ESSAYS ON ELECTORAL POLITICS: AN EXAMINATION OF VOTING TECHNOLOGY, VOTER PARTICIPATION, AND THE PROVISION OF PUBLIC SCHOOLS

A Dissertation

Presented to

The Faculty of the Department

of Political Science

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Indrajit Sinha Ray

December, 2018

ESSAYS ON ELECTORAL POLITICS: AN EXAMINATION OF VOTING TECHNOLOGY, VOTER PARTICIPATION, AND THE PROVISION OF PUBLIC SCHOOLS

An Abstract of a Dissertation

Presented to

The Faculty of the Department

of Political Science

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Indrajit Sinha Ray

December, 2018

Abstract

This dissertation comprises of three essays on electoral politics. The first essay examines the impact of electronic voting machines(EVMs) on electoral outcomes in India. It finds that the introduction of EVMs led to a decline in voter turnout while seeing an increase in the winning margin and a decrease in the vote share of the first losing candidate in an electoral constituency. The results of this essay further suggest significant regional variation in the impact of electronic voting machines on voter turnout. The second essay studies the impact of income on voter turnout in California using the 2016 super bloom in Inyo County as a source of exogenous variation in income. The empirical results of this essay suggest that income has a negative effect on voter turnout. The third essay assesses the impact of the distribution of the male-female electoral ratio in an administrative district on the provision of public schools in India. The empirical results of the third essay demonstrate that a higher male-female electoral ratio leads to higher provision of public schools in an administrative district in India, indicating a male bias in the pattern of provision of public schools.

Acknowledgements

I am deeply indebted to my advisor, Dr. Jim Granato for his support, guidance, and advice. His trust in my ability gave me huge impetus in life to pursue my research projects with enthusiasm. He shaped my academic thought process in my years in the graduate program. Without his support, this dissertation would not have been possible. I am also thankful to other members in my dissertation committee, Dr. Jennifer Clark, Dr. Francisco Cantú, Dr. Pablo Pinto, and Dr. John Antel for their insightful comments. I am thankful to my parents, sisters, Anupda, Soumenda, Anish and Soumik for their constant support for my effort. I express my sincere thanks to Nishithda for his important suggestions for success in graduate studies. I also thank Saptarshi, Anirban, Somdeep, Abhishek, Anurag, Vidyut, Mariella, Alper, Kenicia, and Tina for their important roles as friends, mentors, and well-wishers. Versions of the first essay of this dissertation were presented at the 2016 EITM summer workshop in Houston, 2017 SPSA annual conference in New Orleans and 2017 MPSA annual conference in Chicago. I thank all the participants in this workshop and the conferences for their insightful comments on my paper.

Contents

List of Figures

List of Tables

DEDICATION

In memory of my aunt Hiranmayi Sinha Ray

1 Introduction

This dissertation presents three essays on electoral politics. It covers India and the U.S. which are the largest and oldest democracies of the world respectively. Together the three essays explore research questions on voting technology, voter participation and the provision of public schools. It touches on two large quasi-experimental exercises in India: the introduction of the electronic voting machine and 2008's redistricting. In the nature of the empirical investigation, all three essays share a common quest of causal inference in an observational setting. There is the inner motive to see the world from the prism of an experimental setting.

In the first essay, I investigate how a change in voting technology can affect electoral outcomes in the largest democracy of the world, India. Electronic voting machines(EVMs) were introduced in India in place of the paper ballot system in a phased manner starting from the late 1990s. I exploit the staggered schedule of the introduction of EVMs to run a difference-in-differences(DID) empirical methodology to estimate the causal impact of electronic voting machine on a number of outcomes using data on Indian national elections. The primary outcome of interest is voter turnout. The other electoral outcomes considered in this study are the total number of contestants, total number of female contestants, winning margin for the winner, vote share of the winner, vote share of the first losing candidate, vote share of the second losing candidate, ratio of vote shares of second losing to first losing candidates (sf-ratio). The empirical results show that there is a negative effect of EVMs on voter turnout. Voter turnout decreased by 2.81 percentage points in response to the EVM exposure. There is evidence that the total number of contestants decreased in response to EVMs. There is a positive effect of EVMs on the winning margin of the winner. On the other hand, EVMs have a negative effect on the vote share of the first losing candidate. The heterogeneity analysis suggests that richer states experienced a bigger fall in voter turnout with the introduction of EVMs. The empirical evidence further suggests that the fall in voter turnout has been mainly caused by the constituencies in big metropolises in India. Using constituency level covariates, I matched the EVM constituencies with non-EVM constituencies in a subsample of constituencies employing the Propensity Score Matching(PSM) method. The DID-analysis on this matched sample of data shows that there is no significant effect of EVMs on voter turnout. I run similar matched difference-in-difference(PSM-DID) estimation with a sample of selected northern Indian states. In this case, I find that the EVM effect is positive and is significant both statistically and substantively. These results show that the EVM effect varies significantly over regions of India.

In the second essay, I investigate whether income has a causal impact on voter turnout. In an observational world, it is difficult to identify the causal impact of income on voter turnout owing to various sources of endogeneity. I use a natural phenomenon of the 2016 super bloom in the Inyo County in California as a source of exogenous variation in the income generation process to identify its impact on voter turnout at the county level in California. I use three empirical methods (Difference-in-Differences(DID), Synthetic Control and Nearest Neighbor Covariate Matching) to identify the causal effect using the U.S. presidential elections data from 1992 to 2016. The nature of the empirical results is broadly similar across different empirical methods. The voter turnout, as a percentage of the total voting eligible population, decreased due to the super bloom(a proxy for economic activity). There is evidence that the registration rate experienced a decline due to the super bloom. The fall in voter turnout mostly manifested in terms of a decrease in mail voter turnout.

In the third essay, I study the relationship between the gender composition of the electorate and the provision of public schools in India. The proportion in which male and female voters are electorally distributed in electoral constituencies can possibly impact the extent and direction of public provision of schools. However, in an observational world, it is difficult to identify the causal impact of the male-female relative electoral ratio on public provision of schools. I use the 2008 redistricting exercise in India as a source of exogenous variation in the male-female relative electoral ratio at the administrative district-level to identify its impact on public school provision measured by the total number of government schools. I use a fixed effect regression method with repeated cross-section state election data from the 2001-2012 period. The empirical results demonstrate a possible male bias in the pattern of provision of public schools in India. The number of public schools is more likely to be higher in administrative districts where the average gender electoral ratio favors men.

2 Voting Technology and Electoral Outcomes: The Case of Indian Electronic Voting Machines (EVMs)

2.1 Introduction

In a large democracy like India, where about a million voters vote in the average constituency (electoral district), the use of traditional paper ballot technology poses challenges. First, the counting of the votes becomes a cumbersome and complicated process. Second, owing to the combined possibility of corruption and human error, the perception among voters about the electoral results reflecting the true opinion of the electorate has been suspect. The Election Commission of India, perhaps sharing such concerns, introduced the use of electronic voting machines, henceforth referred to as EVM, in the late 1990s on an experimental basis. In the early 21st century, EVMs were universally adopted in all major elections in India. [1]

Even though this was a very important electoral reform in the largest democracy of the world, the effects of this policy on voter behavior and electoral outcomes in India have not been adequately studied. In this paper, I estimate the causal effect of this program on various electoral outcomes like voter turnout, winning margin, number of candidates, vote shares, etc. There are a few studies on the effect of electronic voting machines in Brazil. Hidalgo (2012) finds that the introduction of the electronic voting machine increased the percentage of valid votes cast by nearly a third in legislative elections in Brazil. And, the adoption of this new voting technology increased the political competition and negatively affected the right-of-center parties. Moraes (2012) also finds a positive impact of electronic voting technology on enfranchisement in Brazil, with an increase in the number of valid votes. Fujiwara (2015)

[1] I define national and state elections as major elections

shows that use of electronic voting machine reduced error in the voting process in Brazil. It also increased the enfranchisement of less educated citizens. This, in turn, led to higher government spending on healthcare, which is particularly beneficial for the poor. Ansolabehere & Stewart (2005) analyze the relative performance of various voting technologies using the U.S. presidential, gubernatorial and senatorial election data from 1988 and 2000. They find paper ballots to be more effective in reducing the number of residual votes in the U.S. presidential elections in comparison to other voting technologies, like optically scanned ballots, mechanical lever machines, and direct register electronic machines. Allers & Kooreman (2009) find a positive impact of electronic voting on voter turnout and a negative impact on the fraction of residual votes in the Netherlands. Card & Moretti (2007) analyze the impact of touch-screen voting machines in the 2000 and 2004 presidential elections in the U.S. They find a negative impact on the voter turnout among Hispanics and a positive effect on the vote share for George W Bush. Alvarez et al. (2011) find that voters in Argentina and Colombia show high levels of confidence in electronic voting. In the case of electronic voting in Belgium, a larger section of the voting population has found comfort with this new technology of voting (Delwit et al., 2005). In analyzing the relationship between voter confidence and voter participation in the U.S., Alvarez et al. (2008) find that voting technology impacts the confidence of voters about the accuracy of the voting process. They find that white voters had relatively more confidence on the voting process when they cast their vote by traditional paper ballot compared to punch card or electronic voting machines. Kimball & Kropf (2008) analyze the effect of voting technology on the ballot measures in U.S. elections. They find that EVMs increase the number of residual votes on the ballot measures. Stein et al. (2008), on the other hand, find that voting technology does not have any significant effect on the voters' confidence about the accurate recording of the votes in the U.S. Pomares et al. (2014) investigate the introduction of electronic voting in Salta, Argentina. They find that the usability of EVMs is more important to both voters and poll-workers compared to the confidence issues related to the EVMs. In the context of this electronic voting experiment in Argentina, Barnes et al. (2017) show that electronic voting can lead to ballot splitting, depending on the ballot structure it follows.

There are two working papers (Debnath et al., 2017; Desai & Lee, 2018) which study the impact of the introduction of electronic voting machines in India. Debnath et al. (2017) investigate the impact of EVMs using the state-level electoral constituency as the unit of study. They find that voter turnout declined at the state-level electoral constituency by 3.5 percentage points in response to EVMs. The winning margin and vote share of the winner also declined due to the introduction of EVMs. The results of this study show that EVMs led to the increased provision of electricity at the level of the state legislative assembly constituency. My paper studies the impact of EVMs on electoral outcomes at the national electoral constituencies in India. National level elections in India are much bigger in terms of the absolute number of electorates as compared to state-level elections. They also have bigger implications in shaping the political process of a populous country like India. Thus, it is a very pertinent exercise to study the impact of EVMs at the national level. Like Debnath et al. (2017), my study also finds a negative impact of EVMs on voter turnout. However, the magnitude of the estimated effect is different. The difference in the contexts(state and national elections) may explain this difference in magnitude. Contrary to Debnath et al. (2017), I find a significant positive effect of EVMs on winning margin and insignificant effect on the vote share of the winner. This indicates that EVMs have differentially affected the nature of electoral competition in the contexts of state and national elections in India. Desai & Lee (2018) study the impact of EVMs in India using national election data. The direction of the effect of EVMs on voter turnout from this paper and mine are the same. However, the magnitude of this effect is distinct in my paper. This difference may be attributable to the different turnout specification used in my study. I have used state-year fixed effects, the number of polling stations in a constituency, and male-female elector ratio for a constituency in my turnout model specification as explanatory variables, which differentiates it from the model specification of Desai & Lee (2018). These three factors are very pertinent explanatory variables in determining the turnout in an election. State level politics is undeniably a very important factor in shaping the contour of national politics in a diverse country like India, which follows a federal structure of governance. Thus, state-year fixed effects can be extremely important control variables in the causal estimation of the effect of EVMs in India. Both these studies (Debnath et al., 2017; Desai & Lee, 2018) do not explore the possible relationship between electronic voting technology and the total number of contestants competing in an electoral constituency. The effect of the electronic voting machine can affect the selection of candidates competing in a constituency, which in turn can affect electoral outcomes like winning margin and vote shares of different candidates. My study fills this gap in the literature by providing empirical support to this theoretical possibility.

My study makes three broad contributions to the understanding of the effect of EVMs in India. First, it shows that electronic voting machines have an impact on the voter turnout and there is significant regional variation in this effect. The results of this study indicate the concentration of the EVM effects is mainly in the metropolitan constituencies of India. Understanding these nuances in the variation of the effect of EVMs in India will be helpful in formulating relevant electoral policies at the micro-level. Secondly, my study contributes to the literature by providing a theoretical linkage between voting technology and the number of contestants in India. The empirical results establish that electronic voting technology reduces the total number of contestants in an electoral constituency in India. Thirdly, my study also provides a theoretical linkage between electronic voting technology in India and electoral variables like winning margin and vote share of the winner, first losing candidate and second losing candidate. My study shows that electronic voting technology has a positive impact on the winning margin and a negative impact on the vote share of the first losing candidate.

2.2 Background

2.2.1 India as a democracy

India is a parliamentary democracy. In the 2014 national election, there were around 800 million voters. The national parliament is bicameral. The lower house is called *Lok Sabha* (translates to Council of People) and the upper house of the parliament is called *Rajya Sabha* (translates to Council of States). The lower house of the Indian parliament has 543 seats. The representatives of this lower house are elected through a first-past-the-post(FPTP) electoral system. The national electoral districts are called parliamentary constituencies. The boundaries of the parliamentary constituencies do not correspond to the boundaries of the administrative districts of India. However, the boundary of a parliamentary electoral constituency is bounded within a state. Thus, no parliamentary electoral constituency is shared between two states or union territories in India.

India's administrative structure is federal in nature. Currently, India has 30 states and 7 union territories. Each state has its own legislative assembly(Vidhan Sabha) and some union territories like the National Capital Region(NCR) and Pondicherry also have their own legislative assemblies. There are three types of elections in India: National, Provincial(state level) and local (municipality and village level). The most important election is the national election where voters throughout India choose their representatives for the national parliament (Sansad). Normally, national elections are held every five years.

2.3 Electronic Voting Machines in India

The genesis of the use of the electronic voting machine in India can be traced to an experimental use in 50 polling stations in the southern state of Kerala in 1982. There was an amendment of the law in 1988 passed by the national parliament to allow the use of EVMs for conducting elections in India. This amendment rule was notified in 1992. From 1992 to 1998, the Election Commission of India conducted an awareness campaign on electronic voting machines among the stakeholders. Since November 1998, EVMs were being used in the state level elections. In the 1999 national election, EVMs were used in 46 parliamentary constituencies. The total size of the electorate of these 46 constituencies was approximately 60 million. The 2004 national election witnessed the use of EVMs in all the parliamentary constituencies of the country. Since then, EVMs have been used for all the constituencies in national elections(such as the 2009 and 2014 national elections).

The maximum capacity of an electronic voting machine(EVM) is 3840 votes. This capacity is more than enough for a polling station in India where the total number of voters do not exceed 1500. A single EVM machine has the accommodation capacity for a maximum of 16 candidates. If the total number of contestants is more than 16, additional EVMs need to be provided in a polling booth. However, there can be a maximum of four EVM machines in a polling booth, which is sufficient to conduct an election with a maximum of 64 contestants. This is a significant capacity in terms of the number of candidates. If the total number of contestants exceed 64 for an electoral constituency, electronic voting machines can not be used for conducting the election. In such cases, a paper ballot will be used. An EVM is battery-powered and runs on the normal 6-volt alkaline battery. Thus, EVMs can be used even in areas where there are no power supplies. These voting machines are not connected to the internet and they do not have the capacity to transmit any data to any remote server.

The electronic voting machine in India is very simple and user-friendly in design compared to EVMs used in other parts of the world. The names and party symbols of the candidates are printed on the upper surface of the machine. A press button is there for each of the candidates name printed on the machine. A voter only has to press the button corresponding to a candidate and the vote is recorded immediately in the voting machine. An EVM has two units, a control unit, and a ballot unit. These two units are joined by a five-meter cable. The control unit is administered by the presiding officer or a polling officer assigned to a polling booth. In the paper ballot system, the polling officer in a booth used to provide ballot papers to the voters. In the electronic voting system, they have to press the ballot button in the control unit under their supervision. This enables a voter in the polling room to cast vote by pressing the required button in the voting machine. The use of EVMs can substantially reduce the chance of booth capturing. The presiding officer or any of the polling officer in a polling station can push a designated close button and all the EVMs in a polling booth will stop registering votes. An EVM is programmed to record a maximum of five votes in a minute of time. Thus, even in the case of a possible booth capture, the miscreants can only enter five votes per minute at maximum. By the time they can put in some substantial number of bogus votes, security forces will have arrived to clear the captured booth.

2.4 Voting Technology and Electoral Outcomes: Theoretical Framework

The impacts of voting technology on electoral outcomes can be many-fold and can operate through several channels. First, we might expect that such a technological advancement affects voters. Second, there may be idiosyncratic effects on the candidates' choice to run. Finally, we might expect outcomes of the election to be affected in terms of the actual results, which are jointly determined by voters and the candidate selection.

2.4.1 Voter Turnout

In the empirical literature, Fujiwara (2015) does not find any impact of electronic voting technology on voter turnout both as a share of registered voters and the voting age population in Brazil. Allers & Kooreman (2009) find a temporary small positive impact of electronic voting on voter turnout in the Netherlands. However, they do not find any evidence of any significant permanent effect of electronic voting on turnout in this context. Thus, there is no definite consensus in the literature regarding the direction and magnitude of the effect of electronic voting technology on voter turnout.

Here, I explain how a change in voting technology can impact voter turnout using the expected utility of voting framework of the economic theory of voting (Downs, 1957; Riker & Ordeshook, 1968). The expected utility from voting is $pB + D - C$. Here, p is the probability of being an instrumental/pivotal voter. B is the instrumental benefit from voting. D represents the expressive/consumption utility of voting. C denotes the cost of voting. The probability of turnout is a monotonic function of this expected utility of voting for an individual voter.

Mathematically,

$$P_i = f_i(E(V)_i) \tag{1}$$

where $E(V)_i = p_i B_i + D_i - C_i$ is the expected utility of voting for an individual i. P_i is the probability of turnout. $f_i' = \frac{dP_i}{df_i} > 0$ which implies that turnout probability is an increasing function of the expected utility of voting. Suppose, there are N eligible voters in an electoral constituency and τ is the change of voting technology. Thus, we want to estimate the following

$$\sum_{i=1}^{N} \frac{dP_i}{d\tau} = \sum_{i=1}^{N} \frac{df(E(V)_i}{d\tau} == \sum_{i=1}^{N} \frac{d}{d\tau} * f(pB_i + D_i - C_i) = \psi \tag{2}$$

The purpose of the empirical exercise of this paper is to empirically identify and quantify the ψ term in the above equation across multiple constituencies. If identified, the empirical analysis will recover the causal estimate (β) as a monotonic function of ψ.

In electoral constituencies in India with a very large number of the electorate, the p term will be close to zero. Hence, the expected utility term pB will be extremely low in magnitude. Therefore, the pB term can be ignored owing to its negligible impact on the total expected utility from voting. For individuals with a great passion for using technology in all spheres of human activity, the electronic ballot system can increase the D term which is the expressive or consumption utility of voting. A technology-savvy individual can assign higher intrinsic utility to the civic duty of voting for its use of electronic technology. Thus, for this type of individuals, the total expected utility of voting can increase with the introduction of electronic voting technology. This argument may hold true for the younger voters in an electoral constituency. For other individuals, the shock of new technology can have an adverse impact on the expected utility of voting. It can increase the psychological cost of the act of voting. This set of individuals may be technology-fearing or technology-averse. For them, the introduction of electronic voting increases the cost of voting term C in the expected utility function, thus reducing the expected utility of voting. The technology-aversion of the Indian people is perhaps more questionable. The rapid expansion of mobile phone technology across India (Gupta, 2015; S. Rai, 2006) in the early years of the 21st century is a testimony to the technology adaptability of the Indian people. However, doubt about the authenticity of new voting technology can dissuade a portion of potential voters from casting their votes. This may be true for the older segment of the electorate. In the U.S., the change of voting technology from punch card to electronic voting machines had a negative impact on the vote share among elderly people (Roseman & Stephenson, 2005) and on the voter turnout of Hispanics (Card & Moretti, 2007). Thus, electronic voting technology can affect the turnout behavior of various segments of electorate differently. Hence, there can be two theoretical possibilities in this scenario depending on the size and direction of the effects of the C and D terms. The size of the negative effect of the C term can dominate over the size of the positive effect of the D term. In that case, the effect of EVMs on voter turnout will be negative. The second possibility is that the size of the positive effect of the D term will be higher than the size of the negative impact the of the C term. Based on these two theoretical possibilities, I propose the following two hypotheses.

Hypothesis H_{1a}: The effect of electronic voting machine on voter turnout will be negative(when $C > D$)

Alternatively, Hypothesis H_{1b}: The effect of electronic voting machine on voter turnout will be positive(when $D > C$)

2.4.2 Total Number of Candidates

The number of candidates contesting in an electoral constituency is high in Indian national elections. Various factors contributed to the presence of a larger number of candidates contesting in an electoral constituency. Party infighting, ethnic fractionization and non-stringent nomination rules for candidacy led to the proliferation of candidates contesting in an electoral constituency, and electoral contests are often characterized by the presence of dummy and spoiler candidates (P. Mayer, 2013). An electoral reform in 1996 involving the deposit requirement and nomination procedure has reduced the number of contestants significantly. However, it is still comparatively higher at 8.74 per constituency in the 1998 election and at 8.6 in the 1999 election[2]. There is empirical evidence that voting technology can affect the prospects of candidates differently given their viability as a strong contender. The introduction of EVMs has encouraged the viable candidates to invest more vigorously in their political communication with voters and increased the vote share of the more viable candidates in Brazil (Schneider & Senters, 2018). It will be an interesting theoretical exercise to investigate how the introduction of voting technology can possibly impact the total number of contestants in an electoral constituency in India. If candidates believe that superior voting technology will lead to a more transparent electoral process, they might respond in two ways. A high-quality candidate(someone who has a genuine chance of winning or securing a certain percentage of votes) who previously felt insecure may now feel encouraged to contest in elections. A low-quality candidate who previously wanted to take advantage of the faulty electoral mechanism may now feel discouraged to contest in elections. Thus, the use of EVMs will lead to the emergence of a set of higher quality contestants in an electoral constituency. However, there are two distinct possibilities about the direction of the effect of electronic voting machines on the total number of contestants in an electoral constituency. Based on this, I propose the following two hypotheses.

Hypothesis H_{1c}: Total number of candidates contesting in an electoral constituency will increase in response to the introduction of the electronic voting machine.

The alternative hypothesis H_{1d}: Total number of candidates contesting in an electoral constituency will decrease in response to the introduction of the electronic voting machine.

[2]Source: Election Commission of India

2.4.3 Total Number of Female Candidates

It will be an interesting exercise to investigate the effects of EVMs on the number of female candidates contesting in an electoral constituency. Women participate in politics at a lower rate than men in India (Chhibber, 2002). The introduction of a more transparent voting technology may encourage potential female candidates to contest in elections with a higher likelihood. Thus, I expect an increase in the number of female contestants in response to the introduction of EVMs. I propose the following hypothesis.

Hypothesis H_{1e}: The total number of female candidates contesting in an electoral constituency will increase in response to the introduction of the electronic voting machine.

2.4.4 Vote Shares and Winning Margin

EVMs can affect the electoral variables, like vote share of the candidates and winning margin through three channels: variation in voter turnout, elimination of invalid voting, and selection of a better set of candidates contesting in an electoral constituency. Theoretically, the effect of EVM on turnout decision of a voter is idiosyncratic in nature. It is unlikely to be correlated with political ideologies of the voters. Thus, variation in turnout is less likely to generate differential impacts across different candidates' expected vote shares. Elimination of invalid voting is definitely one of the channels through which electronic voting technology can impact the vote share of the candidates and the winning margin of the winning candidate. In paper-based voting technology, the invalid votes were part of the total turnout measures. However, it was not possible to assign these votes towards the intended candidates. Thus, these votes were rejected for the purpose of counting against candidates. The design of an electronic voting machine in India practically eliminates the chance of realization of residual or invalid votes. Debnath et al. (2017); Desai & Lee (2018) find empirical support for near elimination of invalid voting after EVMs were introduced in India. In the time period of this study, the "None-of-the-Above" (NOTA) option was not available to Indian voters.[3] Thus, the elimination of invalid votes certainly affected the total votes for political candidates during that time period. However, as the identities of the invalid votes are not observable, it is difficult to map the invalid votes to the size of vote shares of the candidates and winning margin. Fujiwara (2015) finds that invalid voting is mostly associated with poor voters in the case of Brazil. Thachil (2014a,b) show that Indian political parties are well-connected with the poor voters and poor voters often support elite parties. Hence, if the propensity of casting invalid ballots is correlated with the socio-economic conditions of the electorate, then different party candidates' vote shares are less likely be affected differentially (in expectation) by the elimination of invalid voting in India. Thus, variation in turnout and elimination of invalid voting are less likely to generate differential impacts on different candidates' expected vote shares. However, it is possible to make a theoretical prediction about the direction of the effect of EVMs on winning margin and vote shares of the candidates through analyzing the channel of the better selection of candidates. In a previous paragraph, I argued that the introduction of EVMs can lead to a selection of a better set of candidates competing in an electoral constituency. A better selection of candidates will, in turn, affect the vote shares of the contestants through the interaction with Duvergian dynamics. Presence of a better set of candidates can reinforce the Duvergian dynamics in the Indian context. The Indian electoral system is characterized by single-member electoral districts with the first-past-the-post (FPTP) rule to select the winner. Ideally, as per Duverger's law (Duverger, 1954), the first-past-the-post(FPTP) single-member electoral constituency

[3] None-of-the-Above(NOTA) option was introduced in India in 2013. Voting for NOTA implies a rejection of all the available contestants.

in the Indian electoral system should lead to the concentration of the votes to two candidates in an electoral constituency. This can be attributed to the incentive for strategic voting. In an FPTP single-member electoral district, the electorate will vote for the candidate who has the highest chance of winning. Empirically, there is mixed evidence in favor of the validity of Duverger law in India (Chhibber & Kollman, 1998; Chhibber & Murali, 2006; Diwakar, 2007; P. Mayer, 2013). Electronic voting technology increases the transparency of the election process by ensuring the accuracy of counting of polled votes. This fundamental change in the electoral process can induce the best-ranked candidate to exert more effort to connect to the voters. EVM's introduction in Brazil led to increased campaign activities by the more viable candidates (Schneider & Senters, 2018). Better selection of candidates and the incentive of the best candidate to exert more effort in this system can lead to better identification of the best-ranked candidate to the electorate. In such a context, the vote share of the winner is expected to increase while the vote share of the second-ranked (first-losing) and third-ranked(second losing) candidates should decrease. The ratio of vote shares of the third-ranked(second losing) candidate and the second-ranked(first losing) candidate should decrease. Based on this, I propose the following hypotheses.

Hypothesis H_{1f}: Vote share of the winner will increase with the introduction of the electronic voting machine.

Hypothesis H_{1f}: Vote share of the second-ranked(first losing) candidate will decrease with the introduction of the electronic voting machine.

Hypothesis H_{1g}: Vote share of the third-ranked(second losing) candidate will decrease with the introduction of the electronic voting machine.

Hypothesis H_{1h}: Winning margin between the winner and the second-ranked(first losing) candidate will increase with the introduction of the electronic voting machine.

Hypothesis H_{1i}: The ratio of vote share of the third-ranked(second losing) candidate to the second-ranked(first losing) candidate (SF-ratio) will decrease in response to the introduction to the electronic voting machine.

2.5 Data

The main data source for this paper is the online data archive of the Election Commission of India, which is responsible for the conduct of elections in India. This database covers all national and state elections and documents information on all electoral constituencies for these elections. For reasons that I describe in the next section, I look at national elections in 1977, 1980, 1984, 1989, 1991, 1996, 1998, 1999 and 2004. Within this period of 1977-2004, the boundaries of the electoral constituencies remained fixed as per the delimitation of boundaries done in 1976. The next delimitation of the electoral constituencies took place in May 2008. Thus, I could not consider any national election that took place after 2008 (2009 and 2014). I look at outcomes like voter turnout, winning margin, number of contestants, vote share etc. The dataset covers all the 543 parliamentary constituencies in India. In some election years, elections were not held for some specific constituencies. Overall, there are a total of 4848 constituency-election year observations spanning election years 1977 to 2004. However, the main empirical analysis has been done with data from 1977 to 1999. In this dataset, there are a total of 4305 total constituency year observations. Out of these 4305 observations, there were some uncontested elections. Thus, the actual election did not happen in these uncontested elections. Overall there are 4300 constituency-year observations where contested election happened. Information on the vote share of the first losing candidate and second losing candidate has been collected from a publicly available dataset (R. R. Bhavnani, 2018). The source of data for state-level per capita income data is the Handbook of Statistics published by the Reserve Bank of India. Data for constituency level demographic and economic characteristics have been compiled from the publicly available datasets corresponding to two published papers (Asher & Novosad, 2017; Jensenius, 2015).

Table 1 presents the summary statistics for the dataset. These statistics are based on all the national elections in the 1977-2004 period. The average size of the constituency as described by the total number of eligible voters is close to 1 million. Voter turnout is defined as the total number of polling place voters as a percentage of the total number of eligible voters. The average voter turnout per constituency is around 60%, whereas every constituency has 12 candidates(contestants) on average. I define winning margin as the difference in the total number of the votes cast in favor of the candidate who wins the race and the candidate who secures the second position, as a percentage of the total number of valid votes cast. On average, the winning margin is 15.62% in a parliamentary constituency. The vote share of the winner is the number of votes cast in favor of the winner as a percentage of the total number of valid votes cast. On average, the winning candidate secures approximately 50% of the total valid votes. Table 2 compares the parliamentary constituencies which received EVM exposure in 1999 with those that did not receive it. The results show that average voter turnout in the EVM-constituencies was lower compared to non-EVM constituencies from 1977 to 2004. On average, the EVM constituencies had more male voters and more candidates contesting in an election compared to the non-EVM constituencies.

