Democratic Backsliding in the World's Largest Democracy

Sabyasachi Das*
Ashoka University

July 3, 2023

Abstract

Democratic backsliding is a growing concern globally. This paper contributes to the discussion by documenting irregular patterns in 2019 general election in India and identifying whether they are due to electoral manipulation or precise control, i.e., incumbent party's ability to precisely predict and affect win margins through campaigning. I compile several new datasets and present evidence that is consistent with electoral manipulation in closely contested constituencies and is less supportive of the precise control hypothesis. Manipulation appears to take the form of targeted electoral discrimination against India's largest minority group – Muslims, partly facilitated by weak monitoring by election observers. The results present a worrying development for the future of democracy.

JEL Codes: D72, D73, P00, Z12
Keywords: Electoral fraud, precise control, democracy, economics of religion

*Das: Economics Department, Ashoka University, National Capital Region, India. Email: sabyasachi.das@ashoka.edu.in.
The author wishes to thank Lakshmi Iyer, Ajay Shenoy, Sam Asher, Milan Vaishnav, Amrita Dhillon, Sourav Bhattacharya, Siddharth George Ajay Shah, Rohit Lamba, Gaurav Chiplunkar, Raphael Susewind, Jonathan Lehne, Dev Patel, Neelanjan Sircar, Aaditya Dar, Vimal Balasubramaniam and Sugat Chaturvedi for valuable comments. Rakesh Kumar provided excellent research assistance. The author is responsible for errors, if any.

I Introduction

Free and fair elections are cornerstone of a democracy. Yet in many democracies, the fairness of elections is increasingly in doubt. In 2020, for example, the Constitutional court of Malawi declared the Presidential election to be fraudulent.[1] The 2019 Presidential election result in Bolivia is also reported to be have been manipulated (Escobari and Hoover 2020). In the US, one-third of voters believe that Joe Biden won the 2020 Presidential election solely because of voter fraud, even though there is no evidence favoring such a claim.[2] Even before the election, the share of American voters reported to have confidence in the honesty of elections has been declining for several years, and in 2019, stood at only 40%.[3] This figure is 50% for the entire world, according to the Gallup World Poll (2007-13).

The global erosion of trust in electoral institutions coincides with the autocratizing tendencies of several democracies, known as democratic backsliding or deconsolidation (Waldner and Lust 2018, Foa and Mounk 2016, 2017a,b). Freedom House 2021 report points out that global freedom deteriorated for 15 consecutive years, with 75 percent of the world living in a country that experienced deterioration in 2020. Democracy Report (2020) mentions: "For the first time since 2001, democracies are no longer in the majority. [...] The countries that have autocratized the most over the last 10 years are Hungary, Turkey, Poland, Serbia, Brazil and India." While the overall pattern of democratic backsliding is based primarily on subjective evaluation by experts, objective evidence on this trend is lacking (Little and Meng 2023).

I contribute to this important debate by examining objective evidence of democratic backsliding in the form of electoral manipulation in the world's largest democracy – India. India is a somewhat unusual case for electoral fraud as it stands out in terms of the public trust its election authority enjoys. Two-third of its voters reported to have confidence in the honesty of elections in 2019, based on the Gallup Poll survey. Moreover, the confidence in elections is rising in India at least since 2006.[4] The level of confidence is also higher than many democracies with strong institutions, such as Japan (57%), France (57%), UK (61%) etc. The independence and institutional strength of the electoral authority in charge of conducting elections, the Election Commission of India (ECI), is an important factor that can potentially explain such high degree of confidence.[5] This makes the ECI one of the most powerful election management bodies in the world.

In the past few years, however, the credibility of the ECI has been called into question, with allegations of bias in scheduling of elections (Ramachandran 2022) and arbitrary deletion of names of registered Muslim voters (Malhotra 2019, Trivedi 2019, Naqvi 2022), both favoring the ruling party. The recent democracy reports of the V-Dem Institute highlight that various indicators of democracy in India, including the autonomy of the ECI, has been declining. Democracy Report (2021) have consequently classified India as an "electoral autocracy". As the V-Dem report points out, decline in the autonomy of the ECI was one of the important factors contributing to the reclassification of India's regime type. Similarly, Freedom House has changed India's status in 2021 from Free to Party Free (Repucci and Slipowitz 2021). The Supreme Court of India, in a recent judgement in 2023, acknowledged the dangers

[1] https://www.nytimes.com/2020/02/03/world/africa/Malawi-president-election-fraud.html

[2] https://www.monmouth.edu/polling-institute/reports/MonmouthPoll_US_031721/

[3] https://news.gallup.com/poll/285608/faith-elections-relatively-short-supply.aspx

[4] https://news.gallup.com/poll/248495/confidence-key-institutions-high-india-votes.aspx

[5] Section II provides a brief discussion on the independence of election authorities in India and the contextual details of India's general elections.

of a weak ECI and granted it significant autonomy and protection from executive overreach.[6]

In light of these developments, I first document that the 2019 general election in India that reelected the incumbent party shows significant irregularities in the election data – the density of the incumbent party's win margin variable exhibits a discontinuous jump at the threshold value of zero. It implies that in constituencies that were closely contested between a candidate from the incumbent party and a rival, the incumbent party (BJP) won disproportionately more of them than lost. This is known as the McCrary test and is now a standard check for manipulation of running variable in the regression discontinuity design (RDD) method used in analysis of political economy (Prakash et al. 2019, Nellis et al. 2016, Bhalotra et al. 2014). I do not find similar discontinuities in the previous general elections for either BJP or INC (Indian National Congress), the other major national party, as well as for state assembly elections held simultaneously with the 2019 general election and those held subsequently. Moreover, BJP's disproportionate win of closely contested constituencies is primarily concentrated in states ruled by the party at the time of election.

Failure of McCrary test however does not necessarily imply electoral fraud. If the incumbent party, due to its superior electoral machinery, was able to accurately predict and affect win margins in closely contested constituencies – a phenomenon known as *precise control* (Jeong and Shenoy 2020, Vogl 2014), then it could also generate such patterns. The incumbent party in India may have been able to exercise precise control in 2019 since it had significantly built up its organizational capacity in several states, subsequent to its 2014 general election victory. It mobilized active party workers at the level of polling stations who monitored and shaped voter attitudes, backed by centrally managed teams analyzing the collected information and suggesting campaign strategies (Jha 2017). Precise control in this context, therefore, if exercised, is likely to be facilitated by localized and targeted campaigning facilitated by grassroots presence of the party organization. This can explain the patterns described above. In the subsequent analysis I attempt to look for evidence that may distinguish between the two competing hypotheses.[7]

For my analysis, I put together several new datasets in addition to accessing the candidate level general election results for 1977-2019 and state assembly election results for 2019-2021 from standard sources. To examine precise control, I access the well-established post-poll survey – the National Election Survey (NES) of 2019 that gives micro data on election campaigning by political parties. To investigate election manipulation, I compile two different but official versions of constituency level Electronic Voting Machine (EVM) turnout data (for 2019 general election) to directly measure data discrepancy. The ECI initially released in its official website the "final" count of EVM votes polled for each Parliamentary Constituency (PC) for the first four out of seven phases of the 2019 elections (373 out of 543 PCs). Subsequently, it released constituency wise number of votes counted in EVMs, which did not match the initial numbers. When the media pointed out the discrepancy, the ECI removed the earlier figures from its website. I access copies of the earlier turnout data to measure discrepancy. I also put together the list of counting observers assigned to each constituency by the ECI to monitor counting of votes in 2019. The data provides various characteristics of the counting observers such as the state where they work, their cadre (i.e., whether they are part of the central or state bureaucracy), year of joining service etc. Additionally, I compile polling station level election outcomes for the 2019 general

[6] https://www.thehindu.com/news/national/committee-of-pm-lop-cji-to-advice-on-appointment-of-election-commissioners-supreme-court/article66570806.ece

[7] The excess mass of constituencies that BJP barely won could also arise purely due to chance.

election by scraping and parsing the scanned PDFs containing the data, available from the official websites of election authorities in individual states. To examine targeted voter suppression of Muslims as a potential mechanism of manipulation, I compute electorate share of Muslims at the level of Assembly Constituencies (ACs) using a 3 percent representative sample of voter lists and applying a highly accurate religion prediction algorithm on their names, and match it to polling stations and ACs.[8]. Section III describes the datasets and their sources.

I use a new question added to the NES in 2019 to measure campaigning in the form of door-to-door visits by BJP and other political parties in a representative sample of PCs to directly test for precise control. I find that neither BJP nor any other party campaigned significantly harder in constituencies that BJP barely won. Moreover, in BJP ruled states, campaigning by BJP does not exhibit statistically significant discontinuity, while that for the other parties does. This makes precise control less likely to be the primary mechanism.

Electoral manipulation, on the other hand, can take place at the stage of voter registration (registration manipulation) or at the time of voting or counting (turnout manipulation). To examine the mechanisms facilitating manipulation, I focus on Muslim voters who generally do not support BJP (Varshney 2019), and are easily identified in the voter list due to their culturally distant names. Therefore, they are potentially subject to both registration and turnout manipulation.[9] I consider two channels; first, strategic deletion of Muslim names from the list of registered voters or electoral rolls (Lehne 2022). Second, strategic suppression of Muslim votes at the time of voting (or counting) (Neggers 2018). I do not consider the possibility of manipulation of EVMs themselves as a mechanism, as Purkayastha and Sinha (2019) have pointed out that given its technology, it is hard to manipulate them at scale.

To test for registration manipulation, I compute growth rate of electorate (i.e., number of registered voters) for each Parliamentary Constituency (PC) between 2014 and 2019. I show that the growth rate falls discontinuously by 5 percentage points (compared to mean of 0.09) in PCs barely won by BJP, and the fall is concentrated in PCs with higher share of Muslim electorate. To examine turnout manipulation, I first examine the absolute difference between the two official versions of EVM turnout data. The discrepancies could be due to administrative errors during counting of votes. However, the extent of discrepancy, in that case, should not exhibit any discontinuous change with respect to the incumbent's win margin at its threshold value of zero. I however find that there is a large discontinuous increase in the magnitude of data revision at the threshold. Consistent with previous results, the discontinuity is concentrated in BJP ruled states.

I interpret the evidence on turnout discrepancy as indicative of manipulation done *locally* at the polling stations, rather than resulting from aggregation fraud at the constituency level (Callen and Long 2015). It is unlikely that ECI would engage in direct tampering of turnout data ex-post. Moreover, barring one case, the magnitude of data revision is smaller than BJP's absolute margin of victory. I show that polling station level election outcomes in the relevant PCs exhibit irregularities consistent with local

[8]Census data on religious composition of population is not ideal in this case, since the lowest level of geographic unit for which such data is available is *tehsil*, which (a) does not always map to a single AC and (b) hard to map to polling stations, since the map of geographic area covered by a polling station is not available and data on location of polling stations is also error-prone (Hintson and Vaishnav 2021). Additionally, electorate share of Muslims is the ideal measure, which can differ from their population share because of various reasons such as differential fertility and child survival rates etc., which could be correlated with their support for BJP.

[9]Religious identity, especially the Hindu-Muslim divide, is a salient political cleavage in India (Bhalotra, Clots-Figueras, Iyer, and Vecci 2021, Varshney 2003). Moreover, the salience of religion has heightened under BJP's rule since 2014 (Khosla and Vaishnav 2022).

manipulation.

To examine whether turnout manipulation was in part facilitated by weak monitoring of counting of votes, I analyze the assignment of counting observers across PCs. I compute the fraction of counting observers assigned in a PC who are from the State Civil Service (SCS), as opposed to the Indian Administrative Service (IAS).[10] Since SCS officers are appointed by the state government, unlike the IAS officers who are centrally appointed, they more likely to be politically pliable. I also compute the fraction of observers in a PC who are SCS and work in a BJP ruled state.[11] I find that both fraction exhibits large, positive and statistically significant discontinuity at the BJP win margin of zero. For the fraction of SCS officers from BJP ruled states, the discontinuity is larger in magnitude in PCs of BJP ruled states, while it is smaller and statistically insignificant for non-BJP ruled states. Additionally, in PCs won by BJP, the fraction of counting observers who are SCS and come from BJP ruled states positively predicts the extent of turnout data discrepancy in the PC; in PCs that BJP lost, no such relationship holds.

I analyze polling station level election results to test for local manipulation. For each polling station, I compute the vote share of BJP at that polling station relative to its vote share in the PC; I refer to this as the relative BJP vote share. This makes comparison of polling stations across constituencies easier. I show that within a constituency, relative BJP vote share typically hovers around one across polling stations with different turnout, except in closely contested constituencies barely won by BJP in BJP ruled states. In those constituencies, the relative vote share of BJP exhibits a large spike in polling stations with high turnout. The pattern is replicated with a polling station level indicator of BJP's vote share exceeding 95^{th} percentile of its distribution.[12] I compute the distribution of second digit in the polling station level vote tallies of candidates to measure departure from Benford's law at the polling station level. Benford's law (Benford 1938) specifies distribution of digits in naturally occurring numbers, and departures of the observed distribution from Benford's specification is often used as an indicator of manipulation. I show that the departures from Benford's law exhibit the same pattern. Additionally, I perform tests on the shape of the BJP's vote share density, proposed in more recent research on electoral fraud, and find results consistent with fraud. Moreover, the spike in the relative BJP vote share mentioned above is higher in PCs with larger discrepancy in turnout data. While the first couple of results are consistent with both mechanisms, the rest of the results indicate manipulation.

Finally, manipulation in the form of targeted electoral discrimination against Muslim voters would imply that *within* a PC barely won by BJP, high vote shares of the party should be concentrated in areas with higher Muslim presence. On the other hand, if precise control is the appropriate explanation, then we should expect the opposite, as the increase in BJP's vote share in 2019 relative to 2014 came primarily from Hindus, especially from its lower caste groups, while its support among Muslims was low and constant across the two elections (Varshney 2019).

I match the data on AC level electorate share of Muslims (described above) to polling stations to test the above hypothesis. ACs are smaller than PCs, and each PC contains about 7 ACs on average. The matched data therefore provide us *within* PC variation in electorate share of Muslims across polling stations located in different ACs. I find that in PCs that BJP barely lost, its vote share is less likely to exceed the 95^{th} percentile in polling stations located in high Muslim share ACs within the PC. However, this negative relationship gets significantly reduced in PCs barely won by the party; in those PCs, the likeli-

[10] There are typically multiple counting centers in a PC, each of which is assigned a counting observer.

[11] Observers are deployed in a state different from where they work.

[12] In those polling stations, BJP on average received 90% of votes cast.

hood of the event does not fall in ACs with higher Muslim share. This again supports the manipulation hypothesis.

The paper is unable to comment on the overall extent of manipulation in the 2019 general election. It focuses on closely contested constituencies as an empirical strategy to detect the presence of potential manipulation. Back of the envelope calculation shows that in PCs with BJP win margin less than 5%, BJP's "excess" win is in about 11 PCs. Therefore, even if all the disproportionate wins of BJP in closely contested PCs is due to manipulation, it likely would not have changed the government formation.[13] Nonetheless, the results signify a worrying development for the future of democracy in India and consequently, in the world at large.

This paper contributes to our understanding of democratic backsliding in consolidated democracies using objective measures. Little and Meng (2023) argue that subjective evaluation of nature of democracy by experts may be subject to their biases. Therefore, claims about democratic backsliding need to be grounded in more objective evidence. The authors examine objective measures of democracy and do not find any evidence of systematic backsliding across democracies. They conclude, "[...] it may be the case that major backsliding is occurring precisely in ways that elude objective measurement. However, this is an extraordinary claim, which requires a stronger theoretical and empirical basis than has been offered to date."

Additionally, several studies examining democratic backsliding have focused on the "demand side" issues, specifically, voters' willingness to sacrifice democratic principles in the context of increased polarization (Braley et al. 2022, Fishkin et al. 2021, Graham and Svolik 2020), rise of populism (Martinelli 2016) etc., resulting in dismantling of check-and-balances (Şaşmaz, Yagci, and Ziblatt 2022). The paper shows that dilution of electoral integrity is also an important and "supply side" contributor to democratic backsliding. Several consolidated or stable democracies, such as India[14], Mexico[15], Hungary (Scheppele 2022), have witnessed weakening of its electoral institutions in recent times. It is relevant to understand whether and how this weakening contributes to democratic backsliding. There is little evidence in mature democracies of direct electoral fraud typically observed in weaker democracies, such as ballot stuffing, booth capturing or direct manipulation of data by election authorities. Incumbents in these countries are likely to adopt subtler strategies, such as fragmentation of opposition (Arriola, Devaro, and Meng 2021) or voter suppression (Manheim and Porter 2019) etc. My examination of the latest general election in India adds to our understanding of this process.

Empirical analyses of electoral fraud have typically focused on weak democracies such as Afghanistan (Callen and Long 2015), Ghana (Asunka et al. 2019), Nigeria (Onapajo and Uzodike 2014), Russia (Enikolopov et al. 2013, Rundlett and Svolik 2016), Mexico during 1980s (Cantú 2019), nineteenth century Germany (Ziblatt 2009), or local elections in robust democracies such as Japan (Fukumoto and Horiuchi 2011). In these cases, the nature of fraud typically entails aggregation fraud, tampering of election documents at the polling station level etc. In case of India, my paper shows, manipulation took the form of localized and targeted discrimination against a well-identified minority group, via manipulation of voter registration as well as weaker monitoring of the election process.

Previous studies have employed several methods to detect electoral fraud – Cantoni and Pons (2020) use sampled data on proven and suspected fraud cases, Asunka et al. (2019) and Enikolopov et al.

[13]BJP won 303 PCs and it needed 272 PCs to form the government.

[14]https://www.telegraphindia.com/india/election-commission-weak-kneed-say-former-officials/cid/1688448

[15]https://www.bbc.com/news/world-latin-america-64742733

(2013) examine effect of poll observers on incumbent vote share, Christensen and Schultz (2014) analyze turnout behavior of specific voting groups more likely to be targeted for frauds, James and Clark (2020) conduct survey of polling station workers etc. My paper contributes methodologically by analyzing irregularities across polling stations and constituencies with different demographic composition of minority voters and applying regression discontinuity and difference-in-discontinuity designs. Additionally, papers on electoral fraud typically employ one specific method to detect fraud. In contrast, this paper employs a combination of methods to demonstrate consistent results. This is of particular importance given that the nature of irregularities is more subtle, as one may expect in a consolidated democracy, and hence, requires a deeper examination.

