Estimation of the Incumbency Effects in the US State Legislatures: A Quasi-Experimental approach*

Yogesh Uppal
Department of Economics
University of California, Irvine
3151 Social Science Plaza A
Irvine, CA 92697
yuppal@uci.edu

November 15, 2005

Abstract

This paper estimates the incumbency effects in the legislative elections of 46 states in the US during the period 1968-89. I improve upon the existing measures of incumbency by disentangling the effect due to the quality of the candidates from the sole effect due to incumbency. To this end, the research design I use compares the candidates in closely fought elections and shows that the candidates in such contests are similar in quality. As a result, a comparison of the outcomes in the next election of the incumbents and the challengers in such contests identifies the true effect due to incumbency. The incumbency entails a significant advantage on the sitting legislators compared to their challengers. The incumbent candidates are 31.5 percentage points more likely to win the next election and gain 7.1 percent more votes compared to the challenger.

*I would like to thank Amihai Glazer, David Brownstone, Priya Ranjan and Kamal Sadiq for their guidance with this paper. I benefitted immensely from the comments from Bernard Grofman and Christopher Carpenter. All errors are mine.

1 Introduction

There is an abundance of research estimating the incumbency effects in elections in the United States. The methods used have ranged from simple comparison of the propotion of victorious incumbents and the average vote won by the incumbents compared to the open seat candidates (Garand 1991, Jewell and Breaux 1988), sophomore surge and retirement slump (Holbrook and Tidmarch 1991, Breaux 1990) to regression based techniques (Carey, Neimi and Powell 2000; Berry, Berkman and Schniederman 2000; Cox and Morgenstern 1993, 1995; King 1991). Major findings suggest that, on average, incumbent candidates in the United States win with more votes and a higher margin of victory and have greater chances of winning than the nonincumbents. But, with few exceptions, little care has been taken to address the first order concerns of providing unbiased estimates of incumbency effects. This is especially true for the estimates of incumbency at the state level. The main issue is that incumbency effect based on the differential outcomes of the incumbents and the nonincumbents suffers from a sorting problem. For instance, it is highly plausible that the incumbents win due to some inherent characteristic that is not possessed by the nonincumbents and that helped them win the office to begin with. It is also possible that the seats with incumbents contesting keep the high quality challengers off and attract only the weak challengers.

As a result of this sorting ability by the candidates, the process that determines the incumbency status of a candidate is rendered nonrandom. This, in turn, confounds the researchers' estimates of the incumbency effect. For example, in a scenario pointed out above, a failure to account for the incumbents' and the challengers' quality would overestimate the effect due to incumbency. The problem is complicated by the inability of the researcher to find a satisfac-

tory measure of candidate quality. This paper employs an improved empirical methodology, called the Regression Discontinuity Design (RDD), that removes the bias inflictng the previous methods. The RDD overcomes the bias by comparing the contests that were closely fought. The idea is that the candidates in such contests are, on average, likely to be similar. The high unpredictability of the outcome of such elections coupled with similarity of the contestants implies that the assignment of the incumbency status is approximately random. This randomness allows us to identify the effect due to incumbency net of any effect due to candidate specific or any other characteristics such as the quality of the incumbent and the challenger.

I use the RDD to estimate the incumbency effects in the state legislative elections in the United States using a county-level data of election returns of 46 states between 1968-1989 available from the Inter-university Consortium of Political Science Research (ICPSR). I consider only districts having single member seats and multi-member post seats. Since it is not clear how to determine the margin of victory in multi-member free-for-all seats, they have been left out. The exisiting literature is divided between using a vote-denominated measure of incumbency and an outcome-denominated measure of incumbency. Though outcome-denominated measure of incumbency is preferable (Jacobson 1987), I use both measures to provide an estimate of incumbency. The paper also inlcudes the uncontested seats which are normally left out of consideration in the existing literature. Since the RDD estimate is based on the difference in the outcome of winners and losers at margin of victory of zero (explained below), the inclusion of uncontested seats does not affect the estimate of incumbency effects.

Some results of the paper can be summarized as follows. The RDD estimate

of incumbency effect is 0.315 in terms of the probability of winninig. This implies that an incumbent candidate is 31.5 percentage points more likely to win than a challenger. The corresponding effect in terms of the vote share is 7.1% in favor the incumbent candidate implying that the incumbency, on average, adds 7.1% votes to the incumbent candidate's vote share. This is a huge advantage and explains the result why fewer challengers run again after losing once. This is the so called deterrent effect of being an incumbent. The difference in the probability of rerunning between the bare winner and bare loser is about 50 percentage points. The RDD as a research design is valid in the data and provide us with unbiased estimate of incumbency effects as all the characteristics on which I have data seem to be fairly smooth functions of the margin of victory. Various robustness checks further confirm the validity of the RDD.

The brief outline of the paper is as follows. Section 2 lays out the empirical methodology in detail. The complexities of data issues are discussed in section 3. Section 4 talks about the empirical results of the paper and does the robust checks. Section 5 concludes the paper.

2 Why RDD?

There has been a spurt in the papers estimating the incumbncy effects at the state level in the 1990s. Most of it is due to the availability of machine-readable data from the ICPSR. But the inability of the existing methods to provide a reliable estimate of true incumbency effects continues to grab researcher's attention. Jewell and Breaux (1988) find little change in the percentages seeking or winning reelection over time in all the 14 states they consider and a significant

rise in the average margin of victory in all of them. They attribute this finding to an increasing incumbency advantage because the open-seat races in their sample exhibit no such rise in average winning vote. Garand (1991) also finds that mean vote proportions for the incumbents have increased significantly.

Breaux (1990) estimates sophome surge and retirement slump for state legislatures of 19 states between 1968-1986. The sophomore surge is the average vote gain enjoyed by freshman candidates running as incumbents for the first time and the retirement slump is the average falloff in the party's vote when the incumbent retires (Covers and Mayhew 1977). A postive value of the sophomore surge represents an incumbency advantage. The retirement slump should be positive for an incumbency advantage to exist. He finds that the median retirement slump was -6.42% for Democrats and -7.17% for Republicans. The median sophomore surge was 4.96% and 6.67% for the two parties, respectively. Holbrook and Tidmarch (1991) also compute the sophomore surge in 32 states between 1968-1986 and find significant sophomore surge in the most contests.