Figure 1 and Figure 2 show time series plots of the electoral variables, like voter turnout, total voters, total eligible voters in a constituency, winning margin and vote share of the winner, etc. In these diagrams, the EVM constituencies group corresponds to the constituencies that have EVMs in the 1999 national election and the non-EVM constituencies group comprises of the constituencies that did not get EVM intervention in 1999. We see that voter turnout registered a drop in the 1999 national election. Total candidates also declined in EVM constituencies compared to non-EVM constituencies in 1999. Winning margin increased for EVM constituencies in the 1999 national election. The average total voters in a constituency also registered a drop in the year 1999 in the EVM constituencies compared to non-EVM constituencies. On the other hand, average total eligible electors in a constituency have been higher in the EVM constituencies in comparison to the non-EVM constituencies since 1989. Voter turnout dropped in 1999 as evident in Figure 1. The drop in the percentage term can be jointly determined by the drop in the total number of voters and an increase in the number of electors per constituency. Thus, in order to separate out the effect of EVMs on voting outcomes, there is a need to control for the number of eligible electors per constituency.

2.6 Empirical Methodology

The Election Commission of India is the supreme authority in terms of implementing electoral reforms and function, with a considerable degree of autonomy in conducting elections. It gradually rolled out electronic voting machines in the Indian electoral system, with 2004 being the first parliamentary election conducted entirely with EVMs replacing standard paper ballots. However, comparing the outcomes in 2004 when all constituencies in the country received this technology to the earlier elections when no constituency had it would not yield causal estimates because of selection issues as follows. It is possible that elections in 2004 were held on very different issues and agendas compared to 1998 and therefore the difference in outcomes may be due to several other factors apart from EVM.

To deal with this issue, I use the election in 1999 where the election commission introduced this machine to 46 parliamentary constituencies out of total 543 parliamentary constituencies on an experimental basis.[4] However, there was no evidence in favor of the use of any explicit rule of selection of these specific constituencies. Hence, methods like matching and regression discontinuity will not be appropriate methods in this context. The difference of mean method also will not be the right empirical strategy in this case. Comparing these experimental constituencies to the ones without access to EVMs would not identify the true causal effects of the program because of the presence of possible confounders. For example, there may be selection issues like EVMs were introduced to certain constituencies which have historically witnessed low turnout. Comparing turnout in such constituencies to the ones which did not get EVMs would bias our estimates. To address such issues, I use the pre-1999 elections in 1977-1998 period where none of the constituencies received EVMs. Using pre-1999 differences in the EVM-intervened (treated) constituencies and the non-EVM (untreated) constituencies as a control for pre-existing differences, the mean differences in outcomes in 1999 for these constituencies would give us the causal effect of the policy under certain identifying assumptions as described later. Thus, essentially a difference-in-differences(DID) method will be the most appropriate empirical strategy in this context. I have provided a brief conceptual note on the difference-in-differences method in section 6.1 of the appendix.

To estimate the causal effect of the introduction of EVMs on various electoral outcomes, I run the following regression for constituency c in time t:

$$Y_{ct} = \alpha_c + \delta_t + \beta \cdot EVM_{ct} + \gamma \cdot X_{ct} + \rho_{st} + u_{ct} \tag{3}$$

Here, EVM_{ct}=1 if constituency c has EVM in time t. It takes the value 0 other-

[4] The names of the constituencies receiving EVM intervention in 1999 are obtained from a press release of Election Commission of India (Election Commission of India, 1999). This was further confirmed by a response to a Right to Information (RTI) request filed by the author to the Election Commission of India.

wise. α_c represents constituency fixed effects. The constituency fixed effects control for the factors which may vary across constituencies but are constant over time. This may include factors like the area under a constituency, urban status, proneness to electoral crimes, being a continuous stronghold of a particular party or politically powerful clan. δ_t represents election year fixed effects. These year fixed effects control for all the factors which are expected to affect all the electoral constituencies uniformly in a year. This includes the major national level events like the assassination of then Prime Minister of India Mrs. Indira Gandhi in the election year of 1984, the assassination of former Prime Minister of India Mr. Rajiv Gandhi in the election year of 1991 in a poll campaign gathering. It can also include event like India's Kargil war with Pakistan in the 1999 election year. The coefficient of interest is β which gives us the causal effect of the EVM program on electoral outcome Y. ρ_{st} controls for the state-year fixed effects. Indian national politics is characterized by many parties which are often based in a particular state. These state-based parties (often termed as regional parties in India) also compete at the national elections. Essentially, individual states often have their own state-level political currents which can translate into different turnout behavior across the constituencies of different states. State-year fixed effects control for state-specific electoral factors which can affect voting behavior in national elections. This can include state-level political competition, concurrent schedule of state assembly election with the national election in an election year, climatic events like flood or drought that are specific to a state in an election year. These state-year fixed effects also control for factors that vary over states and over time like the emergence of a new political party in a state. For example, a new party All India Trinamool Congress(AITC) was formed in the state West Bengal in the election year of 1998 which significantly changed the political competition dynamics of that state in that election. Similarly, a new party called the Nationalist Congress Party(NCP) was formed in the election year of 1999 which is mainly based in the state of Maharashtra. This new party had a significant impact on the national politics from this state. State-year fixed effects also account for factors like economic development and literacy rate, which may vary by state and over time. The primary coefficient of interest is β, which is the DID-coefficient. The interpretation of β is that it captures the difference in mean of outcome Y in treated constituencies and untreated or control constituencies in the affected year, i.e., 1999 after differencing out these differences in the unaffected years, i.e., pre-1999 election years. Here, X is a set of control variables at the electoral constituency level. The choice of Xs depends on the electoral outcome variable (Y_{ct}) considered in the DID-equation. When the voter turnout is the electoral outcome, I included the natural log of the total eligible voters, natural log of the ratio of eligible male voters and female voters, natural log of the total number of polling stations and natural log of the total number of contestants as constituency level controls. The voter turnout variable may be affected by an increase or decrease of the total number of eligible voters present in a constituency. Male and female eligible voters may have different behavioral patterns regarding turnout. In India, male turnout has generally been higher than female turnout. However, this gender gap in the voter turnout has decreased over the years with the increase in the female turnout ratio (Kapoor & Ravi, 2014; P. Rai, 2011). The total number of candidates competing in a constituency can also affect voter turnout level. A larger number of candidates from a spectrum of ideologies may attract more voters to cast their ballots. Similarly, the absence of candidates for specific parties/ideologies can depress turnout. Voter turnout in a constituency may be impacted by the availability of a requisite number of polling stations in a constituency. I control for these factors by including their natural log values at the constituency-election year level when voter turnout is the electoral outcome in consideration. For electoral outcomes like winning margin, the share of voters and SF-ratio, I control for the natural log of total electors and the total number of candidates at the constituency level.

The DID-estimates of EVM effect on electoral outcomes is less likely to be confounded by the effects of electoral frauds linked to the malpractices like booth capturing and voter intimidation on election days. Elimination of electoral malpractices like booth capturing and voter intimidation are not fully preventable by the use of electronic voting machines. It is dependent on the level of polling booth security. In the 1990s, Election Commission of India introduced a slew of reforms which included deployment of central paramilitary forces for the polling booth security on the election days (Biswas, 2014; Vaishnav, 2017). This reduced the dependence on state police for maintaining law and order around the polling booths. The non-local nature of the central paramilitary force reduces the probability of collusion between local political agents and security forces for organizing electoral frauds. The Central Reserve Police Force (CRPF), one of the federal paramilitary forces of India, played a key role in ensuring the security of the polling booths in the 1999 general election (Sharma & Sharma, 2008). This is a nationwide policy on the election days. This applies both to the EVM and non-EVM constituencies in the year of EVM intervention(1999). In the previous regime of only paper ballot system and less stringent polling booth security, both EVM and non-EVMs constituencies were exposed to the possible incidences of these malpractices in theory. Thus, the election year time fixed effects can account for the general trend in the electoral malpractices in the national elections to a large extent. If some specific constituencies were chronically prone to electoral malpractices, the constituency fixed effects will take care of that. Thus, the estimated DID-coefficients for EVM effects on electoral outcomes are less likely to be confounded by the effect of fraudulent electoral malpractices, like booth capturing. The Indian EVMs are not connected to the internet and do not transmit information to any remote server for counting purposes. Thus, a more sophisticated form of electoral fraud like contamination of vote counts by server capture is not possible in India.

Serial correlation in the difference-in-differences context is a real possibility owing to three possible reasons: limited variation in the treatment variable for units over time, the presence of multiple time periods observations for units and the presence of serial correlation in the dependent variable over time (Bertrand et al., 2004). Thus, the residuals in the difference-in-differences framework can be serially correlated for an electoral constituency over time. In this context, I cluster the standard error at the electoral constituency level following the recommendation of Bertrand et al. (2004). The presence of a large number of clusters(543) is helpful for the asymptotic convergence of the clustered standard error estimate in this case (Cameron et al., 2008).

The key identifying assumption of this framework is as follows: in the absence of EVM policy, β is statistically indistinguishable from 0. Here β is equal to the difference in mean of Y in treated and control constituencies in 1999 minus this difference in 1998 after controlling for Xs. The assumption is that in the absence of this policy, the difference in these differences would be statistically zero. If there is evidence in support of the identifying assumption, then any $\beta \neq 0$ in equation 3 can be attributed to the EVM policy. In the next section, I will describe two falsification exercises that give us confidence in this identifying assumption.

2.7 Falsification Exercise

A potential concern is that the estimated EVM coefficient of equation 3 could have been non-zero even in the counterfactual scenario (i.e. had there been no policy intervention). This could happen if treatment and control constituencies were different over time along dimensions, other than exposure to the EVM program. This potentially violates the identifying assumption. To alleviate such concerns, I propose falsification exercises. However, one must acknowledge that the identifying assumption cannot be tested with the data because it corresponds to a counterfactual scenario which never actually happens. The standard practice, in this case, is to test if the coefficient on EVM will be non-zero either for an unaffected cross-section or during an unaffected period. I have run two falsification exercises for the DID analysis mentioned in the last section. In the first falsification exercise, I take only pre-EVM introduction period data(1977-1999). There were a total 7 election years in this period (1977, 1980, 1984, 1989, 1991, 1996, 1998). I assign 1998 to be the 'fake' EVM intervention year. Then, I run the same DID-specification (as in equation 3) on this pre-EVM period's data. The DID-coefficient will capture the differential trend between EVM constituencies and the non-EVM constituencies in 1998 (EVM-constituencies are those constituencies which used EVMs in 1999). If this 'fake' 1998 EVM DID coefficient is statistically indistinguishable from 0, then it will give confidence to the identifying assumption that in absence of the EVM policy, EVM constituencies and non-EVM constituencies would have zero differential trend.

In the second falsification exercise, I run the following DID specification.

$$Y_{ct} = \alpha_c + \delta_t + \sum_{t=1980}^{1999} \beta_t \cdot D_t \cdot EVM_c + \gamma \cdot X_{ct} + \rho_{st} + u_{ct} \tag{4}$$

In the above specification, D_t is the year specific dummy for the year t. I make the first election year (1977) as the reference year. I take interaction of D_ts for each of the other years with the EVM-treated electoral constituencies group dummy. Each of these interaction coefficients will estimate a DID coefficient with respect to a particular year and the reference year 1977. These year-specific DID-coefficients ($\beta_t s$) will provide information about the evolution path of the possible differential trend between the EVM-intervened constituencies and the rest of the constituencies. This information of year-specific differential trend can be used to assess whether there is sufficient evidence in favor of the causal impact of the EVM introduction on electoral outcomes.

2.8 Results

In this section, I report the estimation results of 3 on a variety of outcomes. I first look at turnout. Then I shift focus to the number of candidates contesting, vote shares and winning margin.

2.8.1 Effects on Voter Turnout

I report these results in Table 3. I find that the voter turnout(the number of polling place voters as a percentage of total eligible electors) declined by 2.81 percentage points in EVM-constituencies compared to non-EVM constituencies in 1999. The falsification exercise results for this electoral outcome have been reported in Table 4 and 5. Table 4 reports the falsification exercise where 1998 has been assigned as a fake EVM-intervention year. The falsification DID-estimation exercise has been done with data in pre-EVM intervention years(1977-1998). The DID estimate for the fake EVM intervention in Table 5 is statistically insignificant and numerically low in value. This falsification result gives confidence to the DID-estimate of the main DID exercise in Table 3. Table 5 reports the falsification exercise estimates of the yearly DID-coefficients which show the possible differential trend between EVM and non-EVM constituencies over the years with respect to the reference year 1977. As per the results of this falsification exercise, there was already a negative differential trend between EVM and non-EVM constituencies before the introduction of EVM in 1999. However, the results also show that the DID coefficient for the year 1999 was numerically larger in absolute value than all other yearly DID-coefficients of the previous years. Thus, there is some evidence(although less robust) as per the results of this falsification exercise that the EVM introduction reduced voter turnout in EVM constituencies compared to non-EVM constituencies. Thus, the empirical results on voter turnout support the validity of hypothesis H_{1a}. Column 1 of Table 16 reports the joint significance of a set of demographic and political explanatory variables at the electoral constituency level for the voter turnout regression estimation. The variables are jointly significant at the 99 percent confidence level. The DID-coefficient for voter turnout is statistically significant for samples with outliers removed (column 4 of Table 6). The outlier observations are defined as observations for which corresponding studentized residual absolute values exceed 2. Table 6 shows the DID estimates of voter turnout with different methods of standard error estimation(heteroscedasticity consistent Huber-White sandwich robust standard error and jackknife resampling method). The effect of EVM on voter turnout is significant in both the cases of robust standard error and jackknife standard error estimation (column 2 and 3 of Table 6).

2.8.2 Effects on Number of Candidates, Winning Margins, Vote Shares

Table 7 reports the DID-estimates for the total number of contestants and the total number of female contestants. The results show that the total number of contestants decreased in EVM constituencies compared to non-EVM constituencies and the effect of EVMs on the total number of candidates is statistically significant. The decline in total female contestants is substantively and statistically insignificant. The falsification exercise estimate reported in Table 8 shows that the total number of candidates already had pre-existing negative differential trend between EVM and non-EVM constituencies in the pre-EVM years. The number of female contestants outcome had no pre-existing differential trend between EVM and non-EVM constituencies which continued even in the post-EVM introduction period. The falsification exercise estimates in Table 9 show that the yearly DID-coefficient for 1999 has a negative sign (although statistically insignificant). However, this coefficient for 1999 is lowest in absolute value compared to the other yearly DID-coefficients from previous years. It is to be noted that the DID-coefficient for the year 1998 is also numerically small than those in the pre-1998 years. However, the coefficient is still positive in value whereas the DID-coefficient for 1999 has negative value. The evidence from the falsification exercises and the main DID-estimates suggest that there had been a modest decline in the number of candidates contesting in EVM constituencies compared to non-EVM constituencies in 1999. The DID-coefficient for the total number of contestants is statistically significant for samples with outliers removed (column 4 of Table 10). Table 10 reports the main DID-estimation results for the total number of candidates with different methods of standard error estimation(heteroscedasticity consistent Huber-White sandwich robust standard error and jackknife resampling method). The effect of EVM on the total number of candidates is significant in both the cases of robust standard error and jackknife standard error estimation (column 2 and 3 of Table 10). Thus, the empirical results provide support to the validity of hypothesis H_{1d}. The DID-estimates reported in Table 11 suggest that vote share of the winning candidate has not undergone any change in EVM constituencies in comparison to non-EVM constituencies. The DID-coefficients for the EVM effect on the vote share of the second losing candidate and the ratio of vote share of the second losing and the first losing candidate(sfratio) are statistically insignificant. The DID-estimates show that the winning margin of the winner has increased in EVM constituencies compared to non-EVM constituencies. The corresponding DID-coefficient is significant at 10 percent level of significance. Table 11 results also show that the vote share of the first losing candidate declined by 2.65 percentage points in EVM constituencies compared to non-EVM constituencies. The DID-estimates of the falsification exercise in Table 12 shows that the trend differential between EVM and non-EVM constituencies were statistically insignificant for the electoral outcome variables winning margin for the winner and vote share of the first losing candidate. The DID-coefficients for these electoral outcomes in the falsification results were numerically also much lower in magnitude compared to the corresponding DID-coefficients in the main DID analysis. For the other electoral outcome variables like vote share of the winner, the vote share of the second losing candidate and the ratio of vote share of second and first losing candidates, the falsification DID-coefficients are statistically insignificant and were similar in value and direction as the DID-coefficients for these outcomes in the main DID-analysis. If we look at the yearly DID-coefficients of the falsification exercise(for the outcome of vote share of the first losing candidate) reported in column 3 of Table 13, we find that the DID-yearly coefficient for 1999 is statistically significant and negative in sign. The DID-coefficients for all the pre-1999 years are statistically insignificant except that of 1980. However, the DID-coefficient for 1980 was positive in magnitude which contrasts with the negative significant value of DID-coefficient for 1999. Thus, compared to 1980, there has been a reversal of differential trend between EVM and non-EVM constituencies which bolsters the evidence that there has been a significant fall in the vote share of the first losing candidate after the introduction of EVM in 1999. The results for winning margin is nearly the same in this falsification exercise. The DID-coefficient for winning margin for 1999 is statistically significant and larger in actual value than the DID-coefficients of other years except 1996. Thus, the empirical results support the hypothesis H_{1g} and H_{1i}. Column 2 and 3 of Table 16 report the joint significance of a set of demographic and political explanatory variables at the electoral constituency level for the winning margin and vote share of the first losing candidate regression estimation respectively. The variables are jointly significant at 99 percent confidence level in both the cases. The DID-coefficients for winning margin and vote share of the first losing candidate are statistically significant for samples with outliers removed (column 4 of Table 14 and column 4 of Table 15). Table 14 and 15 show the main DID-coefficients for winning margin and vote share of the first losing contestant with different methods of standard error estimation(heteroscedasticity consistent Huber-White sandwich robust standard error and jackknife resampling method). The effect of EVM on both these dependent variables is significant in both the cases of robust standard error and jackknife standard error estimation (column 2 and 3 of Table 14 and of Table 15).

The DID-coefficients for 1999 for the electoral outcomes like vote share of the winner, vote share of the second losing candidate and SF-ratio, are statistically insignificant and more or less similar in value and magnitude to the yearly DID-coefficients for the pre-1999 years. Overall, the falsification exercises give confidence to the DID-results in the main analysis for winning margin for winner and vote share of the first losing candidate. The introduction of EVMs in 1999 has increased the winning margin for the winner and reduced the vote share of the first losing candidate in EVM constituencies compared to non-EVM constituencies. Here it is very important to note that the increase in the winning margin is associated with no change in the vote share for the winner. The vote share for first losing candidate decreased. But the vote share of the third-ranked (second losing) candidate remained unchanged. This pattern is interesting. It implies that the loss of vote share of the second-ranked(first losing) candidates got translated into an increase in votes for the candidates who have ranks beyond third. This broadly conforms to the finding of Desai & Lee (2018) that EVMs in India increased vote share of the minor parties. This particular pattern of vote share change may be caused by a number of reasons. First, it may be that better candidate selection had taken place in response to the new voting technology. The decline in the total number of candidates may provide some suggestive evidence in favor of it. This created an uncertainty in the minds of voters about the real ranks of candidates at the lower order(from third onwards). The new electronic voting technology might have encouraged the intense ideological voters to cast their votes to express their ideological solidarity to candidates even if they did not have any realistic chance of winning. Third, it may be that the elimination of previous invalid voting got translated into the higher vote share of the low ranked candidates. It is beyond the scope of this study to test the validity of each of the theories in this context.

2.9 Voter Turnout: Heterogeneity in the EVM Effect

Figure 3 and 4 show time series plots (for the election years 1996 to 2004) of voter turnout for selected states and union territories considered in this study. There were two new states created in the year 2000. These new states have been merged with the original parent states for the purpose of these diagrams. The fall in voter turnout in EVM constituencies in 1999 is very visible for some states like Delhi, Kerala, Goa, Haryana, Punjab and Maharashtra. Thus, there is potential heterogeneity in the effect of EVMs in different states and union territories. This poses an important question as to why the impact of EVMs has potentially differed across states. What can be the underlying factors that contribute to this heterogeneity of the potential impact of EVMs in different states? One such factor may be per capita income. There is evidence in the literature that the rich in India have a lesser propensity to vote while the poor have a higher propensity to vote (Ahuja & Chhibber, 2012; Joshi et al., 2016; Keefer & Khemani, 2004). However, per capita income data is not available for India at the electoral constituency level. Census of India can only provide data on administrative districts for every 10 years (1991 and 2001 census figures may be relevant in our case). However, electoral districts and administrative districts are not congruent in India. Given the lack of year-specific income data at the electoral constituency level, I use income data at the state level. The source of this data is the Handbook of Statistics published by the Reserve Bank of India, which is the central banking authority in India. I use log of the per capita net state domestic product at the factor cost as the measure of per capita state-level income. I look at heterogeneous treatment effects by the log of state per-capita income in the year 1998. It is to be noted that 1998 was the immediate pre-EVM intervention year. I categorize the heterogeneity by considering the treatment effects for log per-capita income greater and less than the mean value of the log per capita income for all the states and union territories in the dataset. A similar analysis of heterogeneity has been done with the median level of log state per capita income in 1998. The purpose of this heterogeneity analysis is to see whether the EVM effect on voter turnout differs for electoral constituencies in states/union territories which had higher per capita income in 1998 than the aggregate level mean (or median) compared to constituencies in states which had lower per capita income in 1998 than the aggregate level mean (or median) per capita income. I estimate the following equation 5 to test for the income heterogeneity in the EVM effect on voter turnout.

$$
\begin{aligned}
Y_{ct} = \alpha_c + \delta_t + \beta_1 after * EVM_c * Incdum + \beta_2 after * EVM_c \\
+\beta_3 after * Incdum + \beta_4 EVM_c * Incdum + \beta_5 EVM_c \\
+\beta_6 after + \gamma \cdot X_{ct} + u_{ct}
\end{aligned} \tag{5}
$$

The primary coefficient of interest, in this case, is β_1 which is the triple difference term. The magnitude and direction of estimated coefficient β_1 inform about how the EVM effect on voter turnout differs across the state per capita income. In one specification, the variable 'Incdum' takes the value 1 if the constituency is in a state which has log per capita income(1998) greater than aggregate log state per capita income(1998), and 0 otherwise. In the other specification, this variable takes the value 1 if the constituency is in a state which has log per capita income(1998) greater than the aggregate log state per capita income(1998) for all states and union territories, and 0 otherwise. In order to get the total effect of EVM in the states with per capita income greater than the aggregate level mean(or median) per capita income, we have to add the values of β_1 and β_2. $after$ is a dummy variable which takes the value of 1 for the year 1999 and assumes the value 0 for all other years. I report the estimation results in Table 17. The sign of the triple difference coefficient is negative in both the specifications. In the case of median per capita specification of characterizing heterogeneity, the triple difference coefficient is also statistically significant at 90%

level of confidence. In both specifications, the sum of estimated $\beta_1 + \beta_2$ are also negative in value. This shows that the magnitude of the decline in voter turnout due to the introduction of EVM is larger in constituencies in richer states than that of constituencies in poorer states.

2.10 Comparison of Turnout: 1999 and 2004 Elections

In 2004, the entire national election was conducted using EVMs whereas, in the 1999 elections, 46 out of total 543 national level constituencies used EVMs. Thus, taking the whole period of 1999 to 2004 into account, we find that 46 electoral constituencies received EVM intervention in national elections twice and rest of the constituencies got the EVM intervention once in the national level elections. We can think of these 46 constituencies receiving an EVM intervention two times as 'intensely treated' and the rest of the electoral constituencies as 'less intensely treated'. Now, I create an EVM treatment dummy which takes the value 1 if the electoral constituency is twice treated by EVM in national elections in the time period of 1999 to 2004 and takes the value 0 otherwise. With this, I run the following difference-in-differences equation 6.

$$Y_{ct} = \lambda_c + \phi_1 after * EVMIT_c + \phi_2 after + \phi_3 EVMIT_c + \psi \cdot X_{ct} + \epsilon_{ct} \tag{6}$$

λ_c is the constituency fixed effect for constituency c. $EVMIT_c$ takes the value of 1 if constituency c is twice treated in 2004 and 0 otherwise. $after$ takes the value of 1 for the year 2004 and 0 otherwise. X_{ct} represents control variables which vary by constituencies and time. I have included the log of total eligible electors, the log of male-female elector ratio, the log of the number of polling stations and the log of the number of candidates as constituency level control. The primary coefficient of interest is ϕ_1 which captures the effect of 'twice EVM treated' compared to 'once

EVM treated' on voter turnout.

Table 18 presents the results. I find that the difference-in-differences estimate on turnout is very close to zero in value and statistically insignificant. This implies that high-intensity exposure of EVMs has no effect on voter turnout as compared to low-intensity exposure.

2.11 EVM and Voter Turnout: Effect of Metropolitan Constituencies

The falsification exercise of Table 5 reveals an interesting pattern of a differential trend among EVM and non-EVM constituencies in the pre-EVM years. In all the pre-EVM election years, there was a negative differential trend between EVM and non-EVM constituencies. This negative differential trend increased after the introduction of EVM in 1999 thus suggesting a possible negative effect of EVM on voter turnout. However, the consistent negative trend differential of pre-EVM years between EVM and non-EVM constituencies suggests that there were some fundamental differences between EVM and non-EVM constituencies in terms of the factors that affect turnout. A number of EVM-constituencies are from the metropolitan cities of India. It may be that metropolitan constituencies follow a differential trend than rest of the constituencies. The EVM constituencies in the metropolitan cities can be affected differently by national events in a given election year. In these highly urban and more affluent constituencies, fatigue about frequent voting can be a reality. There is evidence in the literature that frequent elections reduce voter turnout due to voter fatigue (Garmann, 2017; Rallings et al., 2003). A set of studies (Auerbach, 2015; Chandra & Potter, 2016; Falcao, 2009; Kumar, 2009) report that voter turnout rate is lower in metropolitan constituencies in India. Around the time of introduction of EVMs in India in the late 1990s, three national elections were held within the span of three years (1996, 1998 and 1999) due to coalitional instability in the elected government leading to frequent interim elections. Ideally, these three elections should have happened within a time span of 15 years. It is likely that these frequent elections depressed the turnout in the metropolitan cities disproportionately more than other constituencies. In order to account for this kind of possibilities in our analysis, I run the main DID specification of 3 with the group-specific linear trend for EVM constituencies and another specification with metropolitan constituencies group-year fixed effects. Columns 2 and 3 of Table 19 show the results. In these two specifications, the EVM DID-coefficients have lost both statistical and substantive significance. This starkly contrasts with the main DID-result of the EVM effect in column 1 in Table 19 which is both substantively and statistically significant. Among all these three model specifications, the EVM coefficient is least in value and close to zero in magnitude where I have controlled for metropolis group year fixed effects. These results point out that EVM possibly has no effect on the voter turnout once we control for the factors which distinctly affected the metropolitan constituencies compared to other constituencies.

2.12 Effect of EVM on Voter Turnout: Subsample Analysis

2.12.1 Propensity Score Matching and Difference-in-Differences

The difference-in-differences method does not necessarily require covariate balance among the treated and control groups. However, it may be that secular differences in the electoral constituency characteristics are correlated with a differential trend between EVM-constituencies and non-EVM constituencies. In that case, the DID method alone will not be appropriate. However, DID estimation with constituencies matched by constituency level characteristics will be more appropriate. The underlying idea behind is that EVM and non-EVM constituencies which are similar in observable characteristics are likely to be similar in unobservable characteristics as well(which will help more precise identification of EVM effect in a DID-framework). The exact criteria (if any) on which the selection of these 46 EVM constituencies was done are not in the public domain. The DID estimate does not have the requirement for any particular rule of assignment in the identification strategy. However, by observing the choice of the EVM constituencies in 1999, the probable selection criteria can be surmised. The selected EVM constituencies can be categorized in three types: constituencies from the state capital or state major industrial town, constituencies from India's large metropolitan city like Delhi, Mumbai, Kolkata, Chennai, Hyderabad, and Bangalore. The non-metropolitan and non-state capital EVM constituencies come from the states of Punjab and Haryana. However, these two states have very high agricultural productivity and highly developed rural areas. Thus, effectively the chosen areas were urban in nature or developed rural areas which mimic urban areas in terms of characteristics. Urban areas anywhere in the world now are characterized by a more formal economic structure, greater educational facility, etc. It will be interesting to see how the selected EVM constituencies differ in terms of characteristics along the above-mentioned dimensions in comparison to those constituencies not selected. Unfortunately, the economic and demographic data are not reported at the electoral constituency level in India. They are reported mostly at the administrative district level. Generally, the electoral constituency level data are constructed from the administrative district level data using the Geographic Information System (GIS). Fortunately, there is a publicly available demographic data (Jensenius, 2015) on Indian state-level electoral constituencies for most of the major states of India. However, this data does not have demographic data on all the state level electoral constituencies. I use this dataset for mapping the demographic characteristics for the electoral constituencies for my study which looks at the national electoral constituency(parliamentary constituency) level. The demographic data is available at state electoral constituency level in Jensenius (2015). I construct the demographic data for parliamentary constituencies by averaging across the available data for state assembly constituencies contained within a parliamentary constituency. In this way, I got demographic data for 486 parliamentary constituencies out of total 543 constituencies. In this subsample of 486 electoral constituencies, there are 32 EVM constituencies(which got EVM intervention in 1999). Table 20 shows that the EVM constituencies are different from the non-EVM constituencies on some observed dimensions. The EVM constituencies have less marginal workers than the non-EVM constituencies[5]. This may suggest that the EVM constituencies have a more formal industrial structure compared to the non-EVM constituencies. Similarly, we find that non-EVM constituencies are more rural compared to the EVM constituencies. The literacy rate was also higher in EVM constituencies. The literacy gap for the backward castes and general caste population is also different in EVM and non-EVM constituencies. The non-EVM constituencies have more agricultural laborers as part of the labor force than the EVM constituencies. The percent of agricultural laborers in the backward castes were also higher in the non-EVM constituencies. These demographic data are based on the 2001 Indian census, with 2001 being a post-EVM intervention year. However, this is the closest year for which non-interpolated census data is available with respect to the timeline of the introduction of EVMs. The demographic dimensions on which the EVM and non-EVM constituencies have been compared in Table 20 are least likely to be affected by the introduction of EVMs. Thus, theoretically, we can match the electoral constituencies along these observed characteristics to get a more balanced sample of EVM and non-EVM constituencies. For this, I have used the propensity score matching method. I provided a conceptual note on propensity score matching method in section 6.2 in the appendix. I have executed this matching process in statistical software Stata using a user-written matching algorithm 'pscore'(Becker & Ichino, 2002). In the matching exercise, the

[5]As per the definitions in Indian census, marginal workers are those workers who had not worked for the major part of the reference period (i.e. worked less than 6 months in a reference year)

probit model is used as the model of selection. The estimated common support region is [0.01040202, 0.82924745]. In the propensity score matched sample, I got total 238 parliamentary constituencies (out of which 32 are EVM constituencies as before). Now, we have non-EVM constituencies which are more similar in some observational dimensions to the EVM constituencies. I run the difference-in-differences specifications of section 2.6 with this matched sample and also with the parent sample of 486 constituencies. The falsification exercises are also done with respect to these samples. The main DID-results have been reported in Table 21. The falsification results are reported in Table 22 and Table 23. Table 21 report suggests that effect of EVM on voter turnout is still negative in the parent sample of 486 constituencies. The falsification exercise of Table 22 shows that there was no pre-existing differential trend between the EVM and non-EVM constituencies before the introduction of EVMs. The falsification exercise of Table 23 shows that the yearly DID-coefficient for 1999 was highest in value among all the yearly DID-coefficients, thus suggesting some effects of EVM on voter turnout in 1999. These results are very similar to the results from the DID-analysis done on the main sample of 543 constituencies. The DID-analysis for the matched sample of 238 constituencies shows that there is no substantively and statistically significant effect of EVM on voter turnout (column 2 of Table 21). The falsification exercises for the DID analysis on the matched sample shows that there was no significant differential trend between EVM and non-EVM constituencies before 1999 (second column in Table 22). The other falsification exercise in Table 23 shows that EVM and non-EVM constituencies were having similar differential trends over the years. Thus, I find no evidence of any significant effect of EVM on voter turnout in the matched sample of constituencies.