II Background and Context

Autonomy of the Election Authority in India: Election Commission of India is the central authority in charge of conducting national (and state) elections in India. It was established in 1950. Several scholars have highlighted the exemplary role played by the ECI in ensuring free and fair elections and consequently, in the consolidation of India's democracy, in spite of its challenging social, cultural and economic environment. Banerjee (2017), for example, says: "In contrast to the usual inefficiencies of Indian public institutions, the well-oiled machinery of the Election Commission stands out because of its excellent performance in conducting elections on an unimaginably large scale." (p 410) Sridharan and Vaishnav (2017) point out: "What has emerged over the past six-and-a-half decades is an Election Commission that has significant powers, far greater than what its counterparts in many democracies have at their disposal. [...] According to a 1996 poll conducted by the Centre for the Study of Developing Societies, the ECI was the most respected public institution in all of India with 62 per cent of respondents favourably disposed. A 2008 study found that an even higher percentage – nearly 80 per cent – of Indians surveyed expressed a high degree of trust in the Commission, second only to the army among state institutions." (p 419 of Kapur et al. (2018)) Multiple researchers have shown that redistricting in India does not suffer from gerrymandering, a common phenomenon in the US, thanks to the independence of the Delimitation Commission of India from any political interference (Kjelsrud et al. 2020, Nath et al. 2017, Iyer and Reddy 2013). Eggers et al. (2015) in their study of elections across a number of countries find no evidence of manipulation of election results in India for the period 1977-2004.

Elections in India: India follows a Parliamentary system. The Parliament has 543 legislatures or Members of Parliament (MPs), each of whom is elected from a Parliamentary Constituency (PC) using the first-past-the-post rule. The national or general elections in India are conducted every 5 years, unless there is an early dissolution of the government. There are several parties that field candidates in the general elections. The two main national parties are the BJP (Bhartiya Janata Party) and the INC (Indian National Congress). Apart from the national parties, there are several regional or state parties that are important political actors in specific states.

The ECI also conducts the state elections in India. In a state election, voters from each Assembly Constituency (AC) elect one representative (Member of Legislative Assembly) to the state legislature. The size of the legislature in a state depends on its population. Taken together, there are roughly $4,300$ ACs in India. An AC is always subsumed within a PC. The timing of state elections is not synchronized with the general elections. During every general election, a subset of states has their state assembly elections simultaneously with it. But the subset of states changes over time due to either early dissolution of state government, or the central government or both (Balasubramaniam et al. 2021). In 2019, the states of Andhra Pradesh, Orissa, Arunachal Pradesh and Sikkim had simultaneous general and state elections.

General Election in 2019: The most recent general election in India happened in 2019 that reelected the incumbent coalition (the National Democratic Alliance or NDA), led by the BJP, to power. There are two significant developments related to the 2019 general elections that are worth highlighting. First, the incumbent party, BJP, had built up its grassroots organizational presence significantly in the lead up to the 2019 elections, especially in certain states. The party deployed it efficiently during its 2019 election campaign, as discussed in Jha (2017). The author describes that the party created polling booth level committees who were in charge of connecting with voters enrolled in the booth, organizing membership drives, collecting household level data on various social, demographic and economic indicators, along with their political attitudes. The households were classified according to their intention to vote for the party to decide the party's campaign strategy. This localized campaign, in conjunction with the allocation of abundant campaign resources, gave the party an edge over the other parties.

At the same time, there were reports of mass deletion of voter names of minority groups from electoral rolls (Malhotra 2019, Trivedi 2019, Naqvi 2022). Since the incumbent party enjoys lower electoral support among the minority groups, such deletions may provide an electoral advantage to the party. Additionally, subsequent to the elections, the ECI released two "final" versions of the PC level EVM turnout data that did not match (Agarwal 2019). ECI did not provide any accounting of the data discrepancy.[16] These reports raise fears about possible electoral manipulation during the 2019 elections.

III Data

Aggregate Election Results: I first access the candidate level Parliamentary election results from 1977-2019 and state assembly election results of 2019-2021. It is published by the Election Commission of India, and is compiled and made public by the Trivedi Centre for Political Data (TCPD) at Ashoka University (Bhogale et al. 2019).[17] The data contain for each PC (AC, in case of state election) and each election year, details of candidate names, their party affiliations, votes received by candidates, total turnout and electorate size.

Two Versions of EVM Turnout Data: The Election Commission of India (ECI) initially published "Final Voter Turnout" figures for the first four (out of seven) phases of the 2019 general election.[18] These figures reflect the PC wise number of votes polled in the Electronic Voting Machines (EVMs). These numbers however do not match with the PC wise number of votes counted in the EVMs, as available in the official website of the ECI. This is unusual as votes polled and votes counted in the EVMs should be identical. The news media pointed out this discrepancy in the data, following which

[16]The Association for Democratic Reforms, an independent election watch body, has filed a petition in the Supreme Court of India seeking reconciliation of the data: https://www.nationalheraldindia.com/india/adr-files-petition-in-supreme-court-on-mismatch-in-evm-data.

[17]The data is publicly available from the TCPD's website http://lokdhaba.ashoka.edu.in.

[18]For the rest of the PCs, it released the "estimated" turnout figures, and therefore, are not considered for analysis.

the ECI removed the "Final Voter Turnout" figures from its website. The PDF copies of the data are publicly available here: https://www.scribd.com/docu ment/411811036/EC-s-votes-polled-data-Phase-1. I digitize the data and match it against the (revised) official EVM turnout figures available in the ECI's website: https://eci.gov.in/files/file/10969-13-pc-wise-voters-turn-out/.

Counting Observers in 2019: The ECI appoints officials who are responsible for overseeing and monitoring the counting of votes in each counting center. They are referred to as counting observers. The counting observers have the power to stop the counting process or not declare the results if they find breach of counting procedure or if they suspect that some form of fraud has taken place. They have to report to the ECI if they take such actions. I access the list of counting observers for 2019 general election from the official website of ECI. The data contain the names of officials assigned to each PC, along with their 'office state', i.e., the state where they were currently working as a bureaucrat, their 'home state', i.e., where they were born, whether they are an Indian Administrative Service (IAS) officer or from the State Civil Service (SCS) and their year of joining the service. I am able to match data for 539 PCs (out of 543) containing $1,804$ counting observers.

Polling Station Level Results: I put together polling station level election results for the 2019 general election. Polling station level election records are available in each of the states' Chief Electoral Officer's official website. The format of the data differs from state to state. While in one state the digitized data is available, in most states the data come in the form of scanned PDFs containing the polling station level results for each constituency. I scrape, digitize, clean and compile the results for 22 major states of India covering more than $900,000$ polling stations. For each polling station, the data provide the PC and AC it falls under, candidate-wise vote tallies (along with votes in favor of "None of the Above") and candidate's party affiliation. This allows me to calculate the absolute turnout and vote share of BJP at the polling station level. Except the state of Uttar Pradesh (UP), the data do not contain number of electorates at the polling stations. Therefore, barring UP, it is not possible to calculate turnout rate at the polling station level.

National Election Survey 2019: National Election Survey (NES) is a post-poll voter survey conducted by the Center for the Study of Developing Societies (CSDS). The surveys, conducted right after every general election but before declaration of results, ask a representative sample of voters in a randomly selected sample of PCs questions about their political attitudes, knowledge and activities, among other things. NES has been conducted regularly in India since 1990s and is a credible source of voter preferences and political activities (Balasubramaniam et al. 2021, Banerjee et al. 2019, Thachil 2014). I access the relevant sections of the NES 2019 data to examine the campaigning activities of parties. NES 2019 surveyed 24,236 voters across 208 PCs.

Muslim Electorate Share using Voter List: I create reliable estimates of Muslim electorate share at the AC and PC level to examine electoral discrimination against Muslims as a potential source of manipulation. For this, I use a 3 percent random sample of registered voters, representative at AC level (about 25 million observations). I access this proprietary data from a private organization that has compiled the full list of registered voters for the entire country using the electoral rolls published by the ECI. The data uses electoral rolls published till 2018, and therefore, is not subject to any strategic deletion that may have happened during the 2019 revision prior to the general election. I use a religion prediction algorithm (with 97 percent accuracy) developed by Chaturvedi and Chaturvedi (2023) to predict each voter's religion from their name, which allows me to compute Muslim share in each AC.[19] Appendix Figure A1 plots the local polynomial relationship between vote share received by all Muslim candidates running in an AC in a state assembly election (during 2008-2018) against the electorate Muslim share and finds strong positive relationship. This indicates that my measure of AC level Muslim share is reliable.

IV Irregularities in Aggregate Election Results

I first perform the McCrary test that checks for discontinuity in the distribution of win margin (of any party) at the value of zero (Calonico et al. 2014, McCrary 2008) . The presence of discontinuity would imply that there is disproportionately higher mass of closely contested electoral constituencies where the party has barely won or lost, depending on whether the discontinuity is positive or negative. The idea is that, if elections are fair, then conditional on an election being closely contested between party A and any other party, the party A's chance of winning would be close to 50%. This is because, whether it ends up winning a really close election would effectively be random. I perform the McCrary tests by computing the win margins for the two major national parties of India, namely BJP and INC. Any party A's win margin is defined as:

$$\begin{aligned} \text{Party A win margin} &= \text{(vote share of A - vote share of winner)}, \quad \text{if A loses} \\ &= \text{(vote share of A - vote share of runner up)}, \quad \text{if A wins} \end{aligned}$$

BJP win margin therefore takes negative values in constituencies where it lost and positive values where it won.[20] When the variable takes values close to zero, it implies that BJP either lost or won the election with a narrow margin. Similarly, a large negative (or positive) value would imply that BJP lost (or won) that election with a large margin. Same is true for INC win margin. Figure 1a and 1b plot the densities of BJP and INC win margins, respectively, for the 2019 general election. I observe a large discontinuous jump in the density of BJP Win Margin just right of zero. This implies that conditional on a closely contested election between a BJP candidate and another candidate, BJP was significantly more likely to win that election than lose. I do not observe any discontinuity in the density of INC Win Margin.[21]

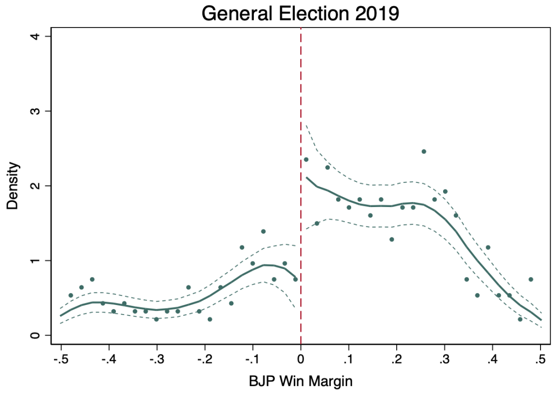
Before commenting on the interpretation of this finding, I wish to point out that failure of McCrary test in electoral context is rare, both in India as well as internationally. I perform the test for past general elections of India using BJP and INC win margins. Table 1 reports the estimated discontinuities in the densities of the two variables for all general elections going back to 1977. I find that the BJP win margin in 2019 is the only case exhibiting statistically significant estimate of the discontinuity. To illustrate this point more clearly, Appendix Table A1 reports the number and percentage of constituencies BJP won and lost in constituencies with small absolute BJP win margins for the past 4 general elections. I consider three narrow win margin bands - within 0.05, 0.03 and 0.02. For each band, the 2019 elections

[19] I thank Sugat Chaturvedi for implementing the algorithm in the data.

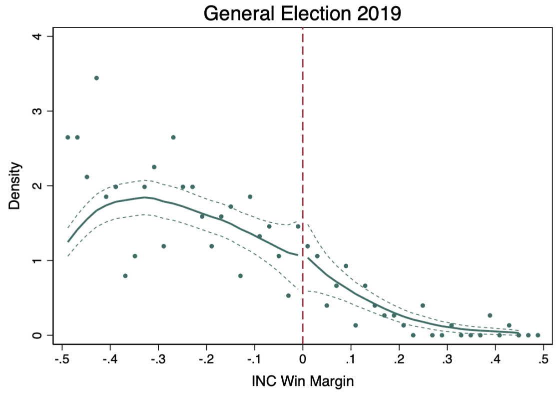
[20] This is a standard definition in this kind of exercise. Nellis et al. (2016), for example, use the same running variable defined for INC to estimate the causal effect of electing an INC politician on violence.

[21] It implies that the disproportionately higher wins of BJP candidates were primarily against regional parties.

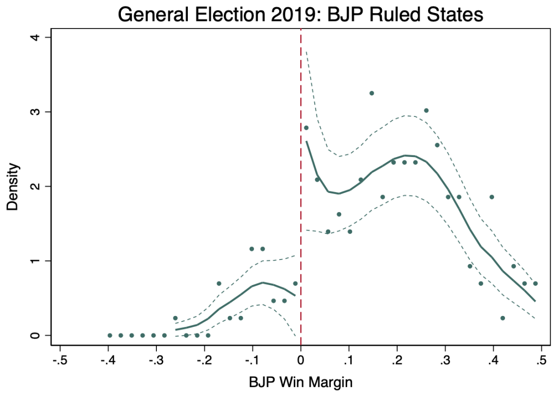
Figure 1—McCrary Tests Demonstrate Discontinuity in BJP Win Margin Distribution



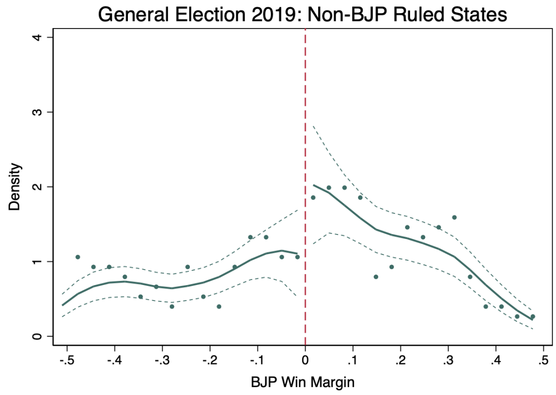
(a) BJP Win Margin: All States



(b) INC Win Margin: All States



(c) BJP Win Margin: BJP Ruled States



(d) BJP Win Margin: Non-BJP Ruled States

show the most lop-sided share of win for BJP; the share of BJP victory is 69-74% in 2019, depending on bandwidth. For each bandwidth, 2019 is the *only* year where the data rejects the null that the likelihood of BJP victory is 0.5. In each case, the null hypothesis is rejected with with p-value less than 0.01, i.e., there is less than 1% probability of observing the patterns with BJP's true probability of victory in close elections being 0.5.

Nellis, Weaver, Rosenzweig et al. (2016) find that INC win margin passes the McCrary test in state assembly elections for the period 1962–2000. Uppal (2009) find the same using incumbent win margin as the running variable for state elections during the period 1975–2003. Moreover, the only evidence of failure of McCrary test that has been documented in a robust democracy, is in the context of elections in the US (Jeong and Shenoy 2020, Vogl 2014, Caughey and Sekhon 2011). Eggers et al. (2015), however, have shown that it is in fact an exception as the test works in a number of countries (including the US and India) and for different time periods. Hence, the failure of the test in the 2019 general election in India warrants notice and additional investigation.

Interpretation: While the result is consistent with possible manipulation of the election results in favor of the BJP, the incumbent party, it is not the only interpretation. Alternatively, it could be that BJP, being the incumbent, was able to exercise *precise control* over win margin, i.e., it was able to precisely predict win margins, especially in constituencies where a close contest was expected, and was able to

Table 1—Estimates of the Discontinuity in the Density of BJP and Congress Win Margins

	2019 (1)	2014 (2)	2009 (3)	2004 (4)	1999 (5)	1998 (6)	1996 (7)	1991 (8)	1989 (9)	1984 (10)	1980 (11)	1977 (12)
Estimate of discontinuity: BJP Win Margin	1.51**	-0.24	-0.83	1.88	2.41	-0.79	-1.20	0.43	-0.01			
	(0.75)	(0.74)	(1.15)	(1.20)	(1.63)	(1.28)	(0.94)	(0.60)	(1.24)			
Estimate of discontinuity: INC Win Margin	0.78	-0.37	1.80	-1.02	-1.37	-0.24	0.66	-1.19	0.36	0.49	-0.54	0.89
	(0.60)	(0.73)	(1.30)	(1.03)	(0.91)	(1.01)	(0.79)	(0.81)	(0.77)	(0.86)	(0.73)	(0.71)

Notes: The table reports the estimates of the discontinuity in the density at the threshold value of zero for two running variables – BJP Win Margin (first row) and INC/Congress Win Margin (second row). The year in each column refers to the general election year. Each estimate, therefore, comes from a separate test for a given of running variable in a given general election year. The estimates are computed using the method proposed by Calonico et al. (2014). The robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

affect it, thanks to its comparative advantage in electoral campaigning and greater access to resources. Notice that it is not enough for BJP to predict the constituencies where it will face a close fight to generate failure of McCrary test. In such a case, it would campaign harder in all the constituencies expected to have a close contest, resulting in a uniform shift of its win margin to the right and hence, no discontinuity would emerge at zero.[22] The party would have to accurately predict whether they are ahead or falling behind in the close contest. Hence, for precise control to be the explanation, BJP had to accurately predict the sign as well as the magnitude of the win margin, to be able to target the set of constituencies where it expects to lose in a close contest. Jeong and Shenoy (2020) have shown that incumbent parties in US state legislative elections do exhibit behavior consistent with precise control and it can explain their ability to consistently win majority of close races. While election prediction in India is still not as sophisticated as in developed countries such as the US, it is possible that BJP, due to its superior electoral machine, was able to precisely predict and affect win margins.

State Assembly Elections: Seven states had their state assembly elections in 2019, including four states where the state elections were held concurrently with the general election.[23] BJP was the incumbent party in the government in three of the seven states. I compute the BJP win margin for state election results for BJP and non-BJP ruled states separately. I find that it does not exhibit failure of McCrary test (Appendix Figures A2a and A2b). Same is true for state elections held in 2020 and 2021 (Appendix Figures A2c and A2d). If precise control is the mechanism responsible for Figure 1a, then we should expect it at work at state level elections as well, at least in 2019. I however do not find that.