Gelman and King (1990) show that the sophomore surge and the retirement slump suffer from a selection problem as their initial baseline is the vote share of a winning candidate which generally exceeds 50%. As a result, the sophomore surge underestimates and the retirement slump overestimates the incumbency effect. They put forward a regression based technique which controls for partisans swings and the previous vote share of the candidates and apply it to the estimation of incumbency effects in the House of Representatives. This technique is later used by King (1991) and Cox and Morgenstern (1993) to estimate incumbency effects in the state legislatve elections. Cox and Morgenstern (1993) consider elections in 24 states over the period 1968-1986. They find that on average incumbency advantage is positive in the states and has increased by over

a quarter of percentage points per election cycle. The general finding is that the incumbency advantage in states is not as large as one would see at the national level. Though an improvment over the previous methods, Gelman and King's approach does not account for the candidate quality and, hence, leads to biased results. So, a measure is still needed that will provide unbiased estimates of incumbency.

The RDD as an approximation to a natural experiment could be the answer. The ideal natural experiment to estimate the incumbency effects would require us to observe a candidate as an incumbent and a non-incumbent at the same point of time and, hence, is not possible. The least requirement would be to have candidates randomly assigned as incumbents and non-incumbents. The latter seems possible but not likely be true in practice because the candidates who received some favorable shocks, say, in election t would become incumbents in election t+1. The resultant unobserved heterogeneity among the candidates will cause the estimates of incumbency effects to be biased. The RDD factors out the unobserved heterogenity by comparing the candidates in close contests and achieves an approximate random assignment of incumbency status.

This is how the RDD works to identify the incumbency effect. First, the incumbency status of a candidate is discontinuous at the margin of victory of zero. The candidates for whom the margin of victory is positive become incumbent and those for whom it is negative lose and are nonincumbent. Second, all other characteristics vary continuously as a function of margin of victory. This amounts to saying that the candidates in the neighborhood around the discontinuity at the margin of victory of zero (bare winners and bare losers) are on average similar to each other in all characteristics and differ only in their incumbency status. So, any difference in their average next period outcome

(either probability of winning or the vote share) provides an unbiased estimate of the incumbency effect. The RDD was first put forward by Thistlethwaite and Cambell (1960) to study the effect of student scholarships on career aspirations, given that students are awarded scholarships only if their test score exceeds a certain threshold. Hahn, Todd and Van Der Klaauw (2001) and Porter (2002) provide a more formal treatment of the RDD technique.

To see an application of RDD to estimation of incumbency effects, let us first consider a naive model specification with no control variables.

$$P(win_{i,t+1} = 1) = f(\alpha_{i,t+1} + \beta * I_{i,t+1}) + \varepsilon_{i,t+1} \tag{1}$$

where $\text{win}_{i,t+1}$ is an indicator variable which is one if a candidate i wins in election t+1 and zero otherwise. $I_{i,t+1}$ is an indicator variable for incumbency status of the candidate such that

$$I_{i,t+1} = \begin{matrix} 1 \text{ if } mov_{i,t} > 0 \\ 0 \text{ if } mov_{i,t} < 0 \end{matrix} \tag{2}$$

$mov_{i,t}$ is the margin of victory for the candidate i in election t. In multi-candidate races, the margin of victory can be defined as follows. If a candidate is a winner in t, her margin of victory in t is the difference between her vote share and the vote share of the second placed candidate in t. Similarly, the margin of victory of a losing candidate in t is the difference between her vote share and the vote share of the winner in t. This construct allows us to have the margin of victory to be positive for the winning candidate and negative for the losing candidate. $\varepsilon_{j,t+1}$ is the stochastic error term that represents all other observable and unobservable characteristics of the candidate. The incumbency

effect from the this specification can be given by[1]

$$E\{win_{i,t+1}=1 \mid I_{i,t+1}=1\}-E\{win_{i,t+1}=1 \mid I_{i,t+1}=0\}=incumb+BIAS_{i,t+1} \tag{3}$$

where $incumb = f'(.) * \beta$ is the true incumbency effect and

$$BIAS_{i,t+1}=E\{\varepsilon_{i,t+1} \mid mov_{i,t} >0\}-E\{\varepsilon_{i,t+1} \mid mov_{i,t} <0\} \tag{4}$$

Instead of incorporating all the control variables, what RDD does is to compare the bare winners and bare losers such that

$$E\{win_{i,t+1}=1 \mid mov_{i,t} =\delta\}-E\{win_{i,t+1}=1 \mid mov_{i,t} =-\delta\}=incumb+BIAS^{*}_{i,t+1} \tag{5}$$

where

$$BIAS^{*}_{i,t+1}=E\{\varepsilon_{i,t+1} \mid mov_{i,t} =\delta\}-E\{\varepsilon_{i,t+1} \mid mov_{i,t} = \text{-}\delta\} \tag{6}$$

and δ represents the closeness of the elections. $BIAS^{*}_{i,t+1}$ goes to zero as δ gets smaller and smaller or as we examine closer and closer elections. Given that

[1] In a linear probability model $P(win_{i,t+1}=1)=\alpha_{i,t+1}+\beta*I_{i,t+1}+\varepsilon_{i,t+1}$, the incumbency effect will be given by $\beta + BIAS_{i,t+1}$. So, if we just run the above regression and take $\widehat{\beta}$ (estimate of β) to be the incumbency effect, our estimate will be biased because of other characteristics of the candidates.

we consider closer and closer elections, *incumb* will give us the true incumbency effect.

$$E\{win_{i,t+1} = 1 \mid mov_{i,t} = \delta\} - E\{win_{i,t+1} = 1 \mid mov_{i,t} = -\delta\} = incumb \quad (7)$$

The aim of this paper is to estimate the left hand side in (7) which is the difference between probability of reelection in $t+1$ of the bare winners and bare losers in t. The only assumption made is that the probability density function of ε, $g(\varepsilon \mid mov)$, is continuously distributed. The latter implies that all other characteristics vary continuously as a function of margin of victory.

Lee (2005) uses RDD to estimate the incumbency advantage at the party level in the elections to the House of Representatives in the United states. He finds that the candidates from the incumbent party are about 40-45 percentage points more likely to win an election than their counterparts from the nonincumbent party. The advantage in terms of the vote share is about 8% in favor of the incumbent party. Linden (2003) uses RDD to find that the incumbent candidates are about 14 percentage points less likely to win an election as compared to the nonincumbent candidates in the national elections in India. Uppal (2005) applies RDD to the state legislative elections in India between 1977-2003 and finds evidence of an incumbency disadvantage to the order of 20-25 percentage points.