2.12.2 Analysis of Selected Northern Indian States

The results of the heterogeneity section and the limited subsample case show that the effect of EVM can differ across a different choice of samples. India is a large country with a highly diverse population. States in India significantly differ in terms of major languages spoken, political culture, and level of socio-economic development. Given this high diversity, it is highly likely that the effects of EVM can have a great variation among different regions of India. Understanding this regional variation of the effect of EVMs is very important to gain a comprehensive understanding of the effect of the EVMs on voter turnout behavior. The sub-regional differences in the effect of EVMs can be extremely useful for formulating more region-centric electoral policies for a large country like India. In this section, I analyze the effect of EVMs in some selected northern Indian states (Bihar, Haryana, Chhattisgarh, Himachal Pradesh, Jharkhand, Madhya Pradesh, Punjab, Rajasthan, and Uttar Pradesh). Uttar Pradesh is the largest state in India in terms of population and contributes the largest number of seats (80) to the national parliament. Punjab and Haryana are major agricultural producers of India. Jharkhand was historically part of the undivided Bihar state. Similarly, Chhattisgarh was part of the undivided Madhya Pradesh. Himachal Pradesh and Haryana were earlier parts of the undivided Indian Punjab state. There are a total of 208 parliamentary constituencies in this group of states, with 13 constituencies which received an EVM intervention in 1999. Thus, this region contributes a large share (nearly 38%) of total parliamentary constituencies of India. With these constituencies, I mapped the demographic characteristics from Jensenius (2015). Additionally, I get constituency level data on the dimension of infrastructure from Asher & Novosad (2017). This dataset has observations on state-level constituencies. I mapped the data for these state-level constituencies to the parliamentary constituency level by averaging over state assembly constituencies contained within a parliamentary constituency. Table 24 compares the EVM and non-EVM constituencies along the demographic and infrastructure variables. Overall, the EVM constituencies have more rural population than that of the non-EVM constituencies. The literacy rate is also higher in the EVM constituencies compared to the non-EVM constituencies. The mean distance to town(based on 1991's census data) was also lower in the case of EVM constituencies compared to non-EVM constituencies. The EVM constituencies also had a higher proportion of villages with electricity connections, paved roads and access to footpath compared to the non-EVM constituencies. I matched the EVM constituencies with the non-EVM constituencies on these observable demographic and infrastructure-related characteristics using the propensity score matching method. A probit model is used as the model of selection. The estimated common support region is [0.0340727, 0.78550932]. Discarding the unmatched non-EVM constituencies, we get a matched sample of total 60 constituencies with 13 EVM constituencies and 47 non-EVM constituencies. I then run the main DID analysis of section 2.6 in both the parent sample and matched sample for these selected northern Indian states. Table 25 reports the DID-estimation results. The results show that EVM interventions had a positive impact on voter turnout for both the parent sample and matched sample. The DID-coefficient is not statistically significant for the analysis with the parent sample. However, the DID-coefficient is statistically significant and substantially larger in the matched sample. The falsification exercise results are reported in Table 26 and Table 27. In both the falsification exercise results, the results give confidence to the DID-estimate for the matched sample. In the falsification exercise where 1998 has been assigned as the fake year of EVM introduction, the DID-coefficient is statistically insignificant and lower in value than the DID-coefficient for the main analysis in the matched sample. In the second falsification exercise results reported in Table 27, we find that the DID-coefficient for 1999 was positive (though statistically insignificant) compared to the negative and larger absolute values of the DID-coefficients of pre-1999 years. The reversal of the direction of the DID coefficient for 1999 strengthens the evidence in favor of the stronger positive effect of the EVM on voter turnout in these selected northern Indian states.

2.13 Conclusion

The introduction of the electronic voting machine is one of the most important electoral reforms in India. The empirical results show that there is a negative effect of EVMs on voter turnout. Voter turnout decreased by 2.81 percentage points in response to the EVM exposure. There is weak evidence that the total number of contestants decreased in response to EVMs. There is a positive effect of EVMs on the winning margin of the winner. On the other hand, EVMs have a negative effect on the vote share of the first losing candidate. The heterogeneity analysis suggests that richer states experienced a bigger fall in voter turnout with the introduction of EVMs. The empirical evidence further suggests that the fall in voter turnout has been mainly caused by the constituencies in big metropolises in India. The effect of EVMs on voter turnout requires tighter scrutiny. The DID-analysis on a propensity-score matched sample of data shows that there is no significant effect of EVMs on voter turnout on a larger subsample of the data. I then run similar matched difference-in-differences estimation with a sample of selected northern Indian states. In this case, I find that the EVM effect is positive and is significant both statistically and substantively. These results show that the EVM effect varies significantly over regions of India.

3 Super bloom in California's Death Valley and Voter Turnout: Evidence from a Natural Experiment

3.1 Introduction

What is the relationship between income and voter turnout? In the U.S., there is an observed puzzle regarding voter turnout (Brody, 1978; Filer et al., 1993). In the cross-section of a particular election year, voter turnout among the rich is higher than the voter turnout among the poor. However, voter turnout has declined with increased economic prosperity over the years. A similar declining turnout pattern over the years has also been observed for other advanced industrialized democracies (Gray & Caul, 2000). This poses a puzzle in the nature of the relationship between income and voter turnout behavior in the U.S. context. Does income have any causal impact on voter turnout behavior? It is not easy to find the answer to this question in an observational world. There are observable differences in the voting turnout behavior between rich and poor people in a given election year. However, this observable difference between rich and poor people cannot be interpreted as the causal impact of income. Rich and poor people may be different in many unobservable characteristics which cause the difference in voter turnout. Over time, voter turnout rate has been found to be decreasing with more economic development. However, this may not actually capture the effect of income on voter turnout because income generation over time is an endogenous process. The factors which are associated with the income generation process (e.g. education, ability, entrepreneurial spirit, tax and redistributive policy of the government) could also impact voter turnout. The lack of exogenous assignment of income is a big impediment in identifying the causal impact of income on voter turnout both in cross-sectional and longitudinal study contexts. In this paper, I attempt to identify the possible causal impact of income on voter turnout using a plausibly exogenous variation in the income generation process induced by a natural phenomenon in a county of California.

3.2 Related Literature

There are multiple strands of literature which directly or indirectly investigate the relationship between income and voter turnout. One strand of literature studies the impact of economic inequality on voter participation. The results of Solt (2008) suggest that income inequality reduces the political participation of people with lower income. Income inequality has a negative effect on voter turnout, although it increases the turnout of people in the higher income bracket (Solt, 2010). Redistribution policy of the political system can be an important determinant of variation in voter participation in the high-income group. Political participation among the rich will be high in a political system where they anticipate taxation (Kasara & Suryanarayan, 2015). Filer et al. (1993) find a negative relationship between voter turnout and real income in the U.S. This study finds that an increase in relative income has a significant impact on the turnout. Increase in relative income first reduces turnout and then increases it with further increases in relative income.

There is a strand of literature on the socioeconomic and resource model of voter participation (Brady et al., 1995; Verba & Nie, 1972). In this model, individual resources (time, money, skills) impact political participation. The individuals with low socioeconomic resources will be less likely to bear the cost of political participation. Therefore, individuals with a lower value of socioeconomic resources will be associated with a lower level of voter turnout. On the other hand, individuals with a higher level of socioeconomic resources will be associated with a higher level of political participation.

Another strand of literature investigates the effect of economic adversity on voter turnout. Overall, the evidence regarding the direction and magnitude of the relationship between the economic condition and voter turnout is mixed (Blais, 2006). Economic adversity can lead to mobilization of people for greater voter participation for expressing their grievances (Schlozman & Verba, 1979). This is one possibility (called the mobilization hypothesis). The other possibility is that stress, uncertainty, anxiety associated with economic hardship can lead to the withdrawal of people from the political participation process (withdrawal hypothesis) (Brody & Sniderman, 1977). There is empirical support for both of the hypotheses in the literature. Worsening economic condition stimulates people to greater political participation which increases turnout (Burden & Wichowsky, 2014). There is evidence in favor of negativity bias in voter behavior, i.e. voters often react disproportionately more to bad situations (Lau, 1982, 1985). Thus, voter turnout can increase in the time of economic difficulty. This effect may be conditioned by the individual characteristics of voters. The individuals who blame the government for economic adversity can mobilize with higher likelihood in times of economic downturns which can raise turnout levels (Arceneaux, 2003). Similarly, there is evidence that economic adversity can lead to lesser political participation and voter turnout (Rosenstone, 1982; Southwell, 1988; Verba et al., 1995). Radcliff (1992) finds that bad economic conditions can depress turnout in developed countries while it mobilizes people for higher political participation in developing countries. There is also another set of studies which find no impact of economic conditions on voter turnout (Arcelus & Meltzer, 1975; Fiorina, 1978). The literature on economic voting sheds light on the relation between the incumbent political party's vote share and the state of the economy (Alvarez & Nagler, 1995; Kinder et al., 1989; Nadeau & Lewis-Beck, 2001). However, this strand of literature does not address the relationship between the condition of the economy and the decision of voters regarding turnout in the election.

Most of the existing studies in the literature draw a conclusion on the basis of correlations; they lack a credible exogenous variation in the income generation process. Therefore, estimated results fail to provide the causal impact of income on voter turnout. This paper contributes to a small body of literature which analyzes the relationship between income/economic activity and voter turnout in a causal framework. Charles & Stephens Jr (2013) investigate the relation between labor market activity and voter turnout in the U.S. using county-level data between 1969 and 2000. The authors exploit variation in employment and wages caused by the exogenous shock in the international oil supply to identify the effect of labor activity on voter turnout. They find a reduction in voter turnout in gubernatorial, congressional, senatorial and state legislative elections in response to the increase in the local wages and employment. However, the effect on turnout in the case of the presidential election is statistically insignificant. Brunner et al. (2011) find that voter turnout decreases in response to positive labor market shocks in California. However, the results of this study pertain to voting in the ballot proposition measures and it does not consider the presidential election outcomes. In this way, my study contributes to understanding how an economic shock can affect political participation in California in the U.S. presidential elections. My study uncovers the negative effect of a positive economic shock on voter turnout in this regard. This contrasts with findings of the null effect of labor market shocks on voter turnout in the U.S. presidential elections in Charles & Stephens Jr (2013). Lunnerdal (2014) identifies the effect of income shock on voter turnout using state legislative election data in the period of 1959-1990 in India. This study uses the variation in rainfall to capture the exogenous variation in income and finds that negative income shock decreases voter turnout while positive income shock increases turnout. An experimental evidence of conditional cash transfer program in Mexico on voter turnout comes from De La & Ana (2013). The findings of this study suggest that the randomized assignment of cash transfer to the recipients have led to an increase in overall voter turnout and an increased vote share of the incumbent. The results of my study are in contrast with the findings of these two studies (De La & Ana, 2013; Lunnerdal, 2014). However, these two studies look at developing countries (Mexico and India respectively). This may explain the difference in the findings. However, a reexamination of the De La & Ana (2013) data by Imai et al. (in press) shows that there is no relationship between income and voter turnout measures. Program-specific components of conditional cash transfer programs can confound the estimate of the effect of income on voter participation (Akee et al., 2018). Thus, use of conditional cash transfer program for inferring the effect of income on voter turnout may be problematic. Bagues & Esteve-Volart (2016) find that incumbent's vote share increases at the provinces of Spain where the winners of the Spanish Christmas lottery are clustered. However, this study did not find any effect of the economic boom generated by lottery winners on provincial voter turnout. Thus, it is an interesting question to study whether similar results hold in case of the U.S.[6] My study finds a significant effect of income shock on voter turnout in the U.S. context. Akee et al. (2018) find that an unconditional cash transfer program in a Native Indian community in the U.S. has a positive intergenerational effect on the probability of children's voter turnout. However, this study does not find any effect of this unconditional cash transfer program on the parents' turnout behavior. The receipts of income through the conditional and unconditional cash transfer programs are more likely to generate a permanent change in the income stream of the individual recipients. The income shock generated by the super bloom is more likely to be a short-term shock because a super bloom is an extremely rare phenomenon. In the economic voting literature, there is evidence that voters reward an incumbent government more on permanent

[6] Doherty et al. (2006) find that lottery-winners in the U.S. express conservative opinions about estate taxes and government redistributive policies. In a similar vein, Ansell (2014) shows that homeowners experiencing an appreciation of house prices become less supportive of redistribution and social insurance policies of the government. However, these studies did not investigate the corresponding effect of exogenous income gain on voter participation behavior.

income gains compared to short-term cyclical gains in income (Alesina & Rosenthal, 1989; Fair, 1978; Kramer, 1971; Suzuki & Chappell Jr, 1996). However, these studies consider the effect of economic conditions on the incumbent's vote share or likelihood of support to political candidates. They do not examine whether economic conditions affect voter turnout. Most importantly, the income shocks (both permanent and short-term) can be attributable to government actions in the context of these studies. The income shock generated by a super bloom is not attributable to government actions. Thus, it is unlikely to affect a voter's evaluation of government performance in managing economic affairs. Any change in voter participation behavior that could accrue to this kind of income shock will not be confounded by voter's perception about management of the economy by government. Uncovering this kind of pure effect of short-term income shock will help better identify the factors which drive political behavior. In this way, it will be extremely useful for the formulation of relevant electoral policies. The super bloom phenomenon provides a unique type of income shock which is very different in some important dimensions from the other kinds of income shocks which can be potentially used for identification of the effect of income on voter turnout. Income shock from lottery win corresponds to more of a concentrated shock among a very few people. Lottery winners can be characteristically very different from an average voter (Brunner et al., 2011). Thus, inference from this kind of income shock can be problematic. Income shocks from government cash transfer programs and oil price fluctuations can possibly be associated with government actions. A super bloom shock is a spread-out shock in the sense that all the residents of a particular county have the equal likelihood to receive this shock a priori. But at the same time, it has no association with any government action. This unique characteristic of the income shock from a super bloom provide compelling reasons for a study on its effect on voter participation behavior. My study makes a contribution in this regard. My paper also makes a contribution for a causal understanding of the mobilization and withdrawal hypothesis regarding voter turnout. The existing literature mostly tested these hypotheses in the context of economic adversity. It is worthwhile to test these hypotheses in the context of economic expansion to broaden our understanding of the relationship between political participation and economic fluctuations. My paper makes a contribution in this direction and complements the work of Charles & Stephens Jr (2013) in this regard. The strength of the work of Charles & Stephens Jr (2013) lies in the coverage of all the U.S. counties in their study. My study is limited only to the state of California. This limitation in the coverage of my study is a source of both strength and weakness. The small coverage may limit the scope of external validity of the results. However, the coverage of only one state adds strength to the internal validity of the empirical design of my paper. It eliminates the bias in the estimate that could have come from the factors which vary among the states in the U.S. This may include various state-level institutional and socioeconomic confounders. Charles & Stephens Jr (2013) studied only one measure of voter turnout (turnout as a percentage of the population of age 20 or above). My study considers a broader set of voter participation outcomes like turnout as a percentage of total eligible voters and voter registration rate. It also examines turnout in finer categories such as turnout for mail voters and polling place voters both as a percentage of voting age population. Analysis of multiple electoral outcomes adds strength to my paper.

3.3 Theoretical Framework and Hypotheses

Here, I present the theoretical arguments linking economic activity and voter turnout from the perspectives of mobilization hypothesis (Schlozman & Verba, 1979) and withdrawal hypothesis (Brody & Sniderman, 1977). As per the mobilization hypothesis, bad economic outcomes mobilize people for a higher level of political participation. The natural corollary to this is that in times of higher economic activity, people will participate less in the political process. Thus, we will observe less voter turnout in times of higher economic activity caused by a super bloom. As per the withdrawal hypothesis, economic adversity will alienate people from the political process which will lead to lesser voter turnout. The regular inference from this hypothesis is that higher economic activity induced by super bloom will stimulate people to higher participation in the political process which will raise the voter turnout. I consider four electoral variables regarding electoral participation: voter turnout, voter registration rate, mail voter turnout, and polling voter turnout. Voter turnout measures are defined as a percentage of the total eligible voting population. Voter registration rate is defined as the number of registered voters as a percentage of the total eligible voting population.

Based on the above arguments, I propose the following hypotheses based on the theoretical expectations emanating from mobilization and withdrawal arguments regarding the effect of income shock on voter participation:

Hypothesis H_{2a}: Voter turnout as a percentage of the total eligible voting population will decrease in response to the super bloom.

Alternatively, H_{2b}: Voter turnout as a percentage of the total eligible voting population will increase in response to the super bloom.

Hypothesis H_{2c}: Voter registration rate will decrease in response to the super bloom.

Alternatively, H_{2d}: Voter registration rate will increase in response to the super bloom.

Hypothesis H_{2e}: Mail voter turnout as a percentage of the total eligible voting population will decrease in response to the super bloom.

Alternatively,H_{2f}: Mail voter turnout as a percentage of the total eligible voting population will increase in response to the super bloom.

Hypothesis H_{2g}: Polling place voter turnout as a percentage of the total eligible voting population will decrease in response to the super bloom.

Alternatively,H_{2h}: Polling place voter turnout as a percentage of the total eligible voting population will increase in response to the super bloom.

3.4 Super bloom in Death Valley, California

Death Valley in California is located at Inyo County which also borders Nevada state. Death Valley is a desert area along the Sierra Nevada mountain range and is one of the hottest zones in the U.S. In this mountainous desert valley, rainfall is extremely meager and highly irregular. If for some climatic idiosyncrasy, the rainfall crosses some high threshold, the seeds of the various wildflower plants which lie dormant for years waiting for sufficient rain, begin to sprout. In the ensuing spring, the entire valley becomes resplendent with the scenic beauty of brilliant bloom of various windflowers in a beautiful mountainous landscape. Death Valley National Park is a popular tourist attraction in California. However, super blooms attract additional tourists. A super bloom in Death Valley National Park is an extremely rare event. In recent history, it has occurred only three times (1998, 2005 and 2016). The 2016 bloom was the highest in scale while the 1998 and 2005 blooms were of a smaller magnitude. It is a great coincidence that this rare event of super bloom happened in the presidential election year of 2016. The super bloom provides a source of exogenous variation in the income generation process. Super bloom gives an exogenous boost to the tourism industry at the county level. Figure 5 plots the number of tourists from 1979 to 2017 visiting Death Valley National Park in the month March (the peak month for super bloom). It shows a very substantial increase in the number of tourists to Death Valley National Park in March 2016 compared to previous years. Figure 6 plots the annual number of tourists visiting the national park from 1979 to 2017. It shows that there is a great increase in the total number of tourists visiting the park in 2016 compared to previous years. The data of the tourist numbers are collected from the National Park Service(NPS) visitor statistics database. This substantial increase in the number of tourists is likely to provide a positive thrust to the tourism industry in Inyo County. Thus, super bloom can be associated with an exogenous increase in economic activity in Inyo County where the Death Valley National Park is located. We can identify the effect of income on voter turnout behavior by exploiting this exogenous variation in economic activity induced by the rare event of the super bloom.

3.5 Data

The electoral data at the county level has been collected from the website of the California Secretary of State. The electoral data includes county-level eligible to vote population, number of registered voters population, the total number of polling place voters and the total number of absentee voters(mail voters). There are 58 counties in California. Data have been collected for all the presidential elections from 1992 to 2016. Three counties (Alpine, Sierra, and Plumas) implemented all mail voters policy within this specified period of 1992 to 2016. I have excluded these three all mail-voters counties from the analysis. All the other counties have both polling place and mail voting facilities.

The demographic data have been collected from the U.S. Intercensal Data at the county level. The data include the proportion of White individuals, proportion of Hispanic individuals, proportion of African-American individuals, male-female population ratio, the proportion of individuals in the age group of 25-64 and proportion of individuals with age 65 and over. The source of data for the county-level median household income is the Small Area Income and Poverty Estimate(SAIPE) maintained by the U.S. Census Bureau. I have used the U.S. Department of Agriculture (USDA) Economic Research Service (ERS) education data for county-level education statistics, like the proportion of individuals with less than a high school education, the proportion of individuals with a high school education only, the proportion of individuals with some college or an associate's degree. This education data is based on decadal statistics assembled from the U.S. census data(1990, 2000) and the American Community Survey(ACS) five-year average data for 2012 to 2016 period. For 1992 and 1996, the education data in my dataset for this study corresponds to 1990 census figures. For 2000, 2004, and 2008, the figures correspond to the 2000 census data. For the years 2012 and 2016, the education statistics come from ACS five-year average data (2012-2016). These education statistics have been calculated for individuals in 25 or more age group. County-election year specific data on voting technology were collected from the California Secretary of State website.

Table 28 and 29 present the summary statistics of county-level electoral outcomes and covariates respectively comparing Inyo County with the rest of California's counties. The electoral outcome summary shows that the Inyo county experienced higher voter turnout compared to the rest of California counties over the time period of this study. Polling place voter turnout was also higher in Inyo County in comparison to the rest of California between 1992 and 2016. From the summary statistics of Table 29, we find that Inyo County had lesser median household income than the average of that in rest of California. The average voting age population in Inyo County was also much smaller to that of rest of California. Inyo County had a higher proportion of old age population (age 65 and over) compared to the rest of California. The county had a smaller proportion of people of Hispanic origin than the rest of California. The Inyo County had a higher proportion of people with only a high school degree compared to the rest of California. But the proportion of people with less than a high school education was much less in Inyo County compared to the counties in the rest of the California group. The proportion of times Inyo County used punch card voting technology is higher compared to the rest of California. The rest of California used optical scan voting technology at a higher proportion of times compared to Inyo County. It is to be noted that touch screen and automatic voting machine technology were not used in Inyo County. The use of these technologies in the rest of California counties is minimal. The older voting technology of automatic voting machine is used only once in a county in the rest of California in the time period of this study.

Figure 7 depicts a time series plot for the electoral outcome variables for Inyo County and the rest of California. Voter turnout as a percentage of eligible voter population was higher for Inyo County compared to rest of California for most of the years. In 2016, both the groups show the nearly similar level of voter turnout. Registration rate of voters was higher for Inyo County compared to the rest of California in the early years (2000, 2004 and 2008). However, it becomes much lower than the rest of California in 2016. Mail voter turnout as a percentage of eligible voter population declined in the super bloom county compared to rest of California in 2016. Polling place voter turnout as percentage eligible voting population was higher in Inyo County compared to the rest of California in all years except 2012. In 2012, the Inyo County and the rest of California had almost similar levels of polling place voter turnout.

3.6 Empirical Methodology

I have employed three different empirical methods to identify the impact of income on voter turnout. These three methods are difference-in-differences(DID), synthetic control, and nearest neighbor covariate matching. The DID results have been described in the main body of this essay. The method details and results corresponding to synthetic control and nearest neighbor covariate matching methods are described in the section 6.3 of the appendix. In this case, Inyo County had super bloom while the rest of the counties did not. If the super bloom is considered as a treatment, then Inyo County forms a treatment group and the rest of counties in California form the control group of this natural experiment. We have observations for both these groups for both before and after treatment periods. Hence, we can apply the difference-in-differences estimation method in this case (see section 6.1 in the appendix for a note on difference-in-differences estimation method).

3.6.1 Difference-in-Differences(DID)

I run the following Difference-in-Differences (DID) regression for county ‘c’ at time ‘t’.

$$Y_{ct} = \alpha_c + \delta_t + \beta \cdot Superbloom_{ct} + \gamma \cdot X_{ct} + u_{ct} \tag{7}$$

$Superbloom_{ct}$=1 if county c has super bloom in time t. It takes the value 0 otherwise. α_c denotes county fixed effects. The county fixed effects control for county-specific factors that are fixed over time. This includes the geographic area of the counties, any time-invariant county specific political culture (e.g. being a traditional bastion of liberal/conservative political ideology). δ_t denotes election year fixed effects. These year fixed effects control for the factors that vary over time but have common impacts over all the counties. This may include the effect of the individual charisma of presidential election candidates. The time fixed effects also control for impacts of any legislative or electoral law covering all the state. The time-fixed effects also control for any common trend in the evolution of the political culture of the state over time. It also takes care of any linear trend in the growth of the economy and demographic factors of the state as a whole over time. There are county-level control variables (X) which vary across counties and over time. u_{ct} is the error term. The coefficient of interest is β which gives us the causal effect of the super bloom on outcome Y. The super bloom acts as a proxy for economic boom/higher economic activity. Thus, β captures the effect of income on electoral outcomes. The unit of analysis is county-presidential election year. Following Bertrand et al. (2004), I cluster the standard error at the county level to account for serially correlated error structure. The number of clusters is fairly high at 55 which is helpful for the asymptotic convergence of the clustered standard error estimate in this case (Cameron et al., 2008).

The key identifying assumption of this framework is as follows: in the absence of a super bloom, β is statistically indistinguishable from zero. However, the counterfactual scenario of no-super bloom (the absence of super bloom) cannot be tested because it never happens. But, we can gather some evidence about how the counterfactual scenario (of the absence of the super bloom) could have evolved by looking at the pre-super bloom years. Here, β_1 is equal to the differences in mean of Y in super bloom county Inyo and 54 control counties in 2016 minus this differences before 2016 after controlling for Xs. The assumption is that in the absence of super bloom, the difference in these differences would be zero. In a later section, I will describe the falsification exercises that give confidence in this assumption.

The control variables (Xs) used in both the main and falsification DID exercises are: natural log of lagged median household income(annual), natural log of total eligible to register population, natural log of male-female population ratio, proportion of population in age group 25-64, proportion of population with age 65 and over, proportion of White population, proportion of Black population, proportion of Hispanic origin population, percent of population with less than a high school education, percent of population with a high school education, percent of population with some college or an associate's degree, and type of voting technology. There is evidence of the causal impact of education on political participation in the U.S. (A. K. Mayer, 2011; Sondheimer & Green, 2010). Political participation can vary by different racial groups in the U.S. (Leighley & Vedlitz, 1999; Uhlaner et al., 1989). There can be gender-specific differences in political participation (Schlozman et al., 1994; Verba et al., 1997). Age is also an important explanatory variable for voter turnout in advanced democracies. The observed relation between age and turnout is curvilinear in general. Voter turnout is lower in youth and old age, and higher in the adult phase of life (Smets & Van Ham, 2013; Bhatti & Hansen, 2012). Age is also correlated with voter registration in the U.S. (Ansolabehere et al., 2012). There is evidence of generational cohort effect on political participation (Lyons & Alexander, 2000). Hence, I control for these factors at the county level in the empirical framework of this paper. Voting technology has an impact on voter turnout behavior in the U.S. (Card & Moretti, 2007). However, it is less likely to be associated with behavior for voter registration (Card & Moretti, 2007). I control for voting technology for voter turnout measures.

3.6.2 DID-Falsification Exercise

I run two falsification exercises for the DID analysis mentioned in the above paragraph. In the first falsification exercise, I take only pre-super bloom period data (1992-2012) and assign 2012 to be the 'fake' super bloom treatment year. Then, I run the same DID-specification in equation 7 on this pre-super bloom period data. The DID-coefficient will capture the trend differential between Inyo County and the rest 54 control counties in pre-super bloom years. If this 'fake' 2012 super bloom coefficient is statistically indistinguishable from 0, then it will give confidence to the identifying assumption that in absence of the policy treatment and control would have zero differential trend.

In the second falsification exercise, I run the following DID specification.

$$Y_{ct} = \alpha_c + \delta_t + \sum_{t=1996}^{2016} \beta_t \cdot D_t \cdot Superbloom_c + \gamma \cdot X_{ct} + u_{ct} \tag{8}$$

In the above specification, D_t is the year specific dummy for the year t. I make 1992 the reference year. I take interaction of D_ts for each of the other years with the super bloom county dummy for Inyo County. Each of these interaction coefficients will estimate a DID coefficient with respect to a particular year and the reference year 1992. These year-specific DID-coefficients ($\beta_t s$) will provide information about the evolution path of the possible differential trend between the super bloom county Inyo and the other 54 counties. This information of year-specific differential trend can be used to assess whether there is sufficient evidence in favor of the causal impact of the super bloom on electoral outcomes.

