The Curious Curse of the Absentee Voter: New Evidence on Political Parties and Candidate Selection from India

Arghya Banerjee[*] Somdeep Chatterjee[†] Manhar Manchanda[‡]

September 4, 2025

Abstract

This paper examines the role of candidate popularity in shaping political parties' candidate selection process. Using data on Indian state elections, we proxy elite voter support using a candidate's performance on absentee (*postal*) ballot performance and exploit a regression discontinuity design comparing candidates who narrowly win versus narrowly lose the postal vote, conditional on overall victory. We find that candidates who narrowly win both in-person and postal ballots are 18–24 percentage points less likely to be re-nominated to contest the subsequent election. This effect is not explained by voter preferences, as re-election rates remain unchanged, indicating a party-driven response. The effect is stronger in centralized parties and for candidates with higher education or a criminal record.

***Keywords*:** Remote Voting; Political Parties; Regression Discontinuity; Assembly Elections

***JEL Classification*:** D72; D73; O12; H77

[*]Economics Group, Indian Institute of Management Calcutta. Email: arghyab23@iimcal.ac.in

[†]Economics Group, Indian Institute of Management Calcutta. Email: somdeep@iimcal.ac.in

[‡]Economics Group, Indian Institute of Management Calcutta. Email: manharmfp21@email.iimcal.ac.in

1 Introduction

The issue of political selection and candidate choice in elections is fundamental to understanding eventual policy positions and decisions within an economy run by democratically elected representatives. However, surprisingly little is known about how candidates are chosen to run in an election. While some studies have looked at candidates' own decisions to contest elections (Dal Bó and Finan, 2018, Besley, 2005), the literature on how political parties select candidates is much thinner (Magesan et al., 2024). The few papers that do study this aspect tend to find that very popular candidates sometimes are considered a threat to party leadership and therefore, despite the fact that parties at the end like to win, such candidates may not be fielded (Chandra, 2016, Mattozzi and Merlo, 2015, Caillaud and Tirole, 2002, Magesan et al., 2024, Galasso and Nannicini, 2011, Besley et al., 2017).

However, the nature of *popularity* of a given candidate that can cause such a large enough *threat* to the centralized party leadership to result in not being nominated to run in the subsequent election under a party ticket, is relatively abstract. For instance, some theories on party organization and internal structures suggest that superior quality of candidates are less likely to share returns from winning (Caillaud and Tirole, 2002) and provide disincentives for other potential nominees to put in efforts (Mattozzi and Merlo, 2015) and therefore selecting such candidates is costly for the party leadership. Other papers such as Magesan et al. (2024), which is the most comprehensive empirical paper in this domain, study the supply-side of candidates in a structural estimation framework to find that parties do select candidates that can win, but prefer those who can win with minimum support, or in other words, are less popular.

This suggests that for a setting like India, which is the world's largest democracy and has a highly centralized party system but no term limits, candidates who seek to remain in office through continually contesting in elections may potentially be taxed if they become too popular. In this paper, we explore an alternative popularity curse for such candidates

in Indian elections, the curse of the *absentee voter*. Absentee or mail-in voting is a widely used form of *convenience* voting– an alternative to precinct-place voting that offers greater accessibility to voters (Gronke et al., 2008). Non-precinct-place voting methods are prevalent across the world, with nearly 40% of democratic nations offering the option of absentee balloting, with varying eligibility restrictions.[1]

From a theoretical and policymaking standpoint, the provision of a convenience voting option is directed at lowering the individual cost of voting to facilitate higher aggregate turnout levels (Riker and Ordeshook, 1968, Aldrich, 1993). While much of the focus with respect to absentee voting has been restricted to electoral participation (Gronke et al., 2007, 2008, Oliver, 1996, Stein and Vonnahme, 2008), little is known about whether convenience voting has meaningful electoral consequences. To assess whether popularity among absentee voters can be potentially detrimental for the candidates seeking to re-contest under party nominations and therefore be a source of potential threat to party leadership's strength, we are interested in whether an incumbent's performance in *postal* ballots– the official term for absentee ballots in India– affects their electoral trajectory in subsequent elections.

On average, approximately 700 postal ballots are cast in Indian state legislative assembly elections, while the vote margin between the top two candidates is around 17,800. Clearly, postal ballots alone cannot numerically alter the election outcome, but they can serve as a critical signal of a candidate's electoral strength. This is a particularly relevant signal in the Indian context, as postal voters primarily consist of government officials, bureaucrats, and army personnel, partially representing the preferences of an elite segment of the electorate.

We estimate how a candidate's *joint* victory in precinct-based and postal ballots influences their subsequent candidacy outcomes. Empirically, the relationship between candidacy outcomes and performance in postal ballots is fraught with endogeneity concerns. In general, candidates who rank first on in-person ballots and also secure the highest number of

[1] See country-wise provisions and regulations around absentee or mail-in ballots here.

postal votes may systematically differ from overall winners who are ranked lower on postal ballots. The unobserved differences between these two types of winners could directly affect subsequent candidacy outcomes, leading to biased estimates. Additionally, joint winners could possess greater political experience, higher precinct-based vote shares, and stronger political party support compared to winners who lost on the postal ballots, leading to a problem of omitted variable bias.

To alleviate these concerns, we employ a regression discontinuity (RD) design using the difference in the postal vote shares of the in-person ballot winner and runner-up candidate as the running variable. By construction, a positive value of this running variable corresponds to a joint winner (which we term as *postal-aligned*), and a negative value implies an overall election winner who ranked lower on the postal ballots (*postal-misaligned*). Therefore, our RD design compares subsequent outcomes of *barely* postal-aligned and *barely* postal-misaligned winners. Note that in our setting, treatment assignment does not correspond to the winning status of the candidate- instead, an overall winner becomes *postal-aligned* (*misaligned*) as the running variable shifts to a positive (negative) value. Conditional on a set of continuity assumptions, the treatment assignment into postal alignment or misalignment is as good as random in a small neighborhood around the zero cutoff of the running variable.

Our RD estimates indicate that a postal-aligned winner is 18–24 percentage points less likely to contest in the next Assembly election. This result remains robust to alternative bandwidth choices, the inclusion of state and election-year fixed effects, and different placebo tests. We additionally find no statistically significant difference in the respective re-election probabilities of the two types of incumbents, suggesting that voters do not differentially perceive barely postal-aligned and barely postal-misaligned candidates. This suggests that political parties may strategically limit the presence of popular candidates- in this case, the postal-aligned winners, in subsequent elections. We conjecture several potential reasons for this result.

First, parties with centralized leadership may perceive popularity among *elite* voters, who cast their votes through postal ballots, as candidates' individual charisma beyond party popularity and, therefore, a potential threat to the party leadership's command over the party. Second, parties may worry that such candidates may defect or not tow the party line and/or demand higher rents in terms of recognition or important positions owing to their increased popularity, which the party leader may not be willing to concede. Third, parties may prefer newer or weaker candidates who marginally win, in line with Magesan et al. (2024), as they may be more likely to attribute their success to party symbol rather than individual appeal and therefore will offer greater allegiance to the party leader.

An alternate explanation for our reduced form empirical results could be that the drop in re-contesting probability is driven by differences in the performance of postal-aligned and postal-misaligned incumbents. Perhaps the postal-aligned candidates actually perform poorly, and therefore, the party simply removes them based on *non-performance*. We are able to rule out this possibility as we find no significant effects on observed economic development outcomes, proxied by nighttime light growth during the electoral term or on the provision of various categories of public goods.

Another alternate explanation would be that these candidates are actually being rewarded with a promotion to higher levels of political office, such as nominating them to contest national elections. We also rule this out as we do not find any evidence that postal-aligned candidates in the assembly constituencies are any more likely to contest in the subsequent general elections to the parliament of India than postal-misaligned incumbents. Furthermore, we find heterogeneous effects based on party and candidate characteristics. We find that our results are more pronounced for parties with a relatively more centralized organizational structure and candidates with (i) a higher educational qualification and (ii) a criminal record.

The fact that postal ballots cannot realistically affect election outcomes is also empirically meaningful for us. Effectively, this implies that our estimates are a local average treatment

effect (LATE) measuring the political parties' response to a "pure" popularity shock for a given incumbent winner. Alternate measures of popularity, such as actual vote shares, can potentially be confounded with issues like winnability and candidates' organizational ability to conduct last-mile get-out-the-vote to increase turnout or effects of last-minute campaigns - none of which can affect absentee votes. Therefore, one major contribution of our paper is to provide one of the first credible causal estimates of the impact of candidate popularity on candidate selection.

Additionally, this paper contributes to two other strands of the literature. *First*, it contributes to the vast literature on candidate selection (Besley, 2005, Jia et al., 2015, Gulzar, 2021, Dal Bó and Finan, 2018). Within this literature, our paper emphasizes the strategic role of political parties in the candidate selection process (Gulzar et al., 2021, Casey et al., 2021, Magesan et al., 2024). Our findings identify a new metric of candidates' popularity and electoral strength as pivotal considerations in the nomination process employed by political parties. *Second*, our paper contributes to the literature on the electoral consequences of alternative voting technologies (Fujiwara, 2015, Ujhelyi et al., 2021, Chatterjee and Kamal, 2021, Desai and Lee, 2021)- in this case, the absentee ballot option (Gronke et al., 2008). To the best of our knowledge, our paper is the first causal study linking performance in absentee ballots to candidacy outcomes.

The rest of the paper is structured as follows. Section 2 briefly outlines the institutional context (Section 2.1), followed by a description of the data sources and a presentation of the summary statistics (Section 2.2). Section 3 details our primary empirical strategy, while Section 4 offers a host of analyses to assess the validity of the RD design. Section 5 presents the main findings on candidacy outcomes, along with several robustness checks (Section 5.1). In Section 6, we explore potential mechanisms underlying our results. Finally, Section 7 concludes.

2 Background and Data

2.1 Institutional Context: Indian Assembly Elections and Postal Ballots

India's electoral process follows a plurality rule (or first-past-the-post) voting system, where the candidate with the highest vote count in a constituency wins, regardless of whether they secure an absolute majority. Citizens directly elect members to the *Lok Sabha* (General Assembly or the Parliament) and the *Vidhan Sabha* (State Legislative Assembly), with the President and members of *Rajya Sabha* (Upper House of the People) being appointed by elected members. The elections are conducted by the Election Commission of India (ECI), which is the apex constitutional body established under Article 324 of the Indian Constitution. Elections across different tiers of government follow a fixed five-year electoral cycle, except in cases of early Assembly dissolution, a vote of no confidence, or constitutional emergencies. The Representation of the People Act, 1951 allows both party-nominated as well as independent candidates to contest elections.

Postal Ballots. The Legislative and General Assembly elections additionally provide voters with the option of 'postal' (or mail-in) voting. Eligible postal voters include the following voter categories: (i) members of the armed forces, paramilitary forces, and government employees deployed on election duty away from their home constituencies, (ii) individuals detained under preventive custody orders during the election period, (iii) government officials and polling staff assigned duties at polling stations other than their own constituencies, (iv) media personnel with authorization letters from the Election Commission of India and those involved in essential services such as metros, railways, and healthcare, and (v) as of October 2019, Persons with Disabilities (PwDs) aged 80 or above. Postal ballots have to be returned to the election officer within a stipulated time period, which are then counted 30 minutes before the commencement of the counting process of in-person votes. Unlike countries like the U.S., mail-in ballots in India can be utilized only by a restricted segment of the electorate.