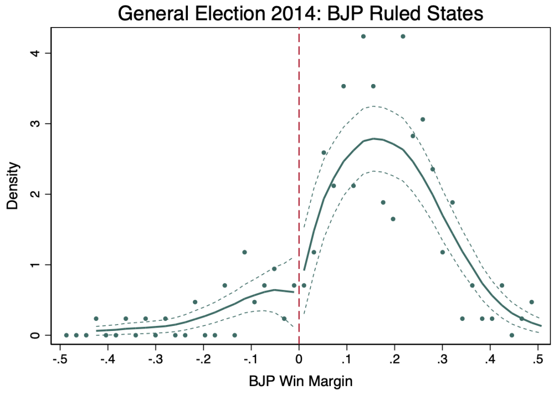
BJP vs. Non-BJP Ruled States: I now perform the McCrary test for the 2019 general election in two sub-samples of constituencies – those in states that were ruled by the BJP at the time of the 2019 election and those in non-BJP ruled states.[24] The two sub-samples have equal number of constituencies. Figures 1c and 1d show the densities of BJP win margin for the two sub-samples respectively. I find that for BJP ruled states, the density shows an even larger discontinuous jump to the right of threshold. For non-BJP ruled states, the jump is muted. Appendix Table A2 reports the estimated discontinuities in the densities for the two sub-samples of states separately. The estimate of the jump for BJP ruled states is highly statistically significant (p-value = 0.007), while it is statistically insignificant (p-value

[22] Lee and Lemieux (2010) refer to this as *imprecise control* over the running variable and argue that the regression discontinuity design remains valid under imprecise control.

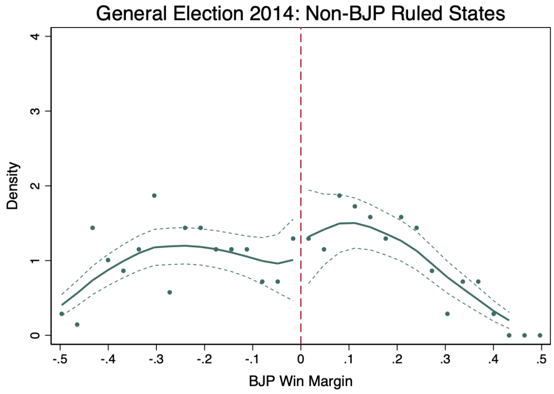
[23] The seven states are Arunachal Pradesh, Haryana, Maharashtra, Jharkhand, Andhra Pradesh, Odisha and Sikkim.

[24] In 2019, the BJP ruled states were Assam, Bihar, Goa, Gujarat, Haryana, Himachal Pradesh, Jharkhand, Maharashtra, Manipur, Nagaland, Tripura, Uttar Pradesh, Uttarakhand.

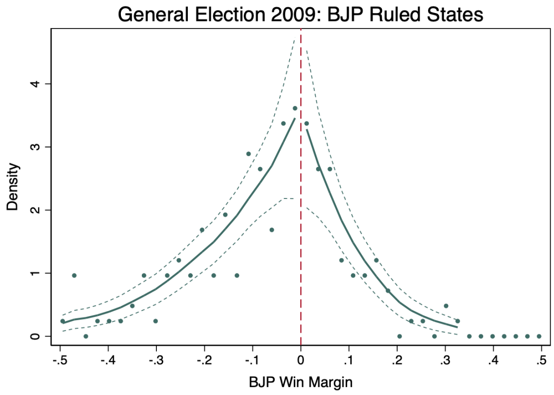
Figure 2—McCrary Test for General Elections in 2014 and 2009



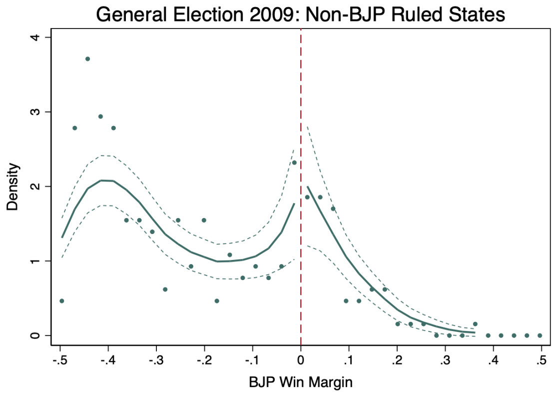
(a) 2014 Election: BJP Ruled States



(b) 2014 Election: Non-BJP Ruled States



(c) 2009 Election: BJP Ruled States



(d) 2009 Election: Non-BJP Ruled States

= 0.84) for non-BJP ruled states. Therefore, the overall failure of McCrary test is primarily driven by constituencies in the BJP ruled states. Those states, on the other hand, do not exhibit differential patterns in the previous two general elections (Figure 2).

The results are consistent with both mechanisms. Having control over the state's bureaucratic machinery can help a party target its manipulation efforts better, especially in a context where widespread manipulation is hard to implement given the intense media attention during elections and vocal rival political parties.[25] It is also consistent with precise control if being in power at the state government helps in mobilizing party workers at the ground. Greater presence of party workers at the ground can generate more precise information about a party's expected vote share vis-a-vis the main rival party, which can facilitate precise control.[26]

Comparability of PCs that BJP Closely Won and Lost: Table 2 reports the estimates of discontinuity of various PC level electoral variables at the BJP win margin threshold of zero. It finds no systematic

[25]Even though bureaucrats (Indian Administrative Service officers) are employees of the central government, their appointment and promotion are influenced by state governments (Iyer and Mani 2012). During general election, they report to the ECI, but the pool of officers available is shaped by the state government, making them pliable to the interests of the incumbent party at the state.

[26]Among the states not ruled by BJP, several of them, such as Madhya Pradesh, Rajasthan, Karnataka, Orissa, have strong presence of the party. Appendix Figure A3 shows the discontinuity for that subsample of states. It does not exhibit a differentially larger discontinuity than the one in Figure 1d.

differences between PCs that BJP barely lost and won. The variables examined are electorate size, turnout rate[27], number of candidates, reservation status for SC/STs, share of female candidates, share of candidates switching political parties (i.e., turncoats), whether the incumbent is running in the election, and whether BJP won the PC in the previous general election. The coefficient on BJP victory in 2014 is large and negative, though is noisily estimated. Therefore, in terms of various characteristics of PCs, the ones that BJP barely lost vs won appear to be comparable.

Table 2—Comparability of PCs across BJP Win Margin Threshold

	Electors	Turnout rate	#Candidates	SC/ST reserved	Female share	Turncoat share	Incumbent rerun	BJP Won in 2014
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BJP Won	0.178	-0.029	-0.931	0.125	-0.004	-0.002	-0.009	-0.170
	(0.172)	(0.036)	(1.273)	(0.154)	(0.016)	(0.010)	(0.172)	(0.181)
Mean Dep. var.	1.66	0.69	14.46	0.30	0.09	0.03	0.41	0.54
Bandwidth (h^*)	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
Observations	189	189	189	189	189	189	189	189

Notes: The table reports RDD estimates using BJP win margin on various PC level variables, such as electorate size in millions (column 1), turnout share (column 2), number of candidates (column 3), Reservation status for SC/ST (column 4), share of female candidates (column 5), share of candidates who switched parties (column 6), whether the incumbent is running (column 7) and whether BJP won the PC in 2014 (column 8). The bandwidth used is the optimal bandwidth used for McCrary test in Figure 1a and Table A2. Standard errors are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

V Evidence on Precise Control via Campaigning

BJP's ability to exercise precise control is likely to arise from the party's superior organizational strength and the related election campaign strategy. It allowed the party to collect and utilize detailed and localized information about voters' attitudes and voting intentions. This can result in precise control, as the party can mobilize voters better than its opponent in closely contested elections, resulting in disproportionate wins. This is similar to Vogl (2014) who argues that in mayoral elections in southern US, Black voters were better mobilized than White voters, resulting in disproportionate wins of Black candidates in closely contested elections. Therefore, for precise control to be the primary explanation of Figure 1, I expect the constituencies barely won by BJP to have significantly more campaigning by BJP *relative to* other parties. It is well-established in the literature on political campaigning that a party's relative campaigning, as opposed to its absolute campaigning activity, matters for its vote share (Bekkouche et al. 2022, Gerber 1998, Levitt 1994). Hence, the ideal test would estimate the discontinuity in the relative campaigning by BJP at the BJP win margin value of zero.

There is no existing data on campaigning by political parties across constituencies in India. To address this, I exploit a new question added to the post poll National Election Survey in 2019. In the 2019 round, the new question asks: "Did a candidate/party worker of the following parties come to your house to ask for your vote in the last one month?" The survey listed the major political parties in each state. The response to this question, therefore, reveals campaigning activity separately by *individual political parties* in each of the sampled PCs.[28] I use this question to define two dummy variables – *home*

[27] This is calculated using the revised turnout numbers as they are last version of data available with ECI.

[28] In previous rounds, the survey asked whether any candidate or party worker visited the respondent's house for campaign-

Table 3—Campaigning by Political Parties in Closely Contested Elections

	Home Visit by Party Worker/Candidate					
	Full Sample		BJP Ruled States		Non-BJP Ruled States	
	BJP	Any Other Party	BJP	Any Other Party	BJP	Any Other Party
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: BJP Win Margin ≤ 0.191						
BJP Won	0.03	0.00	0.05	0.34**	0.04	-0.09
	(0.10)	(0.13)	(0.17)	(0.17)	(0.15)	(0.19)
Mean Dep. Var.	0.41	0.49	0.33	0.37	0.48	0.58
Bandwidth (h^*)	0.191	0.191	0.191	0.191	0.191	0.191
Observation	8945	8945	3927	3927	5018	5018
No. of PCs	76	76	32	32	44	44
Panel B: BJP Win Margin ≤ 0.160						
BJP Won	-0.01	-0.01	0.10	0.38**	-0.01	-0.14
	(0.11)	(0.14)	(0.18)	(0.17)	(0.16)	(0.20)
Mean Dep. Var.	0.41	0.49	0.33	0.39	0.47	0.57
Bandwidth (h^*)	0.160	0.160	0.160	0.160	0.160	0.160
Observation	7897	7897	3297	3297	4600	4600
No. of PCs	68	68	28	28	40	40

Notes: The sample is individual level survey data from the National Election Survey (post poll) 2019. The dependent variable in columns (1), (3), (5) is a dummy variable that takes value one if a BJP party worker or candidate visited the house of the respondent to campaign for general election and is zero otherwise. The dependent variable in columns (2), (4) and (6) is also a dummy variable that indicates whether party worker or candidate from any other party visited the house for campaigning. BJP Won is an indicator of whether BJP is the winner of the Parliamentary Constituency (PC). Sample in Panel A consists of PCs with BJP win margin less than 0.191 – the optimal bandwidth calculated using the MSERD method proposed by Calonico et al. (2014), while Panel B uses the sample of PCs with BJP win margin less than 0.16 – the optimal bandwidth used for McCrary test in Figure 1a and Table A2. Columns (1) and (2) use the full sample of PCs within the respective bandwidths. Columns (3) and (4) restrict the sample to states ruled by BJP during 2019 general election. The last two columns use the sample of non-BJP ruled states. Standard errors are clustered at the PC level and are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

visit by BJP, that takes value one if a BJP party worker or candidate visited the respondent's house, and *home visit by any other party*, that takes value one if any other party visited the house. The mean values of the two variables are 0.38 (BJP) and 0.50 (any other party). Moreover, they are positively correlated ($r = 0.58$), suggesting that parties tend to target similar set of "swing" voters.

I test whether likelihood of home visits by BJP and other parties increase discontinuously at BJP win margin value of zero, and whether the increase for BJP is larger than other parties. Table 3 reports the results. In Panel A, columns (1) and (2) report the RDD estimates for the two outcome variables using the optimal bandwidth of 0.191. We find that both estimates are small in magnitude and statistically insignificant. In columns (3) and (4), I restrict attention to BJP ruled states, as the failure of McCrary test is concentrated in that sample. The coefficient for BJP home visits is 0.05, which is statistically

ing, i.e., it did not ask the question for each party separately.

insignificant, while that for any other party is 0.34, which is statistically significant at 5%. Therefore, in this sample, the estimate for BJP is not larger than that of other parties. For non-BJP ruled states, both estimates again are statistically insignificant. The results remain same if we use 0.16 as the bandwidth (Panel B). Therefore, differentially greater campaigning by BJP relative to other parties cannot be the primary reason for its disproportionate win in closely contested constituencies.

Social Media and Election Outcomes: BJP and other major parties extensively used social media during the 2019 election campaign. While smartphone penetration in India is not widespread, social media could potentially play a pivotal role in shaping voting behavior. There is no micro-data on social media campaigning by political parties. The NES 2019, however, asks individuals about their social media usage. For each social media platform, I define a dummy variable ("social media user") that takes value one if an individual uses that platform at least once every day and zero otherwise. Appendix Table A3 shows how social media usage predicts voting in favor of BJP in the full sample. I find that only Facebook users are statistically significantly more likely to vote in favor of BJP, while users of other social media platforms either vote less for BJP or vote for other parties with equal likelihood.[29] For each PC p, I then run the following regression:

$$I(\text{Voted for BJP})_{ip} \;=\; \alpha_p + \beta_p Fb\_user_{ip} + \theta_p^{'} X_{ip} + \epsilon_{ip} \tag{1}$$

where the vector of controls X_{ip} includes gender, age and caste categories. β_p captures the propensity of Facebook users in constituency p to vote differentially in favor of BJP. The estimate of β_p therefore can be interpreted as a proxy of BJP's differential intensity of Facebook campaigning in the constituency. I use β_p as an outcome variable to test whether its value jumps discontinuously at BJP win margin of zero. Appendix Table A4 column 1 reports the RDD coefficient. It is positive and statistically significant, suggesting that constituencies barely won by BJP exhibited relatively more intense social media campaigning by the party. Columns 2 and 3 report the same coefficient when the RDD analysis is performed on BJP ruled and non-BJP ruled states separately. The estimate in column 2 is small and statistically insignificant, while that in column 3 is positive, comparable in magnitude to column 1 and statistically significant. The evidence therefore cannot explain the patterns observed in the previous section that failure of McCrary test is concentrated in the BJP ruled states.

VI Evidence on Manipulation

Manipulation of elections can take place at one of three stages of elections. First, at the time of voter registration, in the form of targeted deletion of names of voters who are unlikely to vote for the incumbent party. I refer to it as *registration manipulation*. Second, at the time of voting, when polling officers can strategically discriminate against registered voters, who are likely to vote against BJP. Finally, manipulation can take place at the time of counting of votes.[30] Distinguishing between voting and counting

[29] This is consistent with the recent investigative media report that Facebook gave preferential rates to BJP for political ads during 2019 campaign. See here: https://www.aljazeera.com/economy/2022/3/16/facebook-charged-bjp-lower-rates-for-india-polls-ads-than-others.

[30] A fourth possibility is through manipulation of EVMs. However, some commentators have pointed about that widespread manipulation of EVMs may be hard to achieve, given the technology (Purkayastha and Sinha 2019), making it an unlikely mechanism.

manipulation is difficult. Barring one analysis that comments directly on counting manipulation, the rest of the evidence are consistent with both voting and counting manipulation. Hence, I refer to both as *turnout manipulation*. Section II above mentions media reports of potential registration and turnout manipulations. In the sections below, I discuss evidence consistent with each of them.

VI.I Registration Manipulation

To examine the presence of this channel, I compute the growth in the number of electorate (i.e., number of registered voters) in a PC between 2014 and 2019. For each PC p, I define:

$$G_p \equiv \frac{Electorate_{p,2019} - Electorate_{p,2014}}{Electorate_{p,2014}}$$

If names were strategically deleted from the electoral rolls in an attempt to flip closely contested election in favor of BJP, then we should expect the electorate growth rate to fall discontinuously at BJP win margin value of zero. Moreover, if Muslims were the primary target of this strategic deletion, we expect a greater fall in the electorate growth rate in PCs with higher Muslim electorate share. I implement the regression discontinuity design on the full sample of PCs as well as on samples of PCs with Muslim share greater and lower than the median of distribution of Muslim shares across PCs.[31]

Table 4—Electorate Growth Rate Smaller in PCs Barely Won by BJP

	Electorate Growth Rate (G_p)					
	Full Sample (1)	High Muslim Share (2)	Low Muslim Share (3)	Full Sample (4)	High Muslim Share (5)	Low Muslim Share (6)
BJP Won	-0.05***	-0.06**	-0.02	-0.05***	-0.07***	-0.03
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Mean Dep. Var.	0.10	0.10	0.09	0.09	0.09	0.09
Observations	123	72	51	181	101	80
Bandwidth (h^*)	0.107	0.107	0.107	0.16	0.16	0.16

Notes: The data is at Parliamentary Constituency (PC) level. The table reports the regression discontinuity design estimate using BJP win margin as the running variable. The dependent variable in all the columns is the growth rate in the PC electorate between 2014 and 2019. BJP Won is a dummy indicating whether BJP won the PC in 2019. Columns (1) and (3) use the full sample of PCs. Columns (2) and (4) use PCs where the electorate share of Muslims is higher than the median, and column (3) and (6) use PCs where the share is lower than median. Columns (1)-(3) use the optimal bandwidth using the MSERD method specfied by Calonico et al. (2014), while columns (4)-(6) use the optimal bandwidth calculated for McCrary test in Figure 1a. Robust standard errors are are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4 shows the estimates for discontinuity for the three samples. Column (1) reports the RDD estimate for the full sample using the optimal bandwidth 0.125. The estimated discontinuity is -0.05, which is statistically significant at 1%. This implies that constituencies barely won by BJP had a 5 percentage points smaller growth rate in electorate between 2014 and 2019 compared to PCs that it barely lost. This is a large fall, given the mean growth rate of 0.09. Moreover, in PCs with higher Muslim share, the estimated fall is 6 percentage points (column (2)), while it is 2 percentage points (and statistically insignificant) in PCs with lower Muslim shares. The difference between the two estimates

[31] I calculate PC level Muslim electorate share by taking a weighted average of AC level Muslim shares using electorate share of an AC as the weight.

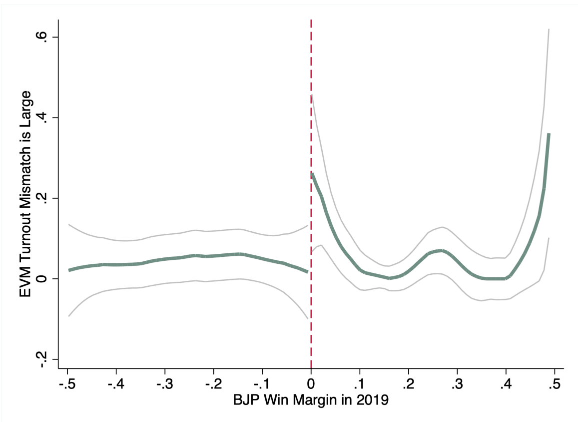
in columns (2) and (3) is statistically significant at 10%. The result remains the same if we use the bandwidth of 0.16. The result, therefore, is consistent with strategic deletion of Muslim names being an important channel of manipulation.