3.7 Results

3.7.1 Difference-in-Differences(DID)

Table 30 reports the difference-in-differences estimation results for four electoral outcomes. The first is voter turnout as a percentage of the total voting eligible population (denoted in the table as poll-p). The second is the percentage of registered to vote individuals out of the total voting eligible population (denoted in the table as regis-vot). The third is mail voter turnout as a percentage of the total eligible voting population (denoted in the table as poll-p-absent). The fourth is the percentage of polling place voters out of the total voting eligible population (denoted in the table as poll-p-prec). The results show that there is a fall of voter turnout as a percentage of total voting age population by nearly 5.94 percentage points in the super bloom county(Inyo) compared to the other 54 counties. The falsification exercise results for this electoral outcome are reported in Table 31 and Table 32. Table 31 reports the falsification exercise where 2012 has been assigned as the fake super bloom year in Inyo County and the DID-estimation was done with data from 1992-2012 (excluding the year 2016 data). As per the falsification exercise reported in Table 31, there was already a pre-super bloom difference in trend between Inyo County and rest of the 54 counties in California. This casts doubt over the robustness of the DID-coefficient of Table 30 for the outcome variable (poll-p) as the causal estimate of the super bloom. However, it is to be noted that the super bloom DID-coefficient (approx. 5.94) is numerically larger than the DID-coefficient (approx. 3.62) in the falsification exercise. Thus, there is some evidence (though less robust) that super bloom reduced the voter turnout (defined as the percentage of eligible voting age population) in the affected county. However, the post-super bloom DID-coefficient(5.94) is an overestimate of the true effect of super bloom. The falsification exercise results of Table 32 is very informative on the year wise pre-2016 differential trend in super bloom and non-super bloom counties. With 1992 as the reference year, the DID-estimate of 2016 (the super bloom year) is highest in numerical value. The estimated value for 2016 (-6.58) is much larger than all other yearly DID-coefficients. The coefficient for 1996, 2000, 2004, 2008 are statistically insignificant and numerically much smaller than the DID-coefficient of 2016 which is the real super bloom year. The 2012 DID-coefficient is statistically significant but almost half in numerical value than the 2016 DID-coefficient. This gives evidence that the super bloom caused a decline in voter turnout as a percentage of voting age population. As a super bloom is a proxy for an economic boom, this implies that the economic boom reduced voter turnout. Thus, there is some support in favor of the hypothesis H_{2a} in this case.

The county-level income, demographic and education variables are jointly significant at 99% level of confidence(column 1 of 33). The main DID-regression has been estimated with other standard error methods (robust and jackknife). The main DID-coefficient for voter turnout is statistically significant for samples with outliers removed (column 4 of Table 34). Here, outlier observations are defined as observations for which corresponding studentized residual absolute values exceed 2. Table 34 reports the estimate of voter turnout with other methods of standard error estimation (heteroscedasticity consistent Huber-White sandwich robust standard error, jackknife resampling method). The statistical significance of the main-DID is robust to these alternative standard error estimation methods.

The DID-coefficient for registration percentage of the eligible voters also declined in the super bloom year (second column in Table 30). However, examining the falsification exercise estimates of Table 31, I find that there was already a very significant negative trend differential between the super bloom and non-super bloom counties before the actual super bloom happened. Table 32 shows that yearly DID-coefficient for 2016 is highest in numerical value compared to other yearly DID-coefficients and this 2016 DID-term is also statistically very significant. Thus, it is possible that the super bloom decreased the registration rate. However, the actual decrease possibly attributable to the super bloom will be smaller considering the negative pre-existing trend differential between the treated and control counties. Thus, we get some support in favor of hypothesis H_{2c}

The county-level income, demographic and education variables are jointly significant at 95% level for the main DID-estimation for voter registration rate(column 2 of 33). The main DID-regression for voter registration rate has been estimated with other standard error methods (robust and jackknife). The main DID-coefficient, in this case, is statistically significant for samples with outliers removed(column 4 of Table 35). Table 35 reports the estimate of voter registration rate with other methods of standard error estimation (heteroscedasticity consistent Huber-White robust standard error, jackknife resampling method). The DID-coefficient, in this case, retains statistical significance across these two alternative standard error estimation methods.

The mail voter turnout as a percentage of the voting eligible population declined plausibly due to super bloom (column 3 in Table 30). The falsification exercise estimates of Table 31 shows that this mail voter turnout outcome had a pre-existing positive differential of the trend between treated and control counties before 2016. Thus, there is a reversal of trend differential after the super bloom happened. This provides strong evidence that the super bloom (exogenous economic boom) decreased mail voter turnout. Looking at the falsification exercise involving yearly DID-estimates in Table 32, it is evident that the 2016 DID-coefficient is largest in numerical value compared to DID-coefficients of other years for both of the mail voter turnout outcomes. This provides some confidence in the main DID-coefficient estimate in this case. Thus, the empirical results provide support to the validity of hypothesis H_{2e}.

There are joint significance for the county level income, demographic and education variables at 99% confidence level for the main DID-estimation for mail voter turnout(column 3 of 33. The main DID-coefficient, in this case, is statistically significant for samples with outliers removed(column 4 of Table 36). Table 36 reports the estimate of voter registration rate with other methods of standard error estimation (heteroscedasticity consistent Huber-White sandwich robust standard error, jackknife resampling method). The DID-coefficient is statistically significant across the alternative standard error estimation methods.

The DID-coefficient for polling place voter turnout measure is not statistically significant (column 4 of Table 30). The year-wise DID-coefficients in the falsification exercise in Table 32 suggest there was already a very strong positive differential of the trend between the super bloom county and control counties in the pre-super bloom years. The magnitude of the 2016 DID-coefficient is very similar to some of the pre-super bloom years DID-coefficient. Hence, we can conclude that polling place voter turnout did not get affected by super bloom. Thus, hypotheses H_{2g} and H_{2h} do not get support from the empirical analysis. This suggests that the decline in aggregate voter turnout is caused mainly by the decline in mail voter turnout.

There are joint significance for the county level income, demographic and education variables at 99% confidence level for the main DID-estimation for mail voter turnout (column 4 of 33. The main DID-coefficient, in this case, is statistically insignificant and substantially low in magnitude for samples with outliers removed (column 4 of Table 37). Table 37 reports the estimate of polling place voter turnout with other methods of standard error estimation (Huber-White sandwich robust standard error, jackknife resampling method). The DID-coefficient is statistically insignificant across the alternative standard error estimation methods. The synthetic control and matching results (in the appendix) conform well with the DID-estimate as far as the direction of the effects is concerned.

3.7.2 Discussion and Conclusion

This study provides evidence in favor of a causal impact of income on voter turnout behavior. There is a fair amount of synergy across the results obtained through different empirical methods. The super bloom in a California county is used as a proxy of exogenous economic boom which causes exogenous variation in income in the affected Inyo County in 2016. The empirical results show that voter turnout as a percentage of eligible voting age population declined in response to this exogenously induced local economic boom. There is strong evidence that mail voter turnout as a percentage of the eligible voting population decreased due to this exogenous economic boom. There is no strong evidence of any change in polling place voter turnout due to the super bloom. As mail-voters constitute the bulk of total voters in 2016, a fall in mail voter turnout caused the decline in aggregate turnout. There is evidence of a moderate decline of voter registration rate in response to the super bloom. The empirical results of this paper overall support the theoretical expectation as per the mobilization hypothesis. According to mobilization hypothesis, voter turnout should increase in the case of economic adversity. Therefore, voter turnout should decrease in response to economic expansion. The overall voter turnout declined as per the empirical results of this paper which supports the mobilization hypotheses. One of the limitations of this study is that the results revolve around only one county in one state in the U.S. The limited coverage enhances the internal validity of the results. However, this limited coverage can have implications for the external validity of the results. The demographic features of the super bloom county Inyo are characterized by a thinly spread white-majority population. It also has a significant proportion of the 65 and over age population. There are numerous counties, particularly in the rural areas within the U.S., which share similar socioeconomic and demographic characteristics of Inyo County. Thus, the results of this paper could have significant external validity in the U.S. context.

4 Redistricting, Male-Female Elector Ratio and Provision of Public Schools: Evidence from India

4.1 Introduction

This paper assesses the impact of gender composition of the electorate on the pattern of public school provision in India at the administrative district level. India is the world's largest democracy and one of the fastest growing major economies of the world. Yet, in terms of school education outcomes, India lags behind the comparable developing countries (Kingdon, 2007). Despite maintaining a robust democratic system in a diverse polity for almost 70 years, the extent of political representation of women has been lower in India (Chhibber, 2002). It may be that these two seemingly puzzling outcomes in the Indian development story may not be uncorrelated. In a highly populous democracy with a high level of electoral competition, it may be possible that the structure of representation could impact the pattern of provision of publicly provided growth input such as school education.

Male and female voters are both important contributors towards electoral democracy. However, men and women can be behaviorally and organizationally different from each other[7]. This difference may emanate from the complex organizational form of a society they are embedded in. The interplay of male-female relations at different levels of life also adds complexity to this issue. Each gender can affect the other's political behavior through the relational structure between them. In a male-dominated patriarchal society, women's political behavior may be highly shaped (maybe dictated) by men's political preferences. Similarly, in matriarchal societies, men's political preferences may be shaped by women's political preferences. The

[7] as suggested by some experimental literature (Andersen et al., 2008; Hannagan & Larimer, 2010; Nowell & Tinkler, 1994)

male-female interpersonal relationship can also affect political preference, such as a husband's influence on his wife's political decisions, and vice versa. In some contexts, male and female political preferences may compete with each other. This can happen in emerging developing economies where men and women vie for the same jobs and aspirations. This interdependence, complementarity or competitiveness between men's and women's political and economic preferences are reflected in a representative electoral democracy. The relative number of male and female voters will shape the preferences of a society. The sociological, economic and political contexts of the society will endow different bargaining power structures to the different set of electorates (with the political dispensation), depending on their gender compositions. The political players are rational actors focused on increasing their probability of winning in elections. As rational actors, the political dispensations will respond to the evolving distribution of male-female voter dynamics. Political players are important mediators in the economic development process as suppliers of important inputs like rule of law, regulation, school education, etc. Thus, the male-female relative electoral distribution could have an impact on the resource allocation process of the economy. The purpose of this paper is to find the empirical evidence for the theoretical possibility of a causal impact of the male-female relative electoral distribution on the provision of public schools. I use 2008 redistricting of electoral constituencies of India as a source of exogenous variation in the male-female relative electoral distribution in administrative districts to identify its impact on public school provision.[8]

There is a strand of literature which studies the representation of women and its policy consequences which are linked to educational and other economic development outcomes (Beaman et al., 2012; Chattopadhyay & Duflo, 2004; Clots-Figueras, 2012; Duflo, 2012). Beaman et al. (2012) find that random assignment of female leadership in Indian local rural governing bodies leads to an increase in adolescent girls' career

[8] Redistricting has been used as a source of exogenous variation in Political Science literature (Ansolabehere et al., 2000; Fraga, 2016; Glazer & Robbins, 1985; Leveaux-Sharpe, 2001)

aspirations and educational attainment. Clots-Figueras (2012) estimates the causal effect of the gender of politicians in India on the educational attainments of individuals they represent. She finds that increased representation of women politicians in an administrative district increases primary school attainment of individuals in the urban areas. Chattopadhyay & Duflo (2004) examine the impact of the random assignment of one-third of the local government seats to women on the public good provision in rural India.[9] These authors find that leaders invest more in the public goods aligned to the preferences of their own genders. Ghani et al. (2014) find that reservation of positions for women in local government elected offices in India has a positive impact on women entrepreneurship by increasing the number of women-owned establishments in the informal sector. Thomas (1991) shows that states in the U.S. with a higher percentage of female legislators pass more bills dealing with issues of women and children. Similarly, Bratton & Ray (2002) find that provision of childcare increased in Norwegian municipalities in response to the increased representation of females in the elected offices.

All these studies concentrate on the gender identity of the legislators and its impact on different policy outcomes. The relationship between gender representation and policy outcomes is analyzed from a top-to-down angle of representation (i.e. from leadership to the electorate). These studies indicate that female representation in the legislator and political executive positions does have a positive impact on various policy outcomes, including education. Thus, the lower level of female representation may partially explain the starting puzzle why India has unimpressive elementary education outcomes compared to countries in the close comparison group. However, this may not fully explain the puzzle as a lower level of female representation is not specific only to India, but observed across different parts of the world (Rosen, 2013). This

[9] One third of local government (village, urban councils, municipalities) seats are reserved for female candidates in India. However, there is no reservation for women in state legislatures and the national parliament

leads our attention to the other dimension of representation- as an elector. Different political, economic and sociological factors can affect the representation problem of an elector as this elector is embedded in a system which is influenced by all these factors. However, these factors can have different natures of interaction with an elector, depending on the gender. This electorate-level, gender-centric differential can have an impact on the provision of any goods and services which are supplied, controlled or regulated by political players. This paper attempts to link this particular aspect of electorate level gender representation with the public school provision outcomes. Essentially, this study takes a bottom-up view of the gender representation problem and analyzes its impact on education outcomes in India. One set of literature shows that there is the validity of adopting this bottom-up perspective. Female participation as voters in politics can have substantial economic consequences. Participation of women in the political process after women suffrage reform in 1920 had a significant impact on the economic outcomes like the size of government spending in the U.S. (Lott & Kenny, 1999; Miller, 2008). These studies indicate that the gender composition of the electorate does have an effect on policy outcomes. However, these studies analyze the relationship in the context of a radical electoral reform. In the contemporary world, men and women have equal rights of voting in all the democracies. It is a worthwhile exercise to examine the relationship between gender representation and policy outcomes in the political environment of equal voting rights across genders. Any significant relationship, in this case, will indicate a more complex relationship between them beyond the realm of entitlements of gender-neutral voting rights. To the best of my knowledge, no other studies have attempted to investigate this aspect of the relationship between representation and education outcomes in the case of India. Thus, my study fills an important gap in the literature. My research also estimates the relationship at the administrative district-level, which is a smaller unit of aggregation compared to more conventional units like state or country. This will strengthen the understanding of the relationship between gender, electoral democracy and public provision of schools at the micro-level. Thus, the implications of this study can help in formulating public policies at the micro level.

4.2 2008 Delimitation (Redistricting) in India

India implemented delimitation of its electoral constituencies(at both the national and state levels) in 2008. This delimitation process was based on population figures of the 2001 census. The goal was to equalize representation across constituencies but the total number of electoral constituencies was kept unchanged both at the national and state level. The redistricting process has been largely politically neutral (Iyer & Reddy, 2013). My study deals with the state level constituencies. The state-level constituencies are contained within administrative districts. Because of the change of constituency boundaries, the electoral constituencies in the pre-redistricting period cannot be directly comparable to the electoral constituencies in the post-redistricting period. The time period of my study is 2001-2017.[10]. This period contains both the pre-redistricting phase and post-redistricting phase. In this period, the administrative district boundaries mostly remained unchanged. Thus, an administrative district provides an ideal unit of analysis for comparing pre and post-redistricting outcomes. Because of 2008 redistricting, the distribution of voters across constituencies changed within the administrative districts. This creates a quasi-experimental setup. The male-female elector ratio was impacted by the redistricting. Using this exogenous variation in male-female elector ratio through redistricting, I want to estimate whether male-female voter composition affects patterns of school provision and whether redistricting has accelerated or decelerated this process. The unit of analysis is the administrative district. It should be noted that redistricting process was de-

[10]Election years, however, limited to 2001-2012 period

lineated, however not instituted in a few states(Arunachal Pradesh, Assam, Jammu and Kashmir, Jharkhand, Manipur, Nagaland) so as not to alter the balance of power within these states (R. Bhavnani, 2018). I, therefore, excluded these states from my analysis. I also excluded three smaller states in northeast India (Tripura, Meghalaya, Mizoram), and excluded the union territories except the National Capital Territory (NCR) from the analysis as most of these union territories do not have elected legislative assemblies. The total population contribution of these excluded states and union territories in total national population is very small. The Tripura and Meghalaya election schedules were such that there were no post-redistricting elections with new electoral boundaries in the 2001-2012 period for these two states. Thus, they can validly be excluded from the dataset. The first state election with the new electoral constituency boundaries happened in Karnataka state in May 2008. All state elections after that also were conducted with the newly instituted constituency boundaries in the states where redistricting was implemented. However, depending on the election cycle, the first election with post-redistricting boundaries happened in different states in different years. This difference in the schedules of state assembly elections in different states was exogenously determined. Redistricting was in no way a determinant of the election schedule of the states. Thus, we find exogenous state-election year variation in the introduction ('activation') of new electoral constituency boundaries in different states. Table 39 shows the years of elections in different states with old electoral boundaries and new electoral boundaries given the 2008 delimitation.

4.3 Theoretical Framework and Hypotheses

Resource allocation process occurs in a political environment. The platform on which the political and economic forces interact with each other is characterized by competition among political actors and scarcity of economic resources. The allocation of the resources for economic development is affected by the objectives of the political players in the institutional setting. Political actors can affect the extent and pattern of resource allocation by manipulating the allocation of government-provided goods and services such as rule of law, roads, bridges, state-funded education and health, and state-controlled banking credit sector. The male-female electoral ratio in an electoral constituency can have an implication for the political actors if men and women have different patterns of political participation behavior. Political actors will align the resource allocation problem with their objective of winning elections. If male-female elector differences impact the probability of winning elections, then it can potentially also impact resource allocation process mediated through the politicians. Although this is a theoretical possibility, it will be difficult to empirically test the validity of it. In the observational world, male-female relative electoral distribution and the resource allocation problem evolve endogenously within the political-economic system. The resource allocation process can shape the male-female relative electoral distribution by impacting the migration rate of people across different regions of a country. The male-female relative electoral ratio can evolve due to demographic reasons of natural population growth. The population growth can affect the resource allocation process, and vice versa. Thus, identification of the causal impact of male-female relative electoral ratio on the allocation of resource inputs will be difficult in an observational world setting. The redistricting exercise in India in 2008 provides an ideal avenue to test the theory. Redistricting caused an exogenous change in the relative electoral ratio of women and men at the electoral constituency level. This translated into an exogenously induced change in the distribution of male-female relative electoral ratio at the administrative district level. After independence in 1947, India adopted a mixed-economy model of economic growth with a larger role of the state in the resource allocation process for economic development. India's Constitution puts education on the list of shared responsibilities between the center and the states. However, state governments are entitled to larger responsibilities in the provisioning of school level education, and Article 246 of the Indian Constitution endows the state governments substantial rights regarding school education policies (Clots-Figueras (2012)). Thus, it is appropriate in this study to consider representation at the state level political constituency.[11] The choice of the administrative district as the unit of analysis is relevant in this study because administrative decisions of the government are executed at the administrative district level in India (not at the electoral constituency level). Due to the large diversity of people in terms of language, culture among the states of India, the internal population migration rates between the states are low (Munshi & Rosenzweig, 2009). Therefore, variation in the composition of the male-female relative electoral ratio is not heavily impacted by the inter-state migration of people for residence. The sheer size of India's one billion plus population makes India a suitable case to study the relationship between male-female relative elector ratio and resource allocation in the area of school provision. The results based on the large population is more likely to be transferable to other contexts, thus increasing the external validity of results. I provide the argument as to why there can be a plausible connection between the electoral ratio of male-female voters and public school provisioning in India with the necessary support from the empirical research. In this context, I will first provide evidence that the process of provision of public goods in India is not devoid of the influences of political actors, and there are imperfections in the actual delivery mechanisms of public goods in India. Secondly, I will show that women are under-represented in the political, economic and social arenas of Indian polity. I connect these two arguments to the lower relative bargaining power of female voters to secure higher public good provision compared to male voters. Based on that, I derive the hypotheses of this study.

In India, there is considerable evidence that crucial inputs for economic devel-

[11]Public schools in Indian states are not provided as local public goods as in the U.S. There is no concept of school districts in the Indian context. Public schools are mostly a state level public good in practice. In this sense, a state is the most comparable counterpart of a school district in the Indian context.

opment are politically allocated. Cole (2009) finds that agricultural credit allocated through state-controlled banks follows an electoral cycle. Similar results have been found in the provisioning of electricity (Baskaran et al., 2015). Asher & Novosad (2017) present evidence of political impact on employment and stock returns of private firms through the regulatory discretion of the government. There is evidence that the local-level allocation of resources for economic development is prone to patronage politics, clientelism, corruption (Besley et al., 2011; Chandra, 2004; Wilkinson, 2006; Witsoe, 2012).

Male voter turnout is higher compared to female voter turnout in India, though this gap is decreasing over time (Kapoor & Ravi, 2014; P. Rai, 2011). Female participation is much lower in local politics in terms of attendance in public meetings (Chhibber, 2002; Prillaman, 2017). Gajwani & Zhang (2014) show that the female leaders in local governments are significantly less connected to higher government officials compared to their male counterparts in the Tamil Nadu state. This study shows that village local governments led by female leaders have a lesser provision of public goods like schools and roads since these two public goods need contact with higher ranked government officials. Women are under-represented at the legislative level of politics in India (Chhibber, 2002; Clots-Figueras, 2011; Jacob, 2014; Kapoor & Ravi, 2014; P. Rai, 2011). Female representation in state legislatures and the national parliament is low, around 10 percent (Beaman et al., 2010). The Economic Survey of India for 2017-18 reveals that only 9 percent of the total of 4118 members of state-level legislative assembly constituencies are women (Masoodi, 2018). As per a 2017 UN report, India ranks low, at 148th among 227 countries, in the representation of women in government and parliament, with 11.8 percent of female members in the lower house of parliament and 11 percent of female members in the upper house of the parliament (Masoodi, 2018). Jacob (2014) finds that legislative activity of women is significantly less than men in the national parliament. However, the same author also finds that the female legislators asked more questions related to women's issues compared to their male counterparts.[12] Similarly, Mishra (2000) shows that the legislative activities of the female legislators are lower compared to the male legislators in the state legislature of Orissa state. There is mixed evidence in favor of substantive representation of women in Indian state-level politics. The general category women legislators are not much supportive of women-friendly policies. It is the female legislators from reserved seats for backward castes who favor more women-friendly policies (Clots-Figueras, 2011).

The male literacy rate is higher than the female literacy rate in India (Katiyar, 2016; Sundaram & Vanneman, 2008). The labor force participation rate among women in India is much less compared to men (Das et al., 2015; Fletcher et al., 2017; Ghani et al., 2013). Indian society follows patriarchal social norms by and large. In such a societal environment, men are more likely to be politically active compared to women. Men in different capacities as voters, party workers, labor union members are in a better position than women to extract a favorable bargain from the political dispensation. The presence of various distortive practices such as patronage, clientelism, corruption in the actual delivery system of the public goods adds a higher premium to the bargaining ability of the electorate with the politicians. The political class in their pursuit of victory in elections will be induced to supply more public schools to the districts where the average electoral ratio favors the male voters. Thus, a constituency will have a higher likelihood of receiving a higher amount of public goods if it has more male voters. Based on the above argument, I propose the following three hypotheses.

[12] Female members of parliament ask fewer questions on economy and defense compared to their male counterparts. Female legislators ask more questions on women's issue than male legislators, often on the topics like dowry, surrogacy, crime against women etc. However, the proportion of questions related to women's issue is much lower in magnitude out of total volume of questions asked by members of parliament (Jacob, 2014)

H_{3a}: An administrative district with a higher mean male-female eligible elector ratio will be associated with a higher number of public schools.

H_{3b}: An administrative district with a higher median male-female eligible elector ratio will be associated with a higher number of public schools.

H_{3c}: An administrative district with a higher proportion of male elector dominated constituencies will be associated with a higher number of public schools.

There is no empirical evidence of gender differences in policy preferences for school education in India at the national level.[13] Chattopadhyay & Duflo (2004) find that men have a higher preference for educational investment from local government compared to women in the Indian states of West Bengal and Rajasthan. The authors also find gender differences in the preferences for public goods like roads and drinking water in these two states. However, the theoretical prediction of this section also holds with an identical preference for educational investment among men and women in India. It can hold even if the women have a more intense preference for the provision of education. The implicit assumption is that the bargaining power of the electorate is a more important instrument to secure public goods provision in the presence of imperfections in the public good delivery mechanism.

4.4 Data Description

4.4.1 School Data

The data on schools are collected from the District Information System for Education (DISE), which is maintained by National University of Educational Planning and

[13] There is evidence of gender-based differences in preferences for public goods in Africa (Gottlieb et al., 2018). There are gender differences in policy preferences in a developed country like the U.S. (Barnes & Cassese, 2017; Schlesinger & Heldman, 2001; Shapiro & Mahajan, 1986).

Administration(NUEPA) in India. This initiative is supported by the Ministry of Human Resource Development, Government of India. The government primary schools refer to the primary schools which are maintained by the Department of Education in state governments, the Tribal and Social Welfare departments, local government bodies, central government, and any other government management. The government schools do not charge any tuition fee from the student. The government is responsible for the salaries of the teachers and building infrastructure for the schools. Private schools include schools under the management of private bodies. In summary, schools in this study include five different categories: 1. Primary only 2. Primary with upper primary 3. Primary with upper primary, secondary and higher secondary 4. Upper primary school 5. Upper primary with secondary and higher secondary. The primary only category includes schools with grade 1 to 4 or 5. The primary with upper primary category includes schools with grades 1 to 7 or 8. The primary with upper primary, secondary and higher secondary category includes schools with grades 1 to 12. Upper primary category includes schools with grades 5 or 6 to grades 7 or 8. The upper primary with secondary and higher secondary schools category includes schools with grades 5 or 6 to grade 12. I have calculated an aggregate measure of schools which is the sum of the total number of schools in these five categories. The dependent variable of this study is the natural logarithm of this aggregate number of schools. DISE data are recorded at the academic year level which spread across two calendar years.

4.4.2 Electoral Data

I use repeated cross-section data comprising India's 19 states and the National Capital Territory (NCR) between 2001 and 2012. These states and the NCR comprise close to 90% of India's population. The source of the state-level election data is Election Commission of India. I have collected the relevant electoral information on state-level legislative elections which took place within the 2001 to 2012 period. I have collected the male and female eligible voter numbers of each of the electoral constituencies for the relevant state election years. In this dataset, I have retained only those states for which we can observe elections both pre and post-redistricting in 2001-2012. For two states (Tripura, Meghalaya), there were elections (2003, 2008) before the May 2008 redistricting process between 2001 and 2012. Therefore, I did not include these two states in the dataset. These states are very small in terms of total population in India. I collect the administrative district affiliation information for all the state level constituencies of a state both before and after redistricting. To get the administrative district affiliation information, I have relied on administrative information available on the state election commission websites of the relevant states. Information on district affiliations is also supplemented from the election-related information available at the website (www.myneta.info) maintained by the Association for Democratic Reforms(ADR). Some of the districts in the dataset have gone through a division to form new administrative districts. In some cases, new districts have been created from a single 'parent' district while in some other cases a new district has been created from multiple 'parent' districts. Among the districts which underwent division, I have retained only those districts which were created from a single 'parent' district. In such cases, an accurate estimation of the total number of schools is possible by adding up the corresponding numbers for the new 'child' district/districts in the DISE school dataset. The new 'child' districts with multiple parent districts and the corresponding 'parent' districts have been omitted from the study because in such cases accurate estimation of the total numbers of schools is not possible from the DISE data. The district-level database covers the total 468 administrative districts of India.[14] The data on cabinet membership of the elected representatives come from sources like administrative records in the state legislative assembly websites, and

[14]There were total 640 administrative districts in India as per 2011 census

newspaper reports of swearing-in ceremonies of the cabinet ministers in the leading national newspapers (Times of India, Hindustan Times, The Hindu, The Tribune). The data on the party affiliation of the elected representatives have been collected from the Election Commission of India database. The data on the gender of the election winner in an electoral constituency has been collected from the election database maintained by the Trivedi Center of Political Data of Ashoka University in India. The data on reservation status of the constituencies have been collected from the Election Commission of India database. The information regarding division of an administrative district and creation of new administrative districts has been obtained from the administrative documents available in the websites of Chief Electoral Officer(CEO) of the states, the Village Dynamics of South Asia database maintained by International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) and reports of district creation published in the leading English-language newspapers in India.

4.4.3 Main Independent Variable : Male-Female Relative Electoral Ratio

I have constructed three measures of the male-female relative electoral ratio at the administrative district level. The first is the mean male to female eligible elector ratio in an administrative district. I first calculate the ratio of male to female eligible voters in each of the state level electoral constituencies contained in an administrative district. Then, I take the mean of all these constituency level male-female elector ratios to get the administrative district level mean male-female elector ratio. Similarly, I calculate the median male-female elector ratio of an administrative district. In this case, I take the median value of male-female elector ratios for all the electoral constituencies in an administrative district. The third is the proportion of male electors dominated constituencies in an administrative district. A male-dominated electoral constituency refers to a constituency where eligible male voters are more in number than eligible female voters. The proportion variable is calculated by dividing the total male-dominated constituencies in an administrative district by the total number of electoral constituencies in that administrative district.

Table 40 describes the summary statistics of the dataset. On average, there are 1801 government schools and 480 private schools in an administrative district. The mean of male-female elector ratios of electoral constituencies in an administrative district is approximately 1.09 on average. The corresponding median value is also approximately 1.09 on average. This implies that the administrative districts are male-dominated on average. The proportion of male-dominated constituencies in an administrative district is approximately 81% on average. The total number of eligible voters in an administrative district is nearly 1.5 million on average. There are nearly 7 state-level electoral constituencies within an administrative district on average. Nearly a quarter of total constituencies in an administrative district falls into the reserved category on average (reserved to be contested only by backward caste candidates). The proportion of female winning candidates from an administrative district is small on average(close to 8 %). Nearly 63% of the winning candidates from an administrative district were associated with the ruling coalition at the state government level. On average, approximately 13% of the elected representatives from an administrative district served as ministers in the state government cabinet of ministers. It is true that the average of all the measures of the male-female relative electoral ratio is highly in favor of the male electors. However, this study uses only the within-variations in those measures to identify their effects on the education outcomes. If there are plausible exogenous within-variations in the electoral variables, one can still identify the causal effects even if the average values of them favor the male electors.

