ, the direct personal incumbency advantage. If so, doing the math, we obtain

---

[16] We get to this value by setting the scareoff equal to its assumed floor, $\rho(\theta + Q_I) = \frac{0.76}{(1+\rho)}$, which leads to $(1+\rho)(\theta + Q_I) = \frac{0.76}{\rho}$, and then using equation (7a) to replace the left-hand side with 5.14.

$$\theta = 5.68$$
$$\sigma_I = 1.03$$
$$\sigma_C = 1.27$$
$$\xi = -3.27$$
$$Q_I = -1.57$$
$$Q_I^* = -1.93$$
$$Q_C = -0.61$$

These estimates are imprecise since they depend on specific assumptions about the relative degrees that a return to normal scareoff and normal luck at time $t+1$ account for the t to $t+1$ incumbent gains after winning a close election. Still, we can now ask, what have we learned from this exercise? First, when an incumbent is in a close reelection battle, the winner (incumbent or challenger) is of lower quality than the winner would be if the seat were open. We see this because at time t+1, neither surviving incumbents nor successful challengers lead *Par* (based on district partisanship and year effects) by as much as would the typical winner of a closely contested open seat.

Second, we see compelling evidence of a scareoff effect. Recall that successful challengers are worse candidates than open seat winners. This is so even though when incumbents are threatened, a stronger-than-usual challenger is a likely contributing factor. If stronger than usual challengers are worse than open seat winners of close contests, their weakness must be due to scareoff. The average scareoff under normal circumstances (e.g., $t+1$) is likely to be larger. We set a plausible estimate of 0.25 of the combined quality-related vote (Q_I) plus incumbency advantage (θ) as the total scareoff.

Third, bad luck must be partially responsible for incumbents ensnared in close contests. By luck, we mean short-term local forces that fade by the time of the next election. We see this because incumbent survivors of close elections bounce back (controlling for year effects) in their next election. This bounce is too large to plausibly be accounted for by increased scareoff alone (or, as argued above, by increased luck alone).

In sum, close incumbent races provide a window on candidate quality. Given reasonable assumptions and the pattern in the data both embattled incumbents and their challengers have negative quality, which means that they are less attractive than the average candidate would be if the seat were open. For an incumbent to become threatened with a loss, it helps to be unattractive, as a partial offset of their incumbency advantage. Their challengers are scared off some by their incumbency but encouraged by their incumbent opponent's poor quality.

6 Comparison with Other Estimates

Based on our regression discontinuity analysis for open seat winners, we estimated a combined direct and indirect incumbency advantage of about 7 percentage points. We now study how this figure compares with other estimates based on the same general sample—non-southern districts, 1968-2008. Table 5 shows the results. For sophomore surge and retirement slump, the calculations are averages based on the residual vote. For the Gelman-King method, we use the regression model from Gelman and King (1990).

For the full sample (not simply close races), the mean sophomore surge is a "mere" 5.60, smaller (5.15) for open seat winners than for successful challengers (6.19). This is expected. In general (unlike at the 50-50 threshold), open-seat winners at t owe their victories in part to their losing opponents' below average quality and also good short-term luck; these factors dissipate at $t+1$, so that apart from their newly earned incumbency advantage, open seat winners should do worse at $t+1$. Successful challengers gain more than open-seat winners in their sophomore race, but this is because they overcome their opponent's incumbency advantage, which somewhat offsets their generally poor quality as incumbent losers.

[Table 5 Goes Here]

The retirement slump is almost 8 percentage points, larger than our RD estimate. This also is as expected. Incumbents generally tend to be of positive quality, so that their margin relative to *Par* exceeds their incumbency advantage. They tend to retire when their popularity is at low ebb, but still in positive territory. The result is a slight inflation of the incumbency advantage.

