

The Disadvantaged Incumbents: Estimating Incumbency Effects in Indian State Legislatures*

Yogesh Uppal[†]

October 29, 2005

Abstract

This paper estimates how the incumbency status of a candidate affects his or her chances of winning, known as the incumbency effects, using a large dataset on legislative elections of 25 states in India between 1975 and 2003. I use an innovative research design that disentangles the effect due to candidate quality from that due to incumbency by comparing the candidates in closely fought elections. This is a major improvement over previous methods and provides unbiased estimates of the incumbency effects. I find that incumbency has a significant negative effect on the fortunes of incumbent candidates. The probability of an incumbent winning the next election is 18 percentage points lower during 1975-91 and 25 percentage points lower during 1991-2003 than that of a nonincumbent. The variation in the incumbency effects across Indian states depends on the differences in degree of competitiveness of elections and, to a lesser extent, on the public provision of educational facilities by the government.

*I thank Amihai Glazer for his guidance and encouragement in writing this paper. I have benefitted immensely from discussions with Bernard Grofman, David Brownstone, Priya Ranjan, Kamal Sadiq and Christopher Carpenter. I also thank Sunil Menon for his help with the automation of a huge dataset from the Election Commission of India (ECI). All errors are mine.

[†]Mailing Address: 3151 Social Science Plaza A, University of California, Irvine, CA 92697.
Email: yuppal@uci.edu

Introduction

On average, incumbent candidates in the United States win more votes and are more likely to win than non-incumbent candidates. There is evidence that the incumbency advantage is greater after the mid-1960s (Cover 1977; Erikson 1971, 1972; Gelman and King 1990; Levitt and Wolfram 1997). However, the evidence from outside the United States does not support an intrinsic advantage to incumbency. For example, Gaines (1998) examines British Parliamentary elections during 1950-1992 and finds no evidence of any significant incumbency advantage. In Southwest Pacific countries of Papua New Guinea, the Solomon Islands, and the Republic of Vanuatu, the rate of turnover for the individual members of Parliament is very high (Steeves 1996; Trease *n. d.*).¹ In India, some anecdotal evidence suggests a disadvantage to incumbency. To quote the words of a losing Chief Minister (highest ranked executive) of Karnataka after the state assembly elections in 2004²

I think it (economic reforms) was ahead of its time, and therefore, the people did not understand it. But this time, it was plain anti-incumbency. Look at what happened to Digvijay Singh in MP (Madhya Pradesh), Ashok Gehlot in Rajasthan, or for that matter, Naidu in AP (Andhra Pradesh). They have all followed different growth paths. Naidu has done excellent work. Gehlot, for one, was very rural-focused. I think people just want change every five years.

The view that an anti-incumbency bias is present in Indian elections is reflected in the following quote about the national elections of 2004. ³

¹I am indebted to Jon Fraenkel, University of the South Pacific, Fiji for calling these references to my attention.

²Chakravarty, Saswati. 2004. "This is anti-incumbency, people just want change." *The Economic Times*, India, May 14th.

³Waldman, Amy. 2004. "Premier of India is forced to quit after vote upset." *The New York Times*, USA, May 14th.

The Bharatiya Janata Party had constructed an American-style presidential campaign around Mr. Vajpayee's perceived popularity, adopting a slogan of "India Shining". But their strategy ran aground on the realities of the Indian parliamentary system, in which voters turned on incumbent legislators who they felt had done little to deliver. Indian voters are known for their anti-incumbent attitudes, and the majority of sitting legislators were rejected in the three-week election.

The evidence above is informal and lacks a systematic investigation of incumbency effects for Indian elections. Using data on election results of 25 states in India between 1975 and 2003, this paper methodically studies the incumbency effects in elections to the state legislatures in India. More specifically, I am interested in whether the incumbency status of a candidate in Indian state legislatures raises or decreases his or her chances of winning. The importance of a study of Indian elections lies in India being the largest democracy in the world. Moreover, a finding that the incumbency effects are negative there provides a dramatic contrast to what we find in the United States.

A major methodological contribution of this paper is its application of the Regression Discontinuity Design (RDD) to estimation of the incumbency effects in Indian state elections. This methodology allows us to approximate a natural experiment, and gives us an unbiased estimate of the incumbency effects. RDD considers closely fought contests, and premises that candidates in such contests (bare winners and bare losers) are *ex ante* comparable, on average, in all candidate-specific characteristics such as experience and constituency-specific (district-specific) factors such as partisan effects, number of candidates contesting the election and so on. The only difference between candidates in such contests is in their incumbency status. The winning candidates become incumbents and the losing candidates will be non-incumbents. Moreover, the outcome

of such contests is highly unpredictable, which coupled with the comparability of candidates, brings about a random assignment of the incumbency status. So, any difference in their outcome in the next election will identify what is essentially an unbiased estimate of the true incumbency effect. The validity of the assumption about comparability of bare winners and bare losers can be readily checked with the data, as is shown below.

RDD is a new methodology, which has only recently been applied to empirical election studies. Lee (2005) uses the RDD to estimate partisan incumbency effects in the United States House of Representatives and finds that incumbent party is 40-45 percentage points more likely than non-incumbent party to win the next election. In terms of the vote shares, the effect is to raise the vote share of incumbent party by 8 percentage points as compared to non-incumbent party. In contrast to the literature in the United States, which defines incumbency advantage at the candidate level, Lee estimates the incumbency advantage at the party level. Linden (2003) is the only formal work prior to my paper estimating incumbency effects in Indian elections. He uses a non-parametric RDD to estimate the incumbency effects in Parliamentary elections at the national level (Lok Sabha) in India. He finds that incumbents in the national elections are at an advantage of about 6.5-9.75 percentage points of probability as compared to non-incumbents between 1980 and 1989, and starting in 1991, they suffer a disadvantage of 14 percentage points. He attributes this switch in incumbency effects in India to a decline in the dominance of Indian National Congress (INC), which is one of the largest political parties at the national level.

My results for state elections (Vidhan Sabha) are different from those at the national level by Linden. First, both the pre-1991 and the post-1991 periods have a negative incumbency effect (incumbency disadvantage). Incumbent candidates are 18 percentage points less likely than non-incumbent candidates to win the next election in the pre-1991 period, and the corresponding figure

for the post-1991 period is about 25 percentage points. Second, the magnitude of the incumbency effect at the state level is lower than that found by Linden at the national level. This implies a greater incumbency disadvantage in the state elections than in national elections. Third, I compare bare winners and bare losers on various candidate and constituency characteristics, which is required for RDD to be valid. Linden checks for the continuity of much fewer characteristics than I do. Fourth, I employ additional tests to check for the robustness of my estimates. Lastly, the comparative analysis across states suggests that the incumbency disadvantage is driven by degree of competitiveness in state elections and variation in the state governments' inability to provide public goods.

The outline of the remainder of the paper is as follows. The next section briefly lays out the empirical methodology. Various methodological complexities are discussed in the third section. The fourth section talks about the empirical results of the paper and provides the robustness checks. The fifth section compares the RDD estimates with those based on other commonly used methods. The sixth section seeks an explanation for variation in incumbency disadvantage across Indian states. The final section concludes and offers a brief discussion of the significance of the findings of this paper.

Methodology

Scholarly work by various political scientists in the 1970s started the debate on incumbency effects in the United States.⁴ Two measures of incumbency ad-

⁴There was a general agreement that the incumbency advantage increased significantly after the mid-1960s. The debate focused more on the factors causing an increase in the incumbency advantage. Among various factors given for increased incumbent security are incumbents' control over redistricting plans (Tuftes 1973), increased franking privileges (Mayhew 1974), increased identification with the candidate rather than the party (Erikson 1971, 1972; Cover 1977; Ferejohn 1977) and increased bureaucratic resources available to incumbents (Fiorena 1977). Jacobson (1985, 1987), however, argued that the incumbency advantage did not increase after the mid-1960s as the reelection chances of incumbents had not increased.

vantage, namely the sophomore surge and the retirement slump, were widely used in the literature. 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 (Cover and Mayhew 1977). Gelman and King (1990) reignited the debate by showing that both these measures were biased. They show that the sophomore surge underestimates and the retirement slump overestimates the effect due to incumbency. They use a regression-based approach in which they control for the national partisan swings that were missing from previous measures. The main limitation of Gelman and King's approach, acknowledged by them, is that their measure does not account for candidate quality. Levitt and Wolfram (1997) point out in relation to the elections in the United States that first, an incumbent is likely to be of higher quality, on average, compared to an open seat candidate, and second, seats contested by incumbents will attract weaker challengers as compared to the open seats. A failure to control for candidate quality, hence, will overestimate the incumbency effect.

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 next best research design would be to have candidates randomly assigned as incumbents and non-incumbents. This seems hard to accomplish in practice due to the presence of unobserved heterogeneity among the candidates that will bias the estimates of incumbency effects. For example, candidates who received some favorable shocks, say, in election t would become incumbents in election $t + 1$. In particular, the candidates who win at t may be better qualified for office than the losers. As a result, their reelection in $t + 1$ may be due to their being better candidates rather than any inherent incumbency advantage.

The Regression Discontinuity Design (RDD) is a simple methodology that

aims to approximate a random assignment of the incumbency status from the real world of nonrandom data. Thistlethwaite and Cambell (1960) used RDD 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. Lee et al (*n. d.*) uses the RDD on roll-call data for the United States House of Representatives during 1946-1995 to investigate whether there is a partial convergence or complete divergence between the announced policies of the candidates.