3 Data Description

This paper uses revised county-level data on election returns from 1968 to 1989 available from Inter-university Consortium of Political Science Research (ICPSR). The data provides names of the candidates, their respective vote shares, party affiliation and incumbency status, the number fo people turned out to vote and number of candidates. This revised data adjusts for any mistakes in the names of the candidates that are present in the previous release of the same data from the ICPSR. The incumbency effects are estimated for the State House (the Lower Chamber) of 46 states and only general election results are considered. The analysis inlcudes all the single member and multi-member post district elections. There is no straight forward way to compute the margin of victory in other types of multi-member districts leading to their omission from the analysis.

The state elections suffer from frequent redistriting. The problems associated with comparison of the election preceding the change of the district lines with the one succeeding it are quite well known. Since the estimation of incumbency effect requires the comparison of electoral outcome at t and $t+1$, the years in which redistriting takes place are left out. The candidates recorded as scattering candidates are also left out. After stacking up the elections in the pairs of consecutive elections at t and $t+1$, I have 26,159 elections and the final count of total candidate-level observations is 49,874. The uncontested seats are included in the anlaysis as the RDD estimate of incumbency seems insensitive to what happens at the extreme ends.

The RDD requires that the bare winners and bare losers be comparable

on all the other characteristics. A check based on all possible characteristics is, however, constrained by the available data. But the original data file from the ICPSR can be used to derive some measures of candidate and challenger quality which are standard in the literature. I compare the incumbents and the non-incumbents on the following variables: the electoral experience at t defined as the number of times a candidate has been nominated before t; the political experiece at t defined as the number of times a candidate has won an election; the vote share at $t-1$; the indicator variable representing victory in $t-1$; the proportion of candidates belonging to the Democratic party; the proportion of candidates belonging to the Republican party; the number of people turned out to vote and the number of candidates.

4 Estimation of the Incumbency Effects

The elections to the state legislatures in the US show a significant advantage to incumbency as can be seen from Figure 1. Panel (a) plots the probability of winning in $t+1$ of a candidate against the margin of victory in t (mov_t). The probability of winning is derived using two methods: the local averages and a parametric fit. Each point in the scatter is the average of an indicator variable representing victory for a candidate in $t+1$. The average is taken over an interval of 0.5% of margin of victory providing us with the proportion of winner in each interval (or the raw probability of winning). A clear discontinuity is evident at the threshold margin of victory of zero. A parametric fit is also used to estimate the probability of winning shown by the solid curve through the scatter. The parametric fit is a logistic regression of the indicator variable for victory in $t+1$ on an indicator variable for victory in t, a fourth order polynomial of margin of victory in t, their interactions and the state-year fixed effects. The incumbency

effect is the size of the discontiuity which is the difference in the predicted probabilities between the losers and the winners evaluated at $mov_t = 0$. The data shows a big discontinuity at the threshold. The bare winners are about 45 percentage points more likely to win the next election compared to the bare losers. The incumbency advantage in terms of the vote share is about 30% in favor of the incumbent candidates as seen in the lower panel (b).

⟨Figure 1 about here⟩

The estimates in Figure 1 are, however, biased upwards. For example, the incumbency effect of 30% of vote share seems rather too high. This is due to the decision of some candidates not to run in the next election, a behavior particularly true for the losers. Figure 2 plots the probability of rerunning in $t+1$ against mov_t again using the local averages and the same parametric fit as described above. There is huge difference between the probability of rerunning for winners and losers. The bare winners are about 50 percentage points more likely to rerun than the bare losers. The bias in the incumbency estimates of Figure 1 arises because of the assumption that the candidates who do not rerun in the next election lose. Given that the winners are more likely to rerun that the losers, this assumption introduces an upward bias in the estimates of incumbency above.

⟨Figure 2 about here⟩

Ideally, the solution to this problem would require one to model the re-running decisions of the candidates independently of the prospects of winning. Since the data to achieve this task is not available and the decision to run is heavily determined by the chances of winning, it is difficult to correct for this

bias in the incumbency effects at the candidate level. In what follows, I condition my estimates on the candidates who rerun in the next election. This solution, however, does not come without any price. The estimates based on this conditional sample may suffer from a sample selection bias. For example, it is highly plausible that the losers who rerun are stronger than the losers who do not rerun resulting in my estimates of incumbency to be biased downwards. I show below that there is no such bias present in my estimates as the losers who rerun and who do not reun are similar.

Figure 3 plots the probability of winning and vote share in $t+1$ of a candidate against the margin of victory in t conditional on the candidates who rerun. As seen from the panel (a), the estimate of incumbency is about 0.3. This implies that a bare winner is about 30 percentage points more likely to win the next election than the bare loser. This is a huge advantage and might force the challengers of similar quality not to run against the incumbent. In the panel (b), the incumbency advantage translated in the vote shares is about 7% implying that the incumbent gains about 7% votes in the next election compared to the nonincumbent.

⟨Figure 3 about here⟩

After having estimated the incumbency effects, we need to check if the RDD is a valid exercise and indeed provides us with unbiased estimates as claimed above. This requires us to check if all the characteristics at t vary continously with margin of victory in t. If these characteristics show any significant discontinuity at the threshold margin of victory, the above estimates of incumbency in Figure 3 will be affected from the distortion present due to the differences in the characteristics other than the incumbency status, and hence, will be biased.

Figure 4 (panels a-h) plots various characteristics against the margin of victory. I compare the incumbents and nonincumbents on the eight characteristics mentioned above. All the characteristics vary smoothly with margin of victory. Any discontinuities present are highly insignificant. This validates the the RD design and reinforces the claim that RDD estimates of incumbency effect are unbiased.

The first two rows in Table 1 provide numerical estimates of incumbency effects. Columns (2)-(4) compare all winners against all losers by taking the average of the variables in the Column (1). For example, the probability of winning the next election for all winners is 75 percentage points and the vote share is 39.4% more than the all the losers. In the rows below, various characteristics are compared between the winners and the losers. The winners, on average, win more votes in the previous election, are more likely to win, run in more contests, win more contests, more likely to belong to the Democratic party, less likely to be a Republican, run from districts with less turnout and less number of candidates than the losers. All these are significant at the conventional levels of significance. These differences illustrate the main idea behind the RD design that the simple comparison of t and $t+1$ election outcome (e.g. vote shares) as an estimate of incumbency is fraught with bias as the candidates differ in other characteritics.