2.2 Data and Summary statistics

Elections Data. We collect electoral information from the "Detailed Results" statistical reports of the Election Commission of India (ECI). These reports provide constituency-level election data such as the number of electors, number of polling booths, number of contesting contesting candidates, the overall voter turnout, and the total number of postal votes polled. In addition, we make use of candidate-level information from these reports which allows us to identify the overall winner in a particular constituency, and their corresponding postal vote shares. We additionally utilize the unique candidate identifiers provided in the *Lok Dhaba* elections dataset of the Trivedi Centre for Political Data (TCPD), Ashoka University to track candidates across elections. Our sample period spans all state Assembly elections conducted between 2013 to 2017. We additionally make use of information from the 2008-12 electoral cycle for balance checks, and the 2018-22 election cycle for examining subsequent candidacy outcomes.[2]

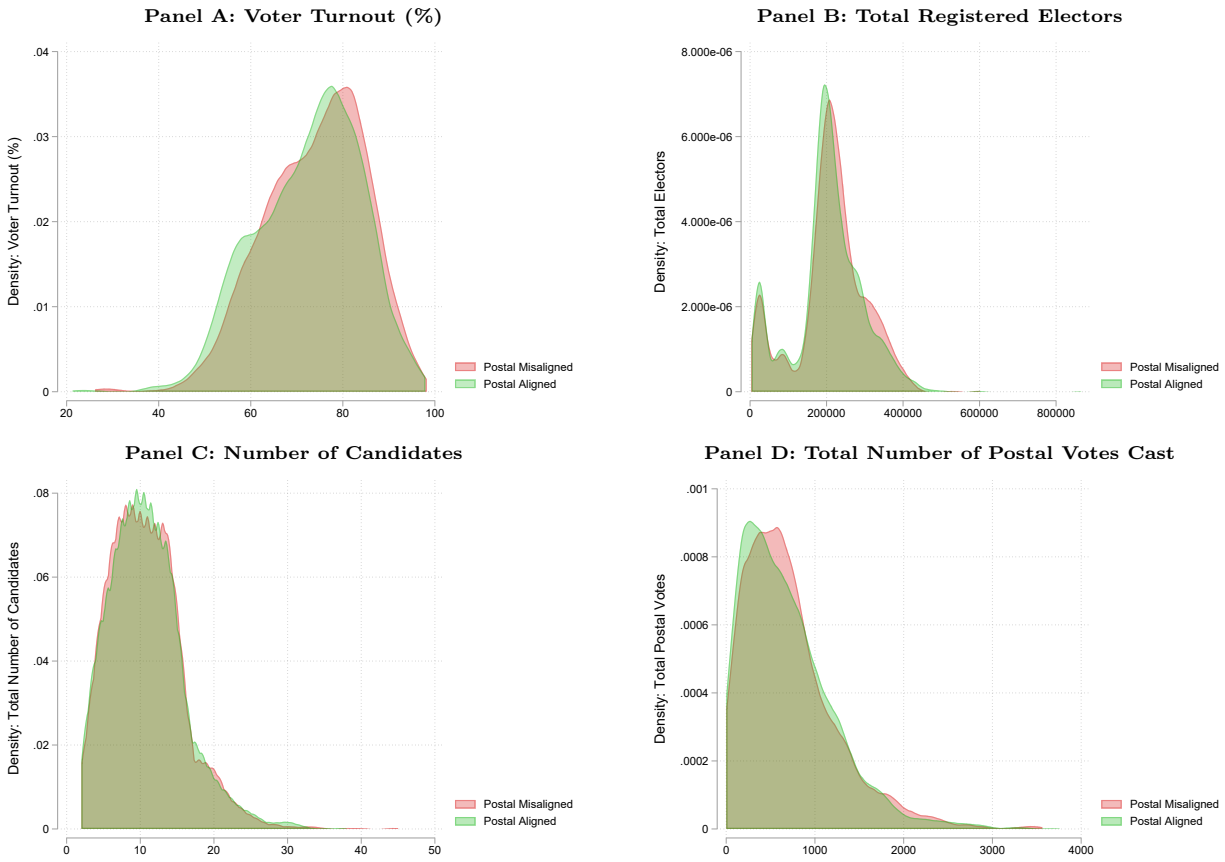
Incumbent Performance. To examine incumbent performance, we utilize granular satellite imagery data on nighttime lights (NL) to proxy for economic activity within the constituency boundaries. The data for this comes from the Socioeconomic High-Resolution Rural Urban Geographic (SHRUG) data portal (Asher et al., 2021), which provides NL data at the Assembly constituency level for the period 2012-21. Given our election sample period of 2013–2017, this allows us to observe NL trends over the full electoral term for each constituency. Further, we use information on village-level public goods from the Mission Antyodaya (MA) survey conducted by the Government of India in 2020. The MA initiative is an effort of the Government of India that surveyed nearly 650,000 villages across India to conduct gap analyses and provide inputs to participatory plans of village councils (Gram Panchayats or GPs), the lowest level of government in the country. This survey is routinely utilized as a replacement for the village-level directory component of the national-level

[2] A key reason for restricting the sample period is the nationwide delimitation exercise conducted in 2008, which redrew constituency boundaries, making it nearly impossible to compare constituencies before and after delimitation.

Census, the last round of which was conducted in 2011.[3]

Descriptive Statistics. We present the summary statistics of the key electoral variables in Table A1, disaggregated by constituencies with and without a postal-aligned winner. We find that most of our variables are fairly balanced across the two types of constituencies, especially when we further disaggregate the descriptive statistics by constituencies where the winner was barely postal-aligned and those where the winner was barely postal-misaligned. Figure 1 displays the sample distributions of key electoral variables, disaggregated by the postal alignment status of the winning candidate. The mean postal votes cast in a constituency during our sample period is 698.98, while the average number of valid in-person votes cast stood at 14,463.

Figure 1: Densities of Key Electoral Variables: By Postal Alignment



[3]For the analysis with respect to public facilities recorded in 2020, we use the 2015-19 electoral period- that is, the elected representative holds office in 2019-20 for each constituency

3 Empirical Strategy

We are interested in estimating the impact of a candidate's joint victory in in-person as well as mail-in ballots on their subsequent candidacy outcomes. As outlined in the Introduction, this question is empirically challenging for several reasons. This is primarily due to the fact that the relationship between candidacy outcomes and performance in postal ballots is potentially mired in endogeneity concerns. In general, candidates who rank first on in-person ballots and also secure the highest number of postal votes may systematically differ from winners who are ranked lower on postal ballots. The unobserved differences between these two types of winners could directly affect subsequent candidacy outcomes, leading to biased estimates. Additionally, joint winners could possess greater political experience, higher precinct-based vote shares, and stronger political party support compared to winners who lost on the postal ballots, leading to a problem of omitted variable bias.

We address these concerns by employing a regression discontinuity (RD) design, comparing winners who narrowly secured the highest number of postal ballot votes with winners who narrowly fell short. The running variable in our setting is then defined as follows:

$$postal\_alignment_{i,j,t} = v^{P}_{1,j,t} - v^{P}_{2,j,t} \tag{1}$$

where $v^{P}_{1,j,t}$ is the postal vote share of the overall winner in constituency j in election t, and $v^{P}_{2,j,t}$ is the postal vote share of the overall runner-up.[4] By construction, this variable varies between -100 and $+100$, and a positive value corresponds to a joint winner (postal-aligned) while a negative value implies a postal-misaligned winner. Therefore, our primary RD specification is as follows:

$$Y_{i,j,t+1} = \alpha + \beta\tau_{i,j,t} + f(postal\_alignment_{i,j,t}) + \epsilon_{i,j,t+1} \tag{2}$$

[4] In our context, a candidate is considered a *winner* if their total votes (in-person and postal combined) are the highest among all contenders. In contrast, the *overall runner-up* is the candidate who secures the second-highest total vote count in the final election results. These labels are irrespective of the candidates' isolated rank on the postal ballots.

where

$$\tau_{i,j,t} = \begin{cases} 1 & \text{if} \quad postal\_alignment_{i,j,t} \geq 0 \\ 0 & \text{if} \quad postal\_alignment_{i,j,t} < 0 \end{cases} \tag{3}$$

In equation (2), $Y_{i,j,t+1}$ is the candidacy outcome in the election held at $t+1$ for winner i elected from constituency j at t. Further, $f(postal\_alignment_{i,j,t})$ is a flexible polynomial function of the forcing variable, as defined in (1). The probability that the winner i in Assembly constituency j is postal-aligned, $\Pr(\tau_{i,j,t} = 1 \mid postal\_alignment_{i,j,t})$, faces a discontinuous jump at $postal\_alignment_{i,j,t}) = 0$. The treatment assignment rule in our setting is then completely described by equation (3), with $\tau_{i,j,t}$ serving as our main treatment variable. Our primary identifying assumption is that no other variable witnesses a discontinuous jump at this zero cutoff. Under a continuity assumption on the forcing variable, the assignment of a winner into treatment (postal alignment) can be considered to be as good as random.

We non-parametrically estimate (2) using a local linear regression technique with a triangular kernel and employ the optimal bandwidth selection procedure of Imbens and Kalyanaraman (2011) to select the optimal neighborhood, h, around the cutoff. As robustness, we also report results for half the optimal bandwidth ($h/2$), double the optimal bandwidth ($2h$), and the optimal bandwidth derived using the procedure by Calonico et al. (2014).

4 Validity of the RD design

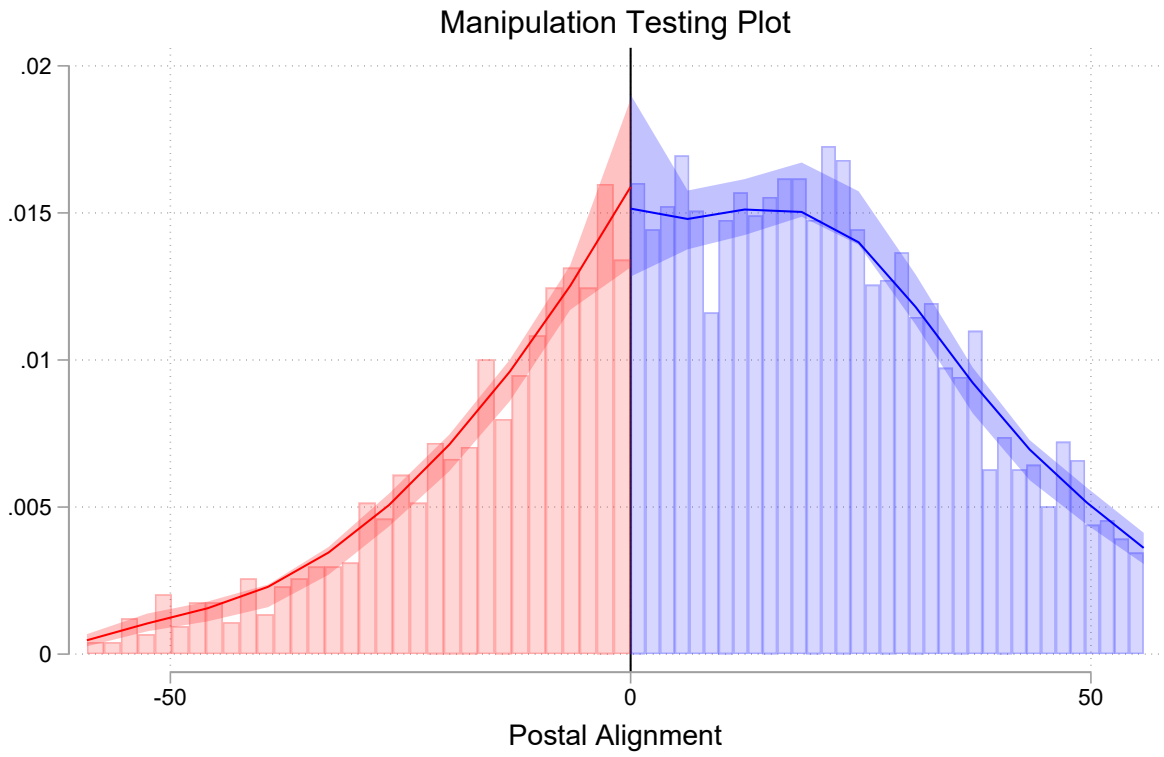
In this section, we conduct a wide range of tests to assess the validity of our RD design. These include examining (i) the continuity of the forcing variable, and (ii) the smoothness of pre-determined covariates around the cutoff.