The RDD starts with the premise that the electoral outcome is a deterministic function of the vote share, which, in turn, is determined by the candidate-specific and constituency-specific (district-specific) characteristics. But there is always some random chance component to the final outcome. For example, there is uncertainty about how many voters turn out to vote, or postal ballots are received or not. So, if we compare the candidates in election t who are marginally above the threshold of winning (where the margin of victory equals zero) i.e. bare winners and marginally below the threshold i.e. bare losers, the assignment of incumbency status is essentially random because such candidates are comparable in other characteristics.

The main identification strategy is that incumbency status changes discontinuously at the threshold of margin of victory (mov) of zero, whereas all other (observable and unobservable) characteristics vary continuously as a function of margin of victory. The candidates for whom mov is positive will become incumbents, and for whom it is negative will be non-incumbents. If incumbency has any effect, the next period election outcome (such as probability of winning in $t + 1$ and vote share in $t + 1$) as a function of margin of victory in t will be discontinuous at the threshold, and the size of the discontinuity will give us an unbiased estimate of the incumbency effect. This can be illustrated with an

application of the RDD to estimation of the incumbency effects. Let us consider a naive model specification with no control variables. I consider a simple linear probability model due to the ease of exposition.

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

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

$$I_{i,t+1} = \begin{cases} 1 & \text{if } mov_{i,t} > 0 \\ 0 & \text{if } mov_{i,t} < 0 \end{cases} \quad (2)$$

where $mov_{i,t}$ is the margin of victory for candidate i in election t . In multi-candidate races, as found in India, the margin of victory of a candidate who wins the election in t is the difference between his or her vote share and the vote share of the second-place candidate in t . Similarly, the margin of victory of a losing candidate in t is the difference between his or her vote share and the vote share of the winner in t . This construct allows the margin of victory to be positive for the winning candidates, and negative for the losing candidates. $\varepsilon_{j,t+1}$ is the stochastic error term that represents all other observable and unobservable characteristics of the candidate. The incumbency effect from this specification can be given by

$$E\{win_{i,t+1} = 1 \mid I_{i,t+1} = 1\} - E\{win_{i,t+1} = 1 \mid I_{i,t+1} = 0\} = \beta + BIAS_{i,t+1} \quad (3)$$

where β is the true incumbency effect and $BIAS_{i,t+1} = E\{\varepsilon_{i,t+1} \mid mov_{i,t} > 0\} - E\{\varepsilon_{i,t+1} \mid mov_{i,t} < 0\}$. Note that the bias arises because the winning and losing candidates differ in their observable and unobservable characteristics. RDD compares the bare winners and the bare losers such that

$$E\{win_{i,t+1} = 1 \mid 0 \leq mov_{i,t} \leq \psi\} - E\{win_{i,t+1} = 1 \mid -\psi \leq mov_{i,t} \leq 0\} = \beta + BIAS_{i,t+1}^* \quad (4)$$

where $BIAS_{i,t+1}^* = E\{\varepsilon_{i,t+1} \mid 0 \leq mov_{i,t} \leq \psi\} - E\{\varepsilon_{i,t+1} \mid -\psi \leq mov_{i,t} \leq 0\}$ and ψ represents the closeness of the elections. $BIAS_{i,t+1}^*$ goes to zero as ψ gets smaller and smaller or as we examine closer and closer elections. Given that we

consider closer and closer elections, β will give us the true incumbency effect.

$$\lim_{\psi \rightarrow 0^+} E\{win_{i,t+1} = 1 \mid 0 \leq mov_{i,t} \leq \psi\} - \lim_{\psi \rightarrow 0^-} E\{win_{i,t+1} = 1 \mid -\psi \leq mov_{i,t} \leq 0\} = \beta \quad (5)$$

This paper estimates the left hand side in (5), which is the difference between probability of winning in $t+1$ of the bare winners and bare losers in t conditional on the candidate and the constituency characteristics. The only assumption made is that the conditional density function of ε , $g(\varepsilon \mid mov)$, is continuously distributed.⁵ This assumption implies that all other characteristics vary continuously as a function of margin of victory.

Data Description

The source of election data is the Statistical Reports on General Election to Legislative Assembly of States published by the Election Commission of India (ECI).⁶ Due to the huge task of collecting and cleaning up the data, I only used data on the elections held between 1975 and 2003.⁷ Another reason for using this time period is that district boundaries were constitutionally fixed between the mid-70s and 2001 and the data prior to 1975 suffered from frequent redistricting. I consider all the states except the state of Jammu and Kashmir, where elections were disrupted during the most of the sample period. The electoral system in the states is characterized by the single member simple plurality (SMSP), same as the elections at the national level. Table 1 provides information on the years of elections and total number of seats for each state in my data. Uttar Pradesh

⁵For detailed properties of the RDD estimates, please see Lee (*n.d.*).

⁶The Election Commission was established under the Constitution as a semi-autonomous permanent body with advisory jurisdiction and quasi-judicial powers. The Election Commission is responsible for preparation, maintenance and periodic revision of the electoral roll, supervising the nomination of candidates, registering political parties, monitoring the election campaign, including candidates' funding, facilitating coverage of the election process by the media, organizing the polling booths and undertaking the counting of votes and the declaration of results (*Source*: www.eci.gov.in).

⁷The source of the data is ECI's website at www.eci.gov.in. The data is in acrobat reader form and not directly readable by statistical software. A program was used to convert the data from the Acrobat format to text format and stack the data under various variable names.

has the largest number of seats (425) and Sikkim the lowest (32). There are on average 5 elections per state and 4,230 constituencies for all states taken together.⁸ In all, I have data on 24,592 elections over the period 1975-2003. This amounted to a dataset of 220,726 candidate-level observations. The unit of observation is a candidate in an election.

⟨Table 1 about here⟩

The data set provides information on the names of candidates contesting the election and their respective vote shares, sex and party affiliation. There is also information on the number of registered voters, the number of registered voters who turned out to vote, the rate of voter turnout, and the number of constituencies reserved for the scheduled casts (SC) and the scheduled tribes (ST) candidates.⁹ A major problem with the data is that the ECI does not record the names of candidates consistently. First, a candidate might be reported as last name followed by his or her first name or vice versa. The order of first and last names is switched in a subsequent election. Second, the middle names are omitted in some elections and included in some other. Third, full names and initialled names used interchangeably over different elections. Lastly, the spellings of the names are incorrectly reported in some elections. This made it extremely difficult to track the candidates over time given the size of the dataset.

I overcome this problem in two ways. First, I drop the observations that have a vote share of less than 5% in any election. In Indian elections, the number of candidates in any election tends to be large, and many candidates perform poorly and are not expected to have any effect on the eventual outcome.¹⁰ Most

⁸There were no elections held in following seventeen constituency codes in the state of Assam in 1983: 32-35, 65-66, 71-72, 75-78, 81, 99 and 118-120.

⁹In India, some seats are reserved for scheduled casts (SC) and scheduled tribes (ST) in an effort to safeguard the interest of certain disadvantaged minority groups, who might otherwise find themselves unrepresented.

¹⁰There are a great number of "non-serious" candidates standing for the elections. In a constituency named Modakurichi in Tamilnadu, 1033 candidates stood for election in 1996. Out of 1033, 1030 candidates won a combined vote share of 5.81.

of these candidates do not belong to any recognized party and, thus, are difficult to track. Second, I match the remaining candidates within a constituency across elections checking for different placement of first and last names, missing middle names, spelling mistakes, and so on. Though the data avoids any major redistricting issues, district boundaries were reset in some small states like Arunachal Pradesh, Delhi, Goa, and Mizoram in 1983-1984. As a result, I exclude these years from the analysis. This leaves me with 75,688 candidate-level observations. These observations are, then, stacked up in pairs of consecutive elections to compare the outcome in t with that in $t + 1$ to give me total of 63,954 observations.

Estimation of Incumbency Effects

The incumbents in India fare worse than their counterparts in the United States. The average vote share and the average margin of victory of a winner are about 48% and 15%, respectively, in India. The same for the United States are 60% to 70% and 20% to 30% (Lee *n.d.*). The simple probability of a winner in t (incumbent) becoming a candidate in the election at $t + 1$ (proportion of incumbents running for reelection in the next election) is 0.55 in India (0.88 in the United States). The probability of an incumbent winning the next election conditional on rerunning is 0.5 (0.9 in the United States). The probability that a loser in t becomes a candidate in election at $t + 1$ is 0.21 (0.2 in the United States). The probability of a loser in t winning the election at $t + 1$ conditional on rerunning is 0.38 (0.15 in the United States).

Figure 1 plots the probability of winning in election $t+1$ against the margin of victory (mov) in election t . The probability of winning is estimated by running a logistic regression of the indicator variable for victory in election $t + 1$ on a dummy that takes a value of one if the candidate won in election t and zero

otherwise, a fourth order polynomial in margin of victory, their interactions, and the state and the time fixed effects. This regression was used to plot the predicted probabilities separately for winners ($mov > 0$) and losers ($mov < 0$). The estimate of difference between the right hand side and left hand side limits of probability of winning at the threshold ($mov = 0$) determines the incumbency effect. The scatter of raw probability of winning (local averages), proportion of winners with in an interval of 0.5% of margin of victory, is also plotted. The top panel plots the probability of winning during the pre-1991 period and the bottom panel does the same for the post-1991 period. There is a slight disadvantage in the pre-1991 period. In the post-1991 period, the incumbency effect is about -0.08 implying bare winners are about 8 percentage points less likely to win the next election than bare losers.