In the remaining columns, I compare only closer and closer contests. In Columns (5)-(7), only the contests within an absolute margin of victory of 25% or less compared. Similarly, in Columns (8)-(10), the contests within an absolute margin of victory of 5% or less are considered. The differences particularly in various characteristics at t become smaller and smaller, strenghtening the idea that in the limit at the threshold all the differences would vanish. In Columns (11)-(13), the estimate of the differences from the parametric fit (dis-

cussed above) evaluated at the $mov_t = 0$ are presented. The standard errors are given in the parentheses and are clustered at the state level. The only significant differences are in the probability of winning and the vote share in t+1 and the differences in other characteristics are insignificant. The estimate of incumbency effect in terms of probability of winning is 0.315 and in terms of vote share is 7.1%.

The arguement in favor of RDD is heavily based on the continuity of characteritics other than the incumbency status of the candidates. The design also provides us with some robustness checks to ascertain further if it is valid. A particularly important check is to include all the characteristics at t as the covariates and see if the estimate of incumbency changes significantly. The idea is that if the effect of incumbency is unbiased and has been stripped off any effect due to the characteristics at t, their inlcusion in the parametric fits used above (to estimate the incumbency effect) should not affect the estimate of incumbency. Table 2(a) carries out the robustness checks based on different specifications. In Column (2), I run a logistic regression of an indicator variable for the victory in $t + 1$ on an indicator variable for victory in t, a fourth order polynomial of margin of victory in t and their interactions and the state and time fixed effects and the probability difference evaluated at the $mov_t = 0$ from this regression is recorded. This is the same specification used above to estimate the incumbency effect. The effect due to incumbency as in Table 1 is 0.315. In Column (3), I include the electoral experience and political experience at t as the covariates in the logistic regression above and then evaluate the probability difference at the $mov_t = 0$. The estimate of incumbency does not change by much. Similalrly, In Column (4), the vote share and the indicator variable for a victory in the previous election are included as the covariates and the incumbency effect practically remains unchanged. The same happens in the remaining columns.

include all the characteritics in column (5).

As a last check of robustness in Column (7), I run the same regression with all the characteristics included as above but the only difference is that the dependent variable is the indicator variable for victory in t-1. The idea is to check if the outcome in t-1 was equally likely as the candidates are comparable. Also, the outcome in t-1 is already determined, and hence, should not be affected by the charcteristics in t. The probability difference in Column (6) is very small and highly insignificant providing further support for my estimates of incumbency. Table 2(b) does the robustness checks with the vote share as the dependent variable. The results are the same and pass all the robustness issues.

The estimates above were conditional on the sample of candidates who re-run in the next election. As suggested above, this might give rise to a sample selection bias in my estimates of incumbency. To check this, I compare the losing rerunners and the losing nonrerunners on various characteristics I have data on. In Table 3, I regress each characteristic on an indicator variable representing whether a candidate reruns in t+1, a fourth-order polynomial of margin of victory and their interactions, state and time fixed effects. The predicted differences are evaluated at $mov_t = 0$. As can be seen, the losing rerunners and the losing nonrerunners are comparable on all the characteristics. None of the differences are significant. This suggests that the sample selection is not present in my estimates.

5 Conclusions and Further Extensions

This paper uses an imporved empirical technique, called Regression Discontinuity Design (RDD) that removes the bias that is present in the previous methodologies. The RDD attempts to remove the effect due to the candidate quality by comparing candidates which are similar in characteristics. Due to the differential in the probability of rerunning of the winners and the losers, the incumbency effect is estimated for a party. But this is an issue which also limits other papers on the issue in addition to the above problems. Another advantage of using RDD is that the incumbency effect is not affected by the inclusion of the uncontested seats as the estimate depends on the discontinuity at the margin of victory of zero.

There is a large advantage to the incumbent candidates in the state legislative elections in the US. The incumbent is 31.5 percentage points more likely to win an election and gains 7.1% votes in the next election compared to the challenger. The chellengers are detered to run again as the difference between the probability of rerunning for the incumbents and the nonincumbents is about 50 percentage points. The approximate random assignment is attained as the other characteristics are comparable between the bare winner and bare loser.

A natural question that arises next is what explains this large incumbency advantage. Traditionally at the federal level, among various factors given for increased incumbent security are incumbents' control over redistricting plans (Tufte ,1973), increased franking privilleges (Mayhew ,1974), increased identification with the candidate rather than the party (Erikson 1971, 1972), (Ferejohn, 1977), (Cover, 1977) and increased bureaucratic resources available to

incumbents (Fiorina, 1977). At the state level, a related factor called professionalization is attributed to determine the incumbency advantage. The professionalization measures how professionalized a state is. Diiferent measures of professionalization are said to affect the incumbency advantage. These include personal staff and trips back home (Holbrook and Tidmarch 1991), operating budgets available to the legislator (Berry, Berkman and Schniederman 2000, Cox and Morgenstern 1993, King 1991, Chubb 1988) and salary (Carey, Neimi and Powell 2000). As an extesion of this paper, I intend to explain the variation in incumbency effects across states and weigh the relative effect of these different measures of professionalization and other factors such as the coattail effects frm the presidential, senatorial, congressional and gubernatorial offices; national economic conditions; the district size and the length of the legislative term

References

[1] Berry, William J., Michael B. Berkman, and Stuart Schniederman (2000). "Legislative Professionalism and Incumbent Reelection: The Development of Institutional Boundaries", *American Political Science Review*, 82(1), 133-54.

[2] Breaux, D. (1990). "Specifying the Impact of Incumbency on State Legislative Elections: A Destrict-level Analysis", *American Politics Quarterly*, 18(3), 270-86.

[3] Carey, John M., Richard G. Niemi, and Lynda J. Powell (2000). "Incumbency and Probability of Reelection in State Legislative Elections", *Journal of Politics*, 62(3), 671-700.

[4] Chubb, John E. (1988). "Institutions, the Economy and the Dynamics of State Elections", *American Political Science Review*, 94(4), 859-74

[5] Cover, A. (1977). "One Good Term Deserves Another: The Advantage of Incumbency in Congressional Elections", *American Journal of Political Science*, 21, 523-541.

[6] Cover, A. D. and David R. Mayhew (1977). "Congressional Dynamics and the Decline of Competitive Congressional Elections", In *Congress Reconsidered*, ed. Lawrence C. Dodd and Bruce I. Oppenheimer, New York, Praeger.