4.1 Continuity of the Running Variable

First, we conduct a density test to check for a discontinuity at the cut-off in the density of the forcing variable. If winning candidates have control over choosing which side of the threshold they end up on- i.e., in our setting, they can *choose* to be postal-aligned or misaligned-, this would invalidate the RD design by introducing a selection bias, owing to potential manipulation of the forcing variable. In other words, we test whether winners whose *postal* margins are close to zero become disproportionately 'postal-aligned'. We use the manipulation density test proposed by Cattaneo et al. (2020) to test for any manipulation around the cutoff.

Figure 2: Density Test à la Cattaneo et al. (2020)

Robust t-stat $= -0.2390$ (p-value $= 0.8111$)



Notes: This figure presents the density plot using the local polynomial density method suggested by Cattaneo et al. (2020) to estimate the continuity of the running variable at the threshold (zero). The robust bias-corrected density test yields a small t-stat of -0.2390 (with an associated p-value of 0.8111), which fails to reject the null of manipulation in the running variable.

The graphical plot for this test, along with the robust t-stat and associated p-value, is presented in Figure 2. We find that the discontinuity estimate is small and statistically insignificant: t-statistic $= -0.2390$ (p-value $= 0.8111$). This provides evidence against any potential manipulation in the running variable around the threshold of zero. As a robustness exercise, we additionally present results from the density test of McCrary (2008) in the Appendix (see Figure A1). We continue to find no evidence of a discontinuity in the running variable at the zero cutoff.

4.2 Balance on Pre-Determined Characteristics

Next, we examine whether pre-determined constituency-level characteristics exhibit a discontinuity at the cut-off. A key identifying assumption of our RD design is that no other pre-determined variable, apart from the treatment, should exhibit a discontinuous jump at the cut-off of our running variable. To verify the validity of this assumption, we estimate equation (2) using a set of pre-determined constituency- and candidate-level characteristics from the previous election ($t-1$) as outcome variables. Table 1 presents the RD esimtates for the impact of the postal alignment status of the winner on covariates from the previous election at $t-1$. All RD estimates on our pre-determined covariates are small and statistically indistinguishable from zero. Of particular importance are the previous election's winner's postal alignment status and postal vote share difference. We find no evidence of a discontinuity in the running variable from the previous election, nor a significant effect on the previous winner's postal alignment status at $t-1$. This alleviates concerns that constituencies featuring a postal (mis-)aligned winner in the current election have historically exhibited a skewed postal vote distribution in favor of (or against) eventual winners. Figure A2 presents the graphical analog of these RD estimates, wherein we continue to find no evidence of a discontinuity in these characteristics around the cutoff.

Following the recent literature on covariate balance in RD setups, we additionally employ the permutation-based test by Canay and Kamat (2018), which compares the empirical

Table 1: Balance Test: RD Estimates

	Conventional RD Estimation				Canay and Kamat (2018)	
Variable	RD Estimate	h	N_h	p-value	N_q	p-value
Log(Electors) $(t-1)$	-0.121 (0.135)	3.662	414	0.368	138	0.594
Voter Turnout $(t-1)$	-2.413 (1.821)	7.182	801	0.185	138	0.267
Number of Candidates $(t-1)$	-0.475 (0.938)	6.365	705	0.613	138	0.249
Male Winner $(t-1)$	0.012 (0.069)	2.397	275	0.857	138	0.413
Winning Margin $(t-1)$	-1.637 (1.352)	7.384	828	0.226	138	0.139
Effective # of Candidates $(t-1)$	-0.083 (0.268)	3.843	430	0.755	138	0.415
SC/ST Reserved Constituency	0.147 (0.117)	2.882	344	0.209	144	0.252
# of Turncoats $(t-1)$	0.126 (0.195)	2.668	291	0.518	136	0.864
# of Graduate Candidates $(t-1)$	0.038 (0.195)	4.255	458	0.894	136	0.292
Postal Alignment $(t-1)$ (running variable)	-0.109 (0.090)	2.213	254	0.225	134	0.216
Postal Alignment Status of Winner at $t-1$	-0.084 (0.126)	3.091	335	0.504	136	0.861

Notes: The table presents the RD estimates for pre-determined constituency-level characteristics estimated for the optimal bandwidth (h) employing the bandwidth selection procedure of Imbens and Kalyanaraman (2011), and a triangular kernel. N_h represents the effective number of observations within the selected optimal bandwidth. Additionally, we also present the p-value and the effective sample size from the permutation-based test of Canay and Kamat (2018).

cumulative density function of pre-determined covariates on both sides of the cut-off. This method has the advantage of being asymptotically valid for small sample sizes, and is additionally able to detect discontinuities in situations where features of a covariate's distribution other than the mean may be discontinuous at the threshold. The results from this permutation-based test presented in Table 1 mirror the conventional RD estimates, and

showcase no discontinuities in the general distributions of pre-determined characteristics around the cutoff.

5 Main Results: Candidacy Outcomes in $t + 1$

In this section, we present our main results from estimating equation (2) using two outcome variables: (i) a binary indicator for whether the incumbent recontests the subsequent election at $t + 1$, which takes the value 1 if the winner in the 2013–2017 election cycle recontests in the 2018–2022 cycle, and 0 otherwise; and (ii) a binary indicator for whether the incumbent is re-elected in the subsequent election.[5] Table 2 presents the results from this estimation. Column (1) presents the local linear RD estimates from using the optimal bandwidth, h, derived from the bandwidth selection protocol of Imbens and Kalyanaraman (2011). As robustness, we additionally present estimates from using different factors of this bandwidth in Columns (2) and (3). Finally, in Column (4), we document the RD estimates from employing the optimal bandwidth selection procedure of Calonico et al. (2014).

Across all specifications, we find that a *barely* postal-aligned winner is nearly 20–25 percentage points less likely to contest the subsequent election compared to a *barely* postal-misaligned winner (see Panel A). Relative to the control mean, this effect corresponds to a 35.6% lower likelihood of the postal-aligned incumbent recontesting the next election. The effect remains fairly stable and statistically significant across different factors of the optimal bandwidth. On the other hand, the estimates in Panel B of Table 2 indicate no significant difference between the re-election probabilities of postal-aligned winners and postal-misaligned winners. This suggests that voters may not differentiate between barely postal-aligned and barely postal-misaligned candidates; instead, it appears that political

[5] Following Anagol and Fujiwara (2016), we consider the winner's unconditional re-election probability—i.e., irrespective of their decision to recontest—to address endogeneity concerns related to recontesting decisions. Accordingly, the re-election outcome is coded as 1 if the candidate secures first place in the subsequent election and 0 if they either fail to win or do not contest the election.

Table 2: Effect of Postal Alignment on Recontesting and Re-Election Probability at $t+1$

	Panel A: Recontesting at $t+1$			
	(1)	(2)	(3)	(4)
	h	$0.5 * h$	$2 * h$	CCT(h)
Postal Aligned	-0.264***	-0.215*	-0.201***	-0.085*
	(0.097)	(0.126)	(0.070)	(0.051)
Bandwidth size	2.913	1.456	5.826	12.239
Effective Obs.	347	169	693	1348
Control Mean	0.741	0.819	0.706	0.702
	Panel B: Re-election at $t+1$			
	h	$0.5 * h$	$2 * h$	CCT(h)
Postal Aligned	0.019	0.014	-0.035	-0.058
	(0.118)	(0.169)	(0.079)	(0.057)
Bandwidth size	3.149	1.574	6.299	11.980
Effective Obs.	375	184	740	1323
Control Mean	0.381	0.396	0.347	0.332

Notes: Local linear RD regression results are presented using different factors of the optimal bandwidth, h, à la Imbens and Kalyanaraman (2011) (in Columns (1)–(3)) and the optimal bandwidth, CCT(h), by Calonico et al. (2020). In all columns, the outcome in Panel A is a binary variable, equaling 1 if the incumbent recontests at $t+1$ and zero otherwise; in Panel B, the outcome is a binary variable indicating the incumbent's unconditional probability of re-election in the election at $t+1$. *Postal Aligned* is our treatment variable, $\tau_{i,j}$, from equation (3). Effective Obs. represents the number of observations within the optimally chosen bandwidth. Control mean is the mean of the dependent variable of the control group (left of the cutoff) within the optimal bandwidth.
Significance levels: $^{***}p < 0.01$;$^{**}p < 0.05$;$^{*}p < 0.1$

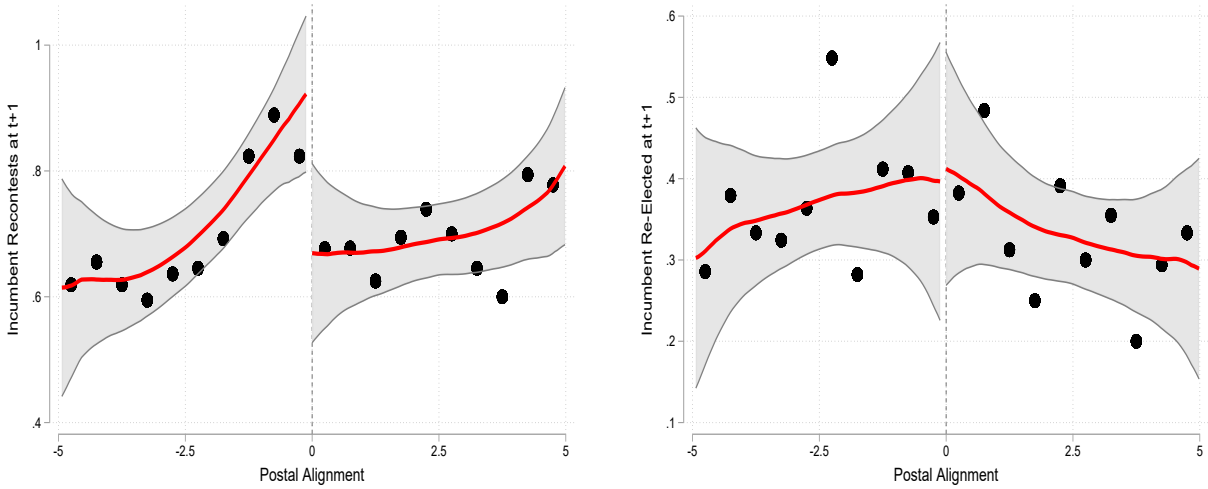
parties treat popular (postal-aligned) incumbents differently, as evidenced by the estimates corresponding to the incumbent's re-contesting decision. Figure 3 presents a graphic analog of the results presented in Table 2.

5.1 Robustness Exercises

In this section, we assess the sensitivity of our main result on the incumbent's recontesting decision presented in Table 2 to a host of alternative specifications. In particular, we

Figure 3: Impact of Postal Alignment on the Incumbent's Candidacy Outcomes at $t+1$

(a) Recontesting at $t+1$ (b) Re-election at $t+1$



Notes: The figure plots the two primary outcomes, the incumbent's recontesting probability and re-election probability at $t+1$, against our running variable, *postal alignment*, defined as the postal vote share difference between the overall winning candidate and the overall runner-up. A positive value of the running variable implies a postal-aligned winner; a negative value implies a postal-misaligned winner. Each dot in the figure depicts the averages over successive bins of 0.5% of the running variable. The curves are local linear regression (with 95% confidence intervals) fit separately for the positive and negative parts of the cut-off. The optimal bandwidth à la Imbens and Kalyanaraman (2011) is used to arrive at the figures.

conduct additional analyses to examine the robustness of our main result to (i) incremental expansions of the optimal neighborhood around the cutoff, (ii) placebo cutoffs, (iii) the inclusion of state, year, and district fixed effects, and control variables, and (iv) the inclusion of particular states in the sample.

Alternative Bandwidths. In Table 2, we presented RD estimates using the optimal bandwidth (h), half the optimal bandwidth ($0.5 * h$), and two times the optimal bandwidth ($2 * h$). While moving farther away from the threshold potentially introduces a higher bias in the estimates, it is important to assess the stability of coefficients across different factors of the bandwidth, particularly to assuage the concern that the results may be specific to a select set of neighborhoods around the threshold. In Figure A3 in the Appendix, we present the point estimates (and the 90% confidence intervals) by varying the bandwidth between 0.1 and 2 of the optimal bandwidth, h, in increments of 0.1. The coefficient patterns remain

fairly stable, statistically significant, and negative for our main outcome as h increases. This provides us with further confidence that our estimates are not particularly driven by a select set of bandwidths around the zero cutoff.