Additionally, Appendix Table A5 reports the RDD coefficients for high and low Muslim share PCs in the BJP ruled and non-BJP ruled states separately. Consistent with previous results, I find that the only statistically significant coefficient is for the sample of high Muslim share PCs in BJP ruled states (Column (1)). The Column (1) coefficient is also larger in magnitude than that low Muslim share in BJP ruled states (Column (2)), though the difference is not statistically significant. The result for non-BJP ruled states is also similar, though both the coefficients are noisily estimated.[32]

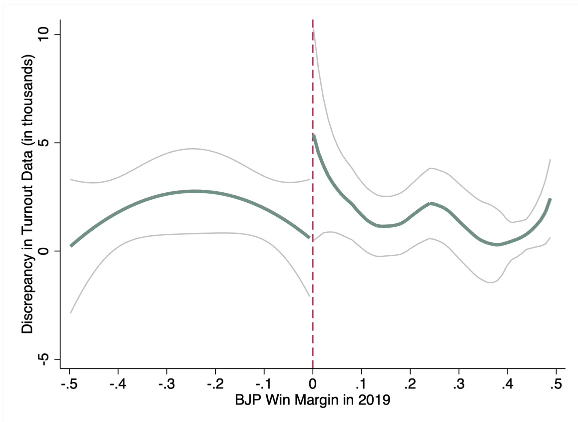
VI.II Turnout Manipulation: EVM Turnout Data Discrepancy

I compile the two different EVM turnout figures described in Section III for the 373 PCs covered in the first four phases of election. In 64% of PCs the turnout was revised up, and in the rest of the cases it was revised down. I compute the absolute difference in vote tallies between the two reports. While the median difference is 358, the 90^{th} and 95^{th} percentiles of the difference are 3302 and 7357, respectively. The largest mismatch is of 57,747 votes in the Gautam Buddha Nagar constituency in Uttar Pradesh. I define a dummy variable called "large" turnout discrepancy: it takes value one if the absolute discrepancy is larger than the 95^{th} percentile and zero otherwise. If the mismatch occurred due to some administrative errors or glitches in the EVM, then we expect the "large" discrepancies to be randomly spread across PCs with different BJP Win Margins.

Figure 3—EVM Turnout Data Mismatch in Closely Contested Constituencies



(a) "Large" Data Discrepancy



(b) Absolute Data Discrepancy

Figure 3a plots the relationship between the dummy variable and BJP win margin separately on the two sides of the threshold value of zero. I find that the probability of "large" discrepancy jumps significantly at zero. The estimate of the jump, using the method proposed by Calonico et al. (2014), is 0.26 (p-value = 0.008), implying that conditional on close election, the PCs that BJP barely won have 26 percentage point larger likelihood of having a "large" mismatch than PCs that BJP barely lost. This is a large effect considering the average value of the dummy variable, by construction, is 0.05. The result implies that the sample of closely contested constituencies that were disproportionately won by BJP

[32]The p-value of the Column (3) coefficient is 0.103.

also has a disproportionately higher likelihood of "large" turnout revision. If the failure of McCrary test demonstrated in the previous section involved manipulation of turnout figures, then we should expect this pattern. Figure 3b plots the same graph directly using the absolute turnout discrepancy (in thousands) and finds a similar pattern, though with a noisier estimate. The estimate of the discontinuity is 5.70 (p-value = 0.09).[33] Hence, the result is consistent with the manipulation hypothesis. Moreover, it is not obvious why precise control would lead to larger turnout revisions by the ECI in the closely contested constituencies won by the BJP.

Table 5—Heterogeneity in RDD Estimates of Turnout Discrepancy

	BJP Ruled States (1)	Non-BJP Ruled States (2)
	Panel A: "Large" Turnout Discrepancy	
BJP Won	0.45**	0.16
	(0.19)	(0.11)
	Panel B: Absolute Turnout Discrepancy	
BJP Won	15.54**	-0.43
	(7.42)	(1.02)
Bandwidth (h^*)	0.153	0.153

Notes: The table reports RDD estimates for two dependent variables – the dummy variable "Large" Turnout Discrepancy (Panel A) which takes value one if the absolute discrepancy in turnout data is larger than the 95^{th} percentile, and the absolute turnout discrepancy, in thousands (Panel B). The running variable is BJP Win Margin. The sample only includes the 373 constituencies for which turnout discrepancy information is available. Column 1 has states ruled by BJP in 2019, while column 2 has the rest of the states. Optimal bandwidth for Panel A is calculated using the MSERD method proposed by Calonico et al. (2014) and is maintained in Panel B. *** p<0.01, ** p<0.05, * p<0.1

BJP vs. Non-BJP Ruled States: Similar to the previous section, I test for heterogeneity across BJP and non-BJP ruled states. Table 5 reports the RDD estimates for the two sub-samples for both outcome variables. We observe in Panel A that the jump in the probability of "large" discrepancy is statistically significant and large in magnitude for BJP ruled states, while it is statistically insignificant and smaller in magnitude in non-BJP ruled states. The estimated jump in column 1 is 0.45 (and the value just to the left of threshold is close to zero), i.e., the likelihood of "large" discrepancy in the PCs barely won by the BJP is *9 times* higher than what it would be under the random chance scenario. It is, on the other hand, 3 times higher for non-BJP ruled states. The results in Panel B are similar. While the estimated jump in absolute discrepancy is more than $15,000$ votes (statistically significant at 5%) in BJP ruled states, it is -430 (statistically insignificant) in non-BJP ruled states. The difference between the coefficients in Panel A is not statistically significant, but in Panel B, it is significant at 5%.

[33] Appendix Figure A4 plots the same relationships using INC win margin and does not find discontinuity at the threshold.

Discontinuity in Turnout Difference: Appendix Figure A6 shows discontinuity in turnout difference at the BJP win margin threshold of zero for the past general elections. Turnout difference for a PC in a general election is the difference between its turnout rates in the current and previous elections. We observe that the discontinuity is positive and statistically significant for 2019. The discontinuity estimate is 0.023 (p-value = 0.05). For previous elections going back to 1989, the discontinuity estimates are statistically insignificant. This is consistent with the result that turnout data discrepancy went up in PCs barely won by BJP.

Interpretation: It would be inappropriate to treat the data revision as aggregation fraud (Callen and Long 2015), i.e., one where higher level ECI officials directly engaged in turnout manipulation while aggregating turnout data from polling station level results. Such acts by the ECI is unlikely. Moreover, in all cases, barring one, the revision is not larger than the win margin. Rather, the revisions are likely indicative of possible manipulations committed locally, at the polling stations. The local manipulations could either be at the time of voting or counting. Most electoral frauds are decentralized in nature, as Rundlett and Svolik (2016) point out. The analysis in the following section suggests that it was at least partly facilitated by weaker monitoring during counting, while the next section provides evidence that local manipulation may explain part of the observed turnout manipulation.

VI.III Counting Manipulation: Assignment of Counting Observers

I examine assignment of counting observers in PCs across the BJP win margin threshold. All counting observers are assigned to a state different from their 'office state' and 'home state', as defined in Section III. 36% of observers are from the SCS cadre. The SCS officers typically work in lower ranked positions in a state bureaucracy, as compared to the IAS officers with same experience (Iyer and Mani 2012). They are also more likely to be politically pliable by the state government, since they are appointed by them, as opposed to the IAS officers who are appointed by the central government. I compute the fraction of counting observers in a PC who come from the SCS cadre. About 50% of PCs have at least one SCS observer assigned. Since I know the 'office state' of each observer, I also compute the fraction of observers who are SCS and work in a BJP ruled state. The mean fraction is 0.13.

Table 6 columns (1) and (4) report the estimates of discontinuity in the two outcomes variables at the BJP win margin threshold of zero. In both cases, we find that the RDD estimate is positive – 0.24 and 0.22, respectively. They are large in magnitude and statistically significant at 5%. Columns (2) and (5) report the results for BJP ruled states and columns (3) and (6) for non-BJP ruled states.[34] All coefficients are positive and 3 out of the 4 coefficients are statistically significant; coefficients for BJP ruled states are larger in magnitude. Specifically, for the fraction of observers who are SCS and come from BJP ruled states, the coefficient is 0.37 and is statistically significant at 1% for BJP ruled states (column (5)) but is 0.17 and statistically insignificant for non-BJP ruled states (column (6)). Appendix Figure A5 depicts the RDD graphs for the four cases. The results indicate that more politically pliant counting observers were assigned in PCs barely won by BJP, and the pattern is concentrated in BJP ruled states.

I regress the absolute data discrepancy in turnout and the indicator for "large" turnout discrepancy computed in Section VI.II above on the fraction of counting observers who are SCS and from BJP ruled states, whether BJP won the PC and their interaction. I include all PCs in the sample to check whether in

[34]Here BJP and non-BJP ruled states refer to the PCs where the observers were deployed.

Table 6—RDD Estimates of Characteristics of Counting Observers

	SCS Observer			SCS Observer from BJP States		
	All States (1)	BJP Ruled States (2)	Non-BJP Ruled States (3)	All States (4)	BJP Ruled States (5)	Non-BJP Ruled States (6)
BJP Won	0.236**	0.335**	0.241*	0.224**	0.371***	0.166
	(0.100)	(0.131)	(0.136)	(0.104)	(0.0739)	(0.154)
Mean dep. var.	0.36	0.31	0.40	0.17	0.17	0.17
Observations	188	83	105	188	83	105
Bandwidth (h^*)	0.16	0.16	0.16	0.16	0.16	0.16

Notes: The table reports RDD estimates for two dependent variables – share of counting observers assigned to a PC who are State Civil Service (SCS) officers (columns (1)-(3)) and share of counting observers who are SCS and work in BJP ruled states (columns (4)-(6)). The running variable is BJP Win Margin. The columns (1) and (4) include all PCs with BJP win margin within 0.16, columns (2) and (5) include PCs with the same win margin and are in BJP ruled states, columns (3) and (6) include PCs with the same win margin and are in non-BJP ruled states. Robust standard errors are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

the full sample, assignment of politically pliant observers is correlated with data discrepancy. Appendix Table A6 reports the results. I find that in PCs that BJP lost, the relationship between data discrepancy and the fraction of SCS observers from BJP ruled states is negative. However, in PCs that BJP won, the relationship turns positive for both outcomes. For the absolute discrepancy measure, the relationship is statistically insignificant (p-value=0.104), while for the "large" discrepancy indicator, the relationship is statistically significant at 5% (p-value = 0.041). This suggests that greater presence of politically pliant counting observers in PCs barely won by BJP may have partly contributed towards discrepancy in turnout data.

VI.IV Irregularities in Polling Station Outcomes

This section examines irregularities in polling station level election results for 2019 from 22 major states of India.[35] For each polling station, the data provide information on the total turnout and candidate wise vote tallies. The data do not mention the number of electorates at the polling station level.[36] If the turnout discrepancy discussed above electorally benefited the incumbent party, then we should expect it to be reflected in its vote share across polling stations. To examine this, I compute vote share of BJP in each polling station in constituencies with a BJP candidate. To make polling station level BJP vote shares comparable across PCs, I then compute the relative vote share of BJP in each polling station j in each PC p:

$$\text{Relative BJP vote share}_{jp} = \frac{\text{BJP vote share}_{jp}}{\text{BJP vote share}_{p}}$$

[35] The states are Andhra Pradesh, Assam, Bihar, Chhattisgarh, Delhi, Goa, Gujarat, Haryana, Himachal Pradesh, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Punjab, Rajasthan, Tamil Nadu, Telangana, Tripura, Uttar Pradesh, Uttarakhand, West Bengal.

[36] The only exception is Uttar Pradesh; it releases the electorate size of polling stations along with vote tallies of parties in the same dataset.

where BJP vote share$_{jp}$ is BJP's vote share in a polling station and BJP vote share$_p$ is its vote share in the entire constituency that the polling station belongs to. I use the revised turnout figures to calculate vote shares, as that is the official data on turnout and are available for all PCs. Hence, *relative BJP vote share* captures the party's vote share in a polling station relative to its vote share in the constituency. If its value is greater than one, then in the polling station, BJP's vote share is higher than that in the constituency. The average value of Relative BJP vote share should be approximately one.

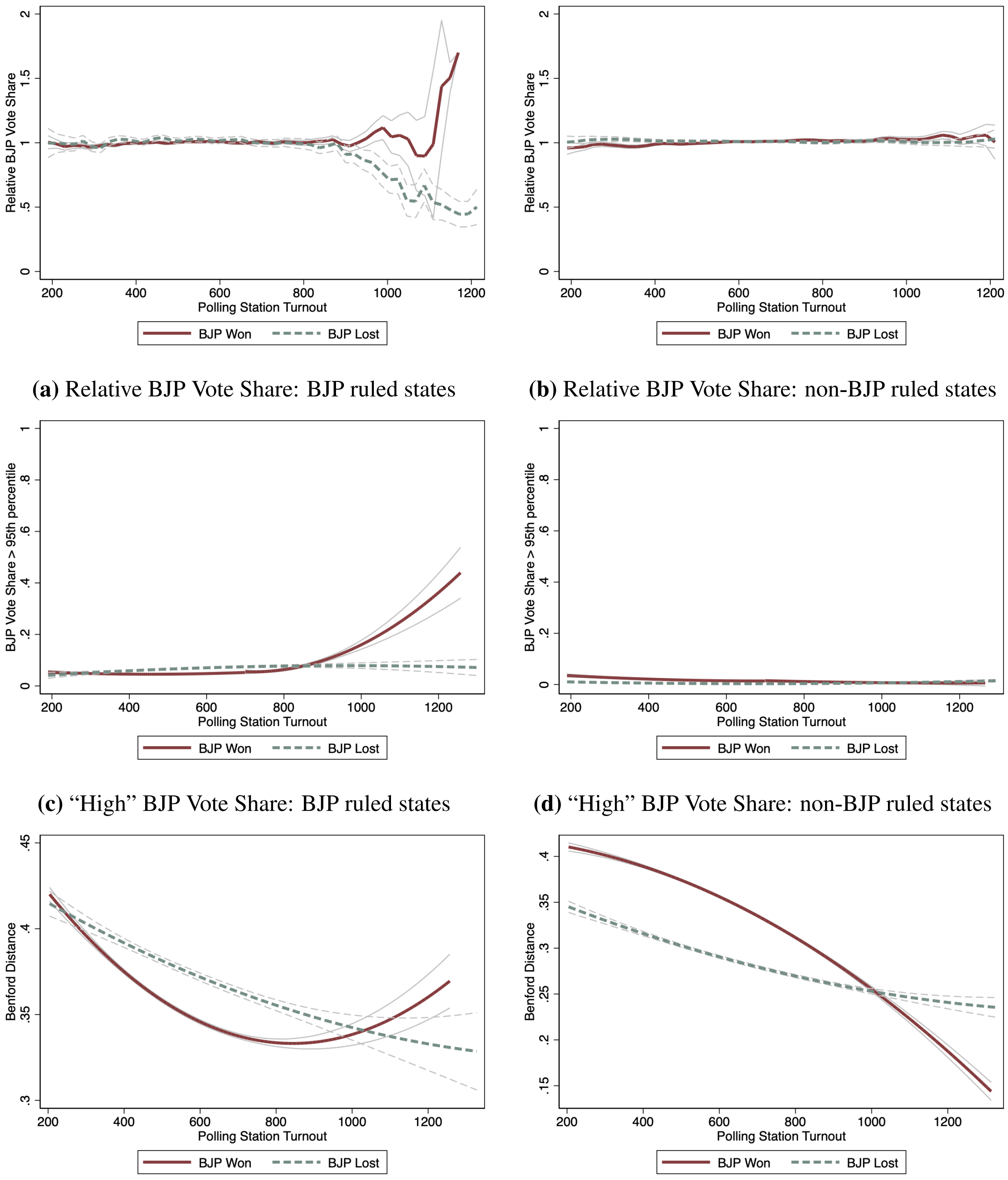
Figure 4a plots, for the states ruled by BJP, the local polynomial relationship between the relative BJP vote share and turnout at polling station level in closely contested constituencies, i.e., constituencies where BJP's win margin was less than 0.16 – the optimal bandwidth for the McCrary test performed in Figure 1a. I estimate it separately for PCs where BJP won and lost, depicted by the solid and dashed lines respectively. For comparison, Figure 4b plots the same graph for the states not ruled by BJP. Figure 4a shows that in both types of constituencies, the estimated relationship hovers around one for polling stations with turnout 800 or below. In polling stations with higher turnout, relative BJP vote share spikes in constituencies where BJP won, and falls in constituencies where BJP lost.[37] In Figure 4b we do not see such striking patterns.

I examine this directly by creating a dummy variable at the polling station level, called *"high" BJP vote share* that takes value one if the BJP vote share in the polling station is higher than the 95^{th} percentile of the BJP vote share distribution in the entire sample. In these polling stations, BJP's vote share on average is 0.90. By construction, the average value of the dummy variable is 0.05. The dummy variable essentially flags polling stations with "extreme" outcomes in favor of BJP. Figures 4c and 4d plot the relationships for BJP ruled and non-BJP ruled states. The sample of PCs is same as before – those with absolute BJP win margin within 0.16. We observe in Figure 4c that in both constituencies won and lost by BJP, the average value of *"high" BJP vote share* is low in polling stations with turnout below 800. However, in larger turnout polling stations, the estimated relationships diverge. In constituencies won by BJP, the average likelihood of "high" BJP vote share increases sharply, going beyond 0.4, while the other graph remains flat. This is consistent with Figure 4a. Moreover, there is no such pattern in the corresponding figure for non-BJP ruled states (Figure 4d). The pattern in Figure 4c is especially noteworthy given the fact that I only consider closely contested constituencies for the estimation. Hence, the rival party in these constituencies have received comparable vote share, which would make it less likely for BJP to get "high" vote shares in any polling station.

Bemford'a Law: I compute second digit distribution of absolute vote tallies across all candidates for each polling station to check departures from the Benford's law. Benford's law specifies the distribution of digits in different positions of naturally occurring numbers (Benford 1938, Raimi 1976). Manipulation of such numbers leads to a different distribution of digits, which allows analysts to detect the manipulation (Hill et al. 1995). This method is used in a variety of contexts to detect fraud (Diekmann 2007, Nigrini 2012) – such as, income tax receipts, financial transactions, as well as elections. In the context of election forensics, analysts usually focus on the distribution of second digits (Mebane

[37] The pattern for PCs lost by BJP is similar in constituencies that BJP lost with margin higher than 0.16 (Appendix Figure A7). Therefore, in all the PCs that BJP lost, it got lower vote share in polling stations with high turnout. These polling stations are likely to be located in urban centers. Since BJP's primary support base is more urban than other parties, less support in urban areas is a good indicator of its performance in a PC.