Table 41 results show how the male-female relative electoral ratio measures have been different in pre and post 2008 redistricting exercise. The regression results suggest that redistricting is associated with a statistically significant increase in all the three measures of male-female relative electoral ratio. The administrative district fixed effects remove the total between-variation among the administrative units. Therefore, the remaining raw variation is undoubtedly within-variation in the male-female relative electoral ratio at the administrative district level. However, this total within-variation may not be totally exogenously determined. Some part of it was determined endogenously. By controlling for the potential factors causing endogeneity, one can isolate the exogenous variation in the male-female relative electoral ratio caused by redistricting. Exploiting the isolated exogenous variation, one can possibly identify the causal impact of male-female relative electoral ratio on the provision of public schools. In this regard, I propose an identification strategy in the next section to estimate this causal impact.

4.5 Identification Strategy

I use the following equation 9 to estimate the effect of male-female relative electoral distribution on the public provision of schools. This follows closely to the identification strategy in R. Bhavnani (2016) where the effect of representation on economic development has been estimated using 2008 redistricting of India as the source of exogenous variation in the representation measure.

$$Y_{d,s,t+5} = \alpha_d + \delta_{t+5} + \beta \cdot MFER_{d,s,t} + \gamma_1 \cdot X_{d,s,t} + \gamma_2 \cdot X_{d,s,t+5} + \psi_s \cdot t + \psi_s + u_{d,s,t} \quad (9)$$

The dependent variable $Y_{d,s,t+5}$ denotes the natural log of the total number of public schools in administrative district d in state s at year $t+5$. The data on schools are reported at the academic year level. For example, if the election year is 2001, then the total number of public schools corresponds to the academic year 2005-06. This timeline closely approximates the electoral cycle of five years in the Indian states. α_d

are administrative district fixed effects and this control for any district characteristics which may vary among districts but are constant over time. These fixed effects may include district specific factors like attitude towards education, the legacy of being a historical education hub, the presence of big cities or metropolis within the boundary, or distance to the nearby big cities. All these factors can have an impact on the number of schools to be built in a district. δ are the time-fixed effects and they control for all the factors which affect all the administrative districts uniformly in a year. The time fixed effects also control for the impact of national education policies of the federal government like Sarva Siksha Abhijaan (Education for All Movement). This also takes care of the effect of macroeconomic shocks that affect the demand and supply of schools. The primary independent variable is $MFER_{d,s,t}$ which is the male-female relative electoral ratio at the administrative district level. There are three measures for it as described in the earlier section. X denotes the administrative district control variables. I have included the ratio of total male eligible voters and total female eligible voters in an election year in an administrative district. It helps to a great extent to isolate the part of the change in the district level male-female elector distribution variable($MFER_{d,s,t}$) that is exogenously changed by redistricting process (thus not because of the natural demographic factors-induced change). I have included the natural log of the total number of eligible voters in an administrative district on the eve of next state legislative assembly election(approximately at $t + 5$ period). The total number of eligible voters can be interpreted as a proxy for the total population of a district. If the total number of population for a district is high, the total number of schools should be higher as well. Hence, we need to control for this variable. Controlling for this also takes care of the possible reverse causality problem. Availability of schools may induce migration of people to different areas. Here, by controlling for the proxy of total population for an administrative district, we are taking care of the variation in the number of schools caused by possible school-availability induced migration. Empirical evidence shows that there is a growth of private schools throughout India (Kingdon, 2007, 2017). Availability of private schools can also affect the number of government schools in an administrative district. I control for the natural log of the total number of private primary schools comprising the same categories as their government counterparts at $t+5$ period. I also control for other possible confounding factors that can be related to the male-female distributional variable and can impact government allocation of development funds or level of economic development in an administrative district. Theoretically, we should expect a positive relationship between these development funds/level of economic development and the total number of government schools in a district. A higher level of economic development may attract new inhabitants which can increase the demand for schools in an area. The controls include the proportion of electoral winners from the ruling coalition in the state out of the total electoral constituencies in an administrative district. The ruling government is the executive authority in a state to allocate government funds to different areas within the state. The ruling government can manipulate this funds allocation process to maximize their chance of reelection. Being a part of ruling political dispensation, an electoral winner can possibly bring more funds to his/her district compared to the opposition party. Asher & Novosad (2017) find that state legislative constituencies aligned with ruling political parties experienced higher private sector job growth in India. Brollo & Nannicini (2012) show that ruling party deprives funds to the non-aligned municipalities in Brazil. Results of some studies (Arulampalam et al., 2009; Khemani, 2003) show that states which are ruled by parties aligned with the central government are likely to receive more funds from the central government in India. A similar pattern of majority favor in government funds transfers has been observed in the U.S. (Albouy, 2013; Ansolabehere & Snyder, 2006). Similarly, I also control for the proportion of elected representatives from the districts who were part of the state government cabinet as ministers. Membership in the ruling government's cabinet of ministers increases an elected representative's ability to secure more funds for his/her constituency (Denemark, 2000). Thus, more cabinet membership from an administrative district can increase the probability of receipt of more development funds from the government for that district. The other two factors I control at the district level are the proportion of reserved seats of backward castes in an administrative district and the proportion of electoral constituencies from an administrative district out of total seats in state legislative assembly. These two proportions have changed as part of the redistricting process. These are institutional factors which change with redistricting process and remain fixed at the set value till the next redistricting occurs. These variables can impact economic development by changing the degree of representation from a district or uplifting the backward castes to a higher development path. Chin & Prakash (2011) find a positive effect of reservation policy for scheduled tribes on economic development in India. R. Bhavnani (2016) shows that increased representation from an administrative district is positively associated with higher economic development in India. Similarly, a positive relationship between increased representation and higher government funds allocation has been observed in the U.S. and Argentina (Ansolabehere et al., 2002; Porto & Sanguinetti, 2001). There is evidence that women as elected leaders have a distinct impact compared to their male counterparts on education outcomes (Beaman et al., 2012; Clots-Figueras, 2011, 2012). Hence, I control for the proportion of seats in an administrative district won by female candidates. The process of division of old administrative districts and the creation of new administrative districts can have an impact on the provision of publicly provided inputs like schools. New schools may be sanctioned by the government to cater to the needs of a new district. I control for this by adding a dummy variable which takes the value of 1 for both the 'parent' district/districts and 'child' district and takes the value of 0 for districts which have not gone through any division within the period of 2001-2017. Lastly, I control for state-level linear trend($\psi_s \cdot t$). This linear trend takes care of the state level confounders which moves approximately linearly over time. It can include state-level demographic and economic factors like population growth, literacy rate, intra-state and inter-state population migration pattern over time and state-GDP growth rate. This can also control for impacts of education policy changes happening at the state level over a particular year. By construct, it also controls for state fixed effects (ψ_s). These state fixed effects may include factors like cultural importance attached to education in a state, men-women relative importance in society, political culture, the efficiency of state bureaucracy etc.

The strength of this identification strategy emanates from three factors: redistricting as a credible source of exogenous variation in male-female elector ratio distribution variable $MFER_{d,s,t}$, exogeneity of the electoral schedules of the states in the dataset and control for relevant confounding variables. The implicit identifying assumption is that exogenous variation in the variable $MFER_{d,s,t}$ has been isolated contingent on controlling for possible confounding factors which could have posed threats to identification. With the isolation of exogenous variation, the primary coefficient (β) will identify the causal impact of the male-female elector ratio distribution variable on the number of public schools.

Next, I use a modified version of equation 9 to estimate the effect of redistricting on the effect of male-female relative electoral ratio on the aggregate school provision measure.

$$\begin{aligned} Y_{d,s,t+5} = \alpha_d + \delta_{t+5} + \beta_1 MFER_{d,s,t} * RS_{s,t} + \beta_2 \cdot MFER_{d,s,t} + \beta_3 \cdot RS_{s,t} \\ + \gamma_1 \cdot X_{d,s,t} + \gamma_2 \cdot X_{d,s,t+5} + \psi_s \cdot t + \psi_s + u_{d,s,t} \end{aligned} \tag{10}$$

$RS_{s,t}$ is the redistricting dummy variable which takes the value 1 for any state-election year(s,t) after May 2008 and it assumes the value 0 otherwise. The magnitude and direction of the coefficient β_1 inform about how the effects of male-female relative elector ratio on school provision differ due to redistricting. The sum of the coefficients

$\beta_1 + \beta_2$ estimates the total effect of male-female relative electoral ratio on the school provision outcome in the post-redistricting periods.

4.6 Results

In this section, I report the regression estimation results of equation 9 on school provision outcome (natural log of the total number of government(public) schools). The regression has been estimated for three measures of the relative electoral distribution of male and female eligible voters within an administrative district. They are- the mean of male-female elector ratios of all the electoral constituencies within an administrative district, the median of male-female elector ratios of the electoral constituencies in an administrative district, and the ratio of male-dominated constituencies out of total electoral constituencies in an administrative district. Table 42 reports the effect of mean male-female electoral ratio on school provision outcome. The results suggest that there is a positive relationship between mean male-female electoral ratio with the total number of government primary schools. As the male-female elector ratio increases, the total number of government primary schools increases in an administrative district. The coefficient is both statistically and substantively significant. Column 1 to 3 report regression with different subsets of the explanatory variables and column 4 represents the specification with a full set of control variables. Column 1 of Table 48 reports the joint significance of the set of demographic and political explanatory variables. The variables are jointly significant at the 99 percent confidence level. The number of schools in an administrative district may not be independent over time. Thus, residuals could be non-independent for a district over time. To account for this kind of possible residual structure, I estimate the regression equation 9 with standard error clustered at the administrative district level. Table 43 reports the estimate of the regression equation 9 with standard error clustered at the administrative district level. The primary coefficient of interest retains significance in this case of standard error estimation method. The full specification regression coefficient is also positive and statistically significant for samples with outliers removed (column 3 of Table 49). The outlier observations are defined as observations with a corresponding studentized residual absolute value exceeding 2. Column 1 and 2 of Table 49 show the estimate of the regression equation 9 with a full set of control variables with different methods of standard error estimation(heteroscedasticity consistent Huber-White sandwich robust standard error, jackknife resampling method). The primary coefficient is statistically significant in the case of the robust standard error estimation. However, it is not statistically significant in the case of jackknife standard error estimation. Thus, overall there is support for hypothesis H_{3a} in this study. Table 51 reports the estimation results of regression equation 10. The interaction term between redistricting dummy and male-female relative elector ratio is negative and statistically significant in all specifications in Table 51. The statistical significance also holds true when standard errors are clustered at the administrative district level (Table 52). The sum of the interaction term and the coefficient of the mean female-male relative elector is positive in all the specifications. However, the total magnitude of the effect of the mean male-female elector ratio reduced in the post-redistricting phase. Table 44 reports the regression estimates of the effect of median male-female elector ratio on the natural log of the total number of public schools. The coefficient of the median male-female relative electoral ratio is statistically insignificant in all the four regression specifications in Table 44. However, the coefficients are positive in magnitude. This coefficient is also statistically insignificant in the specifications with standard error clustered at the administrative district level(Table 45). However, there may be the substantive significance of the primary coefficient. The positive value of the primary coefficient with possible substantive significance indicates some weak support for the hypothesis H_{3b}. The primary variable of interest has joint significance with other political and demographic variables (column 2 of Table 48). The interaction term between the median male-female elector ratio and redistricting dummy is negative for all four regression specifications. The sum of the coefficients of the median male-female elector ratio and interaction terms are negative. Thus, the insignificant positive effect of median male-female relative electoral ratio on public school provision reversed into a negative effect in the post-redistricting periods. The regression results of Table 46 suggest that the proportion of male-dominated constituencies in the administrative district has a statistically significant positive effect on the total number of public schools. This holds true for all four regression specifications with different sets of explanatory variables in Table 46. The coefficient of interest is also statistically significant in all the regression specifications with standard error clustered at the administrative district level(Table 47). There is the joint significance of the primary variable of interest with other demographic and political variables (column 3 of Table 48). Table 50 reports the estimate of regression equation 9 with full set of control variables with different methods of standard error adjustment. The statistical significance of the primary coefficient of interest is robust with respect to other different standard error estimation methods(robust standard error and jackknife). The full specification regression coefficient is also robust with samples with outliers removed(column 3 of Table 50). Thus, the results of this study support the hypothesis H_{3c}. Table 55 regression results show that the interaction term between the male-dominated seat ratio and redistricting dummy is statistically significant. The sign of the interaction term is negative for all four specifications in column 1-4. The sum of the interaction coefficient and male-dominated seat ratio coefficient is positive and close to zero for all these specifications. This shows that the positive effect of male-female relative electoral ratio on public school provision reduced to a large extent in the post-redistricting periods.

4.7 Discussion and Conclusion

The empirical results broadly support the hypotheses derived from the theoretical framework of this paper. The results reveal one particular pattern of public provision of schools in India. An increase in the male-female electoral ratio led to an increase in the public provision of schools in an administrative district. This implies that there is a positive relationship between male-female elector ratio and public provision of schools at the administrative district level. It indicates a form of male bias in the pattern of provision of public schools in India. However, the results of the paper also indicate that this bias has weakened in the post-redistricting phase. This particular pattern of public school provision can have implications on the nature and characteristics of the economic development process. Female-dominated electoral constituencies can have worse educational outcomes which may in turn negatively affect overall educational achievement of India significantly. The empirical results of this paper contribute to the understanding of the connection between two puzzling outcomes of India (relatively lower educational outcomes and skewed gender representation in the political system). Education as a mode of human capital formation acts as crucial input for economic development (Hanushek & Kimko, 2000; Lucas Jr, 1988; Mankiw et al., 1992; Romer, 1990). There is empirical evidence that school education is positively associated with economic growth outcomes in India (Self & Grabowski, 2004). Given the evidence of a male-bias in the provision of this important input of economic growth, the findings can have implications for a pattern of economic growth in India. Male-dominated electoral constituencies can experience different trajectories of economic growth compared to female-dominated constituencies. This, in turn, can affect the overall economic growth pattern of India.

5 Conclusions and Policy Implications

The findings of this dissertation provide a causal understanding of the topics of voting technology, voter participation, and the provision of public schools. The results of the first essay suggest that there is a possible negative effect of electronic voting technology on voter turnout in India. However, a more detailed analysis suggests that the apparent decline in voter turnout is concentrated in the metropolitan constituencies of India. This necessitates electoral policy focus on the major urban centers of India. A more innovative approach is needed for the urban centers of India to increase the voter turnout. The empirical results indicate a substantial variation in the effect of EVMs on voter turnout at the regional level in India. The EVM effect on voter turnout was positive in a group of states in northern India. The findings of the first essay also suggest that vote share of the lower ranked candidates increased in response to EVM introduction. The number of candidates contesting in an electoral constituency also declined in response to EVM introduction. It may indicate that EVM introduction encouraged more efficient candidate selection in the political process.

The empirical results of the second essay suggest that income has a negative causal impact on voter turnout. This finding conforms well to the observed phenomenon of declining voter turnout in the U.S. and other advanced industrialized countries with a higher level of economic development over the years. The negative causal effect of income on voter turnout may be detrimental to the proper functioning of electoral democracy. It indicates that economic development can potentially reduce voter participation in a democracy. However, causal identification of this effect is beneficial for proper policy formulation to ensure higher political participation. A high level of economic development and a high level of political participation are important attributes of a robust democracy. However, high income itself may not guarantee higher voter participation. This calls for policies which strive to build awareness about the civic duty of voting. There is a need to instill the civic duty of voting in the minds of the electorate. Increases in the civic duty of voting can increase voter turnout in advanced industrial countries by offsetting the possible negative effect of income on voter turnout.

The findings of the third essay suggest that there is a positive relationship between male-female elector ratio and the total number of public schools in an administrative district in India. This indicates that the provision of public schools has been higher in the administrative districts where the average electoral ratio favors male voters. This male-bias can have serious consequences for the future path of economic growth in India. The results of this essay can contribute to the current major policy debate in India over implementing reservation of one-third of the seats in state legislatures and the national parliament for women candidates. There is empirical evidence from existing literature that female legislators are associated with higher educational attainment of the electorate in India (Clots-Figueras, 2012). My paper shows that there can be under-provision of public schools in districts where the average electoral ratio favors the female voters. Thus, one of the possible policy suggestions can be that reserving a set number of positions for female candidates should be targeted to female elector dominated constituencies. This may ameliorate the imbalance in the educational investment outcomes through the channel of legislative activity of women elected representatives. Thus, the contour of the policy debate in India about women reservation in legislatures should not be shaped by the idea of arbitrary reservation. Rather, the direction of the debate should be more about targeted reservation to a particular type of constituencies. A targeted reservation policy for women will not only increase the descriptive representation of women but also it can lead to an improvement in the educational policy outcomes in India.

6 Appendix

6.1 Difference-in-Differences

Suppose there are two groups for which we have observations for 2 periods. The first group gets a treatment T at period 2. The other group does not get treatment (which is termed as a control group). Let the pre-intervention outcome for the treatment group be denoted by Y_{T1} and the post-treatment outcome for this group be denoted Y_{T2}. The change in the outcome in pre-post intervention periods for the treatment group is $\Delta Y_T = Y_{T2} - Y_{T1}$. Similarly, let the pre-intervention outcome for the control group be denoted by Y_{C1} and the post-treatment outcome for this group be denoted by Y_{C2}. The change in the outcome in the pre-post periods for the control group is $\Delta Y_C = Y_{C2} - Y_{C1}$. The difference-in-difference(DID) estimator is $DID = \Delta Y_T - \Delta Y_C$. The DID can be estimated with the following regression equation (Cameron & Trivedi, 2005).

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 P_{it} + \beta_3 (T_{it} \cdot P_i) + \epsilon_{it}$$

Here Y_{it} is the outcome variable for unit i at time period t. ϵ_{it} is the error term. P_{it} is equal to 1 for observations in the post-treatment period. In pre-treatment period, this variable assumes the value of 0. T_{it}=1 for observations in the treatment group. This variable takes the value of 0 for observations in the control group. Here β_3 is the difference-in-difference estimate of the treatment effect.

6.2 Propensity Score Matching

The objective of the propensity score matching method is to match treated and untreated observations on the basis of the propensity score. Suppose, Y_{1i} denotes the outcome for a unit i which receives a treatment T. Y_{0i} denotes the potential outcome for a unit i for the situation of non-receipt of the treatment. The propensity score is the estimated conditional probability of being in the treatment group. The propensity score is defined by $P(X) = Pr(T = 1|X)$ where X is a covariate vector. Suppose, potential outcomes are independent of treatment status conditional on the observable multivariate covariate vector (X) and a common support exists ($0 < P(X) = Pr(T = 1|X) < 1$ for X). With the conditional independence and common support assumptions, if $(Y_{1i}, Y_{0i}) \perp T|X$ then $(Y_{1i}, Y_{0i}) \perp T|p(X)$ as per propensity score theorem by Rosenbaum & Rubin (1983). This propensity score theorem implies that if treatment assignment is unconfounded given covariate vector X, then it is unconfounded given p(X) also (Rosenbaum & Rubin, 1983). Thus, it is possible to match treated and untreated observations on the basis of propensity score (see Angrist & Pischke (2008) for more details).

6.3 Additional Methods and Results for Essay 2

6.3.1 Synthetic Control

In the synthetic control method, the idea is to create an artificial counterfactual unit for the actual treated unit by a convex combination of the control units (Abadie & Gardeazabal, 2003; Abadie et al., 2010). The underlying principle for finding the optimal weights for this convex combination is to minimize the difference between treated and untreated units in observed covariates and outcomes over the pretreatment periods. Let Y be the outcome variable of interest. Let us denote Y_{ct}^{N} as the value of Y in the unit (county) c at time t in the absence of intervention(counterfactual scenario) and Y_{ct}^I as its equivalent when intervention does take place. Here, our aim is to obtain an estimate of the effect of the policy in the treated unit , i.e. to obtain an estimate of $\alpha_{ct} = Y_{it}^I - Y_{ct}^N$

Suppose, potential outcome in absence of treatment (the counterfactual outcome) is generated through the following linear data-generating process[15]

$$Y_{ct}^N = X_{ct}\beta + \lambda_t\mu_c + \delta_t + \epsilon_{ct} \tag{11}$$

X_{ct} represents observed time-varying covariates, μ_i denotes unobserved time-fixed factors which do not vary across counties but vary over time. δ_t denotes the year effects which are common across all the counties. The term ϵ_{ct} represents idiosyncratic shock exogenous to the system. Suppose, there is a sample of total $J+1$ counties. Without loss of generality, we can assume that the first county among these $J+1$ counties has got super bloom intervention in period T_0. Define the $J \times 1$ weighting vector W $= (w_2, w_3, ..., w_{J+1})$ such that $w_j \geq 0$ for all j $\in$ (2,....J+1) and $\Sigma_{j=2}^{J+1} w_j = 1$. Let m denotes the total number of covariates and lagged outcome variables. I define Z_0 as the $m \times 1$ vector of observed covariates for the treated unit. Similarly, let us define Z_1 as the covariate vector matrix of order $m \times J$ for the untreated units. The product Z_1W denotes a weighted average of the pre-intervention variables vectors for all counties except the super bloom county of Inyo. The difference between Inyo County and the other 54 control counties is $Z_0 - Z_1W$. The optimal weight matrix W^* is chosen such that $W^* = \text{argmin}_w (Z_0 - Z_1W)'V(Z_0 - Z_1W)$ subject to the weights in W being positive and summing to 1. Here V is a $m \times m$ symmetric and positive semi-definite matrix which weighs the relative contributions of all the explanatory variables. The optimal V matrix is selected to minimize the mean-squared prediction error for pre-treatment outcomes. The optimal weight matrix W^* is used to create a synthetic control unit from the untreated units(donor pool) which approximate the observed

[15]See Abadie et al. (2010); O'Neill et al. (2016)

covariates and pretreatment outcomes corresponding to the treated unit. The underlying idea is to use the synthetic control weight vector to estimate the counterfactual for the treated unit as a linear combination of realized outcomes and covariates in the potential control units, i.e. we construct the estimate for the counterfactual as:

$$Y_{1t}^{N} = \Sigma_{c=2}^{J+1} w_c^* Y_{ct} \text{ for } t \geq T_0$$

Under the assumption of a linear process of data-generating (the way counterfactual data are constructed) and if the number of pre-intervention periods is large compared with respect to the idiosyncratic shock (ϵ_{ct}), then $\alpha_{1t}^* = Y_{1t}^I - Y_{1t}^N$ approximates the average treatment on treated (ATT) effect of the intervention.

The predictor variables (Xs) used in the main synthetic control exercise for electoral outcome[i] are: natural log of lagged median household income(annual), natural log of total eligible to register population, natural log of male-female population ratio, proportion of population in age group 25-64, proportion of population with age 65 and over, proportion of White population, proportion of Black population, proportion of Hispanic origin population, percent of population with less than a high school education, percent of population with a high school education, percent of population with some college or an associate's degree, type of voting technology, electoral outcome[i](1992), electoral outcome[i](2000), electoral outcome[i](2012). I have used 'synth' user-written algorithm for Stata statistical software (Abadie et al., 2014) to estimate the synthetic control results of this paper.

6.3.2 Synthetic Control Placebo Exercise

In order to gain confidence for the estimate of ATT in the above exercise, we can run the same synthetic control exercise in the pre-intervention (i.e. pre-super bloom) years of data. In this study, this implies running the synthetic control exercise for

1992 to 2012 period. The lagged electoral outcome variables, in this case, need to be adjusted in relation to the new shorter year range of the dataset. If in the pre-intervention years, the treated unit and its synthetic unit have a nearly identical path of electoral outcomes over time, then it gives confidence to any possible effect accrued due to the super bloom intervention in the main synthetic control analysis with full dataset. Similar placebo test has been used by Abadie et al. (2015) which studies the effect of German unification of 1990 on the West German economy.

The predictor variables (Xs) used in placebo synthetic control exercise for electoral outcome[i] are: natural log of lagged median household income(annual), natural log of total eligible to register population, natural log of male-female population ratio, proportion of population in age group 25-64, proportion of population with age 65 and over, proportion of White population, proportion of Black population, proportion of Hispanic origin population, percent of population with less than a high school education, percent of population with a high school education, percent of population with some college or an associate's degree, type of voting technology, electoral outcome[i](1992), electoral outcome[i](2000), electoral outcome[i](2008).

6.3.3 Synthetic Control Results

In the synthetic control method, the visual inspections of the graphs of the estimates are a way of inference. The synthetic control algorithm, in this case, creates a synthetic Inyo County(the super bloom county) from the donor pool of other 54 California counties used in this study. The electoral outcomes are then estimated and plotted together for both actual Inyo observed results and counterfactual construct of the outcomes for the synthetic Inyo. Figure 8 shows the synthetic control outcomes for voter turnout as a percentage of the total voting eligible population. The dotted line is the electoral outcome plot for the synthetic Inyo County. Figure 8 shows that the actual Inyo and synthetically constructed Inyo followed almost the same path till

2012. In 2016, a sharp deviation emerged between the actual and synthetic Inyo. The observed voter turnout as a percentage of the total voting eligible population is lower than that of the synthetic Inyo. It is to be noted that there is no super bloom in the synthetic Inyo. Thus, this fall of voter turnout in actual Inyo may be the result of super bloom which created an exogenous economic boom in the county. However, this deviation may happen because of some random chances. Thus, the placebo test results should also be analyzed in this context. The second graph in Figure 8 panel shows the placebo test outcomes graph. The placebo test is conducted for the time period before super bloom (1992-2012) and with this pre-2016 dataset, the main synthetic control algorithm has been conducted with keeping the same covariates and appropriately adjusting the past outcome variables with the reduced time period data. Similar placebo study has been done with respect to other electoral outcomes also. In the placebo study graph of Figure 8 panel, it is evident that evolution path of the electoral outcome variable (voter turnout as a percentage of total voting age population) is almost the same for actual and synthetic Inyo County. This gives further credence to the claim from the main synthetic control analysis that super bloom in 2016 caused the decline in voter turnout in Inyo County.

Figure 9 plots the synthetic control estimates for the registration percentage out of the total voting eligible population. The plot shows that there is a very large decline in registration percentage in 2016. Before 2016(the super bloom year), the actual Inyo and synthetic Inyo follow nearly similar outcome path over time. The second graph in the panel of Figure 9 plots the placebo results with dataset excluding 2016. It shows that there would have been a decline in registration percentage in Inyo County compared to synthetic Inyo in the year 2012. However, the size of this decline in the placebo analysis is much smaller than that of the main analysis. Thus, there is some evidence (although not robust) that super bloom reduced registration percentage in the Inyo County in 2016.

Figure 10 plots the synthetic control estimates of the mail voter turnout as a percentage of the total eligible voting age population. It is evident from the plot that there has been a sharp decline in mail voters turnout percentage in 2016. In the pre-2016 periods, the actual Inyo and synthetic Inyo electoral outcome paths were almost identical to each other. The placebo study with pre-2016 shows that actual Inyo and synthetic Inyo had almost identical mail voter percentage in absence of the super bloom. This gives credence to the argument that super bloom has a negative causal effect on the mail voter turnout as a percentage of total eligible voting age population.

The synthetic control estimates for polling place voter turnout as a percentage of the total voting eligible population has been presented in Figure 11. It shows that there has been an increase in polling place voter turnout in 2016 in actual Inyo County in comparison to its synthetic counterpart. The evolution path of the electoral outcomes for the actual and synthetic Inyo was almost identical in the periods before 2016. However, the placebo study plot shows that there was a positive trend differential in the year 2012 between actual and synthetic Inyo. This differential is nearly the same in magnitude as that from main analysis with the year 2016 included. Thus, the 2016 effect cannot be interpreted as the true effect of super bloom on the polling place voter turnout as a percentage of total eligible voting age population.

6.4 Nearest Neighbor Covariate Matching

In the nearest neighbor covariate matching method, the basic idea is to match treated unit with one or more control units which are closest to the treated unit in terms of observed covariate measures. The control unit/units closest in terms of observed covariates capture the counterfactual unit which is observationally equivalent to the treated unit except for the receipt of treatment. Thus, the difference of outcome between the actual treated unit and the matched unit/units will isolate the effect of the treatment/intervention on the treated unit. The closeness of observed covariates is defined in terms of the normalized Euclidean distance matrix. Suppose,the matching covariate vector is $X = (X_1, ...X_k)'$. The normalized Euclidean distance matrix between ith and jth unit can be expressed as following

$$\|X_i - X_j\| = \sqrt{(X_i - X_j)'\hat{P}(X_i - X_j)}$$

where

$$\hat{P} = \begin{pmatrix} \hat{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \hat{\sigma_2^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\sigma_k^2} \end{pmatrix}$$

Here, $\hat{\sigma_k^2}$ is the sample estimate of the variance of k-th covariate. So, the normalized Euclidean distance can be expressed as following

$$\|X_i - X_j\| = \sqrt{\sum_{h=1}^{k} \frac{(X_{hi} - X_{hj})^2}{\hat{\sigma_h^2}}}$$

The average treatment on the treated effect (ATT) can be computed as following

$$\hat{\alpha}_{ATT} = \frac{1}{N_t} \sum_{T_i=1} \Big[Y_i - \Big(\frac{1}{M} \sum_{m=1}^{M} Y_{j_m(i)} \Big) \Big]$$

N_t is the number of treated units which is one in this case. $T_i = 1$ refers to the receipt of treatment by ith unit. Y_i is the observed outcome for the treated unit i and the corresponding covariate is X_i. $Y_{j_m(i)}$ denotes the outcome for a non-treated unit for which the corresponding covariate $X_{j(i)}$ is closest in Euclidian distance to covariate X_i for the treated unit. I have estimated the ATT by setting M=2. This means the treated (super bloom county Inyo) has been matched with two counties from the control counties which are closest in the distance calculated from the observed covariates (which are the first and second closest in normalized Euclidean distance from the intervention county).