(Figure 1 about here)

Figure 2 plots the probability of rerunning in the next election (local averages and fourth order polynomial fit) against the margin of victory. The bare winners are about 14 percentage points more likely to rerun in the next election than the bare losers in both the pre-1991 and post-1991 periods. Since the candidates who do not run for the election in $t + 1$ are not observed, the incumbency effects in Figure 1 are estimated assuming that such candidates lose the election in $t + 1$. This assumption renders the above estimate of incumbency biased. But, as shown in Figure 2, the candidates marginally below the threshold are less likely to rerun for the next election than the candidates above the threshold, the estimates of incumbency effect in Figure 1 are biased upwards and provide us with an upper bound on the magnitude of the incumbency effect.

(Figure 2 about here)

Figure 3 depicts the probability of winning in $t+1$, conditional on a candidate rerunning in the next election. There is a big discontinuous fall in the probability

of winning at the margin of victory of zero, as we move from the left of the threshold to the right. The discontinuity is not evident at any other level of margin of victory. The bare winners are about 18 percentage points less likely to win the next election than the bare losers in the pre-1991 period. After 1991, the incumbency disadvantage is about 25 percentage points. These findings differ from those in the federal elections by Linden (2003). First, both the pre-1991 and the post-1991 state elections are characterized by a negative incumbency effect. Linden finds that incumbency effect is about 6.5-9.75 percentage points of probability in the pre-1991 period, and about 14 percentage points in the post-1991 period. Second, the magnitude of incumbency disadvantage in state elections is greater than that in elections for the national Parliament.

⟨Figure 3 about here⟩

Figure 4 plots the vote share in $t + 1$ for the conditional sample as another measure for the incumbency effects. The incumbency effect is about -4.5% of the vote share in the pre-1991 period. The effect is about the same in the post-1991 period. The stability of incumbency effect in terms of vote share over the two periods hides the finding that the likelihood of incumbents winning the next election, as compared to the non-incumbents, declined in the post-1991 period and made them more vulnerable to a defeat. This reinforces what Jacobson (1985, 1987) stressed that it is the probability of winning, which is of primary importance, rather than the vote share.¹¹

⟨Figure 4 about here⟩

As emphasized earlier, an important requirement for the RDD estimates of incumbency effects to be valid is that the factors at t other than the incum-

¹¹Jacobson contended the finding by other researchers that incumbency advantage increased in the United States after the mid-1960s. He agreed that House incumbents, on average, won higher vote share in the 1960s as compared to the 1950s. But the probability of losing for the incumbents had not declined rendering incumbents as likely to lose in the 1960s as earlier.

bency status of the candidates be a continuous function of the margin of victory around the threshold. A convincing test of this assumption on the basis of all possible characteristics is constrained by lack of comprehensive data. Lee (2003) uses electoral and official experience as the two measures of the candidate characteristics. I check for continuity of various candidate characteristics such as the *vote share in t-1*, the *electoral experience* of a candidate at t (number of times a candidate has contested the election up to t), the *political experience* at t (number of times a candidate has won an election up to t), the *proportion of female candidates*, the proportion of candidates belonging to Indian National Congress (INC) (*proportion of INC candidates*), and the proportion of candidates belonging to Bhartiya Janta Party (BJP) (*proportion of BJP candidates*). The INC and the BJP are the two largest parties in India. I also check for the following constituency characteristics such as the *number of registered voters*, the number of registered voters who turned out to vote (*# who turned out to vote*), the *rate of voter turnout*, the *number of candidates*, the proportion of seats reserved for the SC candidates (*proportion of scheduled casts*) and the proportion of seats reserved for the ST candidates (*proportion of scheduled tribes*).

Table 2a provides the continuity checks for the pre-1991 period.¹² Columns (2)-(10) show the probability of winning and the vote share in $t + 1$ and other characteristics for winners, losers and their differences for all winners and losers (All), when $|\text{margin of victory}| \leq 25\%$ and when $|\text{margin of victory}| \leq 5\%$. In columns (2)-(4), winners, on average, have a greater vote share in the previous election, have more electoral and political experience, greater proportion of females, slightly less likely to belong to BJP, feature in constituencies with higher voter turnout, have fewer candidates standing for election and are less likely to belong to a constituency reserved for scheduled tribe as compared to the losers.

¹²I shall consider the results of only the conditional sample from hereon. The results based on the entire sample also meet all the continuity conditions, but are not shown due to space constraints. They are available upon request from the author.

The differences become smaller and smaller as the margin of victory gets closer to zero. In column (10), when $|\text{margin of victory}| \leq 5\%$, all the differences in predetermined candidate and constituency characteristics become statistically insignificant implying continuity of the characteristics as a function of margin of victory, whereas differences in the probability of winning and the vote shares in $t + 1$ remain significant.

(Table 2a about here)

Columns (11)-(12) regress all the variables on a dummy variable that takes a value of one for the incumbent and zero otherwise and a fourth order polynomial in margin of victory, their interactions with the incumbency dummy and the fixed effects. The values shown in the table are the predicted values at the margin of victory of zero and the standard errors have been clustered at the state level. All the differences in candidate and constituency characteristics are still insignificant around the threshold. This confirms that my estimate of the incumbency disadvantage is not driven by the characteristics other than the incumbency status of the candidate. Table 2b checks for the continuity assumption for the post-1991 period. The continuity assumption is also valid in this period.

(Table 2b about here)

This research design allows me to test for the robustness of my estimates of incumbency effects. We can further check the claim that the estimate of incumbency effect is not confounded by the candidate and the constituency characteristics by including them in the regression equation used above to find the probability of winning. The resultant estimate of incumbency effect should be insensitive to inclusion of these characteristics as covariates because it is unconfounded by them. In Table 3a, column (2) estimates a regression of the

indicator variable for the victory in $t + 1$ on a fourth order polynomial of margin of victory, an indicator variable for victory in t , their interactions and the fixed effects. The estimate of the incumbency effect is -0.18, which is identical to the estimate in Table 2a. In column (3), I include the candidate characteristics as additional regressors. The estimate of incumbency effect practically remains the same. The estimate does not change by much in column (4), where I include only the constituency characteristics and in column (5), where both candidate and constituency characteristics are included. Finally, in column (6), I regress the indicator variable for victory in $t - 1$ on all the above variables. The estimated difference in probability of winning in $t - 1$ should be close to zero, as it is already determined and cannot possibly be affected by the outcome in t . This difference is -0.0003 and is statistically insignificant. Table 3b performs similar robustness checks for the post-1991 period. The estimated incumbency effect is insensitive to inclusion of the candidate and the constituency characteristics.¹³

⟨Table 3a about here⟩

⟨Table 3b about here⟩

In the above analysis, I conditioned my estimates on the pool of candidates, who rerun in the next election. This could give rise to a sample selection bias in the estimated incumbency effects. This might be the case for example if rerunning losers are systematically different from non-rerunning losers. More specifically, we might have reasons to believe that only those losers, who are stronger than other losers and have higher chances of winning, rerun. In Table 4, I compare losing rerunners with losing non-rerunners on various candidate characteristics around the threshold of winning. I regress each candidate characteristics on a dummy variable that is one if the candidate reruns in election $t + 1$ and zero otherwise, a fourth-order polynomial of margin of victory, their

¹³The robustness checks with vote share in $t + 1$ as the dependent variable also confirm unbiasedness of the incumbency effects. They are available upon request from the author.

interactions and state and time fixed effects for candidates within 5% of margin of victory. All the differences between two sets of candidates are insignificant suggesting that around the threshold, losing rerunners are comparable to losing non-rerunners. So, there is no systematic bias due to the candidate characteristics. This is not to deny what we already know from Figure 2, namely that bare winners are more likely to run in the next election than the bare losers. But the determinant of running decisions of the candidates seems to be exogenous. For example, in India, the top-level leadership of the party (or the party high command) decides whom to nominate for elections (Kochanek 1968, Chhibber 2004, 86).

⟨Table 4 about here⟩

Comparisons to Existing Methodologies

There are three methodologies, namely the sophomore surge (SS), the retirement slump (RS) and the Gelman-King regression technique (GK) that have been widely used in the political science literature to estimate incumbency effects. I estimate the incumbency effects using these three techniques to compare the results with my estimates using RDD. The sophomore surge is defined as the average vote gain enjoyed by freshman candidates running as incumbents for the first time (Cover and Mayhew 1977). The intuition behind this measure is that if incumbency has any effect on the fortunes of the incumbents, it should show up in their vote share in the second election net of any party advantage (vote share of the incumbent in her first election). The retirement slump is the average falloff in the party's vote when the incumbent retires (Cover and Mayhew 1977). This is based on the logic that if incumbency has any effect, then the new party nominee should not be expected to do as well as the retired incumbent did. A negative value of the SS and a positive value of the RS imply

incumbency disadvantage.

Table 5 summarizes the estimates based on the SS, the RS and the Gelman-King method for elections to state legislatures in India. On average, incumbent loses 6.6% of the votes in the pre-1991 and 4.1% of votes in the post-1991 period when he or she runs for his or her first reelection. The estimate is statistically significant at conventional levels of significance. The SS measure suggests an incumbency disadvantage for the freshman incumbents and this disadvantage declines in the post-1991 period. The RS estimate is about -9.1% in the pre-1991 period implying the incumbent party loses about 9.1% of votes if the incumbent retires. The retirement slump for the post-1991 period is -3.0% of votes. According to the RS measure, Indian elections are characterized by the incumbency advantage and this advantage is more in the pre-1991 period than in the post-1991 period. However, this is expected in case of Indian elections. The incumbents who survive multiple terms and retire are likely to be of higher quality. As a result, the RS estimates are biased as the quality of the candidates is not taken into account. Another limitation of the SS and the RS estimates is that they may contradict each other, as is the case here.