Figure 4—Distribution of BJP vote share across Polling Stations – Win margin ≤ 0.16



(a) Relative BJP Vote Share: BJP ruled states

(b) Relative BJP Vote Share: non-BJP ruled states

(c) "High" BJP Vote Share: BJP ruled states

(d) "High" BJP Vote Share: non-BJP ruled states

(e) Benford distance: BJP ruled states

(f) Benford distance: non-BJP ruled states

2008a,b).[38] However, treating a significant deviation from Benford's distribution in a given constituency as evidence of fraud (in that constituency) can lead to misleading conclusions, as researchers have shown that even in cases without fraud, empirical distributions of second digits can deviate from Benford's law (Shikano and Mack 2011). This can happen due to a myriad of reasons as discussed by Mebane (2011). I therefore do not test for deviations from Benford's distribution in each PC individually. I argue that presence (or absence) of patterns in deviations *across* PCs and polling stations is a better statistical test.

[38]There are other digit-based tests of electoral fraud, for example, examining distribution of last digits (Beber and Scacco 2012) etc. However, Benford's law is the most widely used method in this context.

I compute the Euclidean distance between the second digit distribution of the vote tallies of candidates in each polling station and the "ideal" second digit distribution specified by Benford's law. Figure 4e plots the Benford distance for each polling station against turnout for the same sample of PCs as Figure 4c. Figure 4f plots it for the sample of PCs used in Figure 4d. We observe that the Benford distance typically falls with larger turnout, except for PCs won by BJP in BJP ruled states. In those PCs, the Benford distance falls initially with turnout, but then rises for polling stations having turnout higher than 800. This is the same set of polling stations that exhibits irregular patterns, as discussed above. This suggests that the "extreme" outcomes in favor of BJP observed in Figure 4c may have resulted from some form of manipulation of vote tallies in that subset of polling stations.

Shape of vote share density: Some recent works on detection of electoral fraud examine the shape of the density of vote share and turnout distributions using the booth (or precinct) level data. Rozenas (2017), for example, test for presence of excess mass at coarse vote shares using a resampled kernel density method. The method returns as output an estimated fraction of booths exhibiting fraudulent results. I apply this method to the polling station level data from Uttar Pradesh (UP)– the only state for which turnout rate as well as BJP's vote share are available at polling stations, a requirement for the analysis. Also, a significant number of PCs from UP are in the list of PCs with narrow win margin (Appendix Table B1). I find that in the full sample, 0.13% booths are fraudulent. The share increases to 0.19% in PCs with BJP win margin less than 0.08 and won by BJP. The estimates are low but move in the direction that is indicative of fraud. Additionally, for comparison, Rozenas (2017) analyze data from Russian elections in 2011 and 2012 and find estimates of 0.94-0.97%. Klimek et al. (2012) argue that the density of vote share (appropriately calculated and scaled) exhibits high kurtosis in case of fraud and finds that Russian elections in 2011 and 2012 have kurtosis exceeding 10, while in other well-established democracies in Europe, it is typically 5 or lower. In UP, the kurtosis is 29. Appendix Figure A8 shows the density, which looks very similar in shape to those in Russian elections in 2011 and 2012 and unlike those in other countries (Figure 2 in Klimek et al. (2012)).

Interpretation: The patterns observed in Figures 4a and 4c are consistent with both mechanisms. If the incumbent party was able to accurately predict the win margins in closely contested constituencies, and wished to affect them, it might be optimal for the party to target the larger polling stations, as they are fewer in numbers and mostly located in urban areas, making voters easily accessible for campaigning. This may result in high vote shares for the party in large turnout polling stations. Figure 4e and the analysis of the shape of BJP's vote share density, however, advance the manipulation hypothesis over precise control.

VI.V Data Discrepancy and Irregularities at Polling Stations

To further distinguish between manipulation and precise control mechanisms, I utilize the fact that in closely contested PCs barely won by BJP, the EVM data discrepancy is significantly larger. If discrepancies in turnout data and high vote share of BJP in large polling stations are both driven by manipulation, then I should expect the two phenomena to be correlated, i.e., the irregular pattern in polling stations to be primarily driven by constituencies with larger data revisions. However, if precise control is the explanation, then such patterns should be similar irrespective of whether data revision was large or small.

This is because, in that scenario, larger data revisions only reflect administrative errors during counting of votes, which should be uncorrelated with BJP's ability to exercise precise control at the time of elections.

To test this hypothesis formally, I estimate the difference-in-discontinuity specification (Grembi et al. 2016) specified below:

$$\begin{aligned}\text{Rel. BJP vote share}_{jp} \;&=\; \alpha_1 + \gamma_1 BJP\_Won_p + \gamma_2 BJP\_Won_p \times D_p \\ &+\; \beta_1 BJP\_Margin_p + \beta_2 BJP\_Margin_p \times BJP\_Won_p \\ &+\; D_p \times \{\alpha_2 + \beta_3 BJP\_Margin_p + \beta_4 BJP\_Margin_p \times BJP\_Won_p\} + \epsilon_{jp}\end{aligned} \tag{2}$$

where D_p is the absolute discrepancy in turnout data in PC p, measured in unit of 10,000 votes and the rest are as defined before. γ_1 measures the RDD estimate for relative BJP vote share in PCs without any turnout discrepancy. γ_2 measures the differential discontinuity in the relative BJP vote share in PCs with additional discrepancy in 10,000 votes. Our coefficient of interest, therefore, is γ_2.

Table 7—Discrepancy in Turnout Data and Irregularity in Election Results

	Relative BJP vote share		
	All states (1)	BJP ruled states (2)	Non-BJP ruled states (3)
BJP Won	-0.063**	-0.105***	-0.094*
	(0.030)	(0.034)	(0.048)
Absolute Turnout Discrepancy	-0.234*	-1.738	-0.174*
	(0.137)	(1.059)	(0.092)
Absolute Turnout Discrepancy * BJP Won	0.236*	1.737	0.196
	(0.137)	(1.060)	(0.130)
Mean Dep. Var.	0.998	0.998	0.998
Observations	183,275	82,921	100,354
Bandwidth (h^*)	0.160	0.160	0.160

Notes: The data is at polling station level. The dependent variable in all columns is the ratio of BJP vote share in a polling station and BJP vote share in the PC. Absolute Turnout Discrepancy is the absolute mismatch (in unit of 10,000 votes) in EVM turnout data in 2019. Optimal bandwidth calculated for McCrary test in Figure 1a has been used in all specifications. Standard errors are clustered at the constituency level and are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

Before discussing the results, I emphasize that even though the difference-in-discontinuity method is used to estimate heterogeneity in causal effect of some treatment, that is not the appropriate interpretation in this context. The estimation of equation (3) allows us to examine whether irregular outcomes in the polling stations are positively correlated with extent of turnout discrepancy in PCs barely won by the BJP. Table 7 reports the results for the full sample (column (1)), BJP ruled states (column (2)) and non-BJP ruled states (column (3)). We find that estimate of γ_1 is negative and statistically significant in all columns, i.e., PCs barely won by BJP with no data discrepancy exhibits a *fall* in relative BJP vote share. However, estimate of γ_2 in column (1) is positive and statistically significant at 10%. Moreover, the magnitude of γ_2 is about 4 times larger than γ_1, suggesting that in PCs with discrepancy larger than 2500 votes, the relative BJP vote share is higher in PCs won by BJP (relative to PCs lost by BJP). Estimate of γ_2 in column (2) is 17 times larger than γ_1, while in column (3), it is twice as larger. The coefficients in both columns however are noisily estimated. We therefore have weak and suggestive evidence that turnout discrepancy (at the level of PCs) is positively correlated with irregular outcomes at the polling stations in constituencies barely won by the BJP.

VI.VI Turnout Manipulation in High Muslim Share Areas

This section tests whether electoral discrimination of minorities, specifically Muslims, is a potential source of turnout manipulation. Lehne (2022) shows, using individual voter level panel data on electoral rolls from the state of Uttar Pradesh during 2012-2017, that in state assembly constituencies with BJP incumbents (elected in 2012), Muslim voters have a significantly higher probability of being deleted from the electoral rolls in 2017. Neggers (2018) shows using data from the state of Bihar, that polling officers in charge of conducting election in a polling station exercise significant discretion in allowing registered voters, specially from minority communities such as Muslims, to vote. Since Muslim names are culturally distinct, Muslim voters are easily identified in the electoral roll. Therefore, they can be subject to both strategic deletion (discussed above) and strategic discrimination. Moreover, such exercises are easier in states controlled by the incumbent party, since the state government can influence assignment of officials in charge of electoral roll revisions as well as polling officers. Hence, if fraud is the appropriate explanation, I expect the polling station level irregularities to be concentrated in areas within a PC that have high Muslim presence.

Precise control, on the other hand, would predict the opposite. This is because, the party historically enjoyed minimal support among Muslim and consequently, spent significantly less effort in mobilizing Muslim voters. Jha (2017), for example, points out that the party did not focus on areas with significant Muslim presence, since it did not expect to get significant support from them. It instead directed its efforts towards voters who could be converted to vote in favor of the party, especially those belonging to lower castes among Hindus.[39] This is consistent with Varshney (2019) who reports, using NES data, that while support for the party increased substantially between 2014 and 2019, especially among Scheduled Castes (SCs) and Other Backward Classes (OBCs) – two large disadvantaged caste groups among Hindus, it remained constant among Muslims. In both elections, only 8 percent of Muslims are reported to have voted for the BJP. Using the data on home visits by party workers, I also find that BJP is significantly less likely to visit Muslim homes compared to non-Muslim homes, while other parties are more likely to visit them (Appendix Table A7). Hence, if the polling station level irregularities are due to exercise of precise control, we should expect it to be concentrated in areas within a PC that have low Muslim share of the electorate.

Since Muslim electorate share at the polling stations is not known, I map each polling station to the Assembly Constituency it falls under. Each AC is subsumed within a PC and each PC on average contains about 7 ACs. The data on AC level Muslim electorate share (described in Section III) would provide *within-PC* variation in Muslim electorate share across polling stations falling in different AC-segments. The final sample for this analysis contains more than $850,000$ polling stations mapped to 3098 ACs (76$\%$ of all ACs) covering 475 PCs. The mean Muslim share in an AC is 0.14. However, there is

[39] In Uttar Pradesh, for example, during BJP's state-wide membership drive it did not focus on the $13,000$ polling stations with significant Muslim presence, since it did not receive any votes in those areas.

wide variation across ACs, with 5^{th} percentile at 0.01 and 95^{th} percentile being 0.43. Appendix Figure A9 shows the distribution across all ACs. Appendix Table A8 regresses polling station level BJP vote share on AC level Muslim share and finds a sizable and statistically significant negative relationship, both across PCs as well as within a PC. Now, to shed light on the mechanisms discussed above I focus on close election PCs, i.e., those with absolute BJP win margin within 0.16 and run the following specification:

$$\begin{aligned} Y_{jap} &= \phi_p + \gamma BJP\_Won_p * Muslim\_share_{ap} + \delta Muslim\_share_{ap} \\ &+ \beta_1 BJP\_WinMargin_p * Muslim\_share_{ap} \\ &+ \beta_2 BJP\_Won_p * BJP\_WinMargin_p * Muslim\_share_{ap} + \epsilon_{jap} \end{aligned}$$

Table 8—Polling Station Level Irregularities Concentrated in High Muslim Share ACs

	BJP vote share (1)	Rel. BJP vote share (2)	BJP share $\geq 95^{th}$ pctile (3)
	Panel A: All states		
Muslim Electorate Share in AC	-0.745***	-1.811***	-0.194**
	(0.087)	(0.300)	(0.085)
BJP Won * Muslim Electorate Share in AC	0.327**	0.818**	0.199*
	(0.127)	(0.369)	(0.101)
Observations	280,391	280,391	280,391
	Panel B: BJP Ruled states		
Muslim Electorate Share in AC	-0.633***	-1.282***	-0.250**
	(0.173)	(0.373)	(0.124)
BJP Won * Muslim Electorate Share in AC	0.367	0.695	0.269*
	(0.224)	(0.474)	(0.147)
Observations	145,574	145,574	145,574
PC Fixed Effect	YES	YES	YES
Bandwidth (h^*)	0.160	0.160	0.160

Notes: The data is at polling station level. The dependent variables are polling station level BJP vote share (column 1), relative BJP vote share (column 2) and a dummy variable that takes value one if the BJP vote share in a polling station exceeds the 95^{th} percentile (column 3). BJP Won is an indicator of whether BJP is the winner of the Parliamentary Constituency. Muslim Electorate Share in AC is the share of Muslim voters in the Assembly Constituency in which a polling station is located. The sample in Panel A is PCs from all states while that in Panel B is BJP ruled states. Optimal bandwidth calculated for McCrary test in Figure 1a has been used in all specifications. Standard errors are clustered at the PC level and are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

where j denotes polling station, a denotes AC and p denotes PC. Y_{jap} is one of three outcome variables – (i) vote share of BJP in a polling station, (ii) relative BJP vote share, as defined above, and (iii) the indicator *"high" BJP vote share* defined above. ϕ_p is PC fixed effect and $Muslim\_share_{ap}$ is Muslim electorate share in AC a in PC p. The regression implements the difference-in-discontinuity specification with PC fixed effects that subsume the running variable, the treatment BJP Won and their interaction. It compares polling stations within a PC and checks if the BJP's vote share is high or is more like to exceed its 95^{th} percentile in AC segments with higher Muslim share and whether this relationship is different between PCs that BJP barely won and lost. δ estimates the relationship in PCs lost by BJP. γ is the differential estimate for PCs won by BJP and is our coefficient of interest. Precise control hypothesis implies $\gamma < 0$, while manipulation would imply $\gamma > 0$.

Table 8 reports the results. The three columns correspond to the three outcome variables mentioned above. Panel A reports the results for the full sample, while Panel B reports it for the BJP ruled states. As before, I restrict attention to PCs with BJP win margin within 0.16. In Panel A, the estimates of δ in all the columns is negative and statistically significant at 1% or 5%. However, the estimates of γ are positive and statistically significant at 10% or 5%. In Panel B, the estimates of δ are also positive, but in columns (1) and (2) they are noisily estimated. The estimate for the dummy indicating "high" BJP vote share in a polling station (column (3)) is large in magnitude and statistically significant at 10%. The estimates of γ and δ jointly indicate that the Muslim electorate share does not predict "extreme" outcomes in favor BJP in PCs barely won by BJP, even though it strongly negatively predicts such outcomes in PCs barely lost by the party.

Appendix Table A9 partitions the sample used in Panel B column (3) into polling stations with turnout higher and lower than 800 and estimates the same specification. To allow comparison, column (1) of Table A9 reports the same result as column (3) of Table 8. Columns (2) and (3) report the results for the two sub-samples separately. We find that the the estimate of δ is large in magnitude and statistically significant at 1% in column (2), while it is smaller in magnitude and statistically insignificant in column (3). This is consistent with the graphs reported in Figure 4.[40]

VII Concluding Remarks

The paper documents irregularity in India's 2019 general election data by showing that the incumbent party's win margin distribution exhibits excess mass at zero, while no such pattern exists either in previous general elections or in state elections held simultaneously and subsequently. This implies that the incumbent party in 2019 won a disproportionate share of closely contested elections. Moreover, the pattern is concentrated in the states ruled by the incumbent party at that time. While the result is consistent with electoral fraud or manipulation, the incumbent party's superior ability to predict and affect win margin (i.e., precise control), owing to its significant advantage in electoral campaigning over other parties can also explain it. To isolate the two mechanisms, I conduct a series of analyses to check for presence of precise control and manipulation. I do not find that the incumbent party did greater door-to-door campaigning than other parties in constituencies barely won by it. On the other hand, I find evidence consistent with electoral manipulation at the stage of voter registration as well as at the time of voting and counting (turnout manipulation). In both cases, the results point to strategic and targeted electoral discrimination against Muslims, in the form of deletion of names from voter lists and suppression of their votes during election, in part facilitated by weak monitoring by election observers.

The tests are, however, not *proofs* of fraud, nor does it suggest that manipulation was widespread. Proving electoral manipulation in a robust democracy is a significantly harder task that would require detailed investigation of electoral data in each constituency separately. In the 1960 Presidential election in the US, for example, there was reporting of possible fraud in Illinois state that may have resulted in

[40] The result remains the same if I use turnout threshold of 700 or 600, instead of 800.

John F. Kennedy winning that state. Analysis of detailed data on recounting of votes from Cook county showed patterns consistent with fraud, and yet, were not able to conclusively determine its magnitude and whether it caused the result to flip in Kennedy's favor (Kallina 1985). This case also highlights that electoral fraud is often decentralized (Rundlett and Svolik 2016), as opposed to being implemented centrally. Consequently, fraud may occur even in contexts where it would not have mattered for government formation. In 1960, Kennedy would have won the Presidential election even if he had lost Illinois. Similarly, in my context, even if manipulation of election data drives all of the observed irregularities in closely contested constituencies, the aggregate election outcomes in terms of government formation would likely have remained unchanged. Appendix Table A10 reports the number of PCs with "excess" BJP wins in closely contested PCs. It varies from 9-18, depending on the definition of a close contest; the numbers are smaller than the lead of 31 PCs that BJP has over the threshold required to form government. Nonetheless, electoral fraud even in a single constituency would imply that such manipulations by incumbent parties are *possible*. In view of the depletion of trust in electoral processes across the globe and the exceptional integrity of India's electoral institution in its past, the paper presents a worrying development with potentially far-reaching consequences for the world's largest democracy.

References

AGARWAL, P. (2019): "EVM Vote Count Mismatch In 370+ Seats and EC Refuses to Explain," *The Quint*, 31 May, 2019, https://www.thequint.com/news/india/lok–sabha–election–results–2019–mismatch–in–votes–polled–and–counted–in–evm–on–multiple–seats.

ARRIOLA, L. R., J. DEVARO, AND A. MENG (2021): "Democratic subversion: Elite cooptation and opposition fragmentation," *American Political Science Review*, 115, 1358–1372.

ASUNKA, J., S. BRIERLEY, M. GOLDEN, E. KRAMON, AND G. OFOSU (2019): "Electoral fraud or violence: The effect of observers on party manipulation strategies," *British Journal of Political Science*, 49, 129–151.

BALASUBRAMANIAM, V., A. Y. BHATIYA, AND S. DAS (2021): "Behavioral Voters in Synchronized Elections: Evidence from India," *Available at SSRN 3636183*.