The covariates used in the matching process are: natural log of median household income(annual), natural log of total eligible to register population, natural log of male-female population ratio, proportion of population in age group 25-64, proportion of population with age 65 or more, proportion of White population, proportion of Black population, proportion of Hispanic origin population, percent of population with less than a high school education, percent of population with a high school education, percent of population with some college or an associate degree, type of voting technology. The covariates (except the educational variables) are for the year 2015. The education covariates are 5 year average of 2012-2016 (Source is the American Community Survey). The electoral outcome variables are for the year 2016. The educational variables contain some post-super bloom period data. Thus, they contain some post-treatment effect in a strictly theoretical sense. However, from a practical point of view, education variables are least likely to get affected due to super bloom within just a single year. Hence, these variables can be used for valid covariates for matching purpose. I have used 'nnmacth' user-written algorithm for Stata statistical software (Abadie et al., 2004) to estimate the nearest neighbor covariate matching results of this paper.

6.4.1 Results: Nearest Neighbor Covariate Matching

In the nearest neighbor covariate matching exercise for this study, the purpose is to find counties which are very close to the super bloom county in observed covariates. The matching algorithm, in this case, chooses two counties from the 54 in-super bloom counties in the control group. Then, the difference between the electoral outcome of Inyo and the average of the outcome for the two nearest matched counties will pick-up the effect of super bloom on the treated unit Inyo (average treatment on treated effect). The electoral outcomes are of the super bloom year (2016). Under the important assumption that there are no unobservable confounders after controlling for the all the possible observable covariates, the estimated ATT, in this case, will capture the true effect of super bloom on the treated unit Inyo County. Table 38 reports the estimates of the nearest neighbor covariate matching. The matching estimate of Table 38 shows that the voter turnout as a percentage of the total voting age population decreased by 0.7 percentage points in Inyo County because of super bloom. However, this effect is not statistically significant. Voter registration percentage declined in Inyo County by nearly 3 percentage points. However, this decline is not statistically significant. The mail voter turnout as a percentage of the total eligible voting population experienced a high decline of almost 11 percentage points in Inyo county compared to the matched counties. This negative effect of super bloom is highly significant statistically. We find that the polling place voter turnout as a percentage of the total voting eligible population increased substantially by nearly 10 percentage points in Inyo County in response to super bloom (column 4). This effect is statistically significant as well at 99% confidence level.

Table 1: Summary Statistics : 1977-2004 Elections

Variable	Mean	Std. Dev.	Min.	Max.
Number of Eligible Voters	944399	298372	19471	3365902
Total Voters	563349	191872	16480	1553813
Male-Female Elector Ratio	1.096728	.123065	.6264844	2.031356
Voter Turnout(%)	60.11088	11.44238	5.074984	91.50587
Winning Margin(%)	15.62347	13.29791	.0014205	97.18752
Vote Share of Winner(%)	49.95038	10.10418	3.104751	97.6915
Vote Share of first losing candidate (%)	34.32691	8.099699	.5039694	49.83396
Vote Share of second losing candidate(%)	9.632125	8.005228	0	31.5732
Number of Contestants	11.59592	13.33688	1	480
Number of Male Contestants	11.09592	12.37155	0	420
Number of Female Contestants	.5	1.38934	0	60
SF-ratio	.3272953	.3011686	0	.9998002

Table 2: Summary Statistics: EVM and Non-EVM Constituencies (1977-2004 Elections)

	EVM Constituencies	Non-EVM Constituencies
Number of Eligible Voters	980875	940992.3
Total Voters	541312	565408.4
Male-Female Eligible Voter Ratio	1.209402	1.086203
Voter Turnout(%)	57.06892	60.39522
Winning Margin(%)	16.10818	15.57816
Vote Share of the Winner(%)	51.34446	49.82007
Vote Share of first losing candidate(%)	35.23628	34.2419
Vote Share of second losing candidate(%)	8.418439	9.745574
Number of Contestants	18.71014	10.93166
Number of Male Contestants	17.77778	10.47203
Number of Female Contestants	.9323671	.4596301
SF-ratio	.2763121	.332061

Notes: EVM constituencies refer to those parliamentary constituencies which got exposed to EVM in 1999. The non-EVM constituencies refer to the rest of the parliamentary constituencies which did not get EVM exposure in 1999 national election.

Table 3: Effect of Electronic Voting Machine on Voter Turnout

	Voter Turnout
Electronic Voting Machine	-2.812393 **
	(1.119756)
Constituency Fixed Effect	Y
State-Year Fixed Effect	Y
Log of Total Eligible Electors	Y
Log Male-Female Elector Ratio	Y
Log-Number of Polling Stations	Y
Log-Number of Total Contestants	Y
R^2	0.8761
Observations	4,300

Notes: Voter turnout measured by total polling place voters as percentage of total eligible electors(voters). Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Figure 1: A: Election Statistics 1977-2004

Figure 2: B: Election Statistics 1977-2004

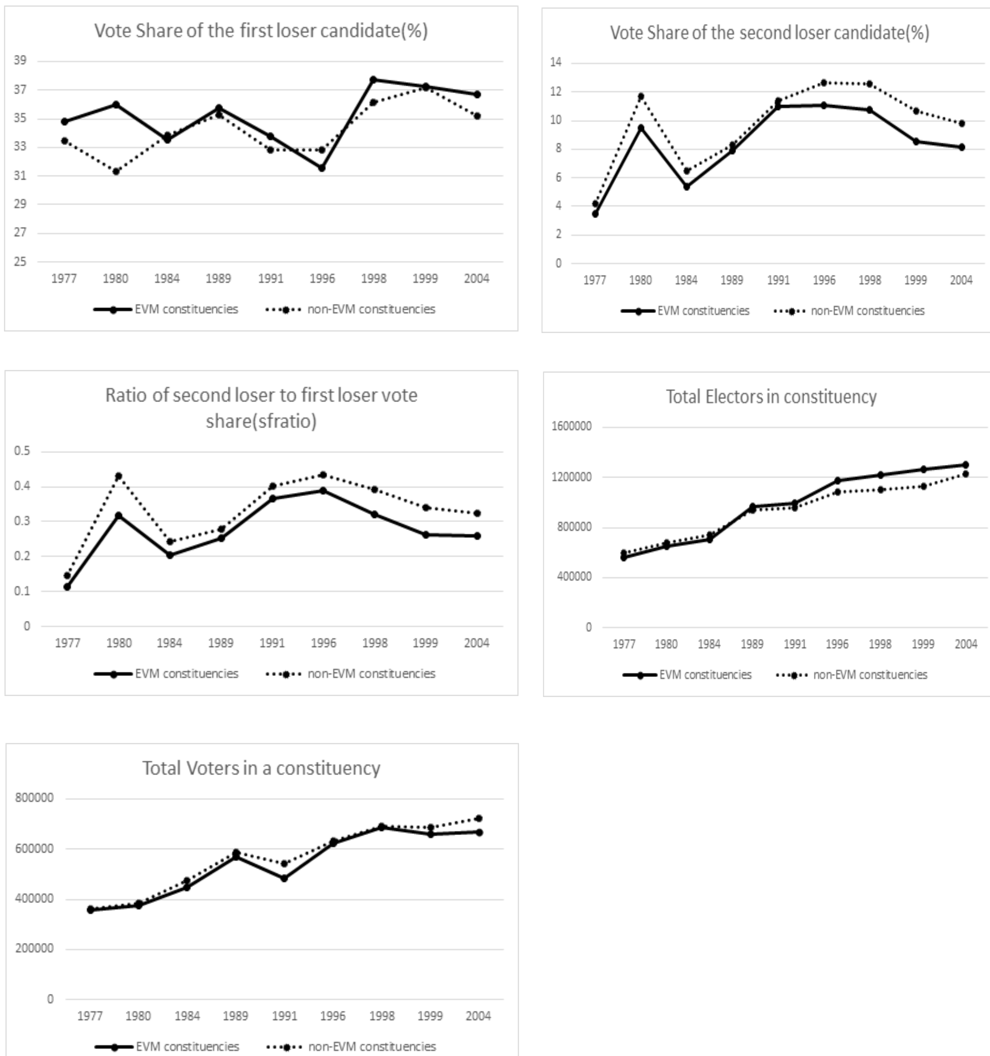


Figure 3: A: Voter Turnout for Selected States and Union Territories

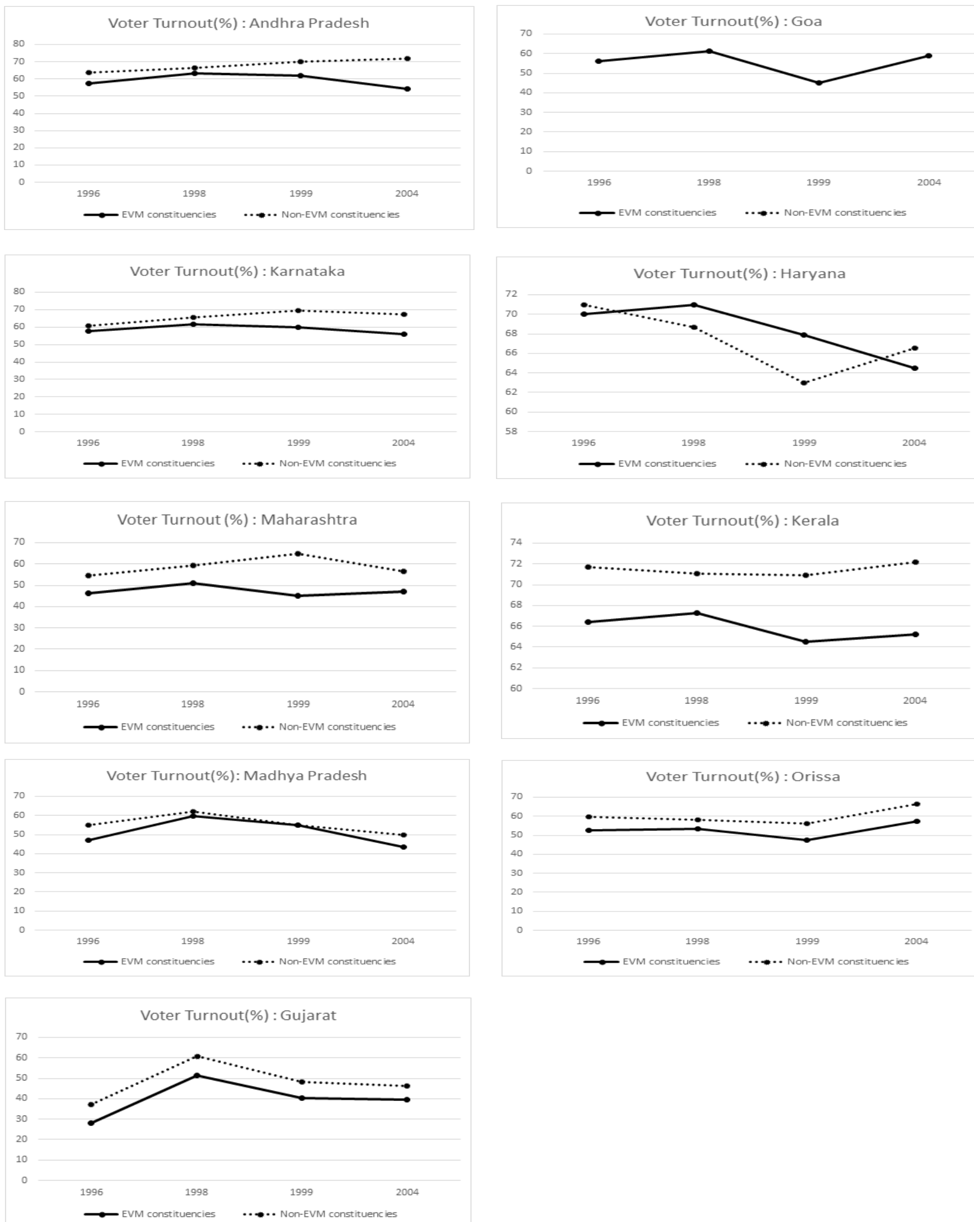


Figure 4: B: Voter Turnout for Selected States and Union Territories

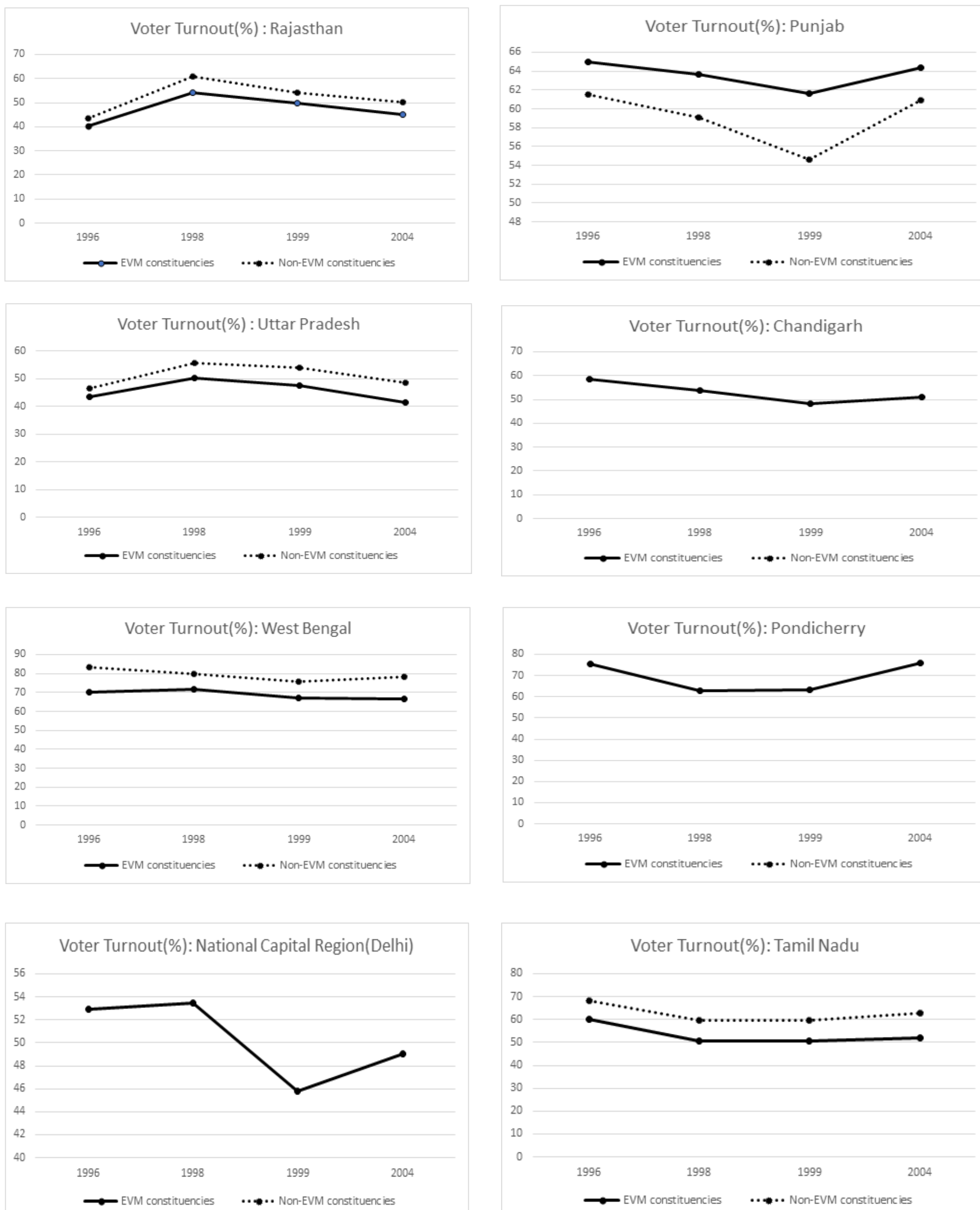


Table 4: DID Falsification Exercise: 1998 as Fake EVM introduction year

	Voter Turnout
Fake-EVM-Policy-98	-.9963338
	(.757213)
Constituency Fixed Effect	Y
State-Year Fixed Effect	Y
Log of Total Eligible Electors	Y
Log Male-Female Elector Ratio	Y
Log-Number of Polling Stations	Y
Log-Number of Total Contestants	Y
R^2	0.8851
Observations	3,757

Notes: Voter turnout measured by polling place voters as percentage of total eligible electors(voters). Standard errors are reported in the parenthesis. Standard error are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 5: DID Falsification Exercise with 1977 as reference year

	Voter Turnout
Fake-EVM*1980	-2.373109**
	(1.17787)
Fake-EVM*1984	-3.195094***
	(1.12919)
Fake-EVM*1989	-3.395018***
	(1.210056)
Fake-EVM*1991	-5.709379***
	(1.478914)
Fake-EVM*1996	-4.921505***
	(1.250796)
Fake-EVM*1998	-4.233553***
	(1.271592)
EVM*1999	-6.25163***
	(1.597093)
Constituency Fixed Effect	Y
State-Year Fixed Effect	Y
Log of Total Eligible Electors	Y
Log Male-Female Elector Ratio	Y
Log-Number of Polling Stations	Y
Log-Number of Total Contestants	Y
R^2	0.8773
Observations	4300

Notes: Voter turnout measured by total polling place voters as percentage of total eligible electors(voters). Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 6: Electronic Voting Machines and Voter Turnout

	Simple OLS	Robust/Sandwich Est	Jacknife	Without Outlier
Electronic Voting Machine	-2.812393***	-2.812393***	-2.812393**	-2.188026**
	(1.024503)	(1.024503)	(1.129427)	(.8673093)
Constituency Fixed Effect	Y	Y	Y	Y
State-Year Fixed Effect	Y	Y	Y	Y
Log of Total Eligible Electors	Y	Y	Y	Y
Log Male-Female Elector Ratio	Y	Y	Y	Y
Log-Number of Polling Stations	Y	Y	Y	Y
Log-Number of Total Contestants	Y	Y	Y	Y
R^2	0.8761	0.8761	0.8761	0.9258
Observations	4300	4300	4300	4,070

Notes: Voter turnout measured by total polling place voters as percentage of total eligible electors(voters). For column 4 estimation, observations with corresponding studentized residual absolute value exceeding 2 are excluded as outliers from the sample. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level at column 4 .***, ** and * represent significance at 1%, 5% and 10% levels respectively..

Table 7: Effect of Electronic Voting Machine

	Total Candidates	Female Candidates
EVM	-5.185775***	-.2433437
	(1.134736)	(.1506662)
Constituency Fixed Effect	Y	Y
State-Year Fixed Effect	Y	Y
Log of Total Eligible Electors	Y	Y
Log Male-Female Elector Ratio	N	Y
R^2	0.5082	0.2444
Observations	4303	4303

Notes: Total number of candidates refers to total number of candidates contesting in an electoral constituency. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 8: DID Falsification Exercise with Fake 1998 EVM Treatment

	Total Candidates	Female Candidates
EVM	-3.952683*	-.3353905
	(1.073534)	(.1763623)
Constituency Fixed Effect	Y	Y
State-Year Fixed Effect	Y	Y
Log of Total Eligible Electors	Y	Y
Log Male-Female Elector Ratio	N	Y
R^2	0.5156	0.2595
Observations	3,760	3,760

Notes: Total number of candidates refers to total number of candidates contesting in an electoral constituency. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 9: DID Falsification Exercise with 1977 as reference year: Number of Contesting Candidates

	Total Candidates	Female Candidates
Fake-EVM*1980	2.724359***	.0676783
	(.8322622)	(.1133664)
Fake-EVM*1984	6.917636***	.3947904**
	(1.311529)	(.1854171)
Fake-EVM*1989	6.001055***	.0255164
	(1.597692)	(.1981141)
Fake-EVM*1991	7.896615***	.3575735
	(1.590768)	(.2488739)
Fake-EVM*1996	7.548929***	.498393
	(2.518762)	(.450276)
Fake-EVM*1998	1.493904*	-.0886627
	(.8292759)	(.1804065)
EVM*1999	-.5234165	-.0636898
	(.9211171)	(.2044847)
Constituency FE	Y	Y
State-Year FE	Y	Y
Log-Total Eligible Electors	Y	Y
Log-Male-Female-Elector Ratio	N	Y
R^2	0.5105	0.2456
Observations	4,303	4,303

Notes: Total number of candidates refers to total number of candidates contesting in an electoral constituency. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 10: Electronic Voting Machines and Total Number of Candidates

	Simple OLS	Robust/Sandwich Est	Jacknife	Without Outlier
Electronic Voting Machine	-5.185775**	-5.185775***	-5.185775***	-5.41506***
	(2.073074)	(1.100258)	(1.202075)	(1.097744)
Constituency Fixed Effect	Y	Y	Y	Y
State-Year Fixed Effect	Y	Y	Y	Y
Log of Total Eligible Electors	Y	Y	Y	Y
R^2	0.5082	0.5082	0.5082	0.8383
Observations	4303	4303	4303	4,265

Notes: Total number of candidates is measured by the total number of contestants in a parliamentary constituency. For column 4 estimation, observations with corresponding studentized residual absolute value exceeding 2 are excluded as outliers from the sample. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level at column 4. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 11: Effect of Electronic Voting Machine: Winning Margin and Vote Shares

	Vsh-Winner	Winmargin	Vsh-2ndwinner	Vsh-3rdwinner	SF-ratio
EVM	.9043411	3.557751*	-2.65341**	-.5447166	.0045943
	(1.260641)	(1.855472)	(1.105085)	(1.147969)	(.0415623)
Constituency Fixed Effect	Y	Y	Y	Y	Y
State-Year Fixed Effect	Y	Y	Y	Y	Y
Log of Total Eligible Electors	Y	Y	Y	Y	Y
Log Total Candidates	Y	Y	Y	Y	Y
R^2	0.6611	0.5524	0.5924	0.6375	0.6205
Observations	4300	4300	4,300	4300	4,300

Notes: Vsh-Win (vote share of the winner as percentage of total valid votes cast in a constituency). Winmar (the difference between the winner and the candidate with second most votes as percentage of total valid votes cast). Vsh-2ndwin(vote share of the candidate with second most votes as percentage of total valid votes cast). Vsh-3rdwin(vote share of the candidate with third most votes as percentage of total valid votes cast). sfratio(ratio of vote share of the candidates with third and second most votes). Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 12: DID Falsification Exercise with Fake 1998 EVM treatment

	Vsh-Winner	Winmargin	Vsh-2ndwinner	Vsh-3rdwinner	SF-ratio
EVM	.7214812	1.283257	-.561775	-.3268831	-.0125411
	(1.181367)	(1.637686)	(.8429523)	(1.109313)	(.0377147)
Constituency Fixed Effect	Y	Y	Y	Y	Y
State-Year Fixed Effect	Y	Y	Y	Y	Y
Log of Total Eligible Electors	Y	Y	Y	Y	Y
Log Total Candidates	Y	Y	Y	Y	Y
R^2	0.6636	0.5539	0.5934	0.6369	0.6185
Observations	3,757	3,757	3,757	3,757	3,757

Notes: Vsh-Win (vote share of the winner as percentage of total valid votes cast in a constituency). Winmar(the difference between the winner and the candidate with second most votes as percentage of total valid votes cast). Vsh-2ndwin(vote share of the candidate with second most votes as percentage of total valid votes cast). Vsh-3rdwin(vote share of the candidate with third most votes as percentage of total valid votes cast). sfratio(ratio of vote share of the candidates with third and second most votes). Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 13: DID Falsification Exercise with 1977 as reference year: Winning Margin and Vote Share

	Vsh-W	Winmar	Vsh-2ndwin	Vsh-3rdwin	SF-ratio
Fake-EVM*1980	-2.356367	-6.169117 *	3.812749**	-.6426499	-.0638965
	(1.774458)	(3.313017)	(1.794498)	(1.071434)	(.0426605)
Fake-EVM*1984	1.017745	1.253695	-.23595	-.5204889	-.0194179
	(1.741472)	(2.917585)	(1.648258)	(1.041277)	(.0489391)
Fake-EVM*1989	2.435876	3.544692*	-1.108816	.317728	-.0049727
	(2.02642)	(3.366727)	(1.606141)	(1.184444)	(.0450506)
Fake-EVM*1991	2.896798*	2.230394	.6664033	-1.516556	-.0669816
	(1.548251)	(2.808528)	(1.663081)	(1.344759)	(.0547718)
Fake-EVM*1996	4.441345***	6.543945**	-2.102601	-.7383277	-.0191777
	(1.567894)	(2.629796)	(1.409622)	(1.318703)	(.0516672)
Fake-EVM*1998	2.140948	2.553365	-.4124168	-.775945	-.0390562
	(1.69688)	(2.907163)	(1.536887)	(1.371104)	(.0483569)
EVM*1999	2.444118	5.024468*	-2.58035*	-1.104331	-.0260452
	(1.650541)	(3.020105)	(1.678715)	(1.24235)	(.0455539)
Constituency FE	Y	Y	Y	Y	Y
State-Year FE	Y	Y	Y	Y	Y
Log-Total Eligible Electors	Y	Y	Y	Y	Y
Log of Total Candidates	Y	Y	Y	Y	Y
R^2	0.6632	0.5561	0.5947	0.5947	0.6209
Observations	4,300	4,300	4,300	4,300	4,300

Notes: Vsh-Win (vote share of the winner as percentage of total valid votes in a constituency) . Winmar (the difference between the winner and the candidate with second most votes as percentage of total valid votes). Vsh-2ndwin(vote share of the candidate with second most votes as percentage of total valid votes). Vsh-3rdwin(vote share of the candidate with third most votes as percentage of total valid votes). sfratio(ratio of vote share of the candidates with third and second most votes). Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 14: Electronic Voting Machines and Winning Margin

	Simple OLS	Robust/Sandwich Est	Jacknife	Without Outlier
Electronic Voting Machine	3.557751*	3.557751**	3.557751*	3.074518*
	(1.91746)	(1.774422)	(1.9434)	(1.739191)
Constituency Fixed Effect	Y	Y	Y	Y
State-Year Fixed Effect	Y	Y	Y	Y
Log of Total Eligible Electors	Y	Y	Y	Y
R^2	0.5524	0.5524	0.5524	0.6644
Observations	4,300	4300	4300	4,079

Notes: Winning margin is the difference of votes for the winner and the candidate with second most votes as percentage of total valid votes cast. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 15: Electronic Voting Machines and Vote Share of the Second Winner

	Simple OLS	Robust/Sandwich Est	Jacknife	Without Outlier
Electronic Voting Machine	-2.65341**	-2.65341**	-2.65341**	-3.417288***
	(1.103819)	(1.048531)	(1.148153)	(.9975505)
Constituency Fixed Effect	Y	Y	Y	Y
State-Year Fixed Effect	Y	Y	Y	Y
Log of Total Eligible Electors	Y	Y	Y	Y
R^2	0.5924	0.5924	0.5924	0.7006
Observations	4303	4303	4303	4,086

Notes: The dependent variable is vote share of the candidate with second most votes as percentage of total valid votes cast. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 16: Joint Significance Test: F-Statistic

	Voter Turnout	Winning Margin	Vote Share of 2nd Winner
	7.30***	5.21***	25.66***
Log of Total Eligible Electors	Y	Y	Y
Log Male-Female Elector Ratio	Y	N	N
Log-Number of Polling Stations	Y	N	N
Log-Number of Total Contestants	Y	Y	Y

Notes: The unit of analysis is parliamentary electoral constituency. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 17: EVM and Voter Turnout: Heterogeneity Analysis with State Per Capita Income of 1998

	Het: Mean Income	Het : Median Income
EVM-Triple	-2.672976	-3.503036*
	(2.349921)	(1.934117)
EVM	-.7023338	-.5696214
	(2.021166)	(1.294249)
Constituency FE	Y	Y
State-Year FE	Y	Y
Log-Total Eligible Electors	Y	Y
Log-Total Polling Stations	Y	Y
Log-Male-Female-Elector Ratio	Y	Y
Log-Number of Total Contestants	Y	Y
R^2	0.8762	0.8763
Observations	4,300	4,300

Notes: Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 18: Effect of Electronic Voting Machine on Voter Turnout: 1999 and 2004

	Voter Turnout
Electronic Voting Machine	-.0104622
	(1.094933)
Constituency Fixed Effect	Y
Year Fixed Effect	Y
Log of Total Eligible Electors	Y
Log Male-Female Elector Ratio	Y
Log-Number of Polling Stations	Y
Log-Number of Total Contestants	Y
R^2	0.9125
Observations	1,086

Notes: Standard error are reported in parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 19: Electronic Voting Machine on Voter Turnout: Effect of Metropolis Constituencies

	Voter Turnout		
Electronic Voting Machine	-2.812393**	-.6625657	-.2930119
	(1.119756)	(.9273307)	(1.076458)
Constituency Fixed Effect	Y	Y	Y
State-Year Fixed Effect	Y	Y	Y
EVM Group Linear Trend	N	Y	N
Metropolis Group-Year Fixed Effect	N	N	Y
Log of Total Eligible Electors	Y	Y	Y
Log Male-Female Elector Ratio	Y	Y	Y
Log-Number of Polling Stations	Y	Y	Y
Log-Number of Total Contestants	Y	Y	Y
R^2	0.8761	0.8769	0.8774
Observations	4,300	4,300	4,300

Notes: Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 20: 2001 Census Demographic Variables Comparison in a limited Sample

	EVM Constituencies	Non-EVM Constituencies
Percent of Marginal Labor	5.152649	9.47478
Percent of Rural Population	37.4514	78.11405
Literacy Rate above age 7	74.97276	62.95197
Literacy Rate GAP (SC and Non-SC)	-13.26232	-9.683165
Percent of Agricultural Labor	4.289863	12.33908
Percent of Agricultural Labor in SC pop	9.074138	20.20336