Placebo Cutoffs. Next, we run our baseline analysis for the incumbent's recontesting decision in $t+1$ using placebo thresholds. In particular, we estimate the treatment effect at placebo cutoffs in increments of 5 percentage points on either side of zero. In our context, the interpretation of a placebo threshold is as follows: Candidates who narrowly secure a postal ballot win- for instance, within 5 percentage points of the overall runner-up's postal vote share- may not be perceived by political parties as particularly popular. In other words, parties might view candidates who barely lost in postal ballots as indistinguishable from those who narrowly won in terms of their popularity. If this is indeed the case, our baseline estimates would fail to capture the true effect of incumbent popularity as measured by a postal ballot win. Figure A4 in the Appendix presents the results from this estimation exercise. We find that in comparison to our baseline estimate of -0.264 (see Table 2), the estimates from using placebo thresholds are substantially smaller in magnitude and statistically insignificant at the 5% significance level. This highlights the salience of the zero cutoff, making it unlikely that the effects we observe on recontesting outcomes are spurious or driven by an unobserved underlying mechanism.

Alternative Specifications. Next, we address the concern that time-invariant unobservables may confound our results. In Columns (1)–(3) of Table A2 in the Appendix, we control for state-specific, election year-specific, and district-specific fixed effects, respectively, in our main RD specification (2). The RD estimates with fixed effects are identical to our baseline estimates presented in Table 2. Further, while we documented balance on pre-determined covariates in Table 1, another way to test for imbalances is simply adding pre-determined variables as controls to the RD specification. Column (4) of Table A2 shows that even after controlling for a host of electoral variables, our results on the incumbent's recontesting outcome remain statistically similar to our baseline estimates. This provides

further credibility to our main results, indicating that our baseline estimates are not a product of potential unobservable and observable confounders.

Sensitivity to Sample Adjustments. Finally, we address the concern that the presence of particular states in our sample could potentially drive the results. For instance, states that have historically high levels of political competition may witness parties enforcing a higher turnover rate of incumbents. To address this concern, we drop one state at a time from our estimation sample and re-estimate equation (2) for our primary outcome variable. Figure A5 in the Appendix presents the graphical analog of this estimation exercise. We find that excluding one state at a time does not alter our results- the estimates for the incumbent's recontesting probability remain statistically significant, negative, and closely aligned with the baseline estimates in Table 2. This lends further confidence to the external validity of our estimates, suggesting that particular states in our sample are not the primary drivers of our results.

6 Mechanisms

In summary, our main findings suggest that an incumbent who narrowly secures a joint victory in both in-person and postal ballots is 26 percentage points less likely to contest the next election compared to an incumbent who barely lost in postal ballots. Moreover, we find no significant difference in the unconditional re-election probability between postal-aligned and postal-misaligned incumbents. In this section, we investigate the underlying mechanisms driving these effects. Specifically, we examine the following potential channels: (i) whether the effects stem from weaker performance by 'popular' incumbents during their tenure (Section 6.1), (ii) whether popular candidates are rewarded with positions in higher-tier political offices by their parties (Section 6.2), (iii) the role of political party dynamics (Section 6.3), and (iv) whether certain types of candidates face a lower probability of renomination in the subsequent election (Section 6.4).

6.1 Incumbent Performance

One possible explanation for why we observe a lower recontesting likelihood of postal-aligned incumbents is simply that they may have delivered a weaker performance during their tenure. Although measuring incumbent performance is challenging, we assess changes in a range of constituency-level development indicators that could have been influenced by the incumbents during their tenure. First, we assess the effect of an incumbent's postal alignment status on the growth of nighttime lights within constituency boundaries, a widely used proxy for economic activity.[6] For this analysis, we examine the effect of postal alignment on the year-on-year annual growth in night-time lights, which is calculated as follows:

$$GR^{NL}_{j,t+1} = 100 * [\log(NL_{j,t+1}) - \log(NL_{j,t})] \tag{4}$$

where $\log(NL_{j,t+1})$ is the natural logarithm of the annual nighttime lights detected within the constituency boundaries. Night-lights growth in year $t+1$, $GR^{NL}_{j,t+1}$, is then computed as the difference between the natural logarithm of nightlights in $t+1$ and t. We calculate this measure for the four years, excluding the election year, that constitute the electoral term. Therefore, we have four observations for each constituency: annual NL growth in $t+1$, $t+2$, $t+3$, and $t+4$, where t is the election year. We use the growth in two measures of night-time lights as the outcome variables: (i) annual mean nighttime lights, and (ii) annual sum of nighttime lights detected within the constituency boundaries.[7]

Table 3 presents the results on the impact of postal alignment on growth in nightlights. Across all columns, we find that the difference in the growth rate of nighttime lights between a postal-aligned and a postal-misaligned incumbent's constituency during their respective electoral terms is small and statistically indistinguishable from zero. This result holds ir-

[6] Night-time lights are routinely used to proxy for economic activity at granular levels, especially in settings where conventional data on economic indicators may be absent. For its use in the political economy literature on India, see Baskaran et al. (2024), Prakash et al. (2019), Jain et al. (2023).

[7] The mean night-time lights measure is computed by dividing the total annual nighttime lights within the constituency boundaries by the number of grid cells comprising the constituency.

Table 3: Effect of Postal Alignment on Growth in Nighttime Lights

	Panel A: Growth in Mean Nightlights			
	(1)	(2)	(3)	(4)
	h	$0.5 * h$	$2 * h$	CCT(h)
Postal Aligned	-0.561	-0.548	-0.666	-0.323
	(0.673)	(0.984)	(0.461)	(0.356)
Bandwidth size	5.672	2.836	11.343	12.239
Effective Obs.	2656	1360	5040	8076
Control Mean	1.801	1.778	1.661	1.678
	Panel B: Growth in Total Nightlights			
	h	$0.5 * h$	$2 * h$	CCT(h)
Postal Aligned	-1.410	-0.682	-1.323	-1.454
	(1.616)	(2.426)	(1.137)	(1.192)
Bandwidth size	7.944	3.972	15.888	14.501
Effective Obs.	3716	1844	6796	6236
Control Mean	6.484	7.431	6.153	6.141

Notes: Local linear RD regression results are presented using different factors of the optimal bandwidth, h, à la Imbens and Kalyanaraman (2011) (in Columns (1)–(3)) and the optimal bandwidth, CCT(h), by Calonico et al. (2014). Standard errors clustered at the Assembly constituency level are presented in parentheses. In all columns, the outcome in Panel A is the annual growth rate in mean nighttime lights; in Panel B, the outcome is the annual growth rate in the total sum of nighttime lights within the constituency boundaries (see Section 6.1 for more details). *Postal Aligned* is our treatment variable, $\tau_{i,j}$, from equation (3). Effective Obs. represents the number of observations within the optimally chosen bandwidth. Control mean is the mean of the dependent variable of the control group (left of the cutoff) within the optimal bandwidth.
Significance levels: $^{***}p < 0.01$;** $p < 0.05$;* $p < 0.1$

respective of the measure of nighttime lights we use. This suggests that postal-aligned incumbents perform just as well as postal-misaligned incumbents during their electoral tenure. To further probe this channel, we utilize information on village-level public facilities from the Mission Antyodaya survey, 2020 (see Section 2.2 for details on the data). We construct the following public goods indices using a principal component analysis technique: (i) healthcare goods index, (ii) school availability index, (iii) drainage infrastructure index, (iv) market availability index, (v) electricity provision index, (vi) roads availabil-

Table 4: Effect of Postal Alignment on Public Goods Provision

	Public Goods Provision						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Healthcare	Schools	Drainage	Markets	Electricity	Roads	Overall
Postal Aligned	-0.047	-0.181	-0.313	0.673*	-0.300	-0.176	-0.125
	(0.268)	(0.278)	(0.314)	(0.395)	(0.263)	(0.298)	(0.261)
Bandwidth size	4.121	4.085	3.781	3.665	4.070	4.523	4.296
Effective Obs.	218	218	200	194	217	242	228

Notes: Local linear RD regression results using the optimal bandwidth, h, à la Imbens and Kalyanaraman (2011) presented in all columns. Each column represents a separate regression with a distinct outcome variable: Column (1): healthcare goods index; Column (2): school availability index; Column (3): drainage infrastructure index; Column (4): market availability index; Column (5): electricity provision index; Column (6): roads availability index; and Column (7) overall public goods index (see Section 6.1 and Footnote 8 for more details). Effective Obs. represents the number of observations within the optimally chosen bandwidth. Control mean is the mean of the dependent variable of the control group (left of the cutoff) within the optimal bandwidth.
Significance levels: $^{***}p < 0.01;^{**}p < 0.05;^{*}p < 0.1$

ity index, and (vii) overall public goods index.[8] We standardize each index so that all RD estimates can then be interpreted in terms of standard deviations. The RD estimates on these indexes using the Imbens and Kalyanaraman (2011) optimal bandwidth are presented in Table 4. With the exception of the markets index, whose estimate is marginally significant, we find no significant difference in the availability of public goods between constituencies with a postal-aligned incumbent and those with a postal-misaligned incumbent. The graphic analog to the results on nighttime lights and public goods is presented in the Online Appendix in Figures A6 and A7, respectively. These results, in conjunction with the estimates on nighttime lights, suggest an absence of a performance differential between postal-aligned and postal-misaligned incumbents. Therefore, we can rule out the possibility

[8]The following variables are used to construct these indices. Healthcare: availability of primary health centre (PHC), community health centre (CHC), maternity health centre, and health sub-centre; Schools: availability of a primary school, a middle school, a high school, and a senior secondary school; Drainage: availability of a closed drainage system, an open covered drainage system, an open uncovered drainage system, and an open *kuchha* drainage system; Markets: availability of a *mandi*, a regular market, and a weekly *haat* (informal market); Electricity: availability of electricity supply to Micro, Small and Medium Enterprises (MSMEs), availability of solar/wind electricity to households, non-electrified villages, and availability of electricity in primary schools; Roads: availability of an internal *pucca* road and an all-weather road. The overall public goods index is constructed using all the mentioned variables. Each index captures the fraction of villages within the constituency that have access to the relevant public facilities.

that the negative effects on re-contesting probability are driven by incumbent performance.

6.2 Rewarding Popular Incumbents?

In the Introduction, we hypothesized that popular candidates are ranked lower in political parties' preference ordering of candidates. However, one possible explanation for the lower recontesting likelihood of postal-aligned incumbents is that they may receive a nomination to contest elections for higher tiers of government. That is, political parties may reward popular candidates by granting them a nomination to contest for higher-echelon political office (in the case of India, the *Lok Sabha* elections). To explore this channel in our setting, we examine the impact of the postal alignment status of an incumbent on the probability that they feature in the subsequent *Lok Sabha* or parliamentary election. Since our primary sample period covers state elections held between 2013 and 2017, we examine candidacy outcomes in the 2019 Lok Sabha election. For this analysis, we conduct two sets of regressions: (i) restricting the sample to incumbents who completed their full electoral tenure as MLAs—specifically, those elected in 2013 and 2014, and (ii) using the complete sample of incumbents elected to the state legislature between 2013 and 2017.[9]

The RD estimates for both samples are reported separately in Table 5. In both specifications, we observe that the effect of postal alignment on the likelihood of contesting the Lok Sabha election is negative and generally not statistically significant. This indicates that political parties do not exhibit a preference for postal-aligned incumbents over postal-misaligned incumbents when selecting candidates for parliamentary elections. If anything, postal-aligned incumbents are less likely to receive a nomination to contest the Lok Sabha election. These findings rule out the possibility that the lower recontesting likelihood reflects a reward for popular incumbents.

[9]In the full sample of incumbents, a fraction of them- particularly those elected between 2015 and 2017- would have their Assembly tenure overlap with the parliamentary election. However, in India, it is not entirely uncommon for MLAs to vacate their positions mid-term to assume roles as MPs.