⟨Table 5 about here⟩

Following Gelman and King (1990), I ran a regression of vote share in the election at $t+1$ on an incumbency dummy after controlling for vote share in the election at t and party effects. The incumbency dummy is 1 if an incumbent belongs to the INC, -1 if she belongs to the non-INC parties and zero if no incumbent runs.¹⁴ The party dummy is 1 if the incumbent belongs to the INC and -1 if he or she belongs to the non-INC parties. The Gelman-King estimate of incumbency effect is -13.1% of votes in the pre-1991 period and -11.1% in the

¹⁴Unlike the United States, India is a multi-party system. The INC is the only party contesting in all the states and other parties feature only in regional contests (Besely and Burgess, 2002). To implement the Gelman-King's method, I clubbed all the non-INC parties together.

post-1991 period. Thus, the Gelman-King method underestimates the incumbency effect for Indian state elections as compared to the RDD estimates. The RDD estimates remove the effect due to various candidate and the constituency characteristics.

Explaining Incumbency Disadvantage across Indian States

The task of carrying out a comparative analysis for the United States and India is beyond the scope of this paper. But it is interesting to look for explanations of contrasting incumbency effects in terms of the differences between the two countries. First, both countries differ in the kind of party systems they have. India is a multi-party system causing the contests to be relatively more competitive (as already pointed out above in terms of lower vote share and margin of victory for the winners). It is quite plausible that in the states, where elections are more competitive, incumbents might find it harder to hold on to their seats. Second, India is a developing country, where people do not have access to even basic necessities of life. The picture is really grim as far as the provision of public goods is concerned. In 1991, only 42.4% of Indian population had access to electricity, 62.3% had safe drinking water, and only 30.4% had both. About 27% villages did not have a primary school and 67% did not have any health infrastructure (Banerjee and Somanathan 2001).¹⁵ Since voters care about the provision of public goods, the states with greater supply of public goods should have higher incumbency effect.

I use the difference between the vote shares of the first-place party and the

¹⁵They provide survey evidence that voters care about the provision of public goods by the government. According to survey results provided by the National Election Study (a post election survey of voters after the Parliamentary elections in India in 1996), four out of ten major problems facing the country are related to physical and social infrastructure such as drinking water, education, health, transport, communication and electricity. These problems would be even more prominent at the state-level elections, where issues tend to be localized and the international issues such the Foreign Policy are not very important.

second-place party as a measure of *political competition* in a state. The higher this difference, the less competitive are the elections in that state and the higher is the incumbency effect (or the lower is the incumbency disadvantage). There is also considerable variation in literacy rate, percentage of population that is literate, across Indian states. In 1991, only 37.5% of the population in Bihar was literate, as compared to 89.8% in Kerala. I proxy the *literacy rate* in a state for the provision of primary educational facilities by the government. The more literate a state is, the higher is the incumbency effect. I use *electricity*, percentage of households with electricity as the source of lighting, as another measure of public good provision. The *rate of voter turnout* is used to represent voter activism¹⁶. I also control for *per capita income* of states, their *population* and the percent of population that is urbanized (*percent urban*).¹⁷

Table 6 summarizes the main results of the comparative analysis across Indian states. The dependent variable is the difference between probabilities of winning of bare winners and bare losers for each state from 1975 to 2003. All the right hand side variables are averaged out for the entire period. For example, the independent variable literacy rate is the average of the literacy rates in 1991 and 2001. In column (2), the probability difference is regressed on literacy and electricity using simple least squares estimation. The coefficient on literacy is positive and significant implying that more literate states have higher incumbency effects (or lower incumbency disadvantage). The coefficient on electricity is negative, but insignificant.¹⁸ In column (3), I include political

¹⁶But the effect of voter turnout is ambiguous. For example, if voters are dissatisfied with the state of affairs, higher voter turnout may work against the incumbent implying a negative relationship. If they rate government's performance as good, the effect might be positive.

¹⁷Literacy rate, electricity, state population and percentage of urban population are taken from the decennial census of India for the years 1991 and 2001. The data on 2001 census is taken from <http://www.censusindia.net/>. (May 24, 2005). Per capita income is per capita net state domestic product in 1980 prices. The source of income data is the Handbook of Statistics for Indian economy published by the Reserve Bank of India. URL: <http://www.rbi.org.in/> (May 24, 2005)

¹⁸The negative coefficient on electricity can be explained on the basis of the poor quality of electricity supply in most parts of India. The electricity supply is mostly erratic and can be off for days even in the big metropolitan cities such as New Delhi. This might cause more

factors such as political competition and the rate of voter turnout that is a proxy for voter activism. The coefficient on literacy is still positive and significant. As expected, the effect of political competition is positive. But it is insignificant at conventional levels of statistical significance. The effect of rate of turnout is negative, indicating greater incumbency disadvantage in states where voter activism is greater. But it is insignificant in explaining incumbency effects.

⟨Table 6 about here⟩

In column (4), I control for state per capita incomes. The coefficient on literacy remains positive and significant. The coefficient on political competition is positive and significant. The richer the state as represented by per capita income, the higher is incumbency disadvantage. This result is counter-intuitive. But per capita income is not a good representative of well being of a state. For example, higher per capita income may be accompanied by highly unequal distribution of income in favor of the elite, and is consistent with lower standard of living for majority of the population. Column (5) is the most general specification, where I control for demographic variables such as population of the state and percent of urbanized population. Since the economic reforms of the 1990s have been concentrated in the urban areas as compared to the rural areas, the rural population of the state is more likely to vote against the incumbent. Thus, the more urbanized a state is, the less is the incumbency disadvantage. The effect of literacy is still positive and significant at 10% level of significance. The effect of political competition is positive and significant. More urbanized states have less incumbency disadvantage. But the effect is statistically insignificant.

Conclusions

In the United States, the incumbent candidates have an advantage, as com-

dissatisfaction if you have electricity than if you do not.

pared to the non-incumbent candidates, due to incumbency. But researchers have struggled to isolate the effect due to incumbency status of a candidate from overall advantage due to candidate-specific and district-specific characteristics. The particular research design allowed by RDD is able to isolate the effect due to incumbency. The comparability of bare winners and bare losers approximates a random assignment of incumbency status. In such a situation, size of the discontinuity in probability of winning at margin of victory of zero gives us an unbiased estimate of incumbency effects. This paper provides evidence of a significant incumbency disadvantage in Indian state elections. This contrasts dramatically with the incumbency effect found in the United States, where the effect of incumbency is positive for the incumbents.

The validity of my RDD estimates is established by comparing bare losers with bare winners at election t . It turns out that all the differences in the candidate and the constituency characteristics between them become insignificant, as we compare closer and closer elections, and thus, any difference in their $t + 1$ election outcome is because of their incumbency status. I check for the robustness of my estimates by considering different specifications of probability of winning in $t + 1$. The estimate passes all robustness checks. Finally, the comparability of the losing candidates, who rerun and those who do not rerun around the threshold, allows me to believe that incumbency effect after conditioning on the candidates who rerun does provide us estimates that are free from the bias due to differences between rerunning and non-rerunning candidates.

The results in the previous section suggest that political competition works to undermine the authority of incumbents. The elections at the state level were competitive much before the decline of the INC at the national level in the late 1980s. Various regional parties rose to challenge its dominance in the state elections in the mid-1970s. INC was, particularly, threatened in Andhra Pradesh, Assam, Haryana, Karnataka, Kerala, Manipur, Punjab, Tamilnadu and

West Bengal, where proportion of seats it won in state assembly fluctuated above and below the majority mark. Chhibber (1999, 183) attributes increased political competition, and hence, low rates of incumbency, to the absence of secondary associations in India, which can help the incumbent mobilize voters. Various regional parties filled in their absence by aligning with the interests of one viable social group or another. He further argues that secondary associations present in the United States help higher rates of incumbency and, as a result, incumbents do not face the same competitive pressures as their counterparts in India do.

Kohli (1990) notes that India faced an increasing crisis of governability since the late 1960s. He argues that INC under the leadership of Indira Gandhi tried to retain control by appointing loyal but ‘weak’ chief ministers in the states, and thus, further damaging itself (1990, 16). He argues that

Weakness in the authority structures makes it difficult to solve precisely those problems whose solutions could strengthen authority. The bulk of political energy is spent fighting one bushfire from another, guided by the central concern of how long to hang on to power.

Furthermore, Economic deprivation of people might make the voters increasingly disenchanted with the government. My results do not imply a democratic failure in India, but only reinforce the working of democracy. Voters’ dissatisfied with their representatives use their electoral franchise to punish them for not delivering the goods. Incumbency advantage in the United States may also be due to the voters’ sense that government is functioning about as well as can be expected, thus ruling out the crisis of governability for the case of India much stressed by Kohli.

References

- [1] Banerjee, Abhijit, and Rohini Somanathan. 2001. "Caste, Community and Collective Action: The Political Economy of Public Good Provision in India." mimeo. MIT.
- [2] Besley, Timothy J., and Robin Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India." *Quarterly Journal of Economics* 117 (November): 1415-52.
- [3] Ministry of Home Affairs, Government of India. 1994. Census of India 1991. Housing and Amenities: A Data Base on Housing and Amenities For Districts, Cities and Towns. Occasional Paper # 5.
- [4] Chhibber, Pradeep K. 1999. *Democracy without Associations*. Ann Arbor, MI: University of Michigan Press.
- [5] Chhibber, Pradeep K., and Ken Kollman. 2004. *The Formation of National Party Systems*. Princeton, NJ: Princeton University Press.
- [6] Cover, Albert D. 1977. "One Good Term Deserves Another: The Advantage of Incumbency in Congressional Elections." *American Journal of Political Science* 21 (August): 523-41.
- [7] Cover, Albert 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.
- [8] Erikson, Robert S. 1971. "The Advantage of Incumbency in Congressional Elections." *Polity* 3: 395-405.