The Gelman-King method (applied to our full sample of non-southern 1968-2008 districts) yields the largest estimate of all, at +8.13. The source of this inflation is subtle. Gelman and King (1990) regress the vote on the lagged vote, year dummies, candidate incumbency, and the incumbent party. The lagged vote is intended to control for sources of the $t+1$ vote other than incumbency. However, this assumption leads to an unbiased estimate only if there is no continuity of candidate effects (personal vote, candidate quality) from one election to the next. This assumption is decidedly untrue, so that the lagged vote is a leaky indicator of the relevant non-incumbency causes of the time $t+1$ vote that it is intended to measure. The consequence is that non-incumbency factors masquerade as part of the incumbency advantage.

We have held up our RD estimate of about an incumbency advantage of about 7 percentage points as the gold standard by which the results of other estimation strategies should be judged. At the same time, we must recognize that our RD estimate, while unbiased, is local—conditional on winning a very close contest for an open-seat. Rival methods, while biased, are intended to be global—measuring the average advantage for incumbents generally. To this distinction we turn next.

7 The Global Incumbency Advantage

Although we can be confident of our estimate of the net incumbency advantage being about seven points for our "laboratory subjects" who win open seats in a tight race, we can be less sure about how far this result generalizes. Do winners of open but safe seats achieve the same gain? Here we veer away from the precise estimation strategy of RD. Although we can measure the incumbency advantage from alternative methods, these methods are biased. We can, however, adjust the sophomore surge for its theoretical bias and rerun the results for all open seat winners.

For open seats in general, the sophomore surge method is biased downward because it fails to adjust for the fact that initial victories are partially due to the draw at t of a poor opponent plus luck factors. As opponent quality and luck return to normal (zero) as an expectation, the freshman incumbent's t+1 vote should decline from its previous base, apart from incumbency.

We can adjust by making some theoretically plausible assumptions. The departure of the vote margin from *Par* at t is equally a function of loser and winner quality. So winner quality should be half of that portion not due to temporary luck. At $t+1$, however, the impact of winner quality at t can become magnified by the scare off factor. Putting the parts together, a justifiable assumption is that the winner's quality contributes about half of the variance of the time t residual vote. Thus, we can adjust the sophomore surge by subtracting only half of the time t residual vote from the residual vote at $t+1$. With this adjusted estimate, the adjusted sophomore surge averaged for all open-seat winners is 6.92 percentage points, close to our RD estimate.[17]

Thus, our test suggests the incumbency advantage at the 50-50 threshold holds approximately for open seat winners generally. However, using the adjusted sophomore surge, the size of the incumbency gain does vary with electoral conditions. The factor that stimulates incumbents to exploit their advantage for the $t+1$ election can best be measured by *Par* at $t+1$. Figure 2 shows how the adjusted surge varies with *Par.* The estimated incumbency-induced gain is greatest for incumbents who are threatened by an adverse *Par* that favors the opposition party. It declines toward zero when *Par* implies the incumbent enjoys a very safe seat. We also see that the when *Par* suggests a competitive race, the variance of the surge is extremely large compared to that for safe seats. Both the greater incumbency advantage and the greater variance are consistent with the idea that where winning or losing is at stake, parties and their candidates exert extra effort.

[Figure 2 Goes Here]

Figure 2 highlights races that were close at time t, where the winner's vote share was under 52 percent. The figure makes clear that close open seat races occur across the range of district partisanship, except for extreme one-party districts. At each level of incumbent-party partisanship, the adjusted vote gains for races that were close at time t are in the middle of the

[17] To review: the unadjusted mean sophomore surge is +5.19. Adjusted by subtracting only half of the lagged residual vote yields +6.92. Adjusting 100 percent (as if there is no candidate history, the implicit Gelman-King assumption), the estimate is +8.52, the average residual vote at t+1 for time t open seat winners.

pack. Thus, close races are not particularly unique in terms of the size of the incumbency advantage, even though we see understandable variability in the incumbency advantage as a function of partisan marginality.