[7] Cox, G. and Scott Mogenstern (1993). "The Increasing Advantage in the US States", *Legislative Studies Quarterly*, 18(4), 495-511.

[8] Cox, G. and Scott Mogenstern (1995). "The Incumbency Advantage in the Multimember Districts: Evidence from the U.S. States", *Legislative Studies Quarterly*, 20(3), 329-349.

[9] Garand, J. C. (1991). "Electoral Marginality in State Legislative Elections, 1968-86", *Legislative Studies Quarterly*, 16(1), 7-28.

[10] Gelman, A. and Gary King (1990). "Estimating Incumbency Advantage without Bias", *American Journal of Political Science,* 34, 1142-1164.

[11] Hahn, J., Petra Todd and Wilbert Van Der Klaauw (2001). " Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design", *Econometrica*, 69, 201-209.

[12] Holbrook, T. and Charles M. Tidmarch (1991). "Sophomore Surge in State Legislative Elections, 1968-1986", *Legislative Studies Quarterly,* 16, 49-63.

[13] Jacobson, G. C. (1987). "The Marginals Never Vanished: Incumbency and Competition in Elections to the U. S. House of Representatives", *American Journal of Political Science*, 34, 126-141.

[14] Jewell, M. and David Breaux (1988). "The Effect of Incumbency on State Legislative Elections", *Legislative Studies Quarterly,* 13, 495-514.

[15] King, G. (1991). "Constituency Service and Incumbency Advantage", *British Journal of Political Science,* 21, 119-128.

[16] Lee, D. (2003). "Randomized Experiments from Non-random Selection in U.S. House Elections", *NBER Working paper* # 8441.

[17] Linden, L. (2003). "Incumbency Disadvantage: The Preference for Non-Incumbents in Indian National Elections", *unpublished manuscript.*

[18] Mayhew, D. R. (1974). "Congressional Elections: The Case of Vanishing Marginals", Polity, 6, 295-315.

[19] Porter, J. (2002). "Asymptotic Bias and Optimal Convergence Rates for Semiparametric Kernel Estimators in the Regression Discontinuity Model", HIER Discussion Paper # 1989.

[20] Thistlethwaite, D., and D. Campbell (1960). "Regression -Discontinuity Analysis: An alternative to the ex post facto experiment", *Journal of Educational Psychology,*