- [9] Erikson, Robert S. 1972. "Malapportionment, Gerrymandering, and part fortunes in Congressional Elections." *American Political Science Review* 66 (December): 1234-55.
- [10] Ferejohn, John A. 1977. "On the Decline of Competition in Congressional Elections." *American Political Science Review* 71 (March): 166-76.
- [11] Fiorina, Morris P. 1977. "The Case of Vanishing Marginals: The Bureaucracy Did it." *American Political Science Review* 71 (March): 177-81.
- [12] Gaines, Brian J. 1998. "The Impersonal Vote? Constituency Service and Incumbency Advantage in British Elections, 1950-1992." *Legislative Studies Quarterly* 23 (May): 167-95.
- [13] Gelman, Andrew. and Gary King. 1990. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34 (November): 1142-64.
- [14] Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69 (January): 201-09.
- [15] Jacobson, Gary C. 1985. "Money and Votes Reconsidered: Congressional Elections, 1972-1982." *Public Choice* 47: 7-62.
- [16] Jacobson, Gary C. 1987. "The Marginals Never Vanished: Incumbency and Competition in Elections to the U. S. House of Representatives." *American Journal of Political Science* 34 (February): 126-41.
- [17] Kochanek, Stanley. 1968. *The Congress Party of India: the Dynamics of One-Party Democracy*. Princeton, NJ: Princeton University Press.
- [18] Kohli, Atul. 1990. *Democracy and Discontent: India's Growing Crisis of Governability*. Cambridge, MA: Cambridge University Press.

- [19] Lee, David. N. d. "Randomized Experiments from Non-random Selection in U.S. House Elections." *Journal of Econometrics*. Forthcoming
- [20] Lee, David, Enrico Moretti, and Matthew Butler. 2004. "Do Voters Affect or Elect Policies? Evidence from the U.S. House." *Quarterly Journal of Economics* 119 (August): 807-59.
- [21] Levitt, Steven D. and Catherine D. Wolfram. 1997. "Decomposing the Sources of Incumbency Advantage in the U.S. House." *Legislative Studies Quarterly* 22 (February): 45-60.
- [22] Linden, Leigh L. 2004. "Are Incumbents Really Advantaged? The Preference for Non-Incumbents in Indian National Elections." *Columbia University Manuscript (January)*.
- [23] Mayhew, David R. 1974. "Congressional Elections: The Case of Vanishing Marginals." *Polity* 6 (Spring): 295-315.
- [24] Porter, Jack R. 2002. "Asymptotic Bias and Optimal Convergence Rates for Semiparametric Kernel Estimators in the Regression Discontinuity Model." *HIER Discussion Paper # 1989 (December)*.
- [25] Steeves, Jeffrey S. 1996. "Unbounded Politics in the Solomon Islands: Leadership and Party Alignments." *Pacific Studies* 19 (March): 115-38.
- [26] Thistlethwaite, D., and D. Campbell (1960). "Regression -Discontinuity Analysis: An alternative to the ex post facto experiment." *Journal of Educational Psychology* 51: 309-17.
- [27] Trease, Howard V. N.d. "The Operation of the Single Non-Transferable Vote System in Vanuatu." *Journal of Comparative and Commonwealth Studies*. Forthcoming.

- [28] Tufte, Edward R. 1973. "The Relationship between Seats and Votes in Two-Party Systems." *American Political Science Review* 67 (June): 540-54.

TABLE 1. Years of Election and Number of Seats

State	Years of Election	Number of Seats
Andhra Pradesh	1978, 1983, 1985, 1989, 1994, 1999	294
Arunachal Pradesh	1978, 1980, 1984, 1990, 1995, 1999	Pre-1984=30; Post-1984=60
Assam	1978, 1983, 1985, 1991, 1996, 2001	126
Bihar	1977, 1980, 1985, 1990, 1995, 2000	324
Delhi	1977, 1983, 1993, 1998	Pre-1983= 56; Post-1983=70
Goa	1977, 1980, 1984, 1989, 1994, 1999	Pre-1984=30; Post-1984=40
Gujarat	1975, 1980, 1985, 1990, 1993, 1998, 2002	182
Haryana	1977, 1982, 1987, 1991, 1996, 2000	90
Himachal Pradesh	1977, 1982, 1985, 1990, 1993, 1998, 2003	68
Karnataka	1978, 1983, 1985, 1989, 1994, 1999	224
Kerala	1977, 1980, 1982, 1987, 1991, 1996, 2001	140
Madhya Pradesh	1977, 1980, 1985, 1990, 1993, 1998	320
Maharashtra	1978, 1980, 1985, 1990, 1995, 1999	288
Manipur	1980, 1984, 1990, 1995, 2000, 2002	60
Meghalaya	1978, 1983, 1988, 1993, 1998, 2003	60
Mizoram	1978, 1979, 1984, 1987, 1989, 1993, 1998	Pre-1984=30; Post-1984=60
Nagaland	1977, 1982, 1987, 1989, 1993, 1998, 2003	60
Orissa	1977, 1980, 1985, 1990, 1995, 2000	147
Punjab	1977, 1980, 1985, 1992, 1997, 2002	117
Rajasthan	1977, 1980, 1985, 1990, 1993, 1998	200
Sikkim	1979, 1985, 1989, 1994, 1999	32
Tamilnadu	1977, 1980, 1984, 1989, 1991, 1996, 2001	234
Tripura	1977, 1983, 1988, 1993, 1998, 2003	60
Uttar Pradesh	1977, 1980, 1985, 1989, 1991, 1993, 1996	425
West Bengal	1977, 1982, 1987, 1991, 1996, 2001	294

TABLE 2a. Probability of Winning and Vote Share in t+1 and the Continuity of Predetermined Characteristics, 1975-1991

(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	0.5 (0.005)	0.4 (0.006)	0.1* (0.008)	0.47 (0.006)	0.44 (0.007)	0.03* (0.009)	0.41 (0.011)	0.53 (0.013)	-0.12* (0.017)	0.39 (0.021)	0.57 (0.018)	-0.18* (0.028)
Vote Share in t+1	39.5 (0.16)	33.0 (0.19)	6.5* (0.25)	38.2 (0.17)	35.0 (0.22)	3.15* (0.27)	36.4 (0.32)	38.5 (0.39)	-2.1* (0.5)	35.6 (0.49)	40.0 (0.51)	-4.4* (0.71)
Electoral Experience	0.65 (0.009)	0.55 (0.009)	0.1* (0.013)	0.66 (0.01)	0.60 (0.01)	0.06* (0.01)	0.65 (0.019)	0.69 (0.022)	-0.04 (0.029)	0.6 (0.026)	0.65 (0.02)	-0.05 (0.033)
Political Experience	0.40 (0.007)	0.31 (0.007)	0.09* (0.01)	0.4 (0.008)	0.35 (0.008)	0.06* (0.01)	0.38 (0.014)	0.4 (0.017)	-0.02 (0.022)	0.33 (0.021)	0.38 (0.022)	-0.05 (0.031)
Vote Share In t-1	18.4 (0.24)	13.4 (0.24)	5.0* (0.35)	18.1 (0.27)	14.8 (0.28)	3.4* (0.4)	16.8 (0.48)	17.3 (0.57)	-0.48 (0.75)	14.8 (0.68)	17.3 (0.82)	-2.5 (1.07)
Proportion of Female Candidates	0.04 (0.002)	0.02 (0.002)	0.02* (0.003)	0.04 (0.002)	0.02 (0.002)	0.01* (0.003)	0.03 (0.004)	0.03 (0.004)	0.0 (0.006)	0.038 (0.006)	0.025 (0.006)	0.013 (0.009)
Proportion of INC Candidates	0.27 (0.005)	0.27 (0.005)	0.0 (0.007)	0.28 (0.005)	0.29 (0.006)	-0.01* (0.008)	0.3 (0.01)	0.3 (0.012)	0.0 (0.015)	0.39 (0.021)	0.35 (0.024)	0.04 (0.032)
Proportion of BJP Candidates	0.07 (0.003)	0.11 (0.004)	-0.03* (0.004)	0.08 (0.003)	0.1 (0.004)	-0.02* (0.005)	0.08 (0.006)	0.1 (0.007)	-0.02 (0.009)	0.004 (0.0005)	0.004 (0.0006)	0.0 (0.0008)
# of Registered Voters	104950 (482)	104422 (566)	527 (739)	106142 (557)	104472 (670)	1670 (865)	103983 (1079)	104679 (1268)	-696 (1659)	103728 (519)	103743 (569)	-15.0 (770)
# who Turned out to Vote	62593 (298)	60355 (340)	2238* (452)	64487 (339)	61629 (399)	2858* (522)	63314 (652)	62968 (753)	346 (996)	61582 (483)	62050 (399)	-468 (626)
Rate of Turnout (%)	61.8 (0.15)	60.7 (0.17)	1.15* (0.23)	63.3 (0.16)	62.2 (0.19)	1.1* (0.25)	64.1 (0.31)	63.6 (0.36)	0.46 (0.48)	61.8 (0.31)	62.3 (0.37)	-0.5 (0.47)
# of Candidates	8.05 (0.06)	8.9 (0.07)	-0.8* (0.09)	8.3 (0.07)	9.0 (0.08)	-0.75* (0.12)	8.6 (0.14)	9.0 (0.16)	-0.39 (0.22)	8.8 (0.12)	8.6 (0.18)	0.2 (0.22)
Proportion of Scheduled Casts	0.13 (0.003)	0.13 (0.004)	-0.0 (0.005)	0.12 (0.004)	0.12 (0.004)	0.0 (0.006)	0.1 (0.007)	0.12 (0.008)	-0.02 (0.01)	0.125 (0.012)	0.14 (0.013)	-0.015 (0.018)
Proportion of Scheduled Tribes	0.12 (0.003)	0.14 (0.004)	-0.02* (0.005)	0.11 (0.004)	0.13 (0.005)	-0.02* (0.006)	0.12 (0.007)	0.12 (0.008)	0.0 (0.01)	0.051 (0.006)	0.046 (0.006)	0.005 (0.008)
# of obs.	16699			12767			3578			16699		

Notes: All comparisons are conditional on rerunning. Standard errors are in the parenthesis. Standard errors are clustered at the state level for the parametric fit. All regressions have a dummy variable indicating incumbency status of the candidate at t+1, a fourth order polynomial of margin of victory, their interactions with incumbency dummy and state and year fixed effects as the right hand side variables. Values with * are significant at 1% level of significance. Not all regressions have the same number of observations. For example, BJP did not exist in some states. As a result they were not used in the regressions.