8 Conclusion

The main virtue of the RD design as Lee (2008) applied it to congressional incumbency was to shift the focus from the incumbent candidate to the incumbent party, avoiding the need to make specific assumptions about how the strategic entry and exit of candidates into the race affects the incumbency advantage estimates. The RD incumbency effect, focusing on the "overall advantage to the party", provided an estimate of the incumbency advantage that entirely sidestepped the methodological difficulties that had been at the center of the incumbency advantage literature for decades. The methodological virtues of the design, however, came at the price of a vague conceptualization of the incumbency advantage, an issue that has gone largely unnoticed despite the growing popularity of the RD design among incumbency advantage scholars. Our paper focused on this issue, and studied the conditions under which the incumbency advantage as traditionally understood by the political science literature can be recovered from an RD design.

Our paper shows how the RD design can be used to identify the personal incumbency advantage--the specific advantage that incumbents obtain as incumbents that they did not have in their initial, non-incumbent race. We show that the necessary ingredients are (a) that in election near the 50-50 vote threshold at time t, winners and losers are of equal quality and (b) that there is little or no strategic retirement affecting reelection decisions at $t+1$. Given these conditions, the RD effect double counts the incumbency advantage, summing the personal vote gains from winners in each party. As we show, in the U.S. House the proper conditions are met for closely contested open seats in non-southern races in the 1968-2008 period. For such districts, we estimate the gain from incumbency to be about seven percentage points.

This estimated seven point gain includes both the average direct effect of incumbency on the candidate's personal vote and the scareoff of higher quality challengers due to their incumbency. We exploit RD for seats with incumbent candidates to search for evidence of scareoff. Leveraging from some plausible assumptions and a sensitivity analysis, we estimate that an appreciable share of the total incumbency advantage is from scareoff. Another relevant inference from this analysis is that in close incumbent-races, both candidates are of poorer quality than if it were an open seat with the promise of greater competition attracting stronger candidates.

Our paper illustrates a common feature of research designs based on natural experiments. These designs are welcome because they can offer a way to avoid strong and problematic assumptions, but they often lead to a redefinition of the quantity of interest (Sekhon and Titiunik 2012). It is up to the researchers who use these kinds of research designs to determine whether the redefined quantity is still conceptually useful and what, if any, is its relationship to the original quantity of interest.

References

Alford, John R, and David W. Brady. 1989. "Personal and Partisan Advantage in US Congressional Elections." In Congress Reconsidered, 4th ed., ed. Lawrence C. Dodd and Bruce I. Oppenheimer. Washington, DC: CQ Press

Ansolabehere, Stephen, James M. Snyder, and Charles Stewart. 2000. “Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage.” *American Journal of Political Science* 44(1): 17–34.

Broockman, David E. 2009. “Do Congressional Candidates Have Reverse Coattails? Evidence from a Regression Discontinuity Design.” *Political Analysis* 17(4):418–434.

Butler, Daniel Mark. 2009. “A Regression Discontinuity Design Analysis of the Incumbency Advantage and Tenure in the U.S. House.” Electoral Studies 28(1):123–128.

Calonico, Sebastian, Matias Cattaneo and Rocío Titiunik. 2013. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". Unpublished manuscript.

Cattaneo, Matias, Brigham Frandsen and Rocío Titiunik, 2013. "Randomization Inference in the Regression Discontinuity Design: An Application to the Study of Party Advantages in the U.S. Senate". Unpublished manuscript.

Caughey, Devin, and Jasjeet Sekhon. 2011. "Elections and the Regression Discontinuity Design: Lessons from Close U.S. House Races, 1942–2008". Political Analysis 19(4):385-408.

Cover, Albert D., and David Mayhew. 1977. "Congressional Dynamics and the Decline of Competitive Congressional Elections." In Congress Reconsidered, Lawrence C. Dodd and Bruce I. Oppenheimer (eds.), New York: Praeger.

Cox, Gary W., and Jonathan N. Katz. 1996. "Why Did the Incumbency Advantage in U.S. House Elections Grow?". *American Journal of Political Science* 40(2):478–497.

Cox, Gary W., and Jonathan N. Katz. 2002. Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution. New York: Cambridge University Press.