BANERJEE, A., A. GETHIN, AND T. PIKETTY (2019): "Growing Cleavages in India?" *Economic & Political Weekly*, 54, 35.

BANERJEE, M. (2017): "Vote," *South Asia: Journal of South Asian Studies*, 40, 410–412.

BEBER, B. AND A. SCACCO (2012): "What the numbers say: A digit-based test for election fraud," *Political analysis*, 20, 211–234.

BEKKOUCHE, Y., J. CAGE, AND E. DEWITTE (2022): "The heterogeneous price of a vote: Evidence from multiparty systems, 1993–2017," *Journal of Public Economics*, 206, 104559.

BENFORD, F. (1938): "The law of anomalous numbers," *Proceedings of the American philosophical society*, 551–572.

BHALOTRA, S., I. CLOTS-FIGUERAS, G. CASSAN, AND L. IYER (2014): "Religion, politician identity and development outcomes: Evidence from India," *Journal of Economic Behavior & Organization*, 104, 4–17.

Bhalotra, S., I. Clots-Figueras, L. Iyer, and J. Vecci (2021): "Leader identity and coordination," *The Review of Economics and Statistics*, 1–50.

Bhogale, S., S. Hangal, F. R. Jensenius, M. Kumar, C. Narayan, B. U. Nissa, and G. Verniers (2019): *TCPD-IED: TCPD Indian Elections Data v1*, Trivedi Centre for Political Data, Ashoka University.

Braley, A., G. S. Lenz, D. Adjodah, H. Rahnama, and A. Pentland (2022): "The Subversion Dilemma: Why Voters Who Cherish Democracy Participate in Democratic Backsliding," *Aletheia*.

Callen, M. and J. D. Long (2015): "Institutional corruption and election fraud: Evidence from a field experiment in Afghanistan," *American Economic Review*, 105, 354–81.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014): "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 82, 2295–2326.

Cantoni, E. and V. Pons (2020): "Strict ID Laws Don't Stop Voters: Evidence from a US Nationwide Panel, 2008–2018," *Quarterly Journal of Economics*.

Cantú, F. (2019): "The fingerprints of fraud: Evidence from Mexico's 1988 presidential election," *American Political Science Review*, 113, 710–726.

Caughey, D. and J. S. Sekhon (2011): "Elections and the regression discontinuity design: Lessons from close US house races, 1942–2008," *Political Analysis*, 19, 385–408.

Chaturvedi, R. and S. Chaturvedi (2023): "It's All in the Name: A Character-Based Approach to Infer Religion," *Political Analysis*, 1–16.

Christensen, R. and T. J. Schultz (2014): "Identifying election fraud using orphan and low propensity voters," *American Politics Research*, 42, 311–337.

Democracy Report, . (2020): *Autocratization Surges – Resistance Grows*, V-Dem Institute.

——— (2021): *Autocratization Turns Viral*, V-Dem Institute.

Diekmann, A. (2007): "Not the first digit! using benford's law to detect fraudulent scientif ic data," *Journal of Applied Statistics*, 34, 321–329.

Eggers, A. C., A. Fowler, J. Hainmueller, A. B. Hall, and J. M. Snyder Jr (2015): "On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races," *American Journal of Political Science*, 59, 259–274.

Enikolopov, R., V. Korovkin, M. Petrova, K. Sonin, and A. Zakharov (2013): "Field experiment estimate of electoral fraud in Russian parliamentary elections," *Proceedings of the National Academy of Sciences*, 110, 448–452.

Escobari, D. and G. A. Hoover (2020): "Evo morales and electoral fraud in bolivia: a natural experiment and discontinuity evidence," *Available at SSRN 3492928*.

Fishkin, J., A. Siu, L. Diamond, and N. Bradburn (2021): "Is deliberation an antidote to extreme partisan polarization? Reflections on "America in one room"," *American Political Science Review*, 115, 1464–1481.

Foa, R. S. and Y. Mounk (2016): "The danger of deconsolidation: The democratic disconnect," *Journal of democracy*, 27, 5–17.

——— (2017a): "The end of the consolidation paradigm," *Journal of Democracy*.

——— (2017b): “The signs of deconsolidation,” *Journal of democracy*, 28, 5–15.

FUKUMOTO, K. AND Y. HORIUCHI (2011): “Making outsiders’ votes count: Detecting electoral fraud through a natural experiment,” *American Political Science Review*, 105, 586–603.

GERBER, A. (1998): “Estimating the effect of campaign spending on senate election outcomes using instrumental variables,” *American Political science review*, 92, 401–411.

GRAHAM, M. H. AND M. W. SVOLIK (2020): “Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States,” *American Political Science Review*, 114, 392–409.

GREMBI, V., T. NANNICINI, AND U. TROIANO (2016): “Do fiscal rules matter?” *American Economic Journal: Applied Economics*, 1–30.

HILL, T. P. ET AL. (1995): “A statistical derivation of the significant-digit law,” *Statistical science*, 10, 354–363.

HINTSON, J. AND M. VAISHNAV (2021): “Who Rallies Around the Flag? Nationalist Parties, National Security, and the 2019 Indian Election,” *American Journal of Political Science*.

IYER, L. AND A. MANI (2012): “Traveling agents: political change and bureaucratic turnover in India,” *Review of Economics and Statistics*, 94, 723–739.

IYER, L. AND M. REDDY (2013): *Redrawing the Lines: Did Political Incumbents Influence Electoral Redistricting in the World’s Largest Democracy?*, Harvard Business School.

JAMES, T. S. AND A. CLARK (2020): “Electoral integrity, voter fraud and voter ID in polling stations: lessons from English local elections,” *Policy Studies*, 41, 190–209.

JEONG, D. AND A. SHENOY (2020): “Can the Party in Power Systematically Win a Majority in Close Legislative Elections? Evidence from US State Assemblies,” *Journal of Politics*, (forthcoming).

JHA, P. (2017): *How the BJP Wins: Inside India’s Greatest Election Machine*, Juggernaut Books.

KALLINA, E. F. (1985): “Was the 1960 presidential election stolen? The case of Illinois,” *Presidential Studies Quarterly*, 113–118.

KAPUR, D., P. B. MEHTA, AND M. VAISHNAV (2018): *Rethinking public institutions in India*, Oxford University Press.

KHOSLA, M. AND M. VAISHNAV (2022): “India@ 75: Religion and Citizenship in India,” .

KJELSRUD, A., K. O. MOENE, AND L. VANDEWALLE (2020): “The political competition over life and death: Evidence from infant mortality in India,” Tech. rep., Graduate Institute of International and Development Studies Working Paper.

KLIMEK, P., Y. YEGOROV, R. HANEL, AND S. THURNER (2012): “Statistical detection of systematic election irregularities,” *Proceedings of the National Academy of Sciences*, 109, 16469–16473.

LEE, D. S. AND T. LEMIEUX (2010): “Regression discontinuity designs in economics,” *Journal of economic literature*, 48, 281–355.

LEHNE, J. (2022): “Incumbents, Minorities and Voter Purges: Evidence from 120 Million Voters’ Registrations in India,” *Unpublished Manuscript*.

LEVITT, S. D. (1994): “Using repeat challengers to estimate the effect of campaign spending on election outcomes in the US House,” *Journal of Political Economy*, 102, 777–798.

LITTLE, A. AND A. MENG (2023): "Subjective and Objective Measurement of Democratic Backsliding," *Available at SSRN 4327307*.

MALHOTRA, A. (2019): "Allegations of mass voter exclusion cast shadow on India election," *AlJazeera*, 30 April, 2019, https://www.aljazeera.com/news/2019/4/30/allegations–of–mass–voter–exclusion–cast–shadow–on–india–election.

MANHEIM, L. M. AND E. G. PORTER (2019): "The Elephant in the Room: Intentional Voter Suppression," *The Supreme Court Review*, 2018, 213–255.

MARTINELLI, A. (2016): "Populism and the crisis of representative democracy," *Populism on the Rise: Democracies Under Challenge*, 13–32.

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

MEBANE, W. R. (2008a): "Election forensics: Outlier and digit tests in America and Russia," in *American Electoral Process conference, Center for the Study of Democratic Politics, Princeton University*, Citeseer.

——— (2008b): "Election forensics: The second-digit Benford's law test and recent American presidential elections," *Election fraud: detecting and deterring electoral manipulation*, 162–181.

——— (2011): "Comment on "Benford's Law and the detection of election fraud"," *Political Analysis*, 19, 269–272.

NAQVI, S. (2022): "The Vanished Minority Voter," *Deccan Herald*, February 14, 2022.

NATH, A., D. MOOKHERJEE, ET AL. (2017): "Resource Transfers to Local Governments: Political Manipulation and Voting Patterns in West Bengal," in *2017 Meeting Papers*, Society for Economic Dynamics, 1266.

NEGGERS, Y. (2018): "Enfranchising your own? Experimental evidence on bureaucrat diversity and election bias in India," *American Economic Review*, 108, 1288–1321.

NELLIS, G., M. WEAVER, S. ROSENZWEIG, ET AL. (2016): "Do parties matter for ethnic violence? Evidence from India," *Quarterly Journal of Political Science*, 11, 249–277.

NIGRINI, M. J. (2012): *Benford's Law: Applications for forensic accounting, auditing, and fraud detection*, vol. 586, John Wiley & Sons.

ONAPAJO, H. AND U. O. UZODIKE (2014): "Rigging through the courts: The judiciary and electoral fraud in Nigeria," *Journal of African Elections*, 13, 137–168.

PRAKASH, N., M. ROCKMORE, AND Y. UPPAL (2019): "Do criminally accused politicians affect economic outcomes? Evidence from India," *Journal of Development Economics*, 141, 102370.

PURKAYASTHA, P. AND B. SINHA (2019): "There is No Ghost in the Indian EVM," *Indian Policy Forum*.

RAIMI, R. A. (1976): "The first digit problem," *The American Mathematical Monthly*, 83, 521–538.

RAMACHANDRAN, S. (2022): "India State Assembly Elections: Where Is the Level Playing Field?" *The Diplomat*, 13 January, 2022, https://thediplomat.com/2022/01/india–state–assembly–elections–where–is–the–level–playing–field/.

REPUCCI, S. AND A. SLIPOWITZ (2021): "Democracy under siege," *Freedom House*.

Rozenas, A. (2017): "Detecting election fraud from irregularities in vote-share distributions," *Political Analysis*, 25, 41–56.

Rundlett, A. and M. W. Svolik (2016): "Deliver the vote! micromotives and macrobehavior in electoral fraud," *American Political Science Review*, 110, 180–197.

Şaşmaz, A., A. H. Yagci, and D. Ziblatt (2022): "How Voters Respond to Presidential Assaults on Checks and Balances: Evidence from a Survey Experiment in Turkey," *Comparative Political Studies*, 00104140211066216.

Scheppele, K. L. (2022): "How viktor orbán wins," *Journal of Democracy*, 33, 45–61.

Shikano, S. and V. Mack (2011): "When Does the Second-Digit Benford's Law-Test Signal an Election Fraud?" *Jahrbücher fur Nationalökonomie & Statistik*, 231.

Sridharan, E. and M. Vaishnav (2017): "Election commission of India," *Rethinking public institutions in India*, 417–63.

Thachil, T. (2014): "Elite parties and poor voters: Theory and evidence from India," *American Political Science Review*, 108, 454–477.

Trivedi, D. (2019): "Missing Voters," *Frontline*, 26 April, 2019, https://frontline.thehindu.com/cover–story/article26781550.ece.

Uppal, Y. (2009): "The disadvantaged incumbents: estimating incumbency effects in Indian state legislatures," *Public choice*, 138, 9–27.

Varshney, A. (2003): *Ethnic conflict and civic life: Hindus and Muslims in India*, Yale University Press.

——— (2019): "Modi consolidates power: Electoral vibrancy, mounting liberal deficits," *Journal of Democracy*, 30, 63–77.

Vogl, T. S. (2014): "Race and the politics of close elections," *Journal of Public Economics*, 109, 101–113.

Waldner, D. and E. Lust (2018): "Unwelcome change: Coming to terms with democratic backsliding," *Annual Review of Political Science*, 21, 93–113.

Ziblatt, D. (2009): "Shaping democratic practice and the causes of electoral fraud: The case of nineteenth-century Germany," *American Political Science Review*, 103, 1–21.

Appendix

A Additional Figures and Tables

Table A1—Number of Constituencies BJP Won vs Lost in Close Elections

	2019 (1)	2014 (2)	2009 (3)	2004 (4)
	Panel A: Absolute BJP Win Margin ≤ 0.05			
# (%) of Constituencies BJP Won	41 (69%)	29 (60%)	49 (51%)	47 (59%)
# (%) of Constituencies BJP Lost	18 (31%)	19 (40%)	48 (49%)	33 (41%)
Total # (%) of Constituencies	59 (100%)	48 (100%)	97 (100%)	80 (100%)
	Panel B: Absolute BJP Win Margin ≤ 0.03			
# (%) of Constituencies BJP Won	28 (74%)	14 (58%)	29 (47%)	22 (56%)
# (%) of Constituencies BJP Lost	10 (26%)	10 (42%)	33 (53%)	17 (44%)
Total # (%) of Constituencies	38 (100%)	24 (100%)	62 (100%)	39 (100%)
	Panel C: Absolute BJP Win Margin ≤ 0.02			
# (%) of Constituencies BJP Won	20 (74%)	10 (53%)	21 (46%)	18 (64%)
# (%) of Constituencies BJP Lost	7 (26%)	9 (47%)	25 (54%)	10 (36%)
Total # (%) of Constituencies	27 (100%)	19 (100%)	46 (100%)	28 (100%)

Notes: The table reports the number and percentage of constituencies that BJP won and lost and total number of constituencies in 2019 (column 1), 2014 (column 2), 2009 (column 3) and 2004 (column 4) general elections where the BJP's absolute win margin was less than or equal to 0.05 (Panel A), 0.03 (Panel B) and 0.02 (Panel C).

Table A2—Heterogeneity in Density Jump in BJP Win Margin

	Panel A	
	BJP Ruled States	Non-BJP Ruled States
	(1)	(2)
Density jump	3.14***	0.21
	(1.15)	(1.03)
Bandwidth (h^*)	0.160	0.160

Notes: The table reports the estimates in the difference in densities of BJP win margin at the threshold value of zero using the McCrary test. Column 1 reports it for the states ruled by BJP in 2019, while column 2 reports for the rest of the states. Optimal bandwidth is calculated using the MSERD method proposed by Calonico et al. (2014). *** p<0.01, ** p<0.05, * p<0.1

Table A3—Social Media Usage and Voting Behavior: NES 2019

	Voted for BJP	
	(1)	(2)
Facebook user	0.035***	0.027**
	(0.013)	(0.013)
Twitter user	-0.081***	-0.078***
	(0.018)	(0.017)
Whatsapp user	0.007	-0.008
	(0.013)	(0.013)
Instagram user	0.020	0.017
	(0.014)	(0.014)
YouTube user	0.001	-0.004
	(0.013)	(0.013)
Constant	0.335***	0.367***
	(0.004)	(0.013)
Observations	22,037	22,037
R-squared	0.002	0.007

Notes: The sample is individual level survey data from the National Election Survey (post poll) 2019. The dependent variable is a dummy indicating whether the individual reported to have voted for BJP in the 2019 election. For any social media platform, an individual is defined to be "user" of the platform if they use it at least once daily. Column 2 controls for the gender, age and dummies for caste categories of individuals. Robust standard errors are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A4—Discontinuity in β_p Estimates in Closely Contested Constituencies

	Dep. Var.: β_p		
	Full Sample	BJP Ruled States	Non-BJP Ruled States
	(1)	(2)	(3)
BJP Won	0.24***	0.09	0.27*
	(0.09)	(0.09)	(0.16)
Bandwidth (h^*)	0.16	0.16	0.16
Observations	68	28	40

Notes: The dependent variable is the estimate of the coefficient β_p from running equation (1) for each PC. The table reports the RDD estimates of a BJP victory using β_p as the outcome variable. Column 1 reports it for the full sample, columns 2 reports it for the states ruled by BJP in 2019, and column 3 reports for the rest of the states. Bandwidth used is the optimal bandwidth used for McCrary test in Figure 1a. *** p<0.01, ** p<0.05, * p<0.1

Table A5—Electorate Growth Rate Smaller in PCs Barely Won by BJP

	Electorate Growth Rate (G_c)			
	BJP Ruled States		Non-BJP Ruled States	
	High Muslim Share	Low Muslim Share	High Muslim Share	Low Muslim Share
	(1)	(2)	(3)	(4)
BJP Won	-0.06*	-0.04	-0.04	-0.01
	(0.04)	(0.03)	(0.02)	(0.02)
Mean Dep. Var.	0.11	0.10	0.08	0.07
Observations	53	30	48	50
Bandwidth (h^*)	0.16	0.16	0.16	0.16

Notes: The data is at Parliamentary Constituency (PC) level. The table reports the regression discontinuity design estimate using BJP win margin as the running variable. The dependent variable in all the columns is the growth rate in the PC electorate between 2014 and 2019. BJP Won is a dummy indicating whether BJP won the PC in 2019. Columns (1) and (2) use the sample of PCs in BJP ruled states, while columns (3) and (4) use PCs in non-BJP ruled states. Columns (1) and (3) use PCs where the electorate share of Muslims is higher than the median, and column (2) and (4) use PCs where the share is lower than median. All columns use the optimal bandwidth calculated for McCrary test in Figure 1a. Robust standard errors are are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A6—Association between Turnout Data Discrepancy and Counting Observer Characteristics

	Absolute Turnout Discrepancy (1)	"Large" Turnout Discrepancy (2)
BJP Won	-475.1	-0.0269
	(623.9)	(0.0229)
SCS Counting Observer from BJP Ruled State	-463.2	-0.0471
	(1,009)	(0.0663)
BJP Won * SCS Counting Observer from BJP Ruled State	4,373*	0.210**
	(2,589)	(0.104)
Constant	1,712***	0.0544***
	(426.3)	(0.0196)
$H_0 : \beta_2 + \beta_3 = 0$ (p-value)	0.102	0.041
Observations	370	370
R-squared	0.012	0.016

Notes: The data is at Parliamentary Constituency (PC) level. The dependent variable in column (1) is absolute discrepancy in turnout data and in column (2) a dummy variable that takes value one when the absolute discrepancy exceeds 95^{th} percentile of its distribution. BJP Won is a dummy indicator whether BJP won the PC in 2019. SCS Counting Observer from BJP Ruled State is the fraction of counting observers assigned to a PC who are SCS cadre and work in BJP ruled states. Robust standard errors are are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A7—Campaigning among Muslim Voters by BJP and Other Parties