FIGURE 1(a). Probability of Winning in t+1

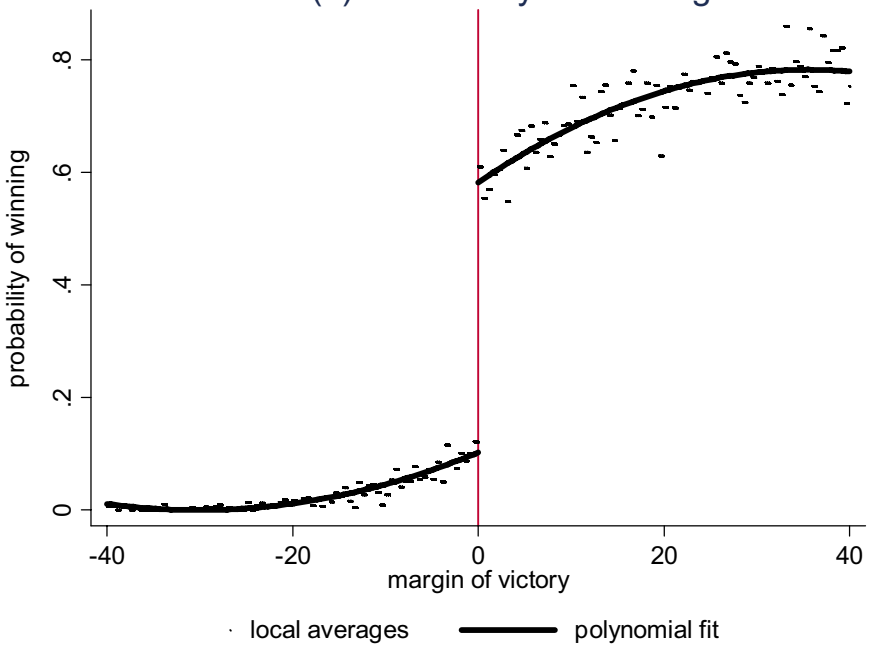

FIGURE 1(b). Vote Share in t+1

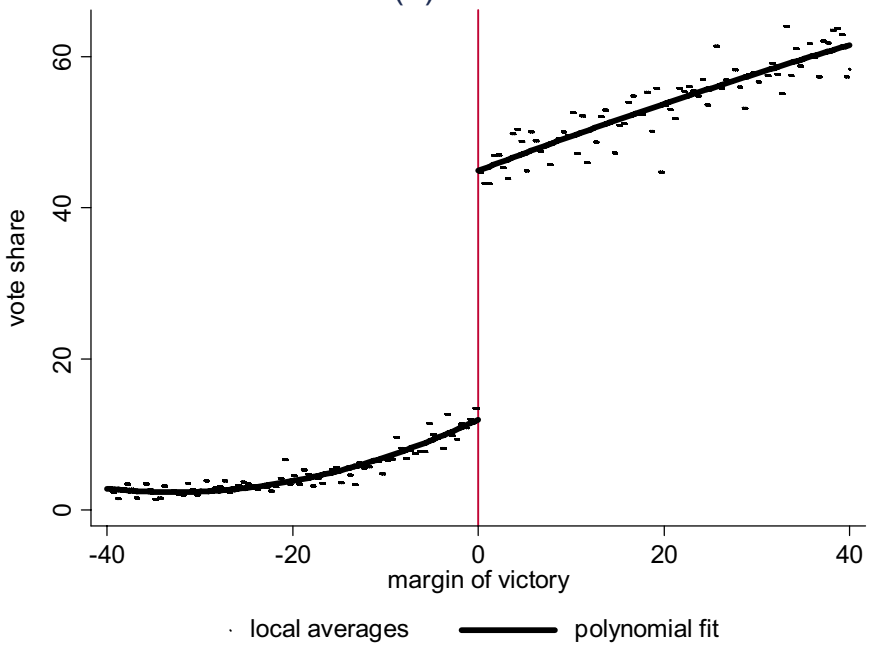

FIGURE 2. Probability of Rerunning in t+1

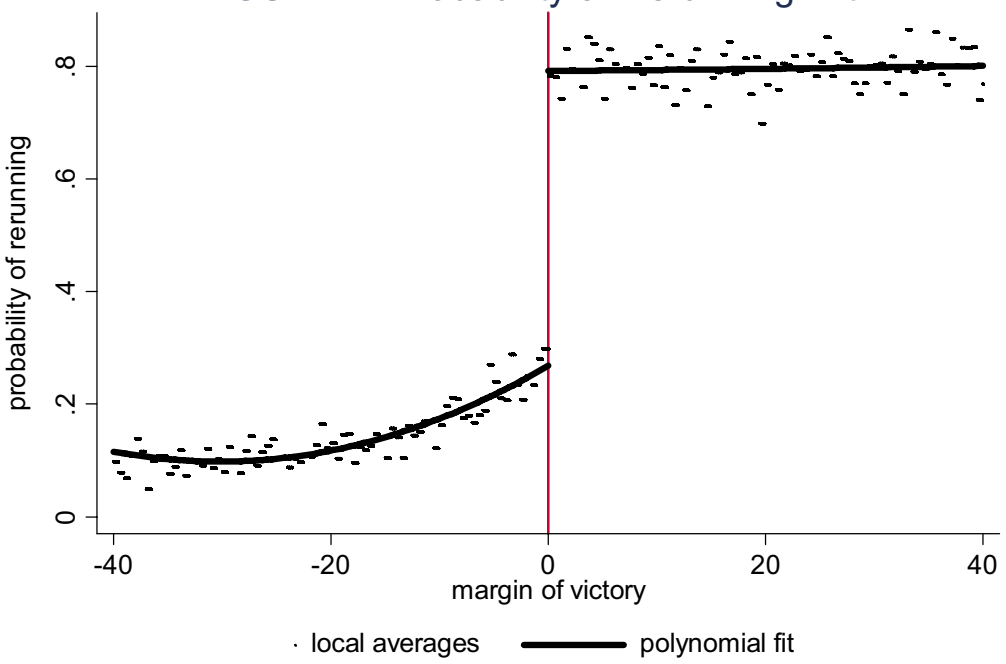

FIGURE 3(a). Probability of Winning in t+1

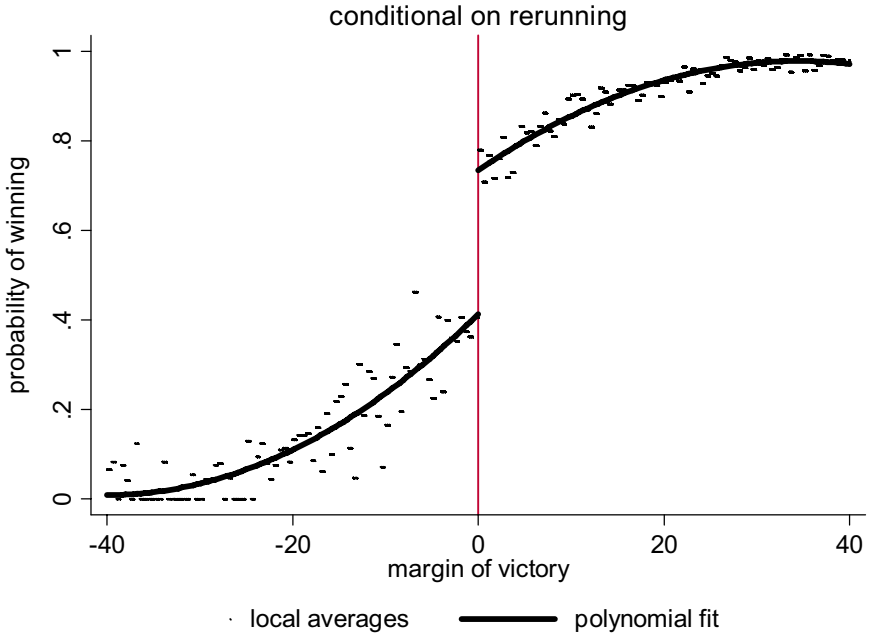

FIGURE 3(b). Vote Share in t+1

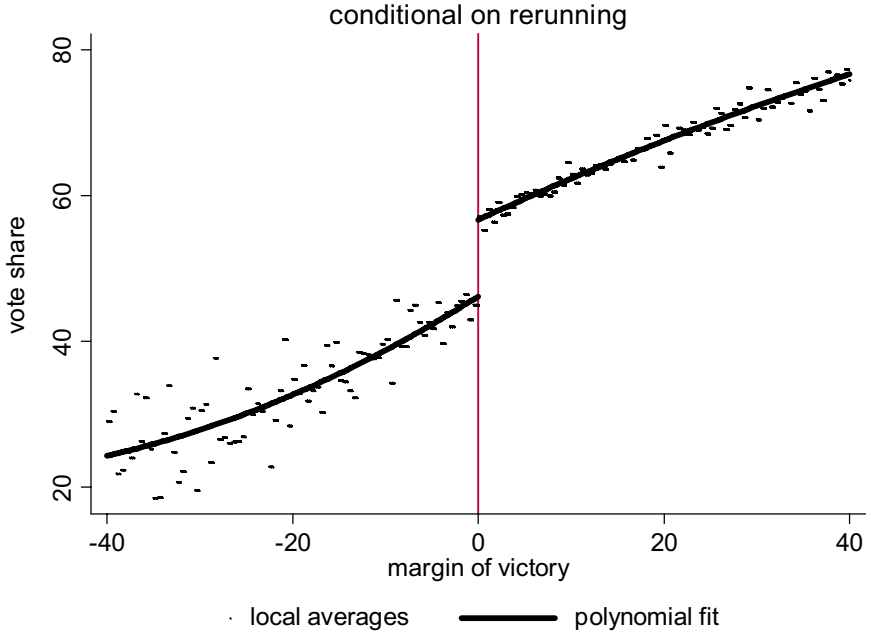

FIGURE 4(a)
Electoral Experience in t

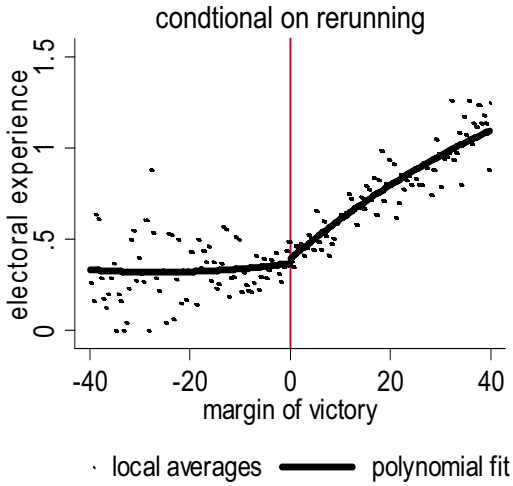

FIGURE 4(b)
Political Experience in t

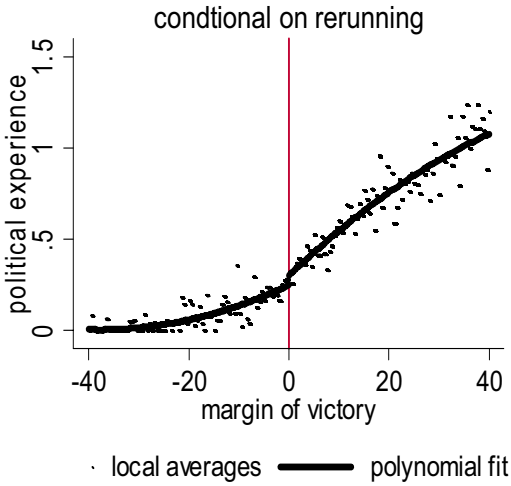

FIGURE 4(c)