Table 21: Matched Difference-in-Differences: Effect on Turnout

	DID	PSM-DID
EVM	-2.249855 **	-.7010042
	(1.209833)	(1.012022)
Constituency FE	Y	Y
State-Year FE	Y	Y
Log-Total Eligible Electors	Y	Y
Log-Total Polling Stations	Y	Y
Log-Male-Female-Elector Ratio	Y	Y
Log-Total Number of Contestants	Y	Y
R^2	0.8788	0.8855
Observations	3,885	1,902

Notes: The dependent variable is voter turnout. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 22: Matched DID in Selected Sample: Falsification Exercise with 1998 as fake EVM year

	DID	PSM-DID
EVM	-.5119529	-.2363379
	(.7809992)	(.807643)
Constituency FE	Y	Y
State-Year FE	Y	Y
Log-Total Eligible Electors	Y	Y
Log-Total Polling Stations	Y	Y
Log-Male-Female-Elector Ratio	Y	Y
Log-Total Number of Contestants	Y	Y
R^2	0.8895	0.8911
Observations	3,399	1,664

Notes: The dependent variable is voter turnout. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 23: Matched DID in Selected Sample: Falsification Exercise with 1977 as reference year

	DID	PSM-DID
Fake-EVM*1980	-1.99438*	-1.946438*
	(1.181765)	(1.129963)
Fake-EVM*1984	-2.924202**	-3.022174**
	(1.194546)	(1.184012)
Fake-EVM*1989	-3.064176**	-2.667131**
	(1.282303)	(1.319959)
Fake-EVM*1991	-4.861572***	-4.521511 ***
	(1.566428)	(1.372728)
Fake-EVM*1996	-3.821512***	-2.966252 **
	(1.292801)	(1.172069)
Fake-EVM*1998	-3.220723**	-2.692493**
	(1.368333)	(1.255253)
EVM*1999	-5.136802***	-3.245185**
	(1.76115)	(1.504182)
Constituency FE	Y	Y
State-Year FE	Y	Y
Log-Total Eligible Electors	Y	Y
Log-Total Polling Stations	Y	Y
Log-Male-Female-Elector Ratio	Y	Y
Log-Total Number of Contestants	Y	Y
R^2	0.8796	0.8867
Observations	3,885	1,902

Notes: Selected Sample: 486 out of total 543 parliamentary constituencies of India. PSM-DID analysis based on 238 matched constituencies among 486 constituencies. The dependent variable is voter turnout. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 24: 2001 Demographic and 1991 Infrastructure Comparison in Northern Indian States

	EVM Constituencies	Non-EVM Constituencies
Percent of Marginal Labor	7.007815	10.2655
Percent of Rural Population	56.85072	79.37324
Literacy Rate above age 7	66.77919	56.3628
Literacy Rate GAP (SC and Non-SC)	-17.16855	-12.75942
Percent of Agricultural Labor	5.708613	10.21552
Avg Distance(km) to town(1991)	13.4685	17.93029
Percent of villages with power supply(1991)	.8822005	.7248634
Percent of villages with paved road(1991)	.7016998	.4423109
Percent of villages with footpath(1991)	.6464099	.6129061

Table 25: Matched Difference-in-Differences: Effect on Turnout in Selected Northern Indian States

	DID	PSM-DID
EVM	1.633398	3.339371**
	(1.535117)	(1.486822)
Constituency FE	Y	Y
State-Year FE	Y	Y
Log-Total Eligible Electors	Y	Y
Log-Total Polling Stations	Y	Y
Log-Male-Female-Elector Ratio	Y	Y
Log-Total Number of Contestants	Y	Y
R^2	0.7993	0.8434
Observations	1,661	478

Notes: The dependent variable is voter turnout. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 26: Matched DID in Northern Indian States: Falsification Exercise with 1998 as fake EVM year

	DID	PSM-DID
EVM	.1203202	1.163913
	(1.39016)	(1.597923)
Constituency FE	Y	Y
State-Year FE	Y	Y
Log-Total Eligible Electors	Y	Y
Log-Total Polling Stations	Y	Y
Log-Male-Female-Elector Ratio	Y	Y
Log-Total Number of Contestants	Y	Y
R^2	0.8118	0.8478
Observations	1,453	418

Notes: The dependent variable is voter turnout. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 27: Matched DID in Northern Indian States: Falsification Exercise with 1977 as reference year

	DID	PSM-DID
Fake-EVM*1980	-3.513531***	-3.773659**
	(1.30583)	(1.664008)
Fake-EVM*1984	-4.521125**	-4.600808*
	(2.215133)	(2.580439)
Fake-EVM*1989	-5.530851***	-3.659339
	(2.387822)	(2.387071)
Fake-EVM*1991	-7.745191***	-6.05814 **
	(2.279165)	(2.905557)
Fake-EVM*1996	-3.747496*	-2.043222
	(2.107753)	(2.43623)
Fake-EVM*1998	-3.96634**	-1.980831
	(1.957224)	(2.517844)
EVM*1999	-2.535722	.2637875
	(2.307402)	(2.561138)
Constituency FE	Y	Y
State-Year FE	Y	Y
Log-Total Eligible Electors	Y	Y
Log-Total Polling Stations	Y	Y
Log-Male-Female-Elector Ratio	Y	Y
Log-Total Number of Contestants	Y	Y
R^2	0.8018	0.8474
Observations	1,661	478

Notes: The dependent variable is voter turnout. Standard errors are reported in the parenthesis. Standard errors are clustered at the parliamentary electoral constituency level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Figure 5: A: Tourist Numbers in Death Valley, CA

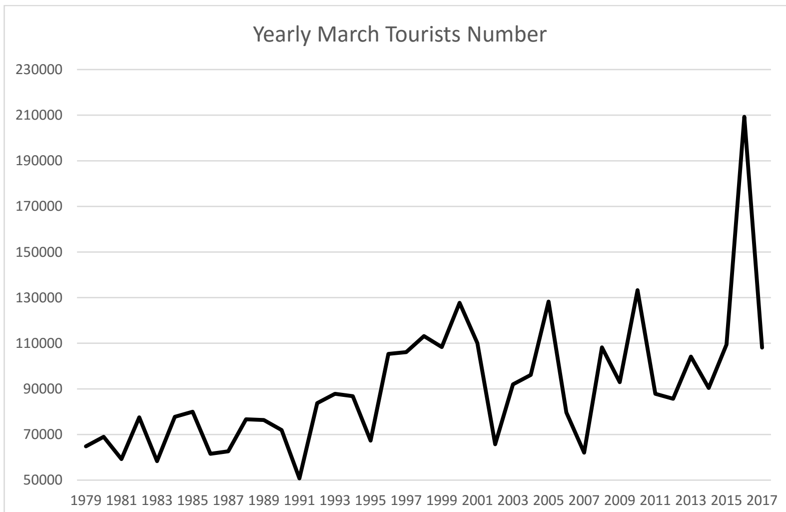


Figure 6: B: Tourist Numbers in Death Valley, CA

Figure 7: Comparison: Electoral Outcomes

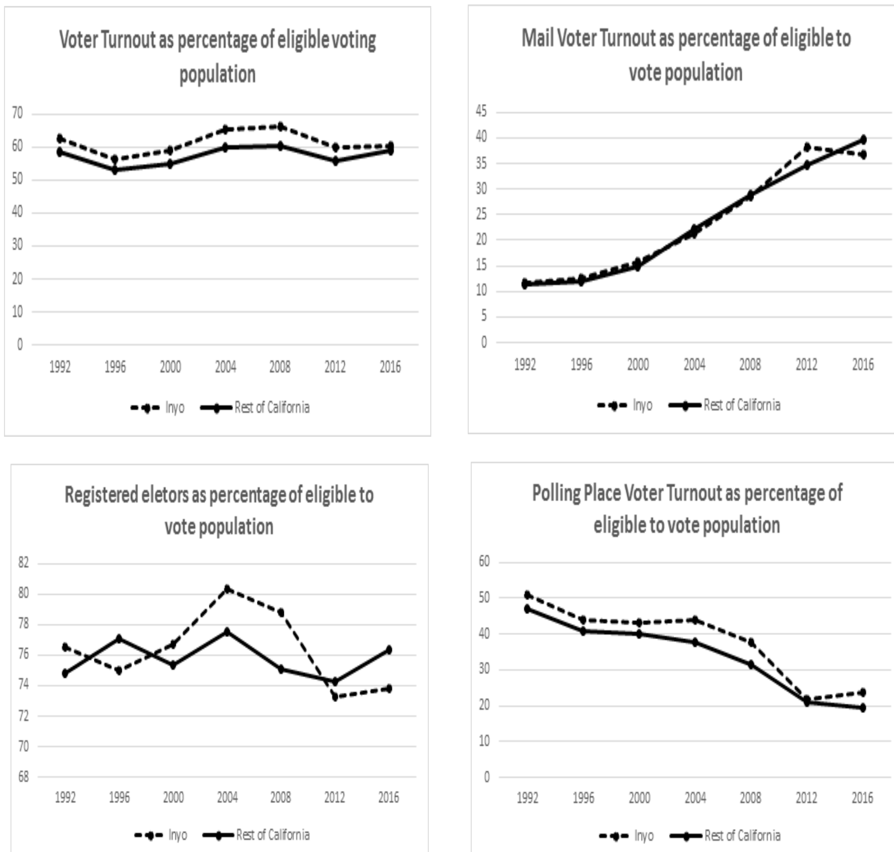


Table 28: Summary Statistics of Presidential Elections 1992-2016 in CA: Comparison of Electoral Outcomes

	Inyo County	Rest of California
Voter Turnout as % of total voting eligible population	61.44328	57.45032
Voter Registration Percentage	76.36008	75.79533
Mail Voters Turnout as % of total voting eligible population	23.51496	23.41119
Polling Place Voter Turnout as % of total voting eligible population	37.92832	34.03913

Notes: Rest of California excludes three mail-only counties : Alpine, Sierra, Plumas. Source of Data : California Secretary of State

Table 29: Summary Statistics(1992-2016): Comparison of Covariates

	Inyo County	Rest of California
Lagged Median Household Income(annual)	38764.71	44536.58
Total Eligible to register population	13505	411600
Proportion of population in age group 25-64	0.5133342	0.5183567
Proportion of population with age 65plus	0.1979343	0.1334852
Male Female Population Ratio	0.9809215	1.022025
Proportion of White Population	0.8480816	0.8440406
Proportion of Black Population	0.0057907	0.0383459
Proportion of Hispanic Population	0.1532121	0.2611937
Percent of population with less than high school education	16.10914	21.04279
Percent of population with high school education	33.162	23.91688
Percent of population with some college or associate degree	32.54543	32.867
Optical Scan Voting Technology	.4285714	.6587302
Punch Card Voting Technology	.5714286	.2962963
Touch Screen Voting Technology	0	.042328
Automatic Voting Machine Technology	0	.0026455

Notes: Rest of California excludes three mail-only counties : Alpine, Sierra, Plumas. Education statistics correspond to age-group 25 and older. Hispanic does not have any racial connotation. Hispanic population can include multiple races. The white population in the dataset include both non-Hispanic and Hispanic white individuals. Median household income is measured in the U.S. dollar. The voting technology figures denote proportion of times a particular voting technology is used in presidential elections in the time period of 1992-2016

Table 30: Effect of Super bloom on Voter Turnout

	poll-p	regis-vot	poll-p-absent	poll-p-prec
Superbloom	-5.939282***	-7.406188***	-7.776196***	1.836914
	(1.036049)	(1.391756)	(1.748252)	(2.122099)
County Fixed Effect	Y	Y	Y	Y
Year Fixed Effect	Y	Y	Y	Y
Voting Technology	Y	N	Y	Y
County level controls	Y	Y	Y	Y
R^2	0.9279	0.7814	0.9211	0.8953
Observations	385	385	385	385

Notes: poll-p (voter turnout as percentage of eligible voting population). regis-vote(registered voters as percentage of eligible voting population). poll-p-absent(mail voters turnout as percentage of eligible voting population). poll-p-prec(polling place voters turnout as percentage of eligible voting population). Standard errors reported in the parenthesis and clustered at the county level. The unit of analysis is county. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 31: DID Falsification Test with 2012 as year of fake bloom

	poll-p	regis-vot	poll-p-absent	poll-p-prec
Fake-Superbloom	-3.618766 ***	-6.746056 ***	2.024811	-5.643574**
	(.8987927)	(1.339649)	(1.834708)	(2.145023)
County Fixed Effect	Y	Y	Y	Y
Year Fixed Effect	Y	Y	Y	Y
Voting Technology	Y	N	Y	Y
Other County level controls	Y	Y	Y	Y
R^2	0.9431	0.8144	0.9062	0.8840
Observations	330	330	330	330

Notes: poll-p (voter turnout as percentage of eligible voting population). regis-vote(registered voters as percentage of eligible voting population). poll-p-absent(mail voters turnout as percentage of eligible voting population). poll-p-prec(polling place voters turnout as percentage of eligible voting population). Standard errors reported in the parenthesis and clustered at the county level. The unit of analysis is county. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 32: Difference-in-Differences(DID) Falsification Test: 1992 as reference year

	poll-p	regis-vote	poll-p-absent	poll-p-prec
Fake-Superbloom*1996	-.6252053	-3.812641***	.3367734	-.9619745
	(.7506856)	(.8699395)	(.5776069)	(.8205758)
Fake-Superbloom*2000	.3924538	-.1087139	-4.849157***	5.241612***
	(1.307017)	(1.985684)	(1.416495)	(1.475586)
Fake-Superbloom*2004	1.305472	-.4695034	-7.667444***	8.972919 ***
	(1.235853)	(1.948814)	(1.698828)	(1.746408)
Fake-Superbloom*2008	-.8111933	-2.249794	-7.379959 ***	6.568764***
	(1.693462)	(2.319964)	(1.791268)	(2.080372)
Fake-Superbloom*2012	-3.29977**	-7.035534 ***	-3.298052	-.0017135
	(1.593738)	(2.077168)	(2.606098)	(3.03482)
Superbloom*2016	-6.579676***	-9.493581***	-11.81751***	5.23784
	(1.734529)	(2.32464)	(2.703149)	(3.268997)
County FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Voting Technology	Y	N	Y	Y
Other County level controls	Y	Y	Y	Y
R^2	0.9284	0.7766	0.9221	0.8969
Observations	385	385	385	385

Notes: poll-p (voter turnout as percentage of eligible voting population). regis-vote(registered voters as percentage of eligible voting population). poll-p-absent(mail voters turnout as percentage of eligible voting population). poll-p-prec(polling place voters turnout as percentage of eligible voting population). Standard errors reported in the parenthesis and clustered at the county level. The unit of analysis is county. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 33: Joint Significance Test: F-Statistic

	poll-p	regis-vot	poll-p-absent	poll-p-prec
	4.15***	2.32**	6.75***	4.03***
Log of Eligible to Register Pop	Y	Y	Y	Y
Prop of White Pop	Y	Y	Y	Y
Prop of Black Pop	Y	Y	Y	Y
Prop of Hispanic Pop	Y	Y	Y	Y
Male-Female Pop Ratio	Y	Y	Y	Y
Prop of Population in Age Group 25-64	Y	Y	Y	Y
Prop of Population in Age Group 65up	Y	Y	Y	Y
Percent of population with less than High School Education	Y	Y	Y	Y
Percent of population with only High School degree	Y	Y	Y	Y
Percent of population with some college or associate degree	Y	Y	Y	Y

Notes: poll-p (voter turnout as percentage of eligible voting population) . regis-vote(registered voters as percentage of eligible voting population). poll-p-absent(mail voters turnout as percentage of eligible voting population). poll-p-prec(polling place voters turnout as percentage of eligible voting population). Standard errors reported in the parenthesis and clustered at the county level. The unit of analysis is county. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 34: Super bloom and Voter Turnout (% of Eligible Voting Population)

	Simple OLS	Robust/Sandwich Est	Jacknife	Without Outlier
Superbloom	-5.939282**	-5.939282***	-5.939282***	-5.835393***
	(2.790457)	(1.110133)	(1.272474)	(.6463903)
County Fixed Effect	Y	Y	Y	Y
Year Fixed Effect	Y	Y	Y	Y
Voting Technology	Y	Y	Y	Y
Other County level controls	Y	Y	Y	Y
R^2	0.9279	0.9279	0.9279	0.9619
Observations	385	385	385	362

Notes: Dependent variable is voter turnout as percentage of eligible voting population. The unit of analysis is county. Standard errors reported in the parentheses. For column 4 estimation, observations with corresponding studentized residual absolute value exceeding 2 are excluded as outliers from the sample. Standard errors clustered at the administrative district level at column 4. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 35: Super bloom and Voter Registration Percentage

	Simple OLS	Robust/Sandwich Est	Jacknife	Without Outlier
Superbloom	-7.014494	-7.014494***	-7.014494***	-8.030493***
	(4.310215)	(1.623283)	(1.85884)	(1.133762)
County Fixed Effect	Y	Y	Y	Y
Year Fixed Effect	Y	Y	Y	Y
Voting Technology	N	N	N	N
Other County level controls	Y	Y	Y	Y
R^2	0.7749	0.7749	0.7749	0.8446
Observations	385	385	385	362

Notes: Dependent variable is registered voters as percentage of eligible voting population. The unit of analysis is county. Standard errors reported in the parentheses. For column 4 estimation, observations with corresponding studentized residual absolute value exceeding 2 are excluded as outliers from the sample. Standard errors clustered at the administrative district level at column 4. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 36: Super bloom and Mail Voter Turnout (% of Eligible Voting Population)

	Simple OLS	Robust/Sandwich Est	Jacknife	Without Outlier
Superbloom	-7.776196*	-7.776196***	-7.776196***	-6.292528 ***
	(4.285722)	(1.792187)	(2.016678)	(1.120755)
County Fixed Effect	Y	Y	Y	Y
Year Fixed Effect	Y	Y	Y	Y
Voting Technology	Y	Y	Y	Y
Other County level controls	Y	Y	Y	Y
R^2	0.9211	0.9211	0.9211	0.9507
Observations	385	385	385	371

Notes: Dependent variable is mail voter turnout as percentage of eligible voting population. The unit of analysis is county. Standard errors reported in the parentheses. For column 4 estimation, observations with corresponding studentized residual absolute value exceeding 2 are excluded as outliers from the sample. Standard errors clustered at the administrative district level at column 4. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 37: Super bloom and Polling Voter Turnout (% of Eligible Voting Population)

	Simple OLS	Robust/Sandwich Est	Jacknife	Without Outlier
Superbloom	1.836914	1.836914	1.836914	.1911308
	(4.564828)	(2.170439)	(2.454391)	(1.23097)
County Fixed Effect	Y	Y	Y	Y
Year Fixed Effect	Y	Y	Y	Y
Voting Technology	Y	Y	Y	Y
Other County level controls	Y	Y	Y	Y
R^2	0.8953	0.8953	0.8953	0.9378
Observations	385	385	385	368

Notes: Dependent variable is polling place voter turnout as percentage of eligible voting population. The unit of analysis is county. Standard errors reported in the parentheses. For column 4 estimation, observations with corresponding studentized residual absolute value exceeding 2 are excluded as outliers from the sample. Standard errors clustered at the administrative district level at column 4. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

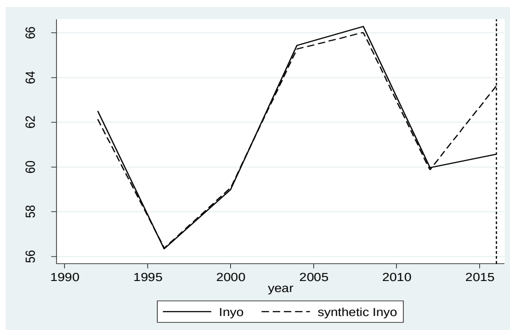
Table 38: Nearest Neighbor Covariate Matching: Effect of Super bloom on Voter Turnout

	poll-p	regis-vot	poll-p-absent	poll-p-prec
Superbloom	-.7089615	-2.9203	-10.77088***	10.06191***
	(.453809)	(2.907794)	(3.290009)	(3.74382)
Observations	55	55	55	55

Notes: poll-p (voter turnout as percentage of eligible voting population) . regis-vote(voter turnout percentage). poll-p-absent(mail voters turnout as percentage of eligible voting population). poll-p-prec(polling place voters turnout as percentage of eligible voting population). The unit of analysis is county. Standard errors reported in the parenthesis.***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Figure 8: Synthetic Control: Super bloom and Voter Turnout as Percentage of Eligible Voting Population

Electoral Variable: Voter Turnout as percentage of Eligible Voting Population



Placebo Study Graph

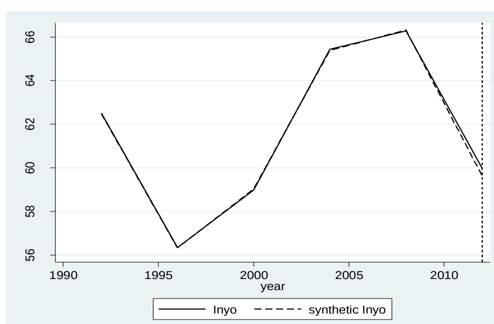
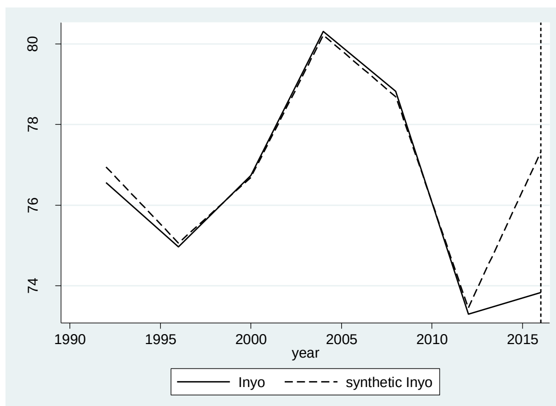


Figure 9: Synthetic Control: Super bloom and Voter Registration Percentage

Electoral Variable: Voter Registration Percentage



Placebo Study Graph

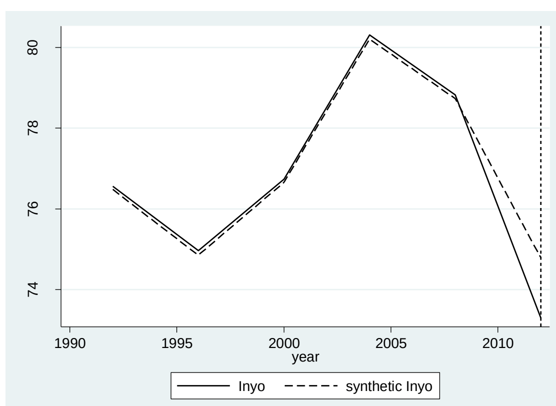
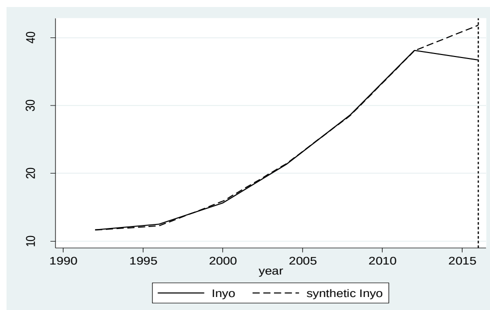


Figure 10: Synthetic Control: Super bloom and Mail Voters Turnout as Percentage of Eligible Voting Population

Electoral Variable: Mail Voter Turnout as percentage of Eligible Voting Population



Placebo Study Graph

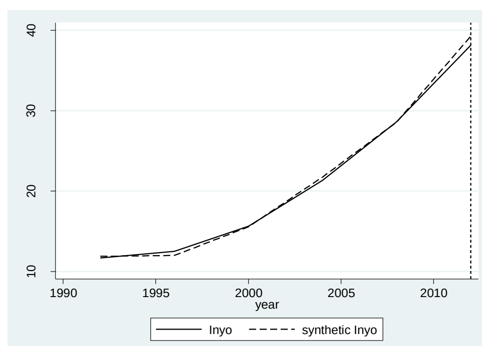
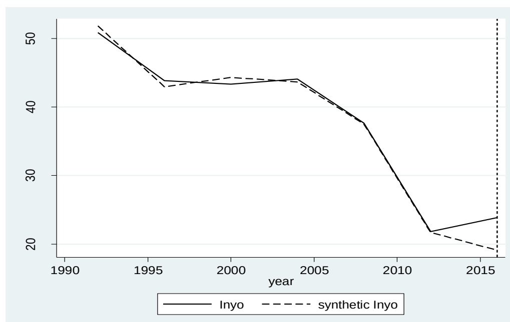


Figure 11: Synthetic Control: Super bloom and Polling Place Voter Turnout as Percentage of Eligible Voting Population

Electoral Variable: Polling Place Voter Turnout as percentage of Eligible Voting Population



Placebo Study Graph

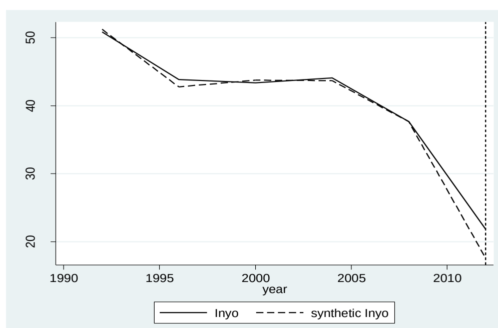


Table 39: Elections in States in 2001-2012

State Name	Elections before Redistricting	Elections after Redistricting
Andhra Pradesh	2004	2009
Bihar	2005	2010
Chattisgarh	2003	2008
Delhi	2003	2008
Goa	2002, 2007	2012
Gujarat	2002, 2007	2012
Haryana	2005	2009
Himachal Pradesh	2003, 2007	2012
Karnataka	2004	2008
Kerala	2001, 2006	2011
Madhya Pradesh	2003	2008
Uttar Pradesh	2002, 2007	2012
Uttarakhand	2002, 2007	2012
Maharashtra	2004	2009
Odisha	2004	2009
Rajasthan	2003	2008
Punjab	2002, 2007	2012
Sikkim	2004	2009
Tamil Nadu	2001, 2006	2011
West Bengal	2001, 2006	2011

Table 40: Summary Statistics

	Mean	Std. Deviation	Min	Max
Total Public(Govt.) Schools($t+5$)	1801.501	1129.057	9	8586
Total Private Schools($t+5$)	480.8304	445.6221	0	2991
Mean Male-Female Elector Ratio	1.093239	.0992199	.7873803	1.565374
Median Male-Female Elector Ratio	1.090761	.0975051	.7755245	1.500104
Male dominated seats ratio	.811411	.333461	0	1
Total Electors in district($t+5$)	1555060	1226546	21089	10699484
Dist. Male-Female Elector Ratio	1.093375	.0979974	.7908394	1.500164
Total Electoral Constituencies within a district	7.082593	4.413245	1	33
Prop of seats from district in State Legislative Assembly	.0415888	.048653	.0041152	.575
Reserved Seats ratio	.2615554	.228348	0	1
Prop of Ruling Coalition Members elected from a district	.6311318	.276573	0	1
Prop of State Cabinet Ministers elected from a district	.1255431	.1345657	0	.6
Prop of Female elected representatives from district	.0774701	.1193178	0	1

Notes: The unit of analysis is administrative district. Prop of seats from district in State Legislative Assembly : total number of state legislative electoral constituency in an administrative district as proportion of total number of seats in the legislative assembly of a state. Prop of Ruling Coalition Members elected from a district : number of elected member who were part of the ruling coalition parties as proportion of total number of electoral constituencies in an administrative district. Reserved Seats Ratio : the number of seats in a district reserved for backward caste candidates as proportion of the total number of electoral constituencies in an district. Prop of total female elected representatives from district : number of elected candidates who are female as proportion of total number of elected candidates from a district. Prop of State Cabinet Ministers elected from a district : Proportion of Cabinet ministers in state government elected from a district as proportion of total number of electoral constituencies in an administrative district. Dist. Male-Female Elector Ratio : Ratio of total number of male electorates to total number of female electorates in an administrative district

Table 41: Effect of Redistricting

	Mean MF Elector Ratio	Median MF Elector Ratio	Prop Male Dom. Seats
2008 Redistricting Dummy	.0184221 ***	.0192695 ***	.0645312***
	(.0017897)	(.0016991)	(.0086673)
Admin. District Fixed Effect	Y	Y	Y
R^2	0.9495	0.9529	0.8951
Observations	1,126	1,126	1,126

Notes: The unit of analysis is administrative district. 2008 Redistricting Dummy takes the value of 1 for all state elections since May 2008 and 0 otherwise. Standard errors reported in the parentheses. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 42: Mean Male Female Elector Ratio and Number of Public Schools

	(1)	(2)	(3)	(4)
Mean Male Female Elector Ratio	1.385667***	1.400139***	1.740281***	1.612723***
	(.4908789)	(.4913752)	(.4977097)	(.4930258)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9956	0.9956	0.9957	0.9959
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 43: Mean Male Female Elector Ratio and Number of Public Schools(clustered standard error)

	(1)	(2)	(3)	(4)
Mean Male Female Elector Ratio	1.385667*	1.400139*	1.740281**	1.612723*
	(.7435718)	(.7487154)	(.8181006)	(.8394496)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9956	0.9956	0.9957	0.9959
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. Standard errors clustered at the administrative district level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 44: Median Male Female Elector Ratio and Number of Public Schools

	(1)	(2)	(3)	(4)
Median Male Female Elector Ratio	.0387317	.0382322	.04724	.0343028
	(.2159263)	(.2160212)	(.2151299)	(.2122581)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9956	0.9956	0.9956	0.9958
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 45: Median Male Female Elector Ratio and Number of Public Schools(clustered standard error)

	(1)	(2)	(3)	(4)
Median Male Female Elector Ratio	.0387317	.0382322	.04724	.0343028
	(.3748864)	(.3762243)	(.3615488)	(.3489226)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9956	0.9956	0.9956	0.9958
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. Standard errors clustered at the district level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 46: Male Dominated Seats Ratio in District and Number of Public Schools