	Home Visit by Party Worker/Candidate					
	from BJP			from Any Other Party		
	(1)	(2)	(3)	(4)	(5)	(6)
Muslim	-0.14***	-0.12***	-0.12***	0.05***	0.04***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Mean Dep. Var.	0.38	0.38	0.38	0.50	0.50	0.50
Fixed Effect	State	PC	AC	State	PC	AC
Observations	24,230	24,230	24,230	24,230	24,230	24,230
No of PCs	208	208	208	208	208	208
R-squared	0.14	0.24	0.31	0.23	0.32	0.39

Notes: The sample is individual level survey data from the National Election Survey (post poll) 2019. The dependent variable in columns (1)-(3) is a dummy variable that takes value one if a BJP party worker or candidate visited the house of the respondent to campaign for general election, and is zero otherwise. The dependent variable in columns (4)-(6) is also a dummy variable that indicates whether party worker or candidate from any other party visited the house for campaigning. Muslim is dummy variable that takes value one if the survey respondent is a Muslim. All columns control for the respondents' age, age squared, gender and education categories. Columns (1) and (4) have state fixed effects, (2) and (5) have PC fixed effects, and (3) and (6) have AC fixed effects. Robust standard errors are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A8—Correlation between AC Muslim Share and Polling Station Level BJP Vote Share

	BJP vote share		
	(1)	(2)	(3)
Muslim Electorate Share in AC	-0.37***	-0.31***	-0.43***
	(0.02)	(0.02)	(0.02)
Mean Dep. Var.	0.46	0.46	0.46
Observations	674,253	674,253	674,253
State Fixed Effect	NO	YES	NO
PC Fixed Effect	NO	NO	YES

Notes: The data is at polling station level. The dependent variable in all the columns is BJP vote share in a polling station. Muslim Electorate Share in AC is the share of Muslim voters in the Assembly Constituency in which a polling station is located. Column (1) has no other controls, while columns (2) and (3) have state and AC fixed effects respectively. Standard errors are clustered at the AC level and are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A9—Polling Station Level Irregularities Concentrated in Polling Stations with High Muslim Shares

	BJP vote share $\geq 95^{th}$ pctile		
	BJP Ruled States	P. S. Turnout ≥ 800	P. S. Turnout < 800
	(1)	(2)	(3)
Muslim Electorate Share in AC	-0.250**	-0.378***	-0.207
	(0.124)	(0.073)	(0.137)
BJP Won * Muslim Electorate Share in AC	0.269*	0.701***	0.224
	(0.147)	(0.174)	(0.158)
Mean Dep. Var.	0.054	0.086	0.052
Observations	145,574	8,310	137,263
PC Fixed Effect	YES	YES	YES
Bandwidth (h^*)	0.16	0.16	0.16

Notes: The data is at polling station level. The dependent variable in all the columns is a dummy variable that takes value one if the BJP vote share in a polling station is larger than the 95^{th} percentile and zero otherwise. BJP Won is an indicator of whether BJP is the winner of the Parliamentary Constituency (PC). Muslim Electorate Share in AC is the share of Muslim electorate in the Assembly Constituency in which a polling station is located. The sample in column (1) is PCs in BJP ruled states with absolute BJP win margin within 0.16. Columns (2) and (3) samples are partitions of the column (1) sample into polling stations with turnout higher than and less than 800, respectively. All columns have PC fixed effect. The optimal bandwidth calculated for McCrary test in Figure 1a has been used in all specifications. Standard errors are clustered at the PC level and are reported in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A10—Extent of "Excess" BJP Wins in Closely Contested PCs

	BJP Win Margin		
	≤ 0.07	≤ 0.05	≤ 0.03
	(1)	(2)	(3)
# Close Election PCs	82	59	38
# "Excess" BJP Wins	18	11	9

Notes: The table reports the number of closely contested Parliamentary Constituencies (PCs) in 2019 and the "excess" number of wins by BJP in those PCs relative to the benchmark of 50% chance of winning. The three columns use three different bandwidths to define a close contest. Column (1) considers BJP win margin within 0.07 while columns (2) and (3) consider 0.05 and 0.03 respectively.

Figure A1—Correlation between Muslim Electorate Share and Vote Share of Muslim Candidates

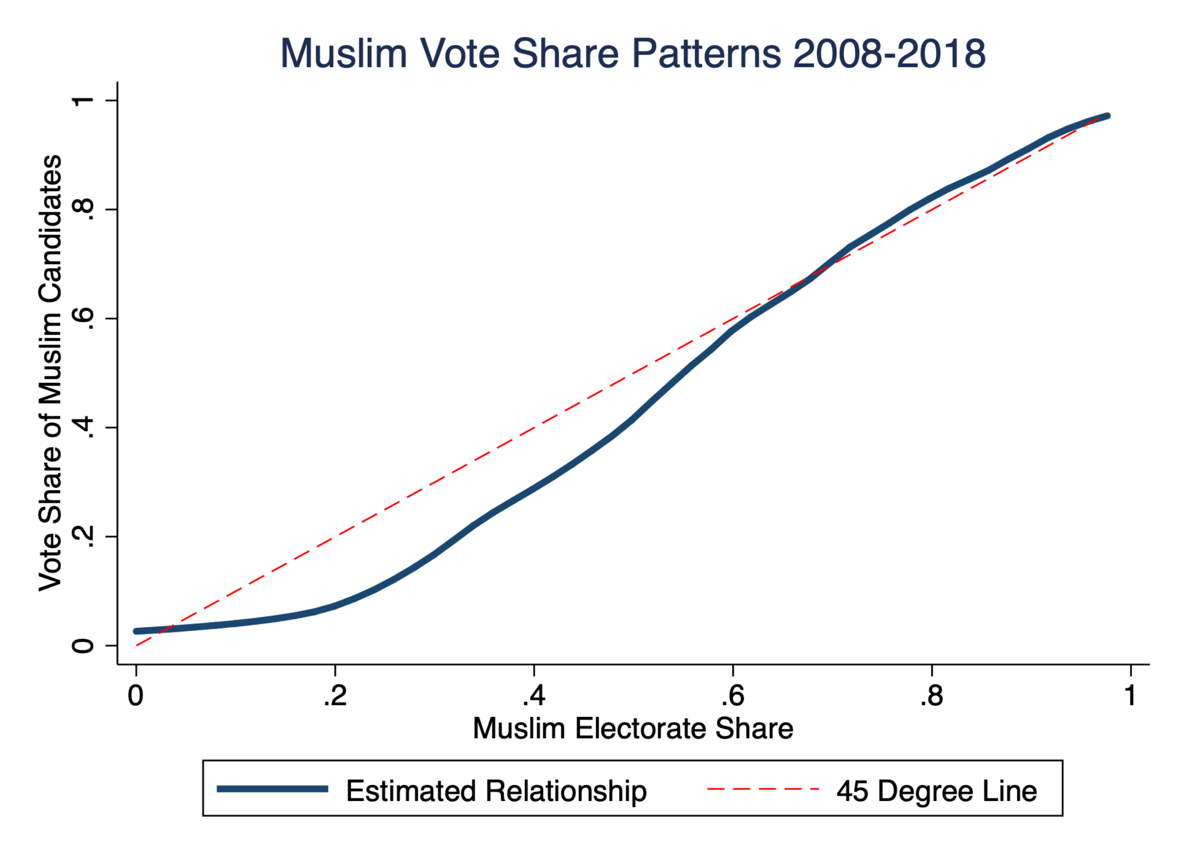
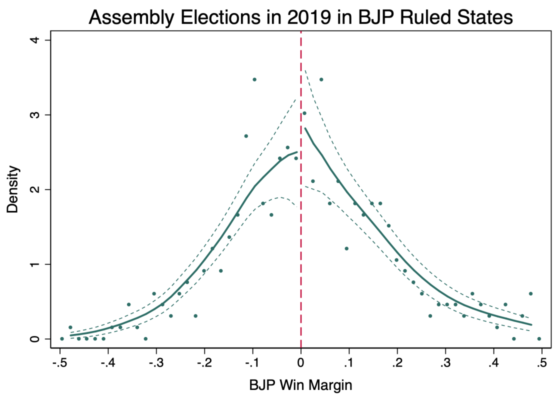
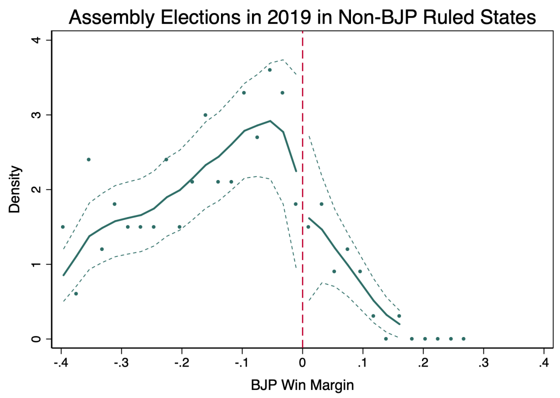


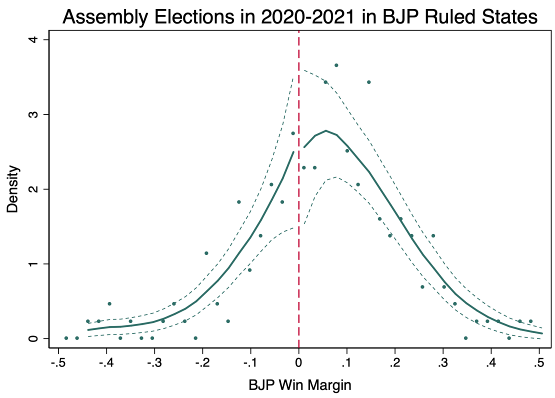
Figure A2—McCrary Test for State Assembly Elections



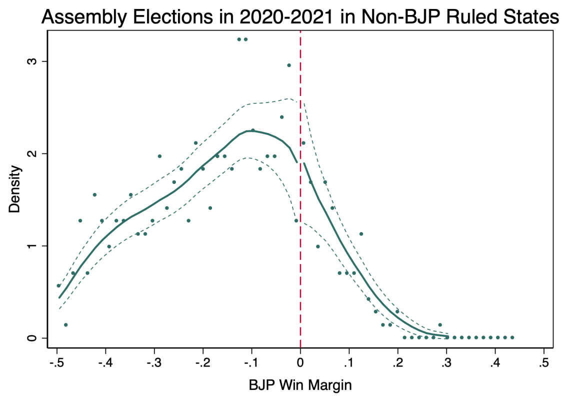
(a) 2019 Elections: BJP Ruled States



(b) 2019 Elections: Non-BJP Ruled States



(c) 2020-21 Elections: BJP Ruled States



(d) 2020-21 Elections: Non-BJP Ruled States

Figure A3—McCrary Test for Non-BJP Ruled States with Strong BJP Presence

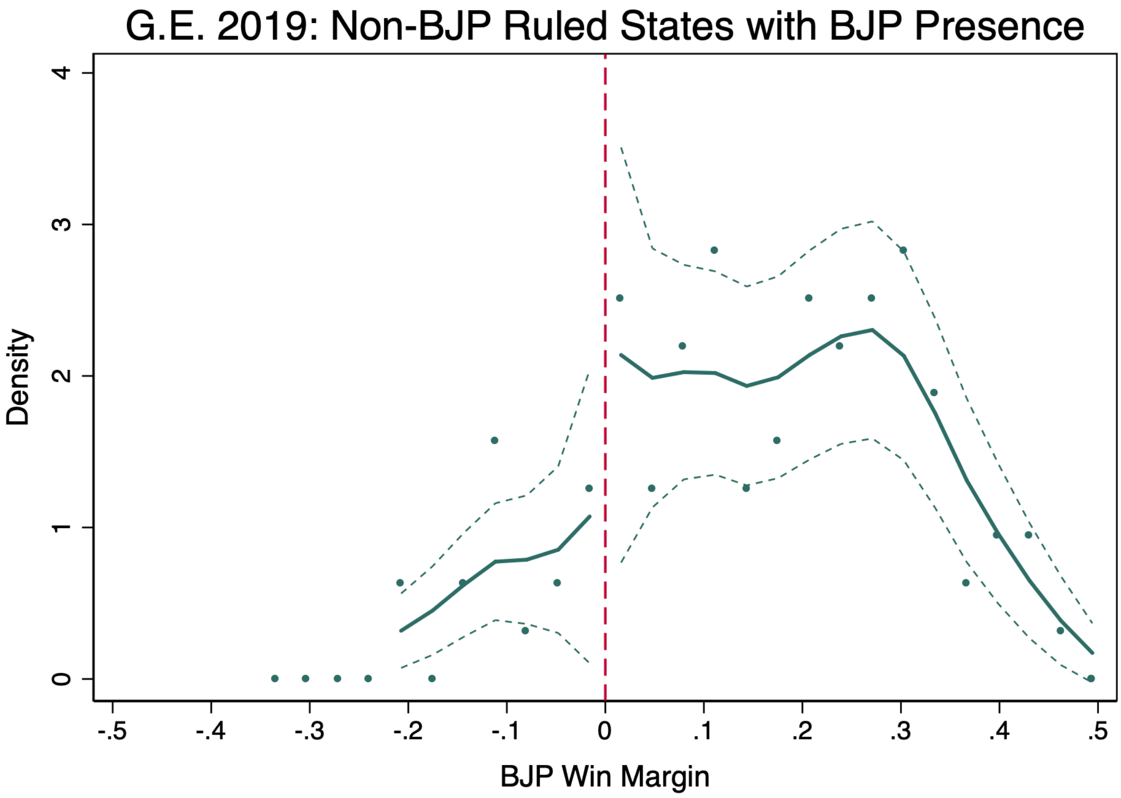
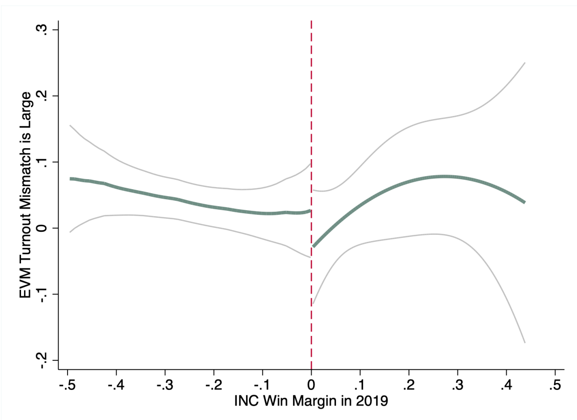
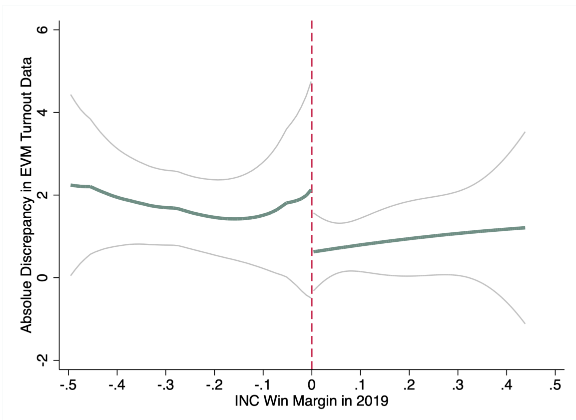


Figure A4—EVM Turnout Data Discrepancy in Closely Contested Constituencies

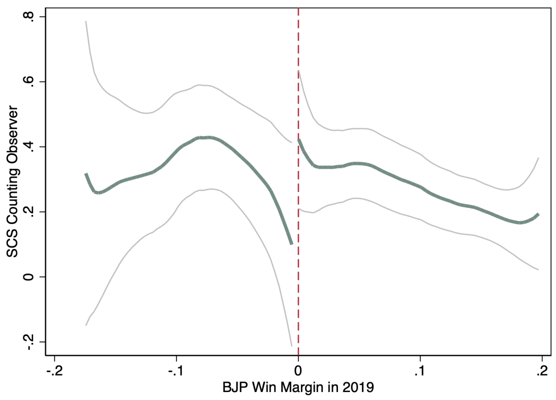


(a) "Large" Data Discrepancy

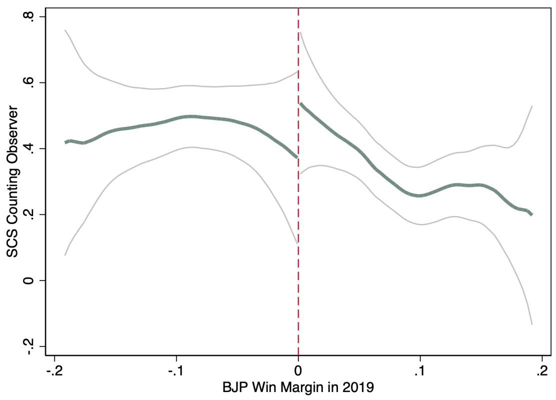


(b) Absolute Data Discrepancy

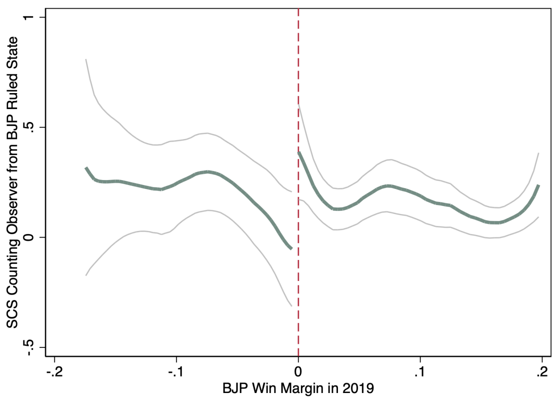
Figure A5—SCS Election Observers in Closely Contested Constituencies