TABLE 2b. Probability of Winning and Vote Share in t+1 and the Continuity of Predetermined Characteristics, 1991-2003

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	All			margin ≤ 25%			margin ≤ 0.25%			Parametric fit		
	Winner	Loser	Diff.									
Probability of Winning	0.52 (0.008)	0.36 (0.008)	0.16* (0.011)	0.49 (0.009)	0.40 (0.009)	0.09* (0.012)	0.39 (0.015)	0.52 (0.017)	-0.13* (0.022)	0.33 (0.033)	0.58 (0.022)	-0.247* (0.04)
Vote Share in t+1	40.1 (0.21)	31.8 (0.25)	8.3* (0.32)	38.9 (0.22)	33.5 (0.26)	5.4* (0.34)	36.8 (0.37)	37.5 (0.46)	-0.7 (0.59)	35.5 (0.82)	40.1 (0.77)	-4.6* (1.13)
Electoral Experience	1.40 (0.021)	1.24 (0.022)	0.16* (0.03)	1.41 (0.023)	1.31 (0.025)	0.1* (0.034)	1.43 (0.041)	1.40 (0.046)	0.03 (0.06)	1.38 (0.06)	1.40 (0.09)	-0.02 (0.11)
Political Experience	0.86 (0.016)	0.69 (0.016)	0.17* (0.023)	0.85 (0.018)	0.74 (0.018)	0.11* (0.026)	0.82 (0.031)	0.75 (0.033)	0.07 (0.046)	0.78 (0.064)	0.75 (0.056)	0.03 (0.085)
Vote Share In t-1	25.5 (0.356)	19.7 (0.353)	5.8* (0.5)	24.9 (0.38)	20.9 (0.4)	4.0* (0.55)	23.2 (0.66)	23.0 (0.74)	0.2 (0.99)	22.4 (1.06)	23.9 (1.28)	-1.5 (1.66)
Proportion of Female Candidates	0.04 (0.003)	0.03 (0.003)	0.01* (0.004)	0.04 (0.003)	0.04 (0.003)	0.0* (0.005)	0.04 (0.006)	0.04 (0.006)	0.0 (0.009)	0.032 (0.009)	0.033 (0.005)	-0.001 (0.011)
Proportion of INC Candidates	0.27 (0.007)	0.30 (0.008)	-0.03 (0.01)	0.28 (0.008)	0.32 (0.009)	-0.04* (0.011)	0.31 (0.014)	0.33 (0.016)	-0.02 (0.02)	0.33 (0.028)	0.31 (0.032)	0.02 (0.042)
Proportion of BJP Candidates	0.22 (0.006)	0.17 (0.006)	0.05* (0.009)	0.23 (0.007)	0.16 (0.007)	0.07* (0.01)	0.21 (0.012)	0.18 (0.013)	0.03 (0.018)	0.145 (0.021)	0.14 (0.021)	0.005 (0.03)
# of Registered Voters	138450 (1058)	139533 (1201)	-1082 (1594)	137678 (1117)	137762 (1289)	-83.6 (1697)	134501 (1802)	130741 (2169)	3760 (2801)	135537 (2032)	137042 (1660)	-1505 (2623)
# who Turned out to Vote	89687 (612)	88404 (630)	1283 (880)	89539 (652)	87478 (714)	2061* (966)	88474 (1160)	84906 (1336)	3568* (1767)	89145 (769)	89814 (865)	-669 (1157)
Rate of Turnout (%)	67.6 (0.21)	67.2 (0.22)	0.39 (0.3)	67.9 (0.22)	67.3 (0.25)	0.6 (0.35)	68.8 (0.4)	68.9 (0.47)	-0.1 (0.62)	68.3 (0.36)	68.4 (0.3)	-0.1 (0.47)
# of Candidates	13.0 (0.29)	14.1 (0.18)	-1.1* (0.35)	12.8 (0.16)	14.3 (0.20)	-1.5* (0.25)	12.8 (0.28)	13.2 (0.34)	-0.4 (0.43)	13.7 (0.48)	13.3 (0.35)	0.4 (0.59)
Proportion of Scheduled Casts	0.13 (0.005)	0.14 (0.006)	-0.01* (0.008)	0.13 (0.006)	0.13 (0.006)	0.0 (0.008)	0.13 (0.01)	0.14 (0.012)	-0.01 (0.015)	0.12 (0.014)	0.13 (0.02)	-0.01 (0.024)
Proportion of Scheduled Tribes	0.11 (0.005)	0.12 (0.005)	0.0* (0.007)	0.11 (0.005)	0.12 (0.006)	-0.01* (0.008)	0.11 (0.009)	0.12 (0.011)	-0.01 (0.015)	0.036 (0.007)	0.036 (0.01)	0.003 (0.012)
# of obs.	7821			6404			1977			7821		

Note: All comparisons are conditional on rerunning. Standard errors are in the parenthesis. Standard errors are clustered at the state level for the parametric fit. All regressions have a dummy variable indicating incumbency status of the candidate at t+1, a fourth order polynomial of margin of victory, their interactions with incumbency dummy and state and year fixed effects as the right hand side variables. Values with * are significant at 1% level of significance. Not all regressions have the same number of observations. For example, BJP did not exist in some states. As a result they were not used in the regressions.

TABLE 3a. Robustness Check of Incumbency Effects based on Alternative Specifications, 1975-1991

(1)	(2)	(3)	(4)	(5)	(6)
Independent Variables	Win Prob. in t+1 for Winners- Win Prob. in t+1 for Losers				Prob. Difference in t-1
	-0.18* (0.026)	-0.176* (0.026)	-0.179* (0.026)	-0.176* (0.042)	-0.0003 (0.0003)
Electoral Experience	No	Yes	No	Yes	Yes
Political Experience	No	Yes	No	Yes	Yes
Vote Share In t-1	No	Yes	No	Yes	Yes
Proportion of Female Candidate	No	Yes	No	Yes	Yes
Proportion of INC Candidates	No	Yes	No	Yes	Yes
Proportion of BJP Candidates	No	Yes	No	Yes	Yes
# of Registered Voters	No	No	Yes	Yes	Yes
# who Turned out to Vote	No	No	Yes	Yes	Yes
Rate of Turnout (%)	No	No	Yes	Yes	Yes
# of Candidates	No	No	Yes	Yes	Yes
Proportion of Scheduled Casts	No	No	Yes	Yes	Yes
Proportion of Scheduled Tribes	No	No	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Obs.	16699	16699	16699	16699	16699

Notes: All comparisons are conditional on rerunning. Standard errors are in the parenthesis. Standard errors are clustered at the state level for the parametric fit. All regressions also have a dummy variable indicating incumbency status of the candidate at t+1, a fourth order polynomial of margin of victory, their interactions with incumbency dummy. Values with * are significant at 1% level of significance.

TABLE 3b. Robustness Check of Incumbency Effects based on Alternative Specifications, 1991-2003

(1)	(2)	(3)	(4)	(5)	(6)
Independent Variables	Win Prob. in t+1 for Winners- Win Prob. in t+1 for Losers				Prob. Difference in t-1
	-0.246* (0.038)	-0.248* (0.038)	-0.245* (0.038)	-0.247* (0.038)	-0.008 (0.03)
Electoral Experience	No	Yes	No	Yes	Yes
Political Experience	No	Yes	No	Yes	Yes
Vote Share In t-1	No	Yes	No	Yes	Yes
Proportion of Female Candidate	No	Yes	No	Yes	Yes
Proportion of INC Candidates	No	Yes	No	Yes	Yes
Proportion of BJP Candidates	No	Yes	No	Yes	Yes
# of Registered Voters	No	No	Yes	Yes	Yes
# who Turned out to Vote	No	No	Yes	Yes	Yes
Rate of Turnout (%)	No	No	Yes	Yes	Yes
# of Candidates	No	No	Yes	Yes	Yes
Proportion of Scheduled Casts	No	No	Yes	Yes	Yes
Proportion of Scheduled Tribes	No	No	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Obs.	7821	7821	7821	7821	7821

Notes: All comparisons are conditional on rerunning. Standard errors are in the parenthesis. Standard errors are clustered at the state level for the parametric fit. All regressions also have a dummy variable indicating incumbency status of the candidate at t+1, a fourth order polynomial of margin of victory, their interactions with incumbency dummy. Values with * are significant at 1% level of significance.