Eggers, Andrew C., Olle Folke, Anthony Fowler, Jens Hainmueller, Andrew B. Hall, James M. Snyder, Jr., 2013. "On The Validity Of The Regression Discontinuity Design For Estimating Electoral Effects: New Evidence From Over 40,000 Close Races". Unpublished manuscript.

Erikson, Robert S. 1971. "The Advantage of Incumbency in Congressional Elections." *Polity* 3(3): 395–405.

Erikson, Robert S. 1972. "Malapportionment, Gerrymandering, and Party Fortunes in Congressional Elections." *American Political Science Review* 66:1234-45.

Erikson, Robert S., and Thomas R. Palfrey. 1998. "Campaign Spending and Incumbency: An Alternative Simultaneous Equations Approach." *Journal of Politics* 60 (May): 355-73.

Erikson, Robert S., and Kelly Rader, 2013. ``Much Ado About Nothing: RDD and the Incumbency Advantage". Unpublished manuscript.

Gelman, Andrew, and Gary King. 1990. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34(4): 1142–64.

Golden, Miriam A., and Lucio Picci. 2012. "Redistribution and Reelection Under Proportional Representation: The Postwar Italian Chamber of Deputies". Unpublished manuscript.

Grimmer, Justin, Eitan Hersh, Brian Feinstein and Daniel Carpenter. 2011. "Are Close Elections Random?". Unpublished manuscript.

Hainmueller, Jens, and Holger Lutz Kern. 2008. “Incumbency as a source of spillover effects in mixed electoral systems: Evidence from a regression-discontinuity design.” *Electoral Studies* 27(2):213–227.

Imbens, Guido and Karthik Kalyanaraman. 2012. "Optimal bandwidth choice for the regression discontinuity estimator." *The Review of Economic Studies*, 79(3): 933-959.

Imbens, Guido W., and Thomas Lemieux, 2008. " Regression discontinuity designs: A guide to practice". *Journal of Econometrics* 142(2):615–635.

Lee, David S. 2008. "Randomized Experiments from Non-random Selection in U.S. House Elections". *Journal of Econometrics* 142(2):675-697.

Levitt, Steven D., and Catherine D. Wolfram. 1997. "Decomposing the Sources of Incumbency Advantage in the U.S. House." *Legislative Studies Quarterly* 22(1):45-60.

Sekhon, Jasjeet S., and Rocio Titiunik, 2012. "When Natural Experiments Are Neither Natural Nor Experiments". *American Political Science Review* 106(1): 35-57.

Trounstine, Jessica. 2011. “Evidence of a Local Incumbency Advantage.” *Legislative Studies Quarterly* 36(2):255–280.

Uppal, Yogesh. 2009. “The disadvantaged incumbents: Estimating incumbency effects in Indian state legislatures.” *Public Choice* 138:9–27.

Uppal, Yogesh. 2010. “Estimating Incumbency Effects in U.S. State Legislatures: A Quasi-Experimental Study.” *Economics and Politics* 22(2):180–199.

Zaller, John. 1998. "Politicians as Prize Fighters: Electoral Selection and Incumbency Advantage". In Party Politics and Politicians, John Geer, ed., Baltimore: Johns Hopkins University Press.

Table 1: Linear regression RD estimates of incumbency advantage
U.S. House elections, 1968-2008
All open seat contests at t where freshman incumbents run again at $t+1$

	Linear regression					Global polynomial
		Dep. Var= Vote $t+1$		Dep Var.= Residual Vote $t+1$	Dep. Var= Residual Surge $t+1$	Dep. Var= Vote $t+1$
Dem. Win t	6.80	7.66	7.85	8.00	7.70	7.29
(Yes=1,No=-1)	(0.60)	(0.54)	(5.30)	(0.60)	(0.56)	(1.35)
Dem, Vote t	0.74	0.76	0.54	0.06	-0.26	-
	(0.05)	(0.04)	(0.06)	(0.05)	(0.50)	
Presidential Vote t	--	--	0.28	--	--	-
			(0.06)			
Adj. R2	.770	.827	.836	.527	.366	0.775
RMSE	8.09	7.02	6.83	8.09	7.53	8.068
N	399	399	399	399	399	402
Year effects	No	Yes	Yes	No	No	No

Note: For midterm years, presidential vote = for previous presidential election in the district.