Vote Share in t-1

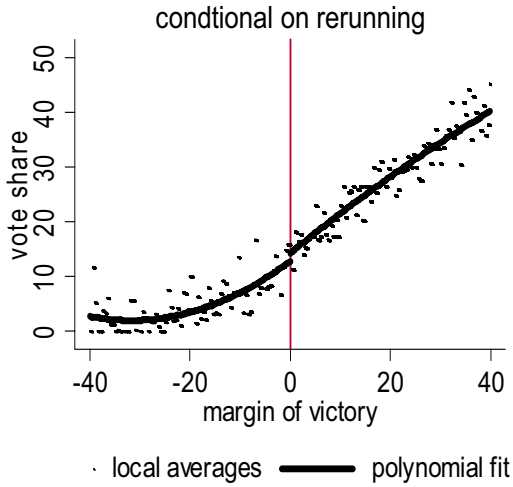

FIGURE 4(d)
Probability of Winning in t-1

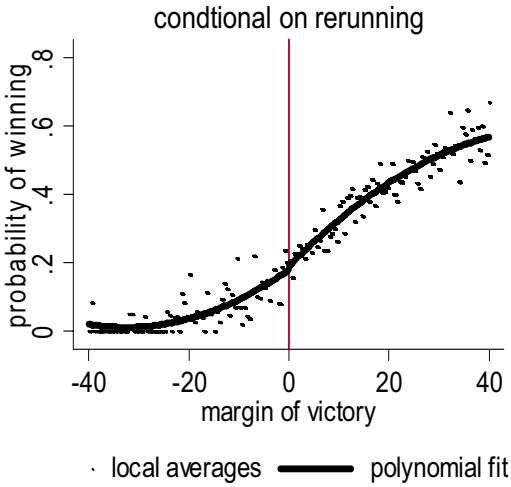

FIGURE 4(e)
Proportion of Democrats in t

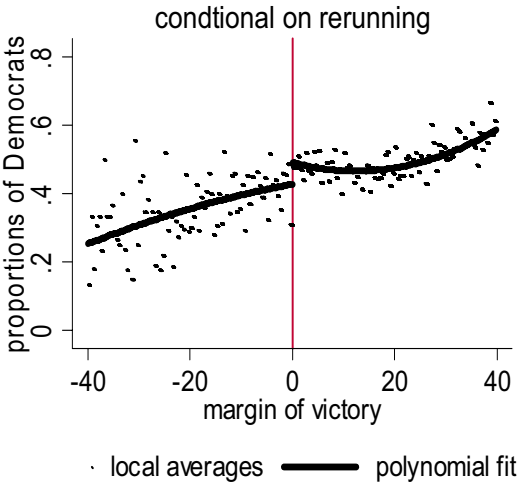

FIGURE 4(f)
Proportion of Republicans in t

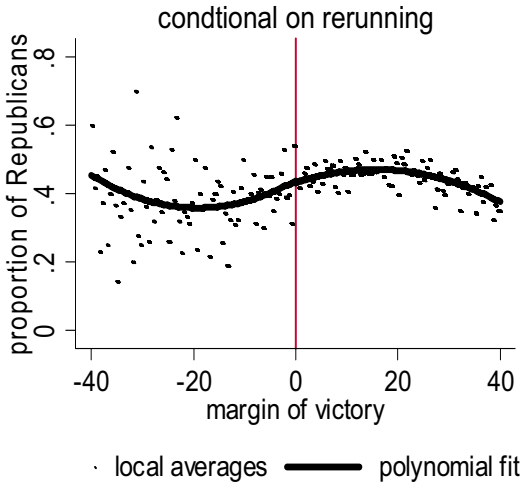

FIGURE 4(g)
Turnout in t

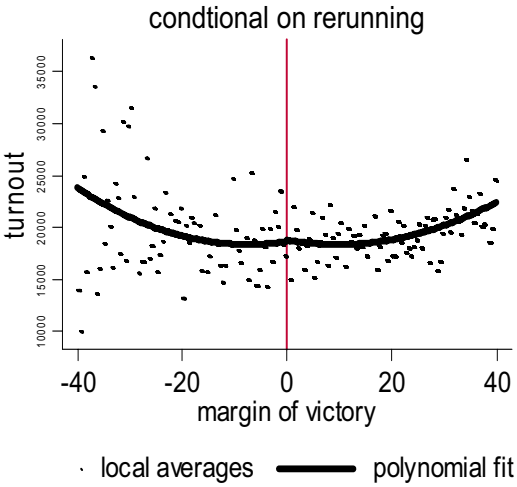

FIGURE 4(h)
Number of Candidates in t

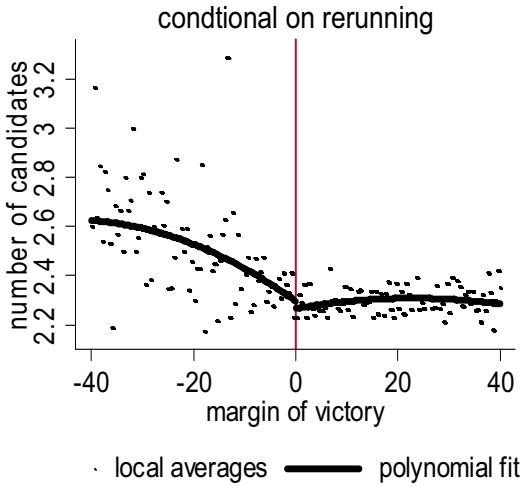

Table 1: Estimates of Incumbency Effects for Democratic Candidates and the Continuity of Predetermined Characteristics