TABLE 4. Losing Rerunners vs Losing Non-Rerunners: Comparison of Candidate Characteristics

(1)	(2)	(3)
	Difference=(Losing Rerunners – Losing Non-Rerunners)	
	1975-1991	1991-2003
Electoral Experience	0.0 (0.12)	0.01 (0.27)
Political Experience	0.07 (0.1)	0.0 (0.23)
Vote Share In t-1	5.0 (3.8)	5.0 (6.3)
Proportion of Female Candidate	-0.003 (0.025)	-0.023 (0.026)
Proportion of INC Candidates	0.1 (0.07)	0.05 (0.12)
Proportion of BJP Candidates	0.0 (0.005)	0.03 (0.05)

Notes: Standard errors are in the parenthesis. Standard errors are clustered at the state level for the parametric fit. All regressions have a dummy variable indicating incumbency status of the candidate at t+1, a fourth order polynomial of margin of victory, their interactions with incumbency dummy and state and year fixed effects as the right hand side variables. Observations with in margin of victory of 5% are considered.

TABLE 5. Comparison with Other Estimates

	1975-1991	1991-2003
Sophomore Surge	-6.6 (0.3)	-4.1 (0.38)
Retirement Slump	-9.1 (0.73)	-3.0 (0.72)
Gelman-King	-13.1 (0.25)	-11.1 (0.23)
RDD	-4.4 (0.71)	-4.6 (1.13)

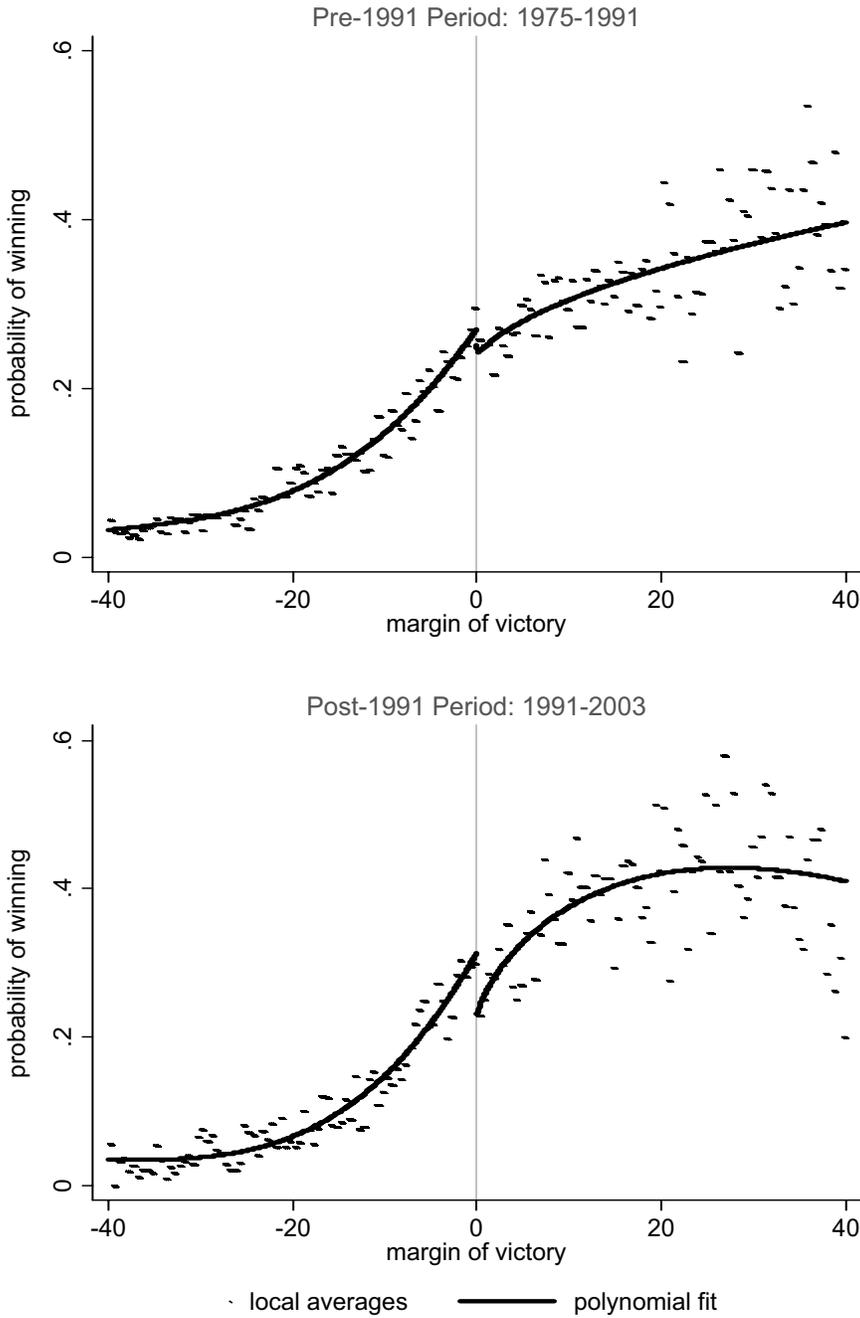
Notes: Standard errors are in the parenthesis. RDD estimates of incumbency effects based on the vote share are presented for the purposes of the comparison.

TABLE 6. Explaining Variation in Incumbency Disadvantage Across Indian States

(1)	(2)	(3)	(4)	(5)
Independent Variables	Dependent Variable: Probability Difference			
Literacy Rate	0.011** (2.17)	0.014** (2.57)	0.009** (2.15)	0.008*** (1.76)
Electricity	-0.004 (-1.55)	-0.005*** (-1.75)	-0.001*** (-0.64)	-0.001 (-0.54)
Political Competition		0.019 (1.57)	0.021** (2.50)	0.023** (2.48)
Turnout		-0.001 (-0.27)	-0.003 (-0.81)	-0.002 (-0.44)
Per Capita Income			-0.00007 (-1.56)	-0.0001*** (-1.88)
Population				-0.0000005 (-0.35)
Percent Urban				0.005 (1.16)
R-squared	0.18	0.28	0.39	0.44
Obs.	25	25	24	24

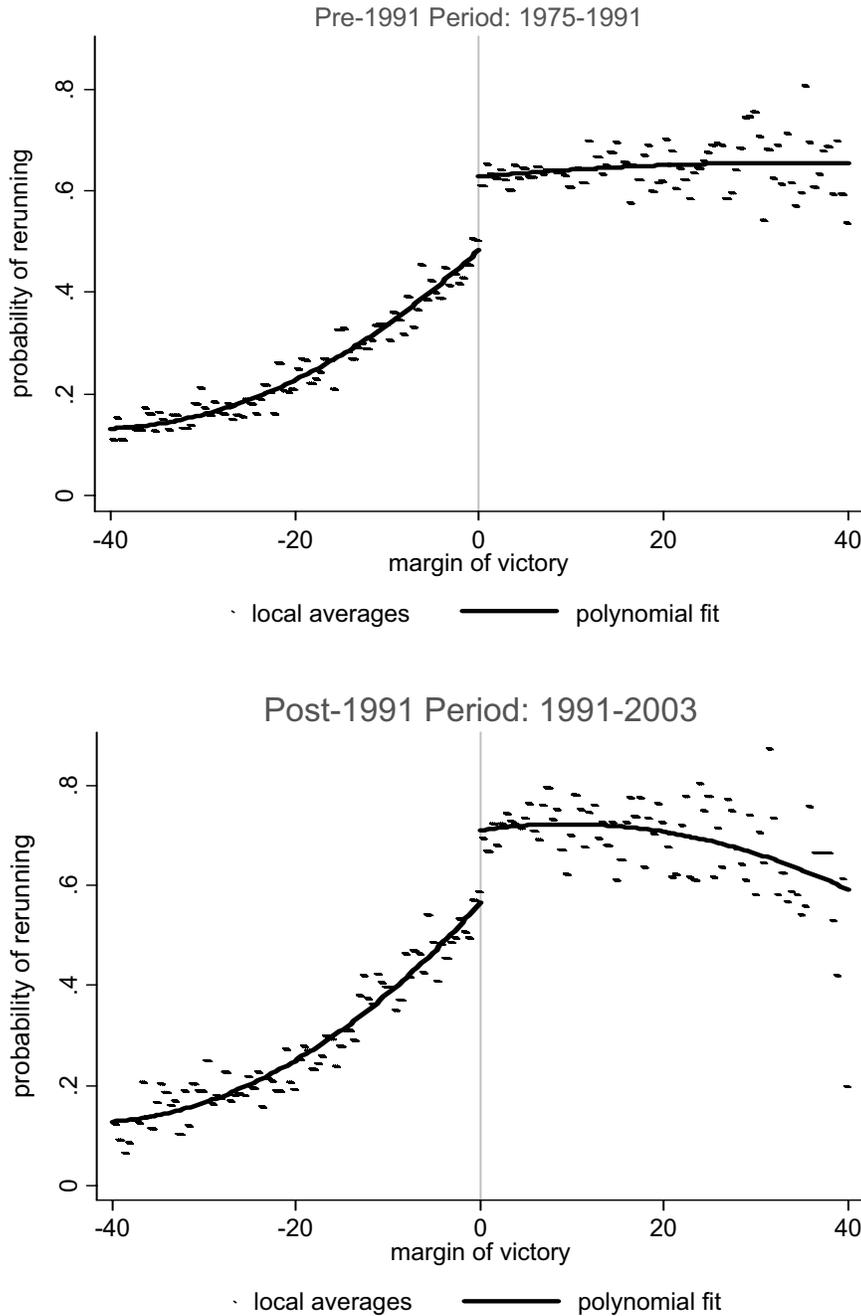
Notes: Dependent Variable is the difference in probability of winning between bare winners and bare losers. The unit of observation is a state. Literacy is percent of literates in a state. Electricity is percent of households having source of lighting as electricity. Political Competition is the difference between the vote shares of first-place and second-place parties. Percent Urban is the state-wise percentage of urban population. The parentheses show the t-ratios of the coefficients. The symbols *, ** and *** indicate statistical significance at 1%, 5% and 10% level, respectively.