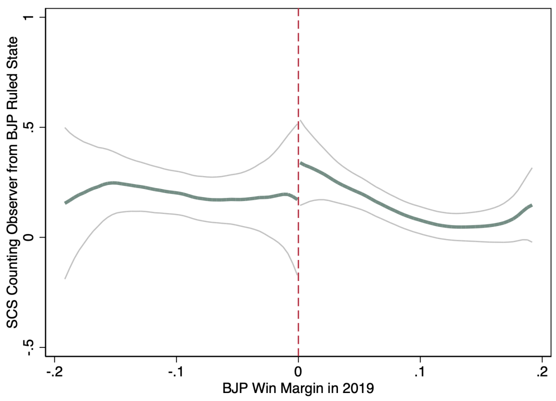
(a) BJP Ruled States



(b) Non-BJP Ruled States

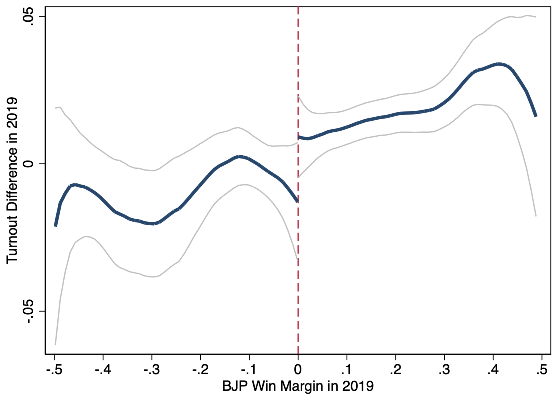


(c) BJP Ruled States

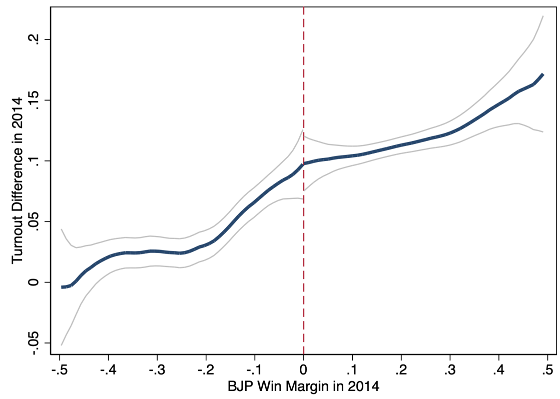


(d) Non-BJP Ruled States

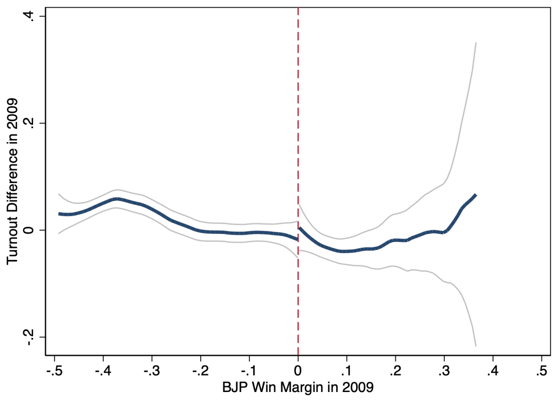
Figure A6—Turnout Difference in Closely Contested PCs in past General Elections



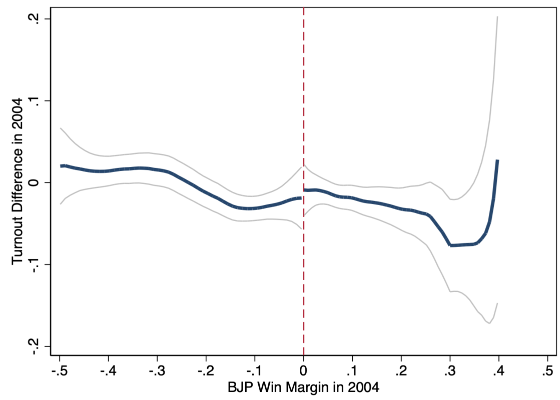
(a) 2019



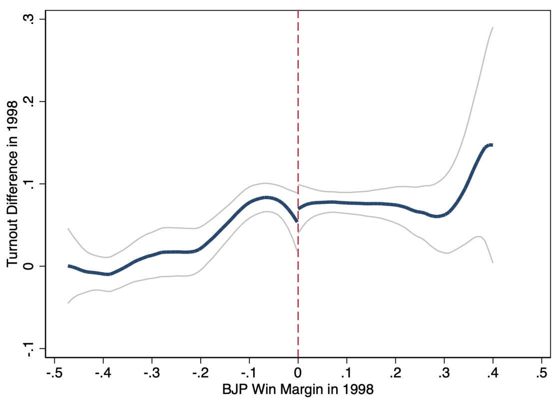
(b) 2014



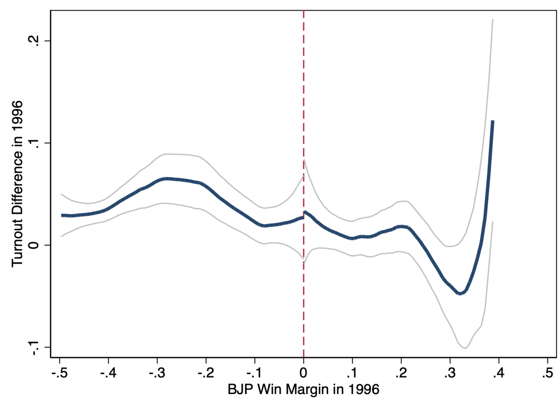
(c) 2009



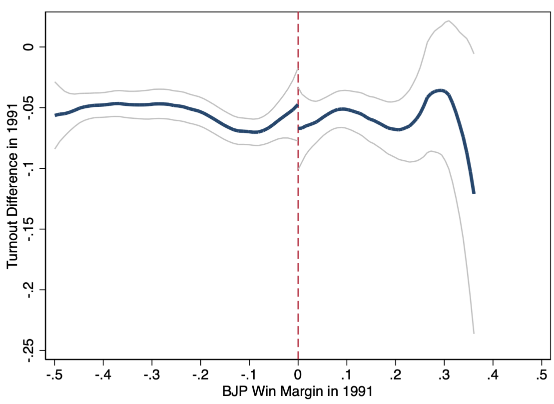
(d) 2004



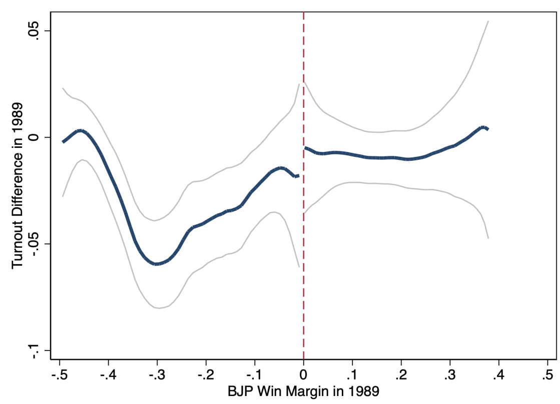
(e) 1998



(f) 1996



(g) 1991



(h) 1989

Figure A7—Differential Pattern only in Closely Contested Elections Won by BJP

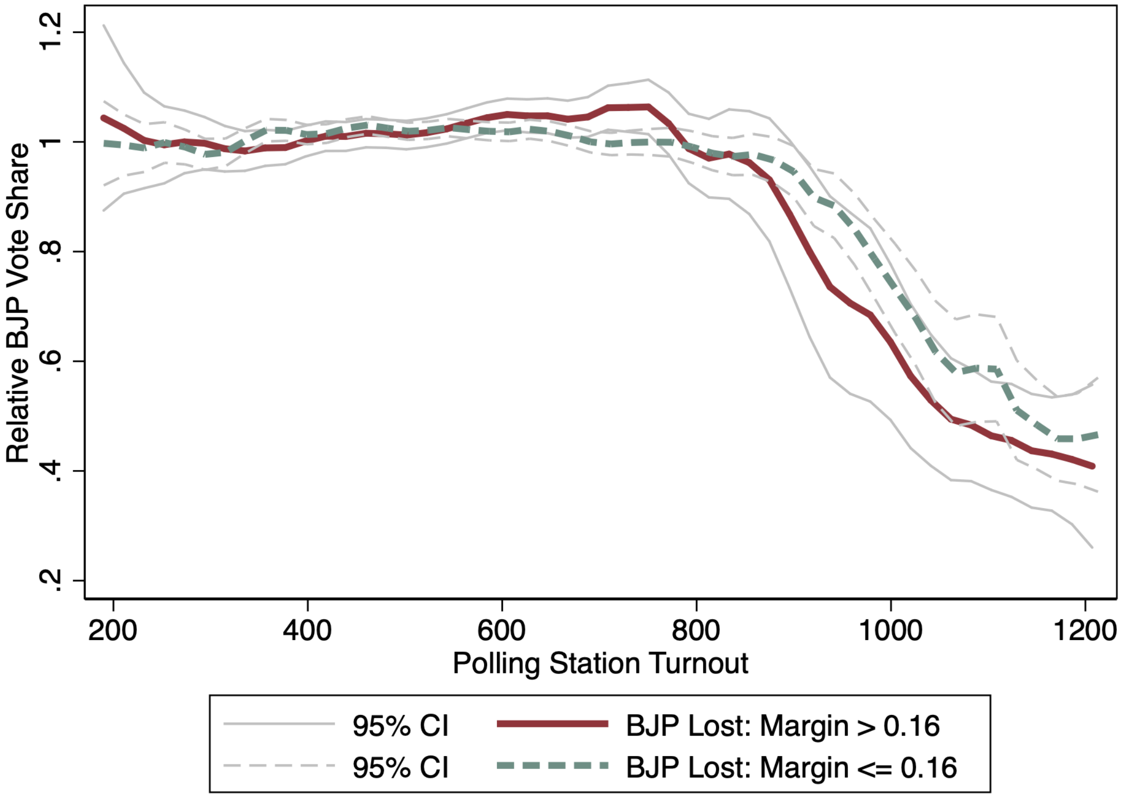


Figure A8—BJP's Vote Rate Density in Uttar Pradesh

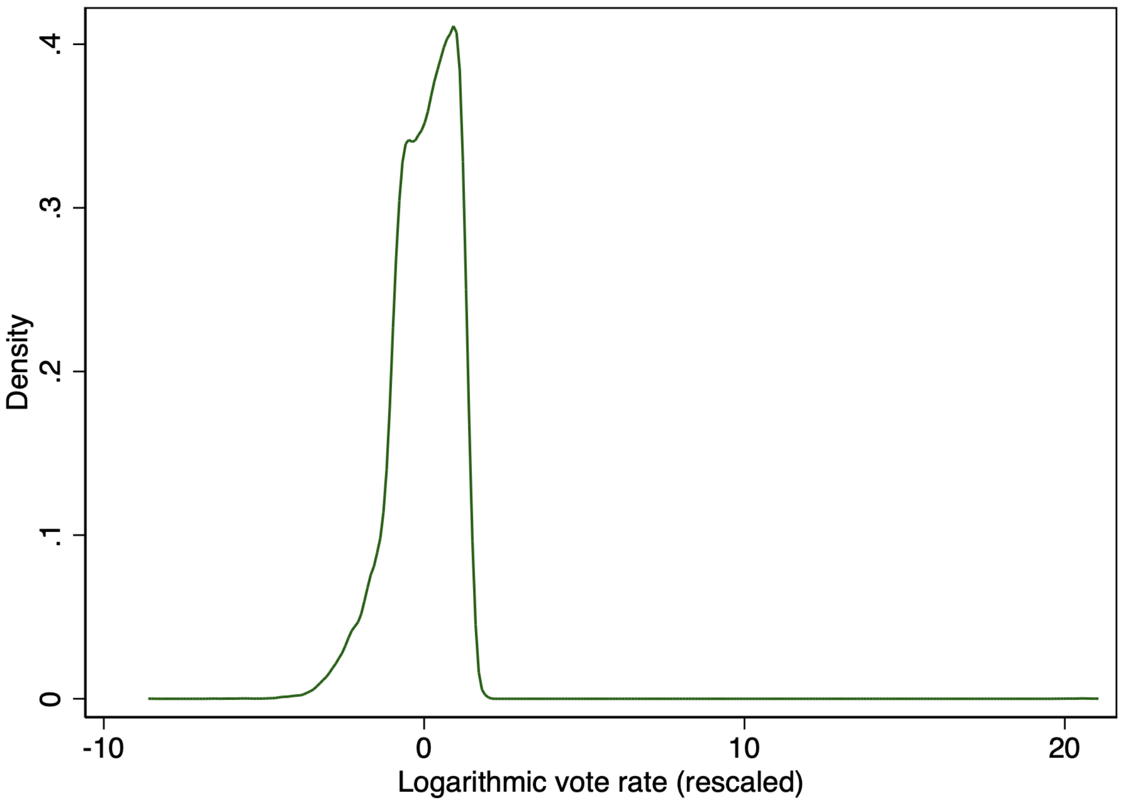
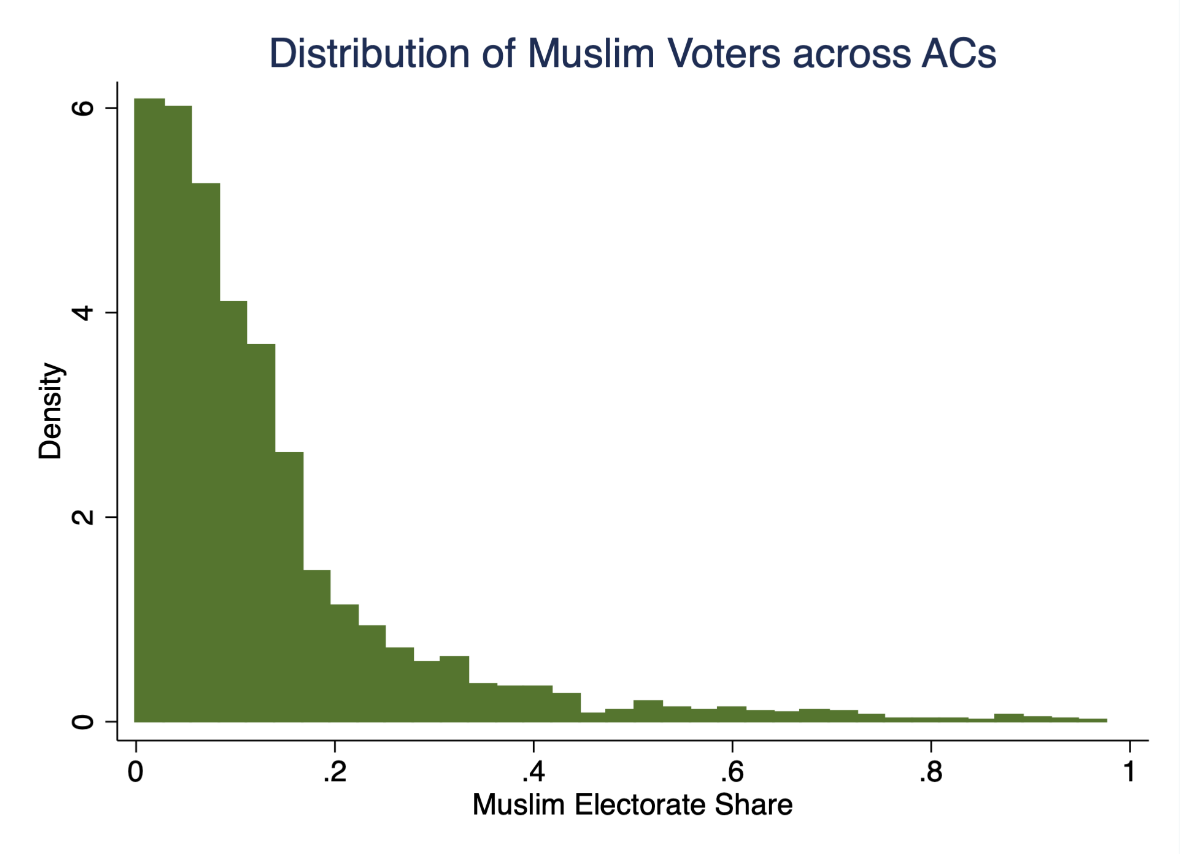


Figure A9—Distribution of Muslim Electorate Share



B Constituency List

BJP Ruled State (1)	State (2)	Constituency (3)	BJP Won (4)
1	Assam	KARIMGANJ	1
1	Assam	NOWGONG	0
1	Bihar	PATALIPUTRA	1
1	Goa	SOUTH GOA	0
1	Haryana	ROHTAK	1
1	Jharkhand	DUMKA	1
1	Jharkhand	KHUNTI	1
1	Jharkhand	LOHARDAGA	1
1	Maharashtra	CHANDRAPUR	0
1	Maharashtra	NANDED	1
1	Manipur	INNER MANIPUR	1
1	Uttar Pradesh	SAHARANPUR	0
1	Uttar Pradesh	MUZAFFARNAGAR	1
1	Uttar Pradesh	MEERUT	1
1	Uttar Pradesh	BAGHPAT	1
1	Uttar Pradesh	FIROZABAD	1
1	Uttar Pradesh	BADAUN	1
1	Uttar Pradesh	SULTANPUR	1
1	Uttar Pradesh	KANNAUJ	1
1	Uttar Pradesh	KAUSHAMBI	1
1	Uttar Pradesh	SHRAWASTI	0
1	Uttar Pradesh	BASTI	1
1	Uttar Pradesh	SANT KABIR NAGAR	1
1	Uttar Pradesh	BALLIA	1
1	Uttar Pradesh	MACHHLISHAHR	1
1	Uttar Pradesh	CHANDAULI	1
1	Uttar Pradesh	BHADOHI	1
0	Andaman & Nicobar Islands	ANDAMAN & NICOBAR ISLANDS	0
0	Chhattisgarh	RAIGARH	1
0	Chhattisgarh	KORBA	0
0	Chhattisgarh	BASTAR	0
0	Chhattisgarh	KANKER	1
0	Dadra & Nagar Haveli	DADRA AND NAGAR HAVELI	0
0	Karnataka	KOPPAL	1
0	Karnataka	BELLARY	1
0	Karnataka	TUMKUR	1

0	Karnataka	CHAMARAJANAGAR	1
0	Madhya Pradesh	CHHINDWARA	0
0	Odisha	SAMBALPUR	1
0	Odisha	MAYURBHANJ	1
0	Odisha	BALASORE	1
0	Odisha	BHADRAK	0
0	Odisha	DHENKANAL	0
0	Odisha	BOLANGIR	1
0	Odisha	KALAHANDI	1
0	Odisha	NABARANGPUR	0
0	Odisha	PURI	0
0	Odisha	BHUBANESWAR	1
0	Punjab	HOSHIARPUR	1
0	West Bengal	COOCH BEHAR	1
0	West Bengal	RAIGANJ	1
0	West Bengal	BALURGHAT	1
0	West Bengal	MALDAHA DAKSHIN	0
0	West Bengal	KRISHNANAGAR	0
0	West Bengal	BARRACKPORE	1
0	West Bengal	DUM DUM	0
0	West Bengal	ARAMBAGH	0
0	West Bengal	JHARGRAM	1
0	West Bengal	BARDHAMAN DURGAPUR	1

Notes: The table lists the 59 constituencies where the absolute BJP win margin was within 0.05. The first column indicates whether the constituency belonged to a state ruled by the BJP during the 2019 general elections. The last column indicates whether the BJP won that constituency in 2019.