Table 5: Effect of Postal Alignment on the Probability of Contesting the *Lok Sabha* Election

	(a) Incumbents elected in 2013-14			
	(1)	(2)	(3)	(4)
	h	$0.5 * h$	$2 * h$	CCT(h)
Postal Aligned	-0.107	-0.173	-0.075	-0.041
	(0.103)	(0.209)	(0.059)	(0.028)
Bandwidth size	2.488	1.244	4.977	14.469
Effective Obs.	172	79	326	854
Control Mean	0.051	0.062	0.042	0.037
	(b) Incumbents elected in 2013-17			
	h	$0.5 * h$	$2 * h$	CCT(h)
Postal Aligned	-0.057	-0.148	-0.021	-0.032*
	(0.079)	(0.169)	(0.036)	(0.018)
Bandwidth size	1.792	0.896	3.584	17.175
Effective Obs.	216	98	442	1880
Control Mean	0.042	0.054	0.046	0.032

Notes: Local linear RD regression results are presented using different factors of the optimal bandwidth, h, à la Imbens and Kalyanaraman (2011) (in Columns (1)–(3)) and the optimal bandwidth, CCT(h), by Calonico et al. (2014). In all panels, the outcome is a binary variable indicating whether the state assembly incumbent contests the 2019 parliamentary election. In Panel (a), the sample consists of incumbents elected between 2013 and 2014; in Panel (b), the sample consists of all incumbents elected between 2013 and 2017. *Postal Aligned* is our treatment variable, $\tau_{i,j}$, from equation (3). Effective Obs. represents the number of observations within the optimally chosen bandwidth. Control mean is the mean of the dependent variable of the control group (left of the cutoff) within the optimal bandwidth. Significance levels: $^{***}p < 0.01$;** $p < 0.05$;* $p < 0.1$

6.3 Nature of Political Parties

Since the lower recontesting probability of popular incumbents is neither driven by their performance nor by their potential promotion to a higher tier of government, we now focus on the characteristics of political parties to explore the underlying mechanism behind these effects. For this analysis, we conduct sub-sample analyses along three dimensions of the party type: (i) national v/s state parties, (ii) mass-based v/s cadre-based parties, (iii) dynastic v/s non-dynastic parties, and (iv) left-leaning v/s right-leaning parties. National

parties are those parties that have a significant seat share in the state Legislative Assemblies as well as the national General Assembly. State parties, on the other hand, are limited only to particular states with limited influence in the national electoral space.[10] Mass-based parties are parties that have a decentralized and democratic organizational structure with a high members-to-voters ratio (Duverger, 1969). The cadre party, in contrast, follows an organizational structure that is loose, with a highly concentrated network of elites centering upon a small circle of "notables" (Clark, 2008).[11] Finally, dynastic parties are those dominated by a specific family across multiple generations, where a majority of office bearers have had a family member previously holding political office.

Table 6: Effect of Postal Alignment on Recontesting: Party-level Hetergeneity

	Recontesting at $t+1$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	National v/s State		Mass v/s Cadre		Dynast v/s Non-Dynast		Left v/s Right	
	National	State	Mass	Cadre	Dynast	Non-Dynast	Left	Right
Postal Aligned	-0.322**	-0.212	-0.257*	-0.299***	-0.330**	-0.226*	-0.292**	-0.355**
	(0.132)	(0.146)	(0.153)	(0.106)	(0.153)	(0.121)	(0.115)	(0.148)
Bandwidth size	3.211	3.590	3.694	3.222	3.301	3.288	3.283	3.505
Effective Obs.	223	144	176	235	135	246	133	204
Control Mean	0.709	0.741	0.750	0.762	0.802	0.684	0.842	0.660

Notes: Local linear RD regression results using the optimal bandwidth, h, à la Imbens and Kalyanaraman (2011) presented in all columns. Each column represents a separate regression with a distinct sub-sample of winners from particular party-types: Column (1): national parties; Column (2): state parties; Column (3): mass-based parties; Column (4): cadre-based parties; Column (5): dynastic parties; Column (6): non-dynastic parties; Column (7): left-leaning parties; Column (8) right-leaning parties (see Section 6.3 for more details). Effective Obs. represents the number of observations within the optimally chosen bandwidth. Control mean is the mean of the dependent variable of the control group (left of the cutoff) within the optimal bandwidth for the particular subsample.
Significance levels: $^{***}p < 0.01;^{**}p < 0.05;^{*}p < 0.1$

[10]The distinction is administratively made by the Election Commission of India, based on the vote share of parties in the parliament. We use the party-type classification provided by the Trivedi Centre for Political Data, Ashoka University in their *Lok Dhaba* dataset.

[11]While in popular discourse, cadre-based organizations are usually defined as having a large set of foot soldiers (cadre), we continue to use the political science distinction outlined above, formulated by Duverger (1969). A primary example of a cadre-based party in India with a concentrated leadership structure centered around a few elites includes the Indian National Congress (INC), while the Bharatiya Janta Party (BJP) much approportiately fits the definition of a mass-based party.

Table 6 presents the results from this sub-sample analysis. Although the low statistical power prevents us from definitively rejecting the null hypothesis that the coefficients across sub-samples are statistically similar, these estimates still offer valuable insights into the relationship between party structure and candidacy outcomes. In terms of the magnitudes of the estimates, we find that popular incumbents from parties with a more centralized organizational structure are less likely to be nominated to contest the subsequent election. In essence, national parties and dynastic parties are relatively less likely to field popular candidates that potentially pose a threat to the central authority of parties. This underscores the key role of political parties and their organizational structure in the candidate selection process.

6.4 Candidate Characteristics

Finally, we examine whether certain types of candidates are less likely to receive party nominations to contest the subsequent elections. Building on our hypothesis that rising incumbent popularity may pose a threat to central party leadership, we analyze treatment effect heterogeneity across three candidate characteristics: (i) education level, (ii) criminal background, and (iii) wealth. For education, we construct two sub-samples: incumbents who have attained at least a college degree, and those who have completed at most 12 years of schooling. For criminal background, we categorize incumbents into two groups: those with a pending criminal case at the time of the election and those without any pending cases. Finally, for wealth, we utilize the candidates' self-disclosed net assets and create two sub-samples based on the median reported net assets within our sample of incumbents.

Table 7 presents the results from the candidate-level sub-sample analysis. The negative effect on the recontesting probability of postal-aligned incumbents is statistically significant, particularly among candidates with higher education, a criminal record, and above-median net assets. While the pairwise estimates do not exhibit significant differences, the findings suggest that political parties are less inclined to renominate candidates of superior

Table 7: Effect of Postal Alignment on Recontesting: Candidate-level Hetergeneity

	Recontesting at $t+1$					
	(1)	(2)	(3)	(4)	(5)	(6)
	Education		Criminality		Wealth	
	Graduate or Above	12th Pass or Below	Criminal	Non-Criminal	Above Median	Below Median
Postal Aligned	-0.311***	-0.191	-0.474***	-0.161	-0.325**	-0.236
	(0.116)	(0.162)	(0.158)	(0.128)	(0.137)	(0.143)
Bandwidth size	3.435	3.128	3.183	3.438	2.971	3.512
Effective Obs.	221	132	128	246	182	170
Control Mean	0.752	0.671	0.714	0.725	0.771	0.711

Notes: Local linear RD regression results using the optimal bandwidth, h, à la Imbens and Kalyanaraman (2011) presented in all columns. Each column represents a separate regression with a distinct sub-sample of winners: Column (1): winners with a graduate degree or higher; Column (2): winners having completed only 12 years of schooling or less; Column (3): winners with a criminal case; Column (4): winners with no pending criminal case; Column (5): criminals with above-median net assets; Column (6): winners with below-median net assets (see Section 6.4 for more details). Effective Obs. represents the number of observations within the optimally chosen bandwidth. Control mean is the mean of the dependent variable of the control group (left of the cutoff) within the optimal bandwidth for the particular subsample.
Significance levels: $^{***}p < 0.01;^{**}p < 0.05;^{*}p < 0.1$

quality (i.e., those with higher educational qualifications). Moreover, criminal politicians who secure elite support (postal-aligned) may pose a challenge to the central leadership structure, leading to their lower likelihood of renomination after a postal ballot victory. These findings highlight a crucial interaction between candidate characteristics and party dynamics in the candidate selection process.

7 Conclusion

In this paper, we examine the interplay between candidate popularity and political party dynamics in shaping the candidate selection process. In particular, we explore an alternative popularity curse for candidates in Indian elections, the curse of the *absentee voter*. We find that candidates who narrowly secure a victory in both in-person and postal ballots are

26 percentage points less likely to be fielded by political parties in the subsequent election. In contrast, we do not find any evidence that voters treat postal-aligned incumbents any differently from postal-misaligned incumbents, as evidenced by the null effects we estimate for their respective re-election probabilities. Further, we are able to rule out the possibility that the negative effect on recontesting likelihood is a product of the incumbent's poor performance during her tenure or a progression to a higher-tier political office. Notably, we find that the negative effects are more pronounced for high-quality candidates and parties with a relatively centralized organizational structure.

References

Aldrich, J. H. (1993). Rational choice and turnout. *American journal of political science*, pages 246–278.

Anagol, S. and Fujiwara, T. (2016). The runner-up effect. *Journal of Political Economy*, 124(4):927–991.

Asher, S., Lunt, T., Matsuura, R., and Novosad, P. (2021). Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in india using the shrug open data platform. *The World Bank Economic Review*, 35(4):845–871.

Baskaran, T., Bhalotra, S., Min, B., and Uppal, Y. (2024). Women legislators and economic performance. *Journal of Economic Growth*, 29(2):151–214.

Besley, T. (2005). Political selection. *Journal of Economic perspectives*, 19(3):43–60.

Besley, T., Folke, O., Persson, T., and Rickne, J. (2017). Gender quotas and the crisis of the mediocre man: Theory and evidence from sweden. *American economic review*, 107(8):2204–2242.

Caillaud, B. and Tirole, J. (2002). Parties as political intermediaries. *The Quarterly Journal of Economics*, 117(4):1453–1489.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *Econometrics Journal*, 23(2):192–210.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-discontinuity Designs. *Econometrica*, 82(6):2295–2326.

Canay, I. A. and Kamat, V. (2018). Approximate permutation tests and induced order statistics in the regression discontinuity design. *The Review of Economic Studies*, 85(3):1577–1608.

Casey, K., Kamara, A. B., and Meriggi, N. F. (2021). An experiment in candidate selection. *American Economic Review*, 111(5):1575–1612.

Cattaneo, M. D., Jansson, M., and Ma, X. (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 115(531):1449–1455.

Chandra, K. (2016). *Democratic dynasties: State, party, and family in contemporary Indian politics*. Cambridge University Press.

Chatterjee, S. and Kamal, J. (2021). Voting for the underdog or jumping on the bandwagon? evidence from india's exit poll ban. *Public Choice*, 188(3):431–453.

Clark, A. (2008). Mass, cadre or franchise parties? assessing scotland's local party organizations. *Journal of Elections, Public Opinion and Parties*, 18(1):27–51.

Dal Bó, E. and Finan, F. (2018). Progress and perspectives in the study of political selection. *Annual Review of Economics*, 10(1):541–575.

Desai, Z. and Lee, A. (2021). Technology and protest: the political effects of electronic voting in india. *Political Science Research and Methods*, 9(2):398–413.

Duverger, M. (1969). The concept of membership. In *Comparative Government: A Reader*, pages 77–85. Springer.

Fujiwara, T. (2015). Voting technology, political responsiveness, and infant health: Evidence from brazil. *Econometrica*, 83(2):423–464.

Galasso, V. and Nannicini, T. (2011). Competing on good politicians. *American political science review*, 105(1):79–99.

Gronke, P., Galanes-Rosenbaum, E., and Miller, P. A. (2007). Early voting and turnout. *PS: Political Science & Politics*, 40(4):639–645.

Gronke, P., Galanes-Rosenbaum, E., Miller, P. A., and Toffey, D. (2008). Convenience voting. *Annu. Rev. Polit. Sci.*, 11(1):437–455.