FIGURE 1. Probability of Winning in t+1



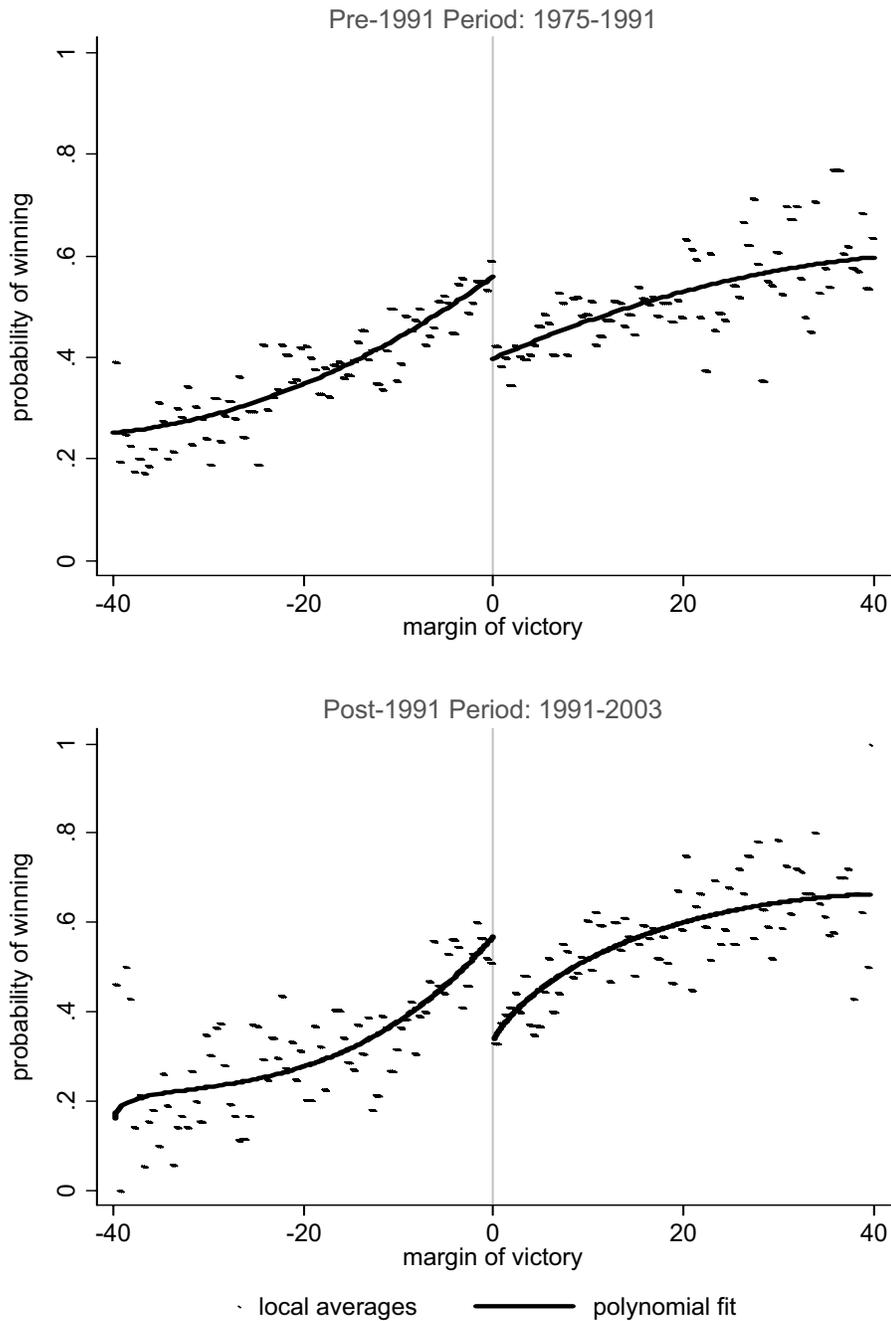
Notes: The local averages are proportion of winners in an interval of margin of victory of 0.5%. The parametric fit is the predicted probabilities from a logistic regression of an indicator variable for a win in t+1 on a dummy variable indicating incumbency status of the candidate at t+1, a fourth order polynomial of margin of victory, their interactions with incumbency dummy and state and year fixed effects. The values are predicted separately for the winners and the losers to find the size of discontinuity at the threshold. The entire sample, assuming the candidates who do not rerun lose, is considered.

FIGURE 2. Probability of Rerunning in t+1



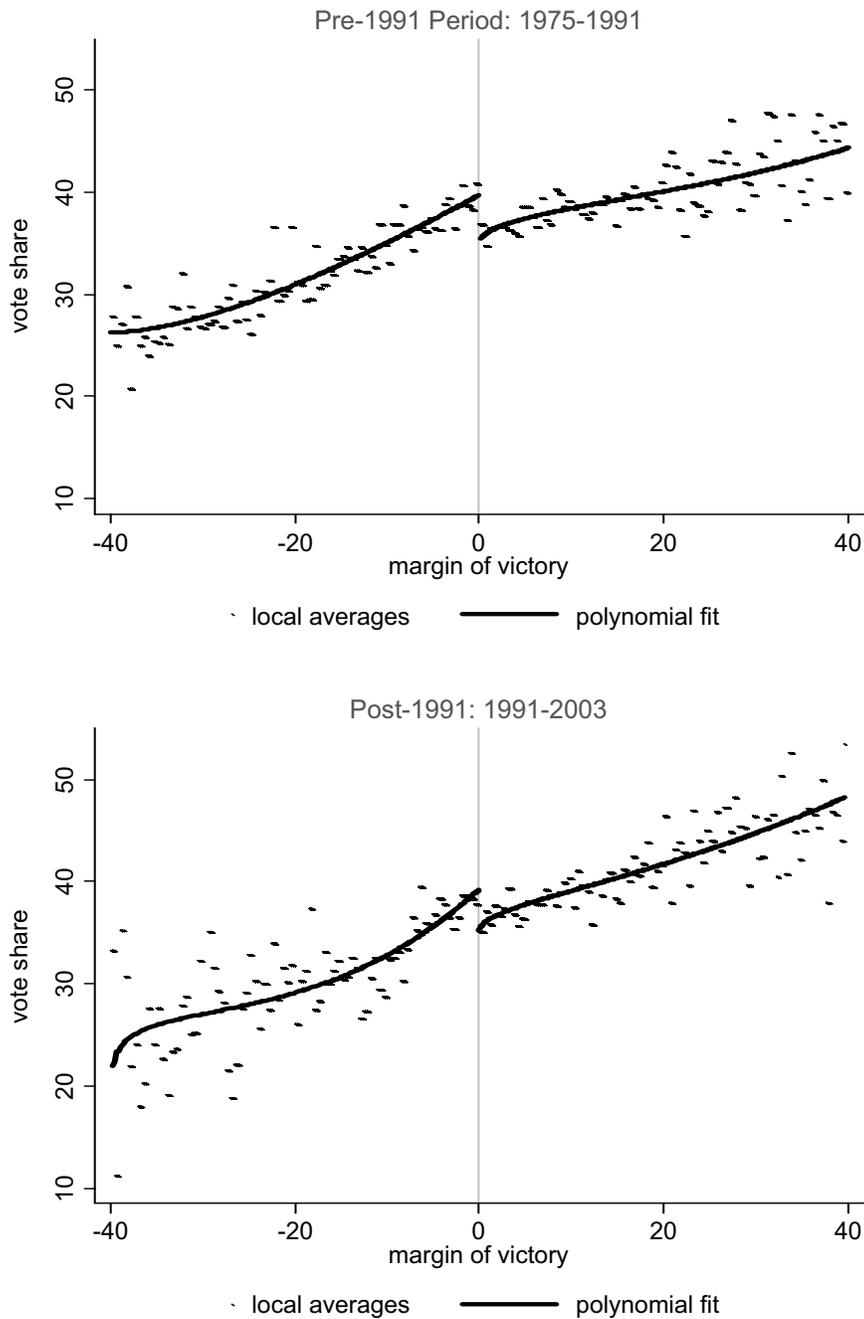
Notes: The local averages are the proportion of rerunners in an interval of margin of victory of 0.5%. The parametric fit is the predicted probabilities from a logistic regression of an indicator variable for rerunning in t+1 on a dummy variable indicating incumbency status of the candidate at t+1, a fourth order polynomial of margin of victory, their interactions with incumbency dummy and state and year fixed effects. The values are predicted separately for the winners and the losers to find the size of discontinuity at the threshold.

FIGURE 3. Probability of Winning in t+1, Conditional on Rerunning



Notes: The local averages are proportion of winners in an interval of margin of victory of 0.5%. The parametric fit is the predicted probabilities from a logistic regression of an indicator variable for a win in t+1 on a dummy variable indicating incumbency status of the candidate at t+1, a fourth order polynomial of margin of victory, their interactions with incumbency dummy and state and year fixed effects. The values are predicted separately for the winners and the losers to find the size of discontinuity at the threshold.

FIGURE 4. Vote Share in t+1, Conditional on Rerunning



Notes: The local averages are average vote shares in an interval of margin of victory of 0.5%. The parametric fit is the predicted vote shares from a regression of vote share in t+1 on a dummy variable indicating incumbency status of the candidate at t+1, a fourth order polynomial of margin of victory, their interactions with incumbency dummy and state and year fixed effects. The values are predicted separately for the winners and the losers to find the size of discontinuity at the threshold.