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10).	(11)	(12)	(13)
	All			|margin| ≤ 25 %			|margin| ≤ 5 %			Parametric fit		
	winner	loser	diff.	winner	loser	diff.	winner	loser	diff.	winner	loser	diff.
Probability of winning in t+1	0.94 (0.002)	0.18 (0.007)	0.75* (0.005)	0.86 (0.004)	0.27 (0.01)	0.59* (0.009)	0.76 (0.01)	0.38 (0.02)	0.38* (0.02)	0.747 (0.017)	0.432 (0.03)	0.315* (0.034)
Vote share In t+1	75.6 (0.13)	36.2 (0.31)	39.4* (0.4)	63.4 (0.19)	42.8 (0.32)	20.6* (0.42)	58.1 (0.34)	46.9 (0.46)	11.3* (0.65)	57.9 (0.62)	50.8 (0.84)	7.1* (1.04)
Vote share In t-1	41.0 (0.27)	7.3 (0.34)	33.3* (0.74)	25.0 (0.35)	9.6 (0.5)	15.4* (0.77)	18.0 (0.65)	13.4 (0.1)	4.4* (1.26)	18.0 (1.05)	16.5 (1.72)	1.5 (2.0)
Probability of winning in t-1	0.54 (0.003)	0.08 (0.005)	0.46* (0.009)	.38 (0.005)	0.13 (0.008)	0.25* (0.0120	0.26 (0.1)	0.19 (0.016)	0.07* (0.02)	0.14 (0.014)	0.13 (0.024)	0.01 (0.028)
electoral exp	1.18 (0.01)	0.25 (0.012)	0.93* (0.03)	0.79 (0.014)	0.29 (0.017)	0.5* (0.03)	0.57 (0.025)	0.37 (0.032)	0.2* (0.048)	0.57 (0.054)	0.44 (0.064)	0.13 (0.084)
political exp	1.14 (0.01)	0.13 (0.009)	1.01* (0.03)	0.72 (0.013)	0.2 (0.014)	0.53* (0.029)	0.47 (0.024)	0.28 (0.028)	0.19* (0.044)	0.46 (0.07)	0.36 (0.1)	0.1 (0.08)
Proportion of Democrats	0.6 (0.003)	0.36 (0.009)	0.24* (0.01)	0.47 (0.005)	0.43 (0.011)	0.04* (0.013)	0.48 (0.011)	0.45 (0.02)	0.03 (0.024)	0.5 (0.029)	0.47 (0.042)	0.03 (0.051)
Proportion of Republicans	0.35 (0.009)	0.45 (0.003)	-0.1* (0.009)	0.46 (0.005)	0.42 (0.011)	0.04* (0.013)	0.44 (0.011)	0.45 (0.02)	-0.01 (0.024)	0.41 (0.023)	0.44 (0.034)	-0.03 (0.041)
Turnout	16469 (144)	19831 (431)	-3361* (417)	18982 (259)	18574 (555)	408 (607)	18568 (602)	19217 (1071)	-649 (1211)	20811 (990)	20622 (1172)	188 (1534)
number of candidates	1.99 (0.005)	2.45 (0.015)	-0.46* (0.016)	2.29 (0.007)	2.35 (0.016)	-0.06 (0.016)	2.28 (0.013)	2.29 (0.022)	-0.01* (0.026)	2.28 (0.033)	2.33 (0.034)	-0.05 (0.047)
# of obs.	24574			9879			2361			24574		

Standard errors are in the parenthesis. Standard errors are clustered at the state level for the parametric fit. All regressions have a fourth order polynomial of margin of victory as the right hand side variables.

* Significant at 1 % level of significance

TABLE 2(a). Robustness Check of Incumbency Effects based on Alternative Specifications

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable	Win Prob. in t+1 for winners- Win Prob. in t+1 for losers					Prob. difference in t-1
	0.315* (0.034)	0.32* (0.035)	0.312* (0.036)	0.311* (0.034)	0.313* (0.035)	0.003 (0.005)
Electoral Exp	No	Yes	No	No	Yes	Yes
Political Exp	No	Yes	No	No	Yes	Yes
Vote Share in t-1	No	No	Yes	No	Yes	Yes
Win in t-1	No	No	Yes	No	Yes	……..
Proportions of Democrats	No	No	No	Yes	Yes	Yes
Proportion of Republicans	No	No	No	Yes	Yes	Yes
Turnout	No	No	No	No	Yes	Yes
number of candidates	No	No	No	No	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	24574	24574	24574	24574	24574	24574

TABLE 2(b). Robustness Check of Incumbency Effects based on Alternative Specifications

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable	Vote Share in t+1 for winners- Vote Share in t+1 for losers					Vote Share difference in t-1
	7.1* (1.04)	7.2* (1.07)	7.0* (1.15)	7.0* (0.99)	7.0* (1.13)	0.029 (0.83)
Electoral Exp	No	Yes	No	No	Yes	Yes
Political Exp	No	Yes	No	No	Yes	Yes
Vote Share in t-1	No	No	Yes	No	Yes	……..
Win in t-1	No	No	Yes	No	Yes	Yes
Proportions of Democrats	No	No	No	Yes	Yes	Yes
Proportion of Republicans	No	No	No	Yes	Yes	Yes
Turnout	No	No	No	No	Yes	Yes
number of candidates	No	No	No	No	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	24574	24574	24574	24574	24574	24574

TABLE 3. Losing Rerunners vs Losing Non-Rerunners: Comparison on the Predetermined Characteristics

(1)	*(2)*
	Difference= (Losing Rerunners – Losing Non-Rerunners)
Electoral Exp	-0.08 (0.05)
Political Exp	-0.04 (0.05)
Vote Share in t-1	0.51 (1.68)
Win in t-1	0.02 (0.03)
Proportions of Democrats	0.05 (0.045)
Proportion of Republicans	0.066 (0.039)
Turnout	113 (1288)
number of candidates	0.037 (0.35)