Gulzar, S. (2021). Who enters politics and why? *Annual Review of Political Science*, 24(1):253–275.

Gulzar, S., Hai, Z., and Paudel, B. K. (2021). Information, candidate selection, and the quality of representation: Evidence from nepal. *The Journal of Politics*, 83(4):1511–1528.

Imbens, G. and Kalyanaraman, K. (2011). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies*, 79(3):933–959.

Jain, C., Kashyap, S., Lahoti, R., and Sahoo, S. (2023). The impact of educated leaders on economic development: Evidence from india. *Journal of Comparative Economics*, 51(3):1068–1093.

Jia, R., Kudamatsu, M., and Seim, D. (2015). Political selection in china: The complementary roles of connections and performance. *Journal of the European Economic Association*, 13(4):631–668.

Magesan, A., Szabó, A., and Ujhelyi, G. (2024). Candidate selection by parties: Crime and politics in india. Technical report, Working Paper.

Mattozzi, A. and Merlo, A. (2015). Mediocracy. *Journal of Public Economics*, 130:32–44.

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

Oliver, J. E. (1996). The effects of eligibility restrictions and party activity on absentee voting and overall turnout. *American Journal of Political Science*, pages 498–513.

Prakash, N., Rockmore, M., and Uppal, Y. (2019). Do criminally accused politicians affect economic outcomes? evidence from india. *Journal of Development Economics*, 141:102370.

Riker, W. H. and Ordeshook, P. C. (1968). A theory of the calculus of voting. *American political science review*, 62(1):25–42.

Stein, R. M. and Vonnahme, G. (2008). Engaging the unengaged voter: Vote centers and voter turnout. *The Journal of Politics*, 70(2):487–497.

Ujhelyi, G., Chatterjee, S., and Szabó, A. (2021). None of the above: Protest voting in the world's largest democracy. *Journal of the European Economic Association*, 19(3):1936–1979.

Appendix

A Additional Tables and Figures

A.1 Descriptive Statistics

Table A1: Summary Statistics

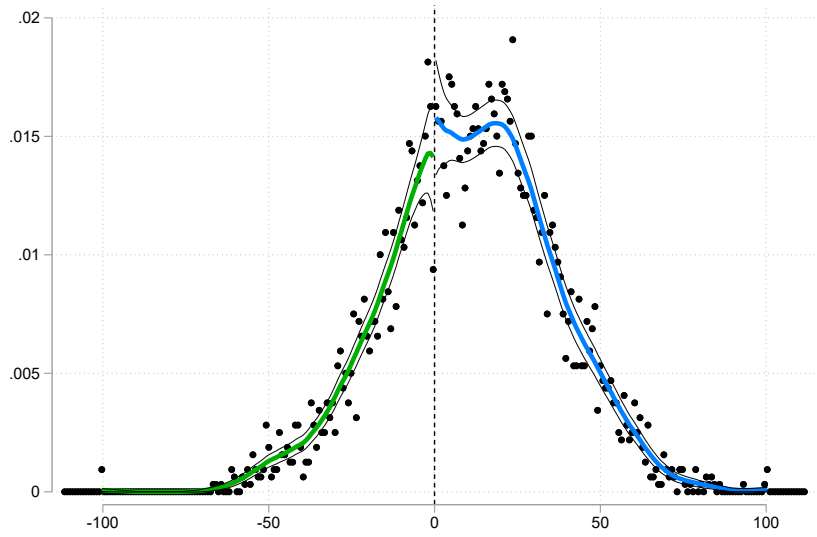
	Postal Aligned	Postal Misaligned	Postal Aligned $[0, 5]$	Postal Misaligned $[-5, 0)$
	(1)	(2)	(3)	(4)
Log (Electorate Size)	12.022	12.075	12.003	11.993
	(0.80)	(0.78)	(0.85)	(0.86)
Number of Candidates	10.77	10.77	10.83	10.5
	(5.14)	(5.07)	(5.32)	(4.76)
SC/ST Reservation Status	0.29	0.27	0.27	0.29
	(0.45)	(0.44)	(0.44)	(0.45)
Winner's Overall Vote Share	46.84	44.06	43.92	44.79
	(9.05)	(7.60)	(8.46)	(8.30)
Total Postal Votes	695.15	711.02	683.62	642.52
	(527.27)	(532.94)	(472.01)	(487.35)
Winner's Gender	0.913	0.91	0.909	0.92
	(0.28)	(0.286)	(0.287)	(0.27)
Winner's Age	51.66	51.59	51.50	51.29
	(10.51)	(10.38)	(10.74)	(10.11)
Voter Turnout (%)	72.32	74.0	72.17	73.79
	(12.22)	(10.91)	(11.93)	(11.01)
ENOP (Effective number of parties)	2.92	2.97	3.06	2.94
	(0.81)	(0.75)	(0.86)	(0.80)
Winning Margin	14.38	8.63	9.95	9.53
	(12.14)	(7.45)	(8.06)	(7.85)

Notes: Data on all variables comes from the ECI statistical reports on Legislative Assembly elections from 2013 to 2017. Column (1) corresponds to constituencies where the overall election winner scores the maximum number of postal votes (*postal aligned*), while Column (2) corresponds to constituencies where the eventual winner did not receive the highest votes in postal ballots (*postal non-aligned*). In columns (3) and (4), we present descriptive statistics for constituencies where the winner was barely postal aligned vis-à-vis constituencies where the winner was barely postal non-aligned.

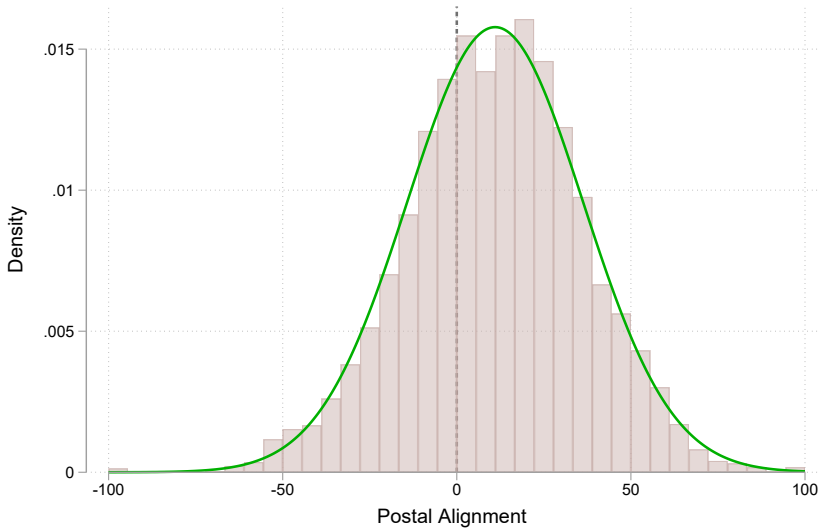
A.2 McCrary Density Test and Histogram Representation

Figure A1: Alternative Density Test: McCrary (2008)

Panel A: McCrary Density Test



Panel B: Density of the Forcing Variable



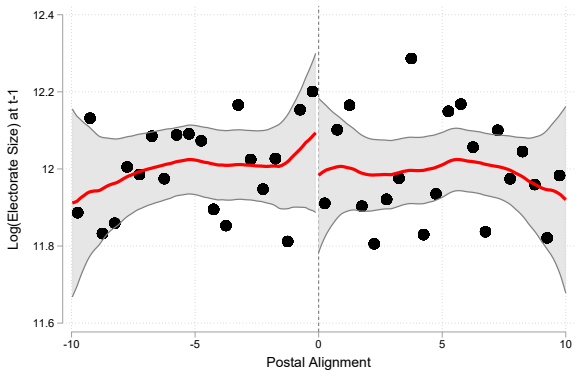
Notes: The figure shows the continuity of the forcing variable, $postal\_alignment_{i,j,t}$ A positive value indicates that the winning candidate received the highest number of postal votes. On the other hand, a negative value implies that the winner lost on the postal ballots. The estimated size of the discontinuity in our running variable (log difference in height) is 0.1139 (SE = 0.124).

A.3 Balance on Pre-Determined Characteristics

Figure A2: Effect of *Postal Alignment* on Pre-Determined Characteristics

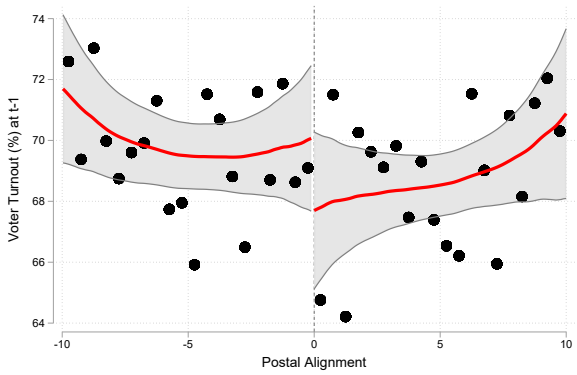
(a) Log(Electorate Size) at $t-1$

p-value $= 0.368$



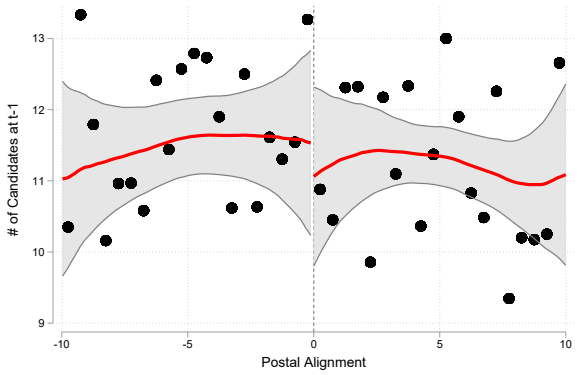
(b) Voter Turnout (%) at $t-1$

p-value $= 0.185$



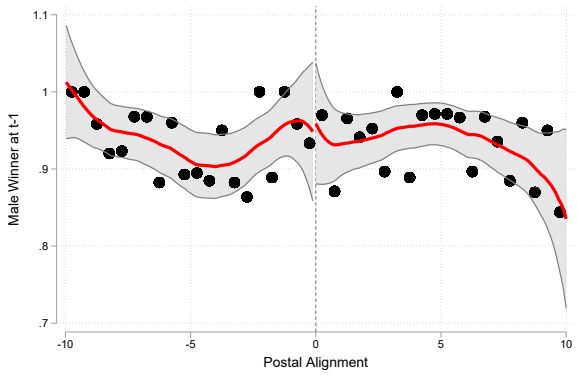
(c) # of Candidates at $t-1$

p-value $= 0.613$



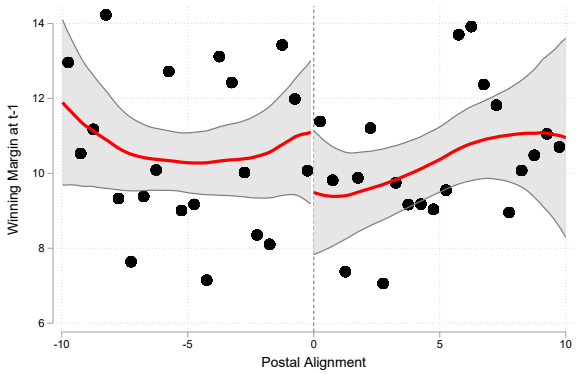
(d) Male Winner at $t-1$

p-value $= 0.857$



(e) Winning Margin at $t-1$

p-value $= 0.226$



(f) Effective # of Candidates at $t-1$

p-value $= 0.755$

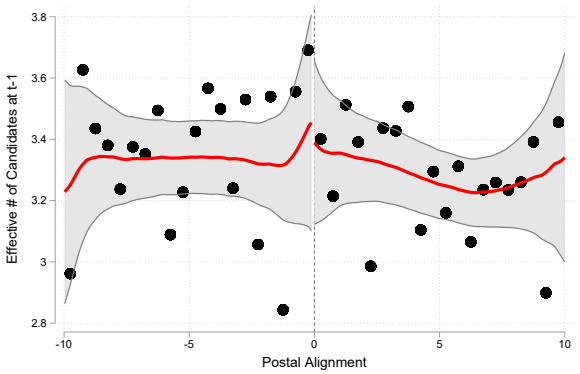
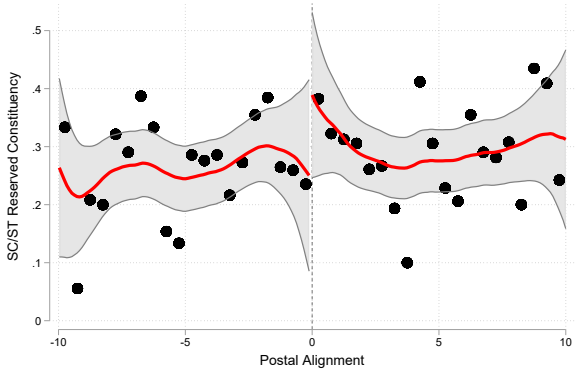