Table 2: Local linear regression RD estimates of incumbency advantage U.S. House, 1968-2008 – All open seat contests at t where freshman incumbents run again at $t+1$
Outcome: Democratic percentage of the vote

RD effect (τ^{RD})	15.12 [9.96, 20.28]
Mean control	40.55
Personal Incumbency Advantage ($\tau^{RD}/2$)	7.56 [4.98,10.14]
Total sample size	399

Note: Results are from local linear regression using Imbens and Kalyanaraman (2012) optimal bandwidth implemented in Stata package `rdrobust` by Calonico, Cattaneo and Titiunik (2013). Estimated optimal bandwidth is 11.23. Sample excludes southern and border states. 95% confidence intervals are in brackets.

Table 3: Linear regression estimates for residual vote for winners of open seats at their next (incumbent) election
U.S. House elections, 1968--2008
Open seat contests at t where freshman incumbents run again at $t+1$ within 48-52% Democratic vote share band

	Party of Open Seat Winner, time *t*		All Open Seats
Mean (S.D.)	Democrat.	Republican	
Residual Vote	*% Dem.*	*% Dem.*	*% time t Winner*
.at t	0.06	-0.59	0.34
	(4.73)	(4.59)	(4.63)
at $t+1$	7.24	-7.61	7.43
	(7.28)	(8.38)	(7.82)
Change	7.19	-7.01	7.10
	(6.93)	(8.34)	(7.65)
Par_t	50.91	49.34	50.72
Par_{t+1}	49.26	49.46	49.99
Percent reelected	82%	77%	80%
N	(33)	(36)	(69)

As a regression Equation: Change in Residual Vote($t+1$)=0.09 + 7.10 Incumbency +e. Standard error of coefficients 9.93 and 0.93, respectively. Adj. R squared=.458. Incumbency = 1 if Dem., -1 if Rep.
*Residual Vote $t+1$ minus.4 x Residual Vote t.
Standard deviations in parentheses.

Table 4: Residual vote for districts where incumbents run at t and the winner at t runs again at t+1, where the major-party vote margin is within 48-52

U.S. House elections, 1968--2008

Mean (S.D.)	Time t incumbent loses	Time t incumbent wins		
		Incumbent Runs Again	Incumbent Retires	All Cases
Time t incumbent vote at t	1.83 (6.36)	1.56 (6.23)	-0.54 (6.32)	1.11 (6.28)
Time t incumbent party vote at $t+1$	-6.34 (6.82)	5.14 (7.25)	0.67 (5.72)	4.19 (7.17)
Change time t incumbent party vote, t+1	-8.17 (6.43)	3.58 (6.24)	1.21 (6.80)	3.08 (6.40)
Time t incumbent Par_t	47.07	49.45	51.22	49.83
Time t incumbent party Par_{t+1}	48.74	49.46	51.53	49.90
Average reelection rate time t incumbent party	15%	76%	64%	71%
Sample size	68	93	25	118

Table 5: Estimated Incumbency Advantage Using Alternative Models

U.S. House elections, 1968-2008

	Sophomore Surge			Retirement Slump	Gelman King
	Open Seat Winners	Successful Challengers	All Cases		
Beta	5.35	6.19	5.60	7.99	8.13
SE	(0.39)	(0.59)	(0.33)	(0.55)	(0.36)
N	401	171	572	378	3827

Note: Each estimate is based on non-southern non-Border state districts, 1968-2008. The sophomore surge estimate is the mean residual-vote gain for freshman incumbents. The retirement slump is the mean residual-vote loss for the party of retiring incumbents. The Gelman-King estimate is the coefficient for the Democratic vote on the lagged vote, incumbent candidate status (1=Dem., -1=Rep.) and the incumbent party.