	(1)	(2)	(3)	(4)
Male Dom Ratio	.055996 ***	.0562451***	.0523803**	.0464273**
	(.0214479)	(.0214589)	(.0214384)	(.0211554)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9956	0.9958	0.9957	0.9958
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 47: Male Dominated Seats Ratio in District and Number of Public Schools(clustered standard error)

	(1)	(2)	(3)	(4)
Male Dom Ratio	.055996 **	.0562451**	.0523803**	.0464273*
	(.0250627)	(.0250689)	(.026237)	(.0251846)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9956	0.9958	0.9957	0.9958
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. Standard errors clustered at the administrative district level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 48: Joint Significance Test: F-Statistic

	(1)	(2)	(3)
	5.60***	4.46***	4.97***
Mean MF-Elector Ratio	Y	N	N
Median MF-Elector Ratio	N	Y	N
Male Dom Seat Ratio	N	N	Y
Log-Total Elig Electors	Y	Y	Y
Log-Total Private Schools	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y
Prop of Reserved Seats	Y	Y	Y
Prop of seats in State assembly	Y	Y	Y
Ratio of ruling coalition legislators	Y	Y	Y
Ratio of ministers	Y	Y	Y
Prop of female legislators	Y	Y	Y
New-District Dummy	Y	Y	Y

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 49: Mean Male-Female Elector Ratio and Public Schools: Estimation with Alternative Standard Error Methods and Samples without outliers

	Robust/Sandwich Est	Jacknife	Without Outlier
Mean Male-Female Elector Ratio	1.612723**	1.612723	1.567154**
	(.6322761)	(1.017488)	(.6922074)
District FE	Y	Y	Y
Year FE	Y	Y	Y
Statewise Linear Trend	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y
Log-Total Private Schools	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y
Prop of Reserved Seats	Y	Y	Y
Prop of seats in State assembly	Y	Y	Y
Ratio of ruling coalition legislators	Y	Y	Y
Ratio of ministers	Y	Y	Y
Prop of female legislators	Y	Y	Y
New-District Dummy	Y	Y	Y
R^2	0.9959	0.9959	0.9959
Observations	1,126	1,126	1,067

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. For column 3 estimation, observations with corresponding studentized residual absolute value exceeding 2 are excluded as outliers from the sample. Standard errors clustered at the administrative district level at column 3. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 50: Male Dominated Seats Ratio and Public Schools: Estimation with Alternative Standard Error Methods and Samples without outliers

	Robust/Sandwich Est	Jacknife	Without Outliers
Male Dom Ratio	.0464273**	.0464273*	.0275736*
	(.214312)	(.0195921)	(.0158895)
District FE	Y	Y	Y
Year FE	Y	Y	Y
Statewise Linear Trend	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y
Log-Total Private Schools	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y
Prop of Reserved Seats	Y	Y	Y
Prop of seats in State assembly	Y	Y	Y
Ratio of ruling coalition legislators	Y	Y	Y
Ratio of ministers	Y	Y	Y
Prop of female legislators	Y	Y	Y
New-District Dummy	Y	Y	Y
R^2	0.9958	0.9958	0.9958
Observations	1,126	1,126	1,070

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. For column 3 estimation, observations with corresponding studentized residual absolute value exceeding 2 are excluded as outliers from the sample. Standard errors clustered at the administrative district level at column 3. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 51: Mean Male Female Elector Ratio and Number of Public Schools

	(1)	(2)	(3)	(4)
Mean Male Female Elector Ratio*redist	-.3713824***	-.3745953***	-.3478294***	-.35671***
	(.0732375)	(.0733032)	(.0736546)	(.0727882)
Mean Male Female Elector Ratio	1.392085***	1.410594***	1.689055***	1.571311***
	(.4815409)	(.4818689)	(.4895977)	(.4842509)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9958	0.9959	0.9958	0.9959
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 52: Mean Male Female Elector Ratio and Number of Public Schools (clustered standard error)

	(1)	(2)	(3)	(4)
Mean Male Female Elector Ratio*redist	-.3713824***	-.3745953***	-.3478294***	-.35671***
	(.1150994)	(.1152739)	(.1260854)	(.1278692)
Mean Male Female Elector Ratio	1.392085***	1.410594***	1.689055***	1.571311***
	(.4853835)	(.4875419)	(.5652813)	(.5760034)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9958	0.9959	0.9958	0.9959
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. Standard errors clustered at the administrative district level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 53: Median Male Female Elector Ratio and Number of Public Schools

	(1)	(2)	(3)	(4)
Median Male Female Elector Ratio*redist	-.3781893***	-.3805926 ***	-.3595789***	-.3636169***
	(.0786058)	(.0786724)	(.0791373)	(.0781071)
Median Male Female Elector Ratio	.0708603	.0704391	.0764435	.0650759
	(.2123388)	(.2123842)	(.2119394)	(.2089318)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9957	0.9957	0.9958	0.9959
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 54: Median Male Female Elector Ratio and Number of Public Schools(clustered standard error)

	(1)	(2)	(3)	(4)
Median Male Female Elector Ratio*redist	-.3781893***	-.3805926 ***	-.3595789**	-.3636169**
	(.1413876)	(.141736)	(.1556697)	(.1565735)
Median Male Female Elector Ratio	.0708603	.0704391	.0764435	.0650759
	(.3572367)	(.3586431)	(.3468283)	(.3339765)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9957	0.9957	0.9958	0.9959
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors clustered at the administrative district level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 55: Male Dominated Seats Ratio in District and Number of Public Schools

	(1)	(2)	(3)	(4)
Male Dom Ratio*redist	-.0554724***	-.0558866 ***	-.0516498**	-.0489053**
	(.0204505)	(.0204632)	(.0205216)	(.0202232)
Male Dom Ratio	.0611124***	.0614237***	.0575808***	.0515501**
	(.0214238)	(.0214337)	(.0214477)	(.0211799)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9957	0.9957	0.9957	0.9959
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

Table 56: Male Dominated Seats Ratio in District and Number of Public Schools(clustered standard error)

	(1)	(2)	(3)	(4)
Male Dom Ratio*redist	-.0554724**	-.0558866 **	-.0516498*	-.0489053*
	(.0258342)	(.0258726)	(.0267569)	(.0266232)
Male Dom Ratio	.0611124**	.0614237**	.0575808**	.0515501**
	(.0242839)	(.0242838)	(.0254532)	(.0242656)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Statewise Linear Trend	Y	Y	Y	Y
Log-Total Elig Electors	Y	Y	Y	Y
Log-Total Private Schools	N	Y	Y	Y
Dist-MF-Elec-Ratio	Y	Y	Y	Y
Prop of Reserved Seats	N	N	Y	Y
Prop of seats in State assembly	N	N	Y	Y
Ratio of ruling coalition legislators	N	N	N	Y
Ratio of ministers	N	N	N	Y
Prop of female legislators	N	N	N	Y
New-District Dummy	N	N	N	Y
R^2	0.9957	0.9957	0.9957	0.9959
Observations	1,126	1,126	1,126	1,126

Notes: Dependent variable is natural log of total number of government schools. The unit of analysis is administrative district. Standard errors reported in the parentheses. Standard errors clustered at the administrative district level. ***, ** and * represent significance at 1%, 5% and 10% levels respectively.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, *105*(490), 493–505.

Abadie, A., Diamond, A., & Hainmueller, J. (2014). *SYNTH: Stata module to implement synthetic control methods for comparative case studies.* Retrieved from `https://EconPapers.repec.org/RePEc:boc:bocode:s457334`

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), 495–510.

Abadie, A., Drukker, D., Herr, J. L., & Imbens, G. W. (2004). Implementing matching estimators for average treatment effects in Stata. *The Stata Journal*, *4*(3), 290–311.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*, *93*(1), 113–132.

Ahuja, A., & Chhibber, P. (2012). Why the poor vote in India: If i don't vote, i am dead to the state. *Studies in Comparative International Development*, *47*(4), 389–410.

Akee, R., Copeland, W., Costello, E. J., Holbein, J. B., & Simeonova, E. (2018). *Family income and the intergenerational transmission of voting behavior: Evidence from an income intervention* (Tech. Rep.). National Bureau of Economic Research.

Albouy, D. (2013). Partisan representation in congress and the geographic distribution of federal funds. *The Review of Economics and Statistics*, *95*(1), 127–141.

Alesina, A., & Rosenthal, H. (1989). Partisan cycles in congressional elections and the macroeconomy. *American Political Science Review*, *83*(2), 373–398.

Allers, M. A., & Kooreman, P. (2009). More evidence of the effects of voting technology on election outcomes. *Public Choice*, *139*(1-2), 159–170.

Alvarez, R. M., Hall, T. E., & Llewellyn, M. H. (2008). Are Americans confident their ballots are counted? *The Journal of Politics*, *70*(3), 754–766.

Alvarez, R. M., Katz, G., & Pomares, J. (2011). The impact of new technologies on voter confidence in Latin America: Evidence from e-voting experiments in Argentina and Colombia. *Journal of Information Technology & Politics*, *8*(2), 199–217.

Alvarez, R. M., & Nagler, J. (1995). Economics, issues and the perot candidacy: Voter choice in the 1992 presidential election. *American Journal of Political Science*, 714–744.

Andersen, S., Bulte, E., Gneezy, U., & List, J. A. (2008). Do women supply more public goods than men? Preliminary experimental evidence from matrilineal and patriarchal societies. *American Economic Review*, *98*(2), 376–381.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Ansell, B. (2014). The political economy of ownership: Housing markets and the welfare state. *American Political Science Review*, *108*(2), 383–402.

Ansolabehere, S., Gerber, A., & Snyder, J. (2002). Equal votes, equal money: Court-ordered redistricting and public expenditures in the American states. *American Political Science Review*, *96*(4), 767–777.

Ansolabehere, S., Hersh, E., & Shepsle, K. (2012). Movers, stayers, and registration: Why age is correlated with registration in the U.S. *Quarterly Journal of Political Science*, *7*(4), 333-363.

Ansolabehere, S., & Snyder, J. M. (2006). Party control of state government and the distribution of public expenditures. *The Scandinavian Journal of Economics*, *108*(4), 547–569.

Ansolabehere, S., Snyder, J. M., & Stewart, C. (2000). Old voters, new voters, and the personal vote: Using redistricting to measure the incumbency advantage. *American Journal of Political Science*, *44*(1), 17–34.

Ansolabehere, S., & Stewart, C. S. (2005). Residual votes attributable to technology. *The Journal of Politics*, *67*(2), 365–389.

Arcelus, F., & Meltzer, A. H. (1975). The effect of aggregate economic variables on congressional elections. *American Political Science Review*, *69*(4), 1232–1239.

Arceneaux, K. (2003). The conditional impact of blame attribution on the relationship between economic adversity and turnout. *Political Research Quarterly*, *56*(1), 67–75.

Arulampalam, W., Dasgupta, S., Dhillon, A., & Dutta, B. (2009). Electoral goals and center-state transfers: A theoretical model and empirical evidence from India. *Journal of Development Economics*, *88*(1), 103–119.

Asher, S., & Novosad, P. (2017). Politics and local economic growth: Evidence from India. *American Economic Journal: Applied Economics*, *9*(1), 229–273.

Auerbach, A. M. (2015). India's urban constituencies revisited. *Contemporary South Asia*, *23*(2), 136–150.

Bagues, M., & Esteve-Volart, B. (2016). Politicians luck of the draw: Evidence from the Spanish Christmas lottery. *Journal of Political Economy*, *124*(5), 1269–1294.

Barnes, T. D., & Cassese, E. C. (2017). American party women: A look at the gender gap within parties. *Political Research Quarterly*, *70*(1), 127–141.

Barnes, T. D., Tchintian, C., & Alles, S. (2017). Assessing ballot structure and split ticket voting: Evidence from a quasi-experiment. *The Journal of Politics*, *79*(2), 439–456.

Baskaran, T., Min, B., & Uppal, Y. (2015). Election cycles and electricity provision: Evidence from a quasi-experiment with Indian special elections. *Journal of Public Economics*, *126*, 64–73.

Beaman, L., Duflo, E., Pande, R., & Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. *Science*, *335*(6068), 582–586.

Beaman, L., Duflo, E., Pande, R., Topalova, P., et al. (2010). Political reservation and substantive representation: Evidence from Indian village councils. In *India policy forum 2010/11 volume* (Vol. 7, pp. 159–191).

Becker, S. O., & Ichino, A. (2002). Estimation for average treatment effects based on propensity scores. *The Stata Journal*, 358–377.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275.

Besley, T., Pande, R., & Rao, V. (2011). Just rewards? Local politics and public resource allocation in south India. *The World Bank Economic Review*, *26*(2), 191–216.

Bhatti, Y., & Hansen, K. M. (2012). The effect of generation and age on turnout to the European parliament–How turnout will continue to decline in the future. *Electoral Studies*, *31*(2), 262–272.

Bhavnani, R. (2016). The effects of malapportionment on economic development: Evidence from India's 2008 redistricting. Retrieved from https://faculty.polisci.wisc.edu/bhavnani/wp-content/uploads/2016/08/Bhavnani-Representation-and-development.pdf

Bhavnani, R. (2018). The effects of malapportionment on cabinet inclusion: Subnational evidence from India. *British Journal of Political Science*, *48*(1), 69–89.

Bhavnani, R. R. (2018). *India national and state election dataset.* Harvard Dataverse. Retrieved from https://doi.org/10.7910/DVN/26526 doi: 10.7910/DVN/26526

Biswas, S. (2014, May 2). Why does India's election take so long? *BBC News*. Retrieved from https://www.bbc.com/news/world-asia-india-27233843

Blais, A. (2006). What affects voter turnout? *Annual Review of Political Science*, *9*(1), 111-125.

Brady, H. E., Verba, S., & Schlozman, K. L. (1995). Beyond SES: A resource model of political participation. *American Political Science Review*, *89*(2), 271–294.

Bratton, K. A., & Ray, L. P. (2002). Descriptive representation, policy outcomes, and municipal day-care coverage in Norway. *American Journal of Political Science*, *46*(2), 428–437.

Brody, R. A. (1978). The puzzle of political participation in America. *The new American political system*, 287–324.

Brody, R. A., & Sniderman, P. M. (1977). From life space to polling place: The relevance of personal concerns for voting behavior. *British Journal of Political Science*, *7*(3), 337–360.

Brollo, F., & Nannicini, T. (2012). Tying your enemy's hands in close races: The politics of federal transfers in Brazil. *American Political Science Review*, *106*(4), 742–761.

Brunner, E., Ross, S. L., & Washington, E. (2011). Economics and policy preferences: Causal evidence of the impact of economic conditions on support for redistribution and other ballot proposals. *The Review of Economics and Statistics*, *93*(3), 888–906.

Burden, B. C., & Wichowsky, A. (2014). Economic discontent as a mobilizer: Unemployment and voter turnout. *The Journal of Politics*, *76*(4), 887–898.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, *90*(3), 414–427.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press.

Card, D., & Moretti, E. (2007). Does voting technology affect election outcomes? Touch-screen voting and the 2004 presidential election. *The Review of Economics and Statistics*, *89*(4), 660–673.

Chandra, K. (2004). *Why ethnic parties succeed: Patronage and ethnic head counts in India*. Cambridge University Press.

Chandra, K., & Potter, A. (2016). Do urban voters in India vote less? *Economic and Political Weekly*, *51*(39), 58–68.

Charles, K. K., & Stephens Jr, M. (2013). Employment, wages, and voter turnout. *American Economic Journal: Applied Economics*, *5*(4), 111–143.

Chattopadhyay, R., & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, *72*(5), 1409–1443.

Chhibber, P. (2002). Why are some women politically active? The household, public space, and political participation in India. *International Journal of Comparative Sociology*, *43*(3-5), 409–429.

Chhibber, P., & Kollman, K. (1998). Party aggregation and the number of parties in India and the United States. *American Political Science Review*, *92*(2), 329–342.

Chhibber, P., & Murali, G. (2006). Duvergerian dynamics in the Indian states: Federalism and the number of parties in the state assembly elections. *Party Politics*, *12*(1), 5–34.

Chin, A., & Prakash, N. (2011). The redistributive effects of political reservation for minorities: Evidence from India. *Journal of Development Economics*, *96*(2), 265–277.

Clots-Figueras, I. (2011). Women in politics: Evidence from the Indian states. *Journal of Public Economics*, *95*(7), 664 - 690.

Clots-Figueras, I. (2012). Are female leaders good for education? Evidence from India. *American Economic Journal: Applied Economics*, *4*(1), 212–244.

Cole, S. (2009). Fixing market failures or fixing elections? Agricultural credit in India. *American Economic Journal: Applied Economics*, *1*(1), 219–250.

Das, M. S., Jain-Chandra, S., Kochhar, M. K., & Kumar, N. (2015). *Women workers in India: Why so few among so many?* (No. 15-55). International Monetary Fund.

Debnath, S., Kapoor, M., & Ravi, S. (2017). The impact of electronic voting machines on electoral frauds, democracy, and development. Retrieved from `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3041197`

De La, O., & Ana, L. (2013). Do conditional cash transfers affect electoral behavior? Evidence from a randomized experiment in Mexico. *American Journal of Political Science*, *57*(1), 1–14.

Delwit, P., Kulahci, E., & Pilet, J.-B. (2005). Electronic voting in Belgium: A legitimised choice? *Politics*, *25*(3), 153–164.

Denemark, D. (2000). Partisan pork barrel in parliamentary systems: Australian constiuency-level grants. *The Journal of Politics*, *62*(3), 896–915.

Desai, Z., & Lee, A. (2018). Technology, choice, and fragmentation: The political effects of electronic voting in India. Retrieved from https://static1.squarespace.com/static/587fef5417bffcfdd148430d/t/5ad4af15aa4a9904672a5607/1523887893560/EVM-paper--Web.pdf

Diwakar, R. (2007). Duverger's law and the size of the Indian party system. *Party Politics*, *13*(5), 539–561.

Doherty, D., Gerber, A. S., & Green, D. P. (2006). Personal income and attitudes toward redistribution: A study of lottery winners. *Political Psychology*, *27*(3), 441–458.

Downs, A. (1957). *An economic theory of democracy.* New York: Harper.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, *50*(4), 1051–1079.

Duverger, M. (1954). *Political parties : Their organization and activity in the modern state.* London : Methuen ; New York : J. Wiley.

Election Commission of India. (1999, June). *Use of electronic voting machines during general elections, 1999.* Press Release. Retrieved from https://eci.nic.in/eci_main1/current/cp19990625.pdf

Fair, R. C. (1978). The effect of economic events on votes for president. *The Review of Economics and Statistics*, *60*(2), 159–173.

Falcao, V. L. (2009). Urban patterns of voting and party choices. *Economic and Political Weekly*, *44*(39), 99–101.

Filer, J. E., Kenny, L. W., & Morton, R. B. (1993). Redistribution, income, and voting. *American Journal of Political Science*, *37*(1), 63–87.

Fiorina, M. P. (1978). Economic retrospective voting in American national elections: A micro-analysis. *American Journal of Political Science*, *22*(2), 426–443.

Fletcher, E., Pande, R., & Moore, C. M. T. (2017). Women and work in India: Descriptive evidence and a review of potential policies. Retrieved from `https://www.hks.harvard.edu/sites/default/files/centers/cid/files/publications/faculty-working-papers/women_work_india_cidwp339.pdf`

Fraga, B. L. (2016). Redistricting and the causal impact of race on voter turnout. *The Journal of Politics*, *78*(1), 19–34.

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

Gajwani, K., & Zhang, X. (2014). *Gender and public goods provision in Tamil Nadu's village governments.* The World Bank. Retrieved from `https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-6854` doi: 10.1596/1813-9450-6854

Garmann, S. (2017). Election frequency, choice fatigue, and voter turnout. *European Journal of Political Economy*, *47*, 19–35.

Ghani, E., Kerr, W. R., & O'Connell, S. D. (2014). Political reservations and women's entrepreneurship in India. *Journal of Development Economics*, *108*, 138–153.

Ghani, E., Mani, A., & O'Connell, S. D. (2013). *Can political empowerment help economic empowerment? Women leaders and female labor force participation in India*. The World Bank. Retrieved from `https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-6675`

Glazer, A., & Robbins, M. (1985). Congressional responsiveness to constituency change. *American Journal of Political Science*, *29*(2), 259–273.

Gottlieb, J., Grossman, G., & Robinson, A. L. (2018). Do men and women have different policy preferences in Africa? Determinants and implications of gender gaps in policy prioritization. *British Journal of Political Science*, *48*(3), 611-636.

Gray, M., & Caul, M. (2000). Declining voter turnout in advanced industrial democracies, 1950 to 1997: The effects of declining group mobilization. *Comparative Political Studies*, *33*(9), 1091-1122.

Gupta, S. (2015). Telecommunications at the crossroads in India. *IIMB Management Review*, *27*(3), 196 - 208.

Hannagan, R. J., & Larimer, C. W. (2010). Does gender composition affect group decision outcomes? Evidence from a laboratory experiment. *Political Behavior*, *32*(1), 51–67.

Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review*, *90*(5), 1184–1208.

Hidalgo, F. D. (2012). *Renovating democracy: The political consequences of election reforms in post-war Brazil* (Doctoral dissertation, UC Berkeley). Retrieved from `https://escholarship.org/uc/item/15h624fw`

Imai, K., King, G., & Rivera, C. V. (in press, 2019). Do nonpartisan programmatic policies have partisan electoral effects? Evidence from two large scale experiments. *The Journal of Politics*.

Iyer, L., & Reddy, M. (2013). *Redrawing the lines: Did political incumbents influence electoral redistricting in the world's largest democracy?* Harvard Business School.

Jacob, S. (2014). Gender and legislative performance in India. *Politics & Gender*, *10*(2), 236–264.

Jensenius, F. R. (2015). Development from representation? A study of quotas for the scheduled castes in India. *American Economic Journal: Applied Economics*, *7*(3), 196–220.

Joshi, B., Pradhan, K. C., & Sidhwani, P. (2016). Urban voting and party choices in Delhi. *Economic and Political Weekly*, *51*(5), 64–72.

Kapoor, M., & Ravi, S. (2014). Women voters in Indian democracy. *Economic and Political Weekly*, *49*(12), 63–67.

Kasara, K., & Suryanarayan, P. (2015). When do the rich vote less than the poor and why? Explaining turnout inequality across the world. *American Journal of Political Science*, *59*(3), 613–627.

Katiyar, S. P. (2016). Gender disparity in literacy in India. *Social Change*, *46*(1), 46-69.

Keefer, P., & Khemani, S. (2004). Why do the poor receive poor services? *Economic and Political Weekly*, *39*(9), 935–943.

Khemani, S. (2003). *Partisan politics and intergovernmental transfers in India* (Vol. 3016). World Bank Publications.

Kimball, D. C., & Kropf, M. (2008). Voting technology, ballot measures, and residual votes. *American Politics Research*, *36*(4), 479–509.

Kinder, D. R., Adams, G. S., & Gronke, P. W. (1989). Economics and politics in the 1984 American presidential election. *American Journal of Political Science*, *33*(2), 491–515.

Kingdon, G. G. (2007). The progress of school education in India. *Oxford Review of Economic Policy*, *23*(2), 168–195.

Kingdon, G. G. (2017). The private schooling phenomenon in India: A review. Retrieved from http://ftp.iza.org/dp10612.pdf

Kramer, G. H. (1971). Short-term fluctuations in US voting behavior, 1896–1964. *American Political Science Review*, *65*(1), 131–143.

Kumar, S. (2009). Patterns of political participation: Trends and perspective. *Economic and Political Weekly*, *44*(39), 47–51.

Lau, R. R. (1982). Negativity in political perception. *Political Behavior*, *4*(4), 353–377.

Lau, R. R. (1985). Two explanations for negativity effects in political behavior. *American Journal of Political Science*, *29*(1), 119–138.

Leighley, J. E., & Vedlitz, A. (1999). Race, ethnicity, and political participation: Competing models and contrasting explanations. *The Journal of Politics*, *61*(4), 1092–1114.

Leveaux-Sharpe, C. (2001). Congressional responsiveness to redistricting induced constituency change: An extension to the 1990s. *Legislative Studies Quarterly*, *26*(2), 275–286.

Lott, J. R., Jr, & Kenny, L. W. (1999). Did women's suffrage change the size and scope of government? *Journal of Political Economy*, *107*(6), 1163–1198.

Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3–42.

Lunnerdal, R. (2014). *Rain, income shocks and voter turnout: Evidence from India* (Master's thesis, Department of Economics, Stockholm School of Economics, Sweden). Retrieved from `http://arc.hhs.se/download.aspx?MediumId=2139`

Lyons, W., & Alexander, R. (2000). A tale of two electorates: Generational replacement and the decline of voting in presidential elections. *The Journal of Politics*, *62*(4), 1014–1034.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, *107*(2), 407–437.

Masoodi, A. (2018, January 30). Only 9% of MLAs in India are women: Economic survey. *Livemint*. Retrieved from `https://www.livemint.com/Politics/NK0j2SUu8WRGgD8ULRBcRJ/Only-9-of-MLAs-in-India-are-women-Economic-Survey.html`

Mayer, A. K. (2011). Does education increase political participation? *The Journal of Politics*, *73*(3), 633–645.

Mayer, P. (2013). Gross violations of Duverger's law in India. *Studies in Indian Politics*, *1*(2), 179-201.

Miller, G. (2008). Women's suffrage, political responsiveness, and child survival in American history. *The Quarterly Journal of Economics*, *123*(3), 1287–1327.

Mishra, R. (2000). *Role of women in legislatures in India: A study.* Anmol Publications Pvt. Ltd.

Moraes, M. F. d. (2012). *Voting technology and political competition: Lessons from overlapping political races in Brazil* (Doctoral dissertation, Universidade de São Paulo). Retrieved from http://www.teses.usp.br/teses/disponiveis/12/12138/tde-16042013-143746/pt-br.php

Munshi, K., & Rosenzweig, M. (2009). *Why is mobility in India so low? Social insurance, inequality, and growth* (Working Paper No. 14850). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w14850 doi: 10.3386/w14850

Nadeau, R., & Lewis-Beck, M. S. (2001). National economic voting in US presidential elections. *The Journal of Politics*, *63*(1), 159–181.

Nowell, C., & Tinkler, S. (1994). The influence of gender on the provision of a public good. *Journal of Economic Behavior & Organization*, *25*(1), 25–36.

O'Neill, S., Kreif, N., Grieve, R., Sutton, M., & Sekhon, J. S. (2016). Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Services and Outcomes Research Methodology*, *16*(1), 1–21.

Pomares, J., Levin, I., & Alvarez, R. M. (2014). Do voters and poll workers differ in their attitudes toward e-voting? Evidence from the first e-election in Salta, Argentina. In *2014 electronic voting technology workshop/workshop on trustworthy elections (evt/wote 14).* San Diego, CA: USENIX Association. Retrieved from https://www.usenix.org/conference/evtwote14/workshop-program/presentation/pomares

Porto, A., & Sanguinetti, P. (2001). Political determinants of intergovernmental grants: Evidence from Argentina. *Economics & Politics*, *13*(3), 237–256.

Prillaman, S. A. (2017). *Strength in numbers: How women's groups close India's political gender gap* (Tech. Rep.). Retrieved from https://www.povertyactionlab.org/sites/default/files/publications/StrengthingNumbers_Oct2017.pdf

Radcliff, B. (1992). The welfare state, turnout, and the economy: A comparative analysis. *American Political Science Review*, *86*(2), 444–454.

Rai, P. (2011). Electoral participation of women in India: Key determinants and barriers. *Economic and Political Weekly*, *46*(3), 47–55.

Rai, S. (2006, September 15). Mobile phone proliferates, a hallmark of new India. *The New York Times*. Retrieved from https://www.nytimes.com/2006/09/15/business/worldbusiness/15cell.html

Rallings, C., Thrasher, M., & Borisyuk, G. (2003). Seasonal factors, voter fatigue and the costs of voting. *Electoral Studies*, *22*(1), 65–79.

Riker, W. H., & Ordeshook, P. C. (1968). A theory of the calculus of voting. *American Political Science Review*, *62*(1), 25–42.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5, Part 2), S71–S102.

Roseman, G. H., & Stephenson, E. F. (2005). The effect of voting technology on voter turnout: Do computers scare the elderly? *Public Choice*, *123*(1-2), 39–47.

Rosen, J. (2013). The effects of political institutions on womens political representation: A comparative analysis of 168 countries from 1992 to 2010. *Political Research Quarterly*, *66*(2), 306–321.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Rosenstone, S. J. (1982). Economic adversity and voter turnout. *American Journal of Political Science*, *26*(1), 25–46.

Schlesinger, M., & Heldman, C. (2001). Gender gap or gender gaps? New perspectives on support for government action and policies. *The Journal of Politics*, *63*(1), 59–92.

Schlozman, K. L., Burns, N., & Verba, S. (1994). Gender and the pathways to participation: The role of resources. *The Journal of Politics*, *56*(4), 963–990.

Schlozman, K. L., & Verba, S. (1979). *Injury to insult : Unemployment, class and political response.* Cambridge, MA : Harvard University Press.

Schneider, R., & Senters, K. N. (2018). Winners and losers of the ballot: Electronic vs. traditional paper voting systems in Brazil. *Latin American Politics and Society*, *60*(2), 41–60.

Self, S., & Grabowski, R. (2004). Does education at all levels cause growth? India , a case study. *Economics of Education Review*, *23*(1), 47–55.

Shapiro, R. Y., & Mahajan, H. (1986). Gender differences in policy preferences: A summary of trends from the 1960s to the 1980s. *Public Opinion Quarterly*, *50*(1), 42–61.

Sharma, M., & Sharma, M. (2008). *Paramilitary forces of India.* Gyan Publishing House.

Smets, K., & Van Ham, C. (2013). The embarrassment of riches? A meta-analysis of individual-level research on voter turnout. *Electoral Studies*, *32*(2), 344–359.

Solt, F. (2008). Economic inequality and democratic political engagement. *American Journal of Political Science*, *52*(1), 48–60.

Solt, F. (2010). Does economic inequality depress electoral participation? Testing the schattschneider hypothesis. *Political Behavior*, *32*(2), 285–301.