Figure A2: Effect of *Postal Alignment* on Pre-Determined Characteristics (*continued*)

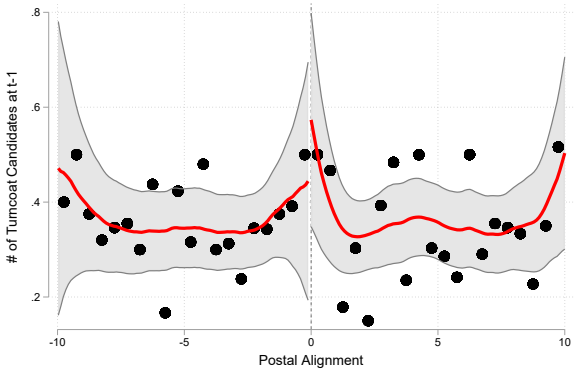
(g) SC/ST Reserved Constituency

p-value = 0.209



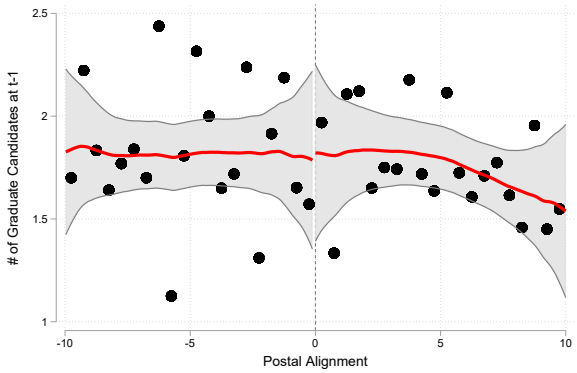
(h) # of Turncoat Candidates at $t-1$

p-value = 0.518



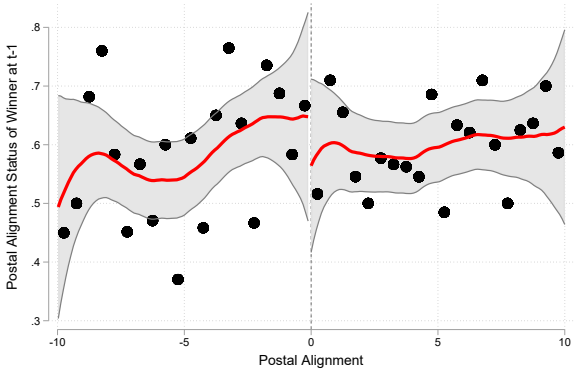
(i) # of Graduate Politicians at $t-1$

p-value = 0.894



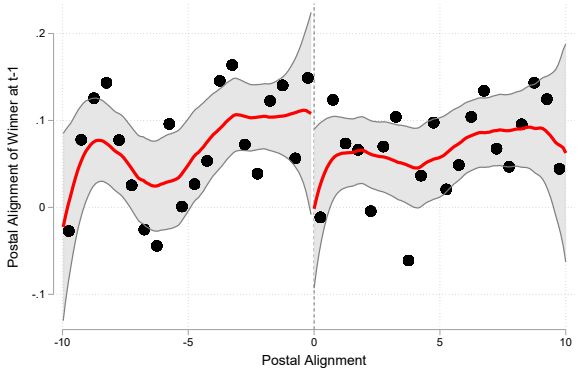
(j) Postal Alignment Status of Winner at $t-1$

p-value = 0.504



(k) Postal Alignment (Running Var.) at $t-1$

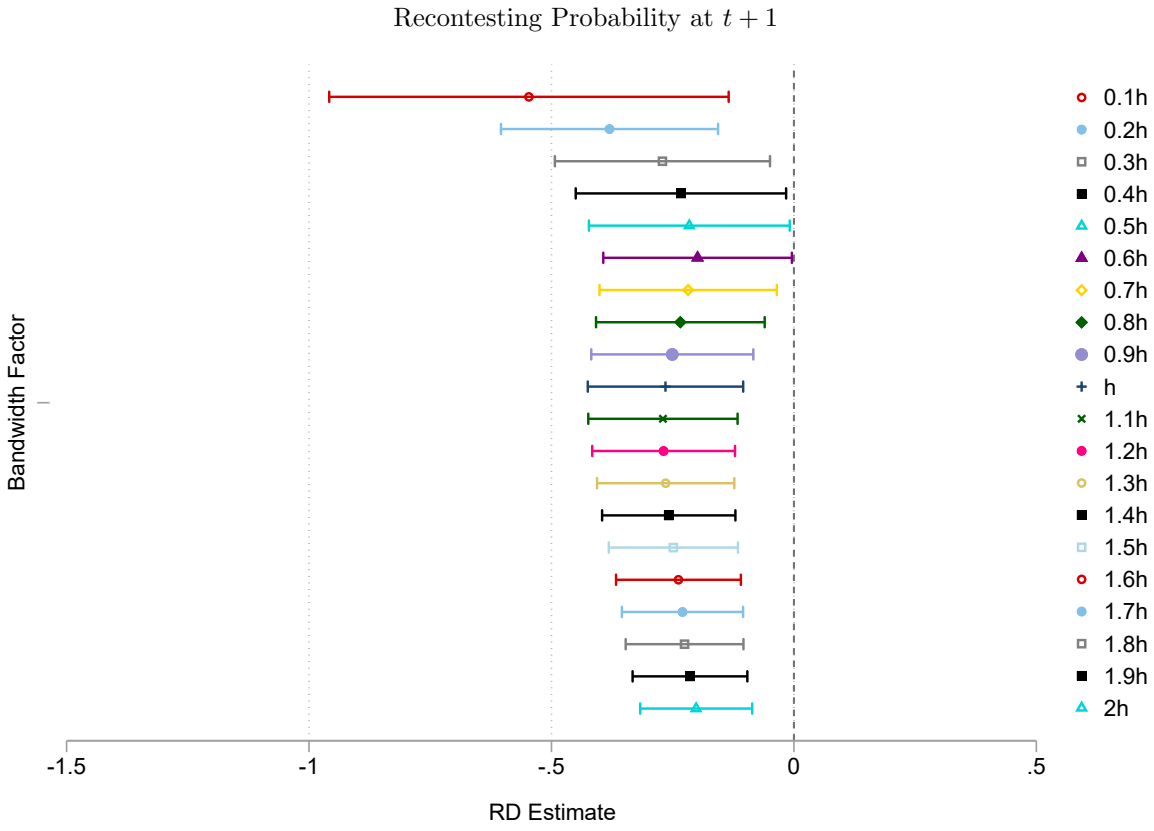
p-value = 0.225



Notes: The running variable, *postal_alignment*, is the difference in the postal vote shares of the winner and the runner-up. Each dot on the scatter plot is an average over successive bins of 0.5% of the running variable. The curves are local linear regressions fit separately for regions above and below the cut-off using a triangular kernel and an optimal bandwidth à la (Imbens and Kalyanaraman, 2011). We also report the p-values corresponding to each RD estimate, presented in detail in Table 1.

A.4 Robustness 1: Alternative Bandwidth Choices

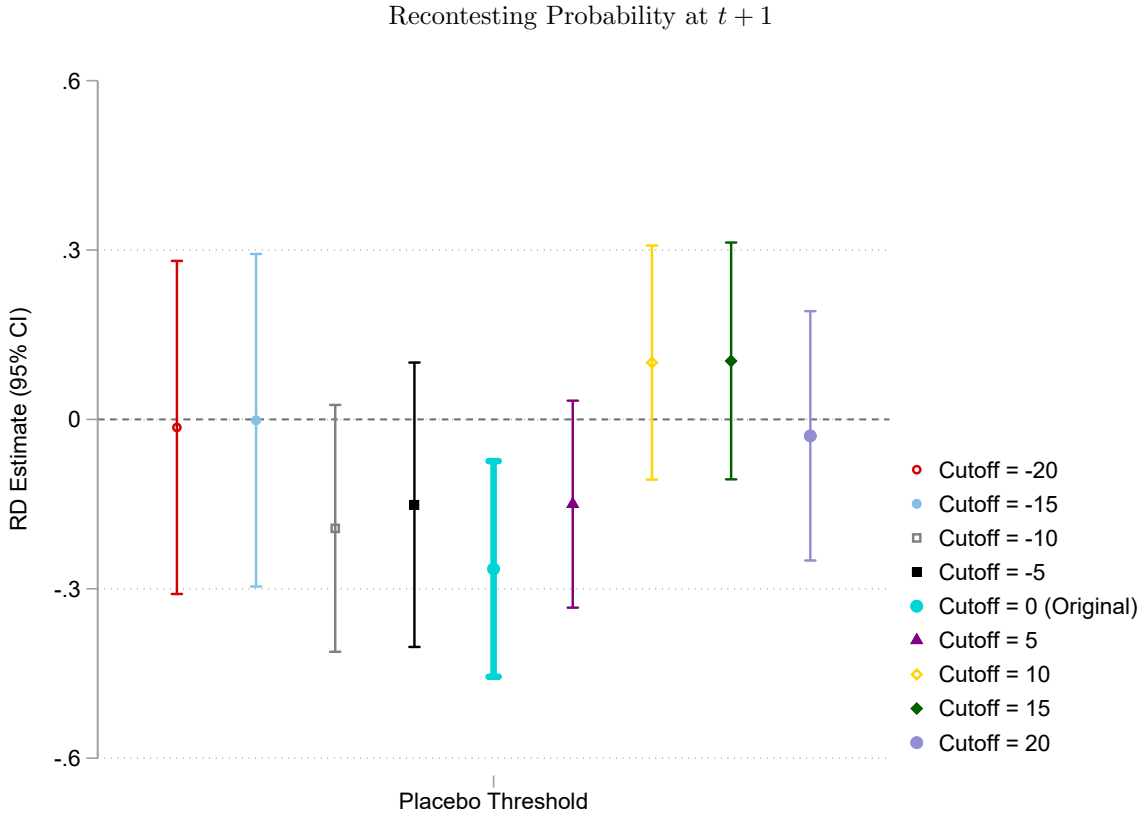
Figure A3: Evolution of Estimates with Incremental Bandwidth Increases



Notes: This presents different point estimates (and the 90% confidence interval) by varying the optimal bandwidth (h) between a factor of 0.1 and 2 of the optimal bandwidth in increments of 0.1. As before, the optimal bandwidth is computed using the selection procedure of Imbens and Kalyanaraman (2011).

A.5 Robustness 2: Placebo Thresholds

Figure A4: RD Estimates with Placebo Cutoffs



Notes: The figure presents the RD estimates (and the 95% confidence interval) from using placebo thresholds on either side of the original zero cutoff. Each CI and point estimate corresponds to a placebo cutoff (see legends) For all placebo cutoffs, the optimal bandwidth around the threshold is computed using the selection procedure of Imbens and Kalyanaraman (2011).