Figure 1: Democratic Vote at $t+1$ as a function of Democratic Vote at t

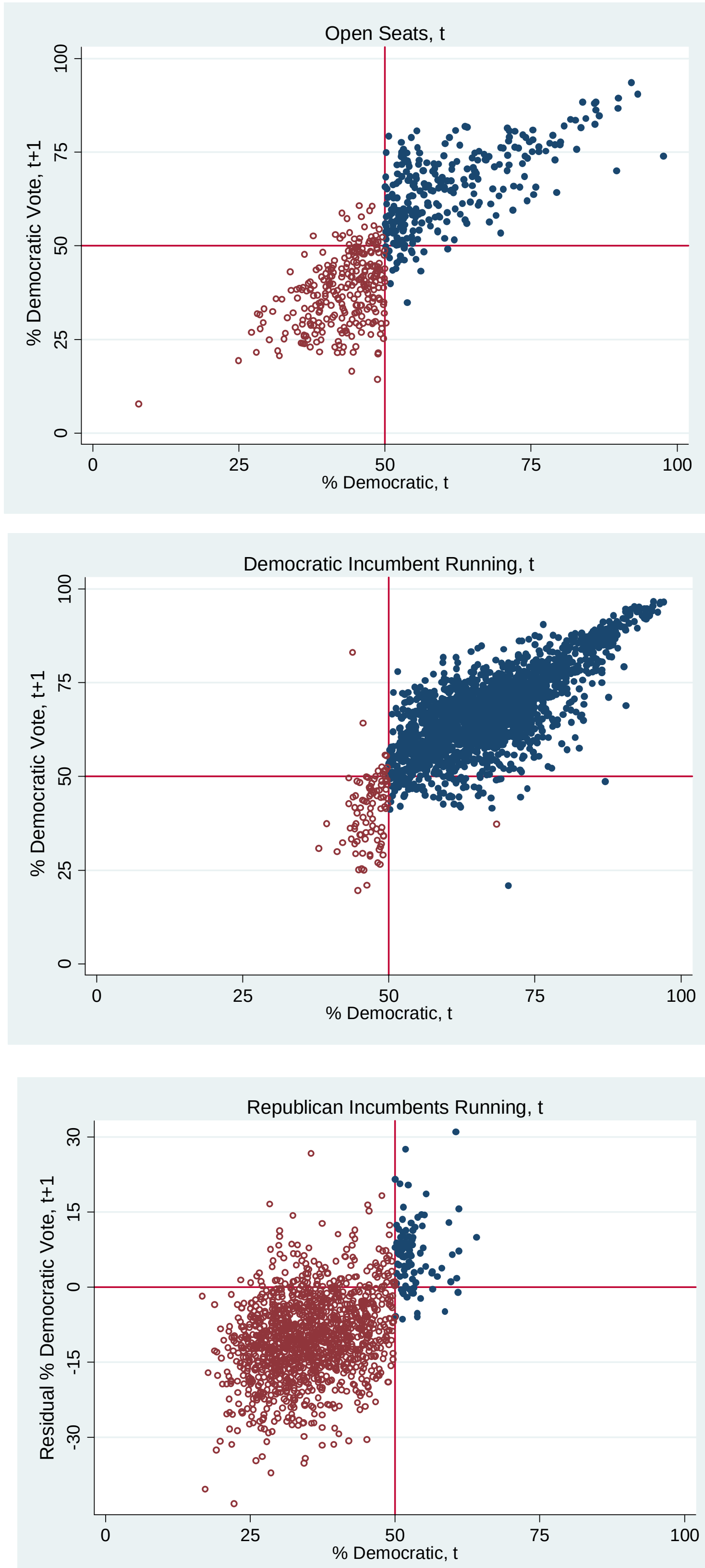
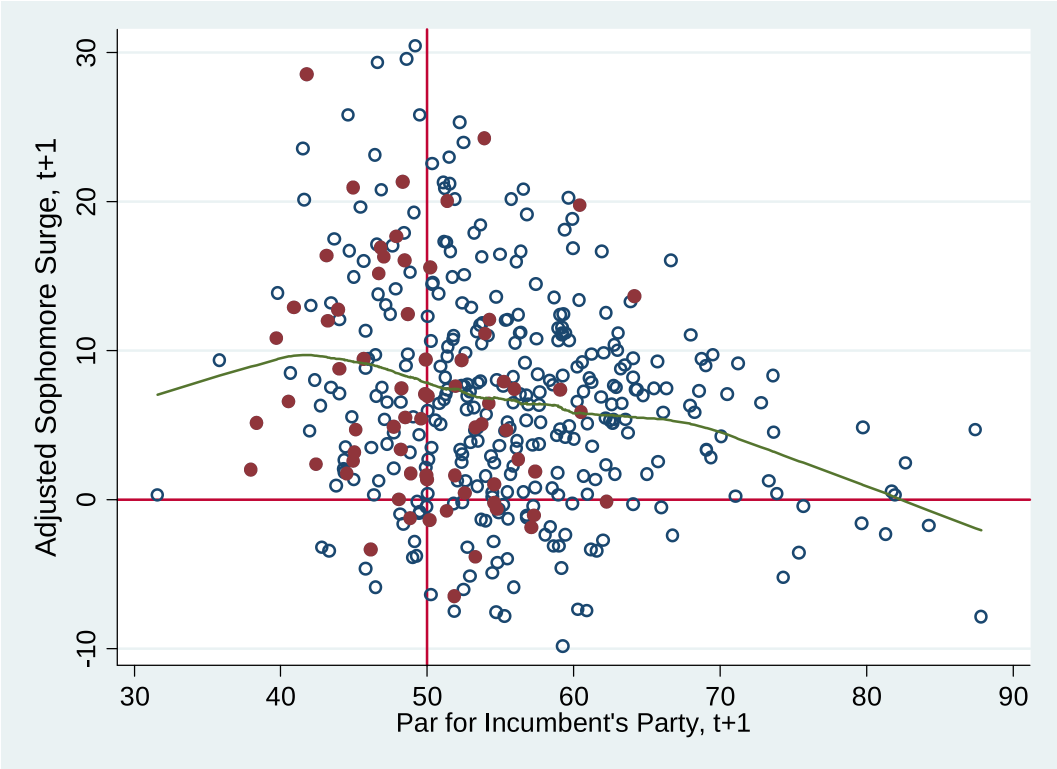


Figure 2: Adjusted sophomore surge for open seat winners, by incumbent-party Par.



Note: The adjusted sophomore surge is the difference between the $t+1$ residual vote and 1/2 times the t residual vote. Close elections where the time t open seat vote margin was within 52.0 percent are highlighted. The curved line is a lowess estimator with bandwidth.80.

Appendix

Table A1 presents a set of probit equations predicting retirements from the incumbent's victory margin in the prior election. The outcome is a dummy variable equal to one if the incumbent elected at t is not a candidate for reelection at $t+1$. It offers partial reassurance about selective retirements. Only three percent of open seat winners fail to run for reelection. And the rare retirements are not disproportionately concentrated among those with minimal winning margins. The coefficient has the expected negative sign but is far from statistically significant. In contrast, veteran incumbents retire at an 11 percent rate. For veteran incumbents, the coefficient is three times greater in magnitude (-0.012) than for freshman incumbents and highly statistically significant. We can conclude that incumbents are somewhat more likely to retire when faced with a close election.

Table A1: Predicting Retirements at t+1 from Winner's Vote Margin at time t. All cases, freshman winners of open seats, and veteran incumbents

	Freshmen winners of open seats		Veteran Incumbents	
Winner Margin, t+1	-0.004 (0.016)	-0.004 (0.017)	-0.012 (0.003)	-0.011 (0.003)
Year Effects?	No	Yes	No	Yes
Proportion retiring	.0337	.0337	.109	.109

Note: Retirements include vacancies from all causes, including primary defeats and death.