A.6 Robustness 3: Fixed Effects and Controls

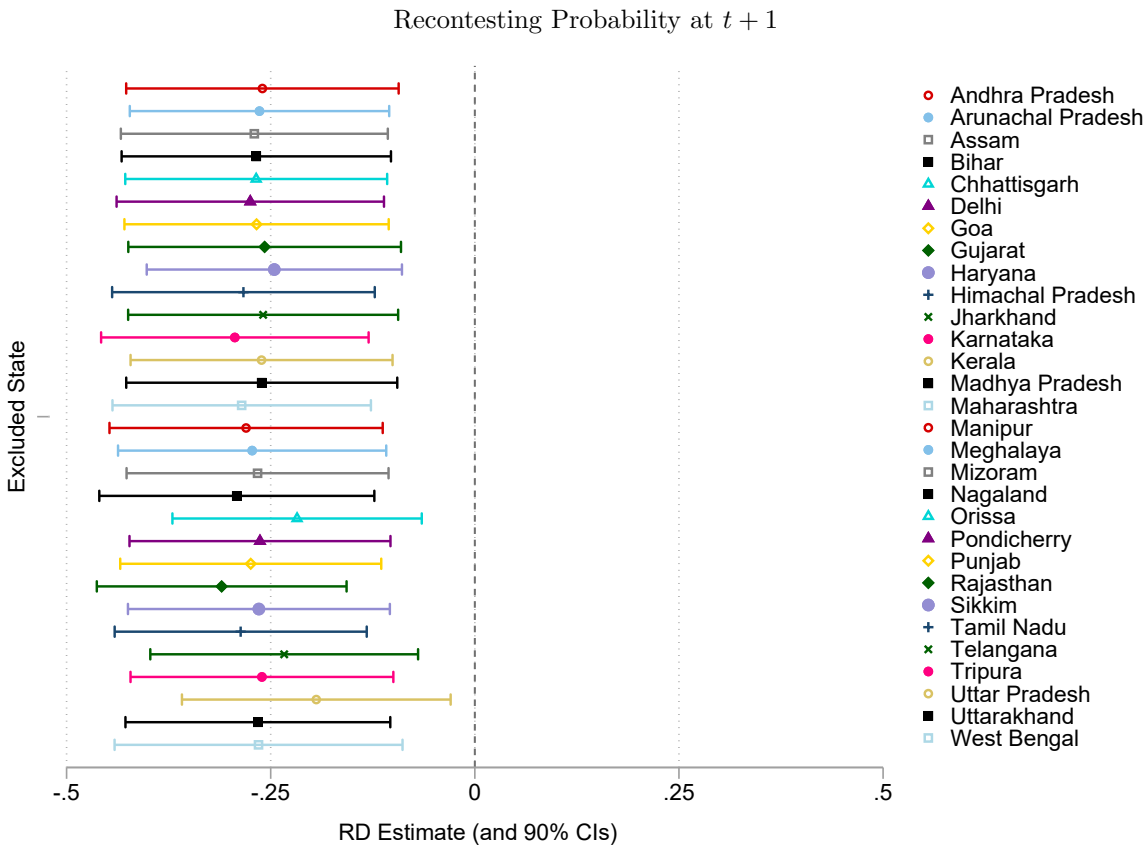
Table A2: Effect of Postal Alignment on Recontesting : Alternative Specifications

	Recontesting at $t+1$			
	(1)	(2)	(3)	(4)
	State FE	Year FE	District FE	Controls
Postal Aligned	-0.2569***	-0.2564*	-0.3543**	-0.2894***
	(0.0973)	(0.0960)	(0.1606)	(0.1021)
Bandwidth size	2.913	2.913	2.913	2.913
Effective Obs.	347	347	347	347
Control Mean	0.741	0.741	0.741	0.741

Notes: Local linear RD regression results using the optimal bandwidth, h, à la Imbens and Kalyanaraman (2011) presented in all columns. In column (1), we add state fixed effects to the RD regression equation (2); in Column (2), we add year fixed effects, in Column (3), we add district fixed effects, and in Column (4), we control for a host of candidate and constituency-level covariates including the log electorate size, total number of contestants, voter turnout, sitting incumbent's gender, winning margin in the previous election, and the effective number of candidates in the previous election. In all columns, the outcome is a binary variable, equaling 1 if the incumbent recontests at $t+1$ and zero otherwise. *Postal Aligned* is our treatment variable, $\tau_{i,j}$, from equation (3). Effective Obs. represents the number of observations within the optimally chosen bandwidth. Control mean is the mean of the dependent variable of the control group (left of the cutoff) within the optimal bandwidth.
Significance levels: $^{***}p < 0.01$;** $p < 0.05$;* $p < 0.1$

A.7 Robustness 4: Leave One-out Analysis by State

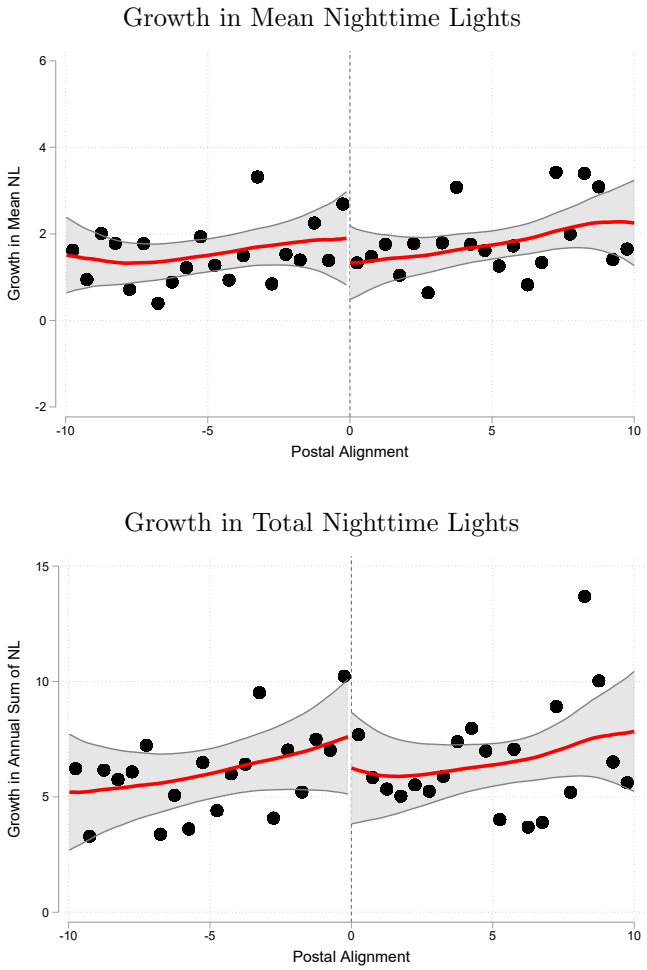
Figure A5: RD estimates from the systematic exclusion of states from the sample



Notes: The figure presents the RD estimates (and the 90% confidence interval) by dropping one year at a time from the sample. Each CI and point estimate corresponds to a specification with a particular state excluded from the sample (see figure legends). For all regressions with sample adjustments, the optimal bandwidth around the threshold is computed using the selection procedure of Imbens and Kalyanaraman (2011).

A.8 Postal Alignment and Growth in Nighttime Lights

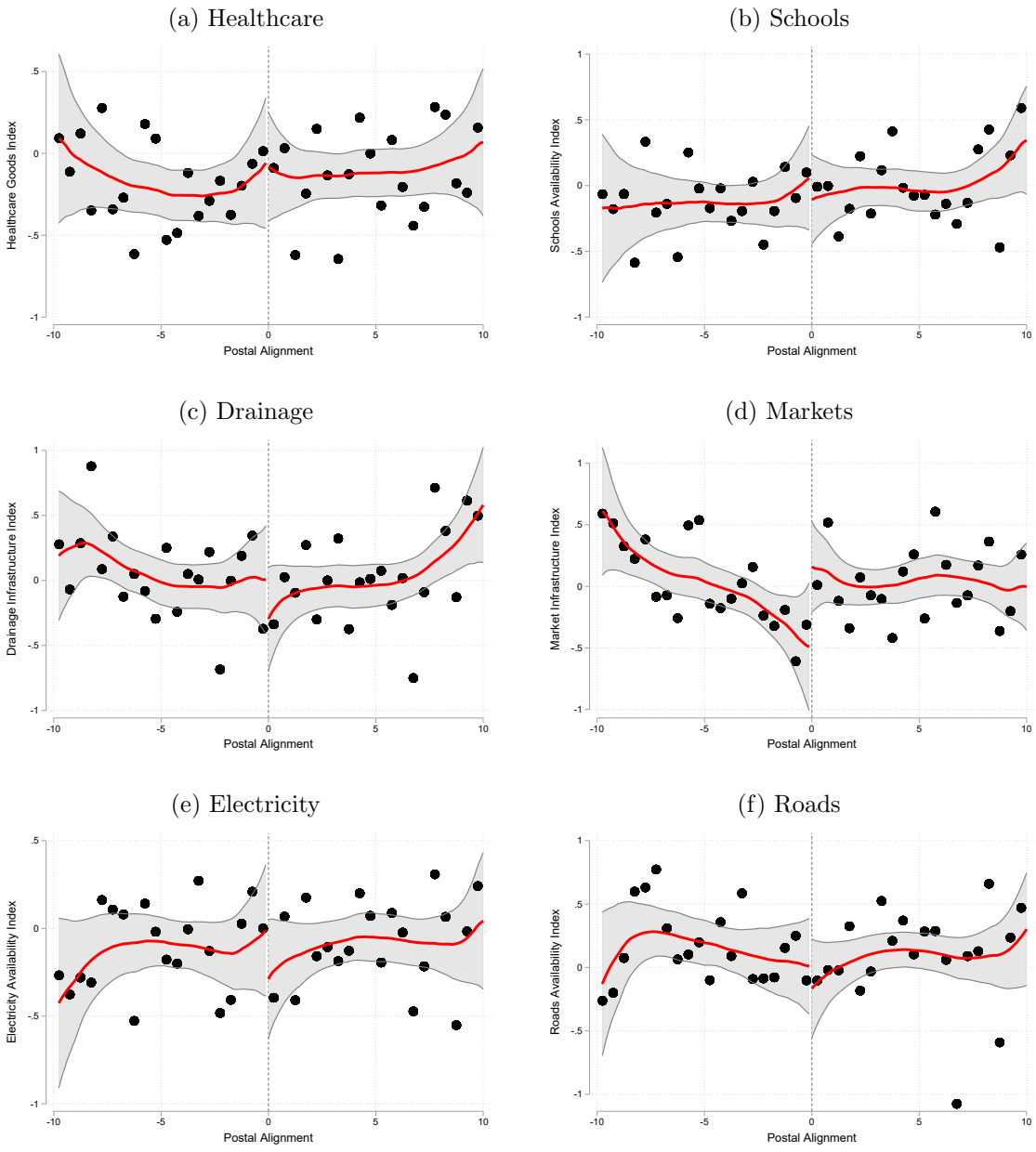
Figure A6: Impact of Postal Alignment on Growth in Nightlights



Notes: The figure plots the growth in nighttime lights against our running variable, *postal alignment*, defined as the postal vote share difference between the overall winning candidate and the overall runner-up. A positive value of the running variable implies a postal-aligned winner; a negative value implies a postal-misaligned winner. Each dot in the figure depicts the averages over successive bins of 0.5% of the running variable. The curves are local linear regression (with 95% confidence intervals) fit separately for the positive and negative parts of the cut-off. The optimal bandwidth à la Imbens and Kalyanaraman (2011) is used to arrive at the figures. The point estimates are presented in Table 3 in the text.

A.9 Postal Alignment and Public Goods Provision

Figure A7: Impact of Postal Alignment on Public Goods Provision



Notes: The figure plots different public good indices against our running variable, *postal alignment*, defined as the postal vote share difference between the overall winning candidate and the overall runner-up. A positive value of the running variable implies a postal-aligned winner; a negative value implies a postal-misaligned winner. Each dot in the figure depicts the averages over successive bins of 0.5% of the running variable. The curves are local linear regression (with 95% confidence intervals) fit separately for the positive and negative parts of the cut-off. The optimal bandwidth à la Imbens and Kalyanaraman (2011) is used to arrive at the figures. The point estimates are presented in Table 4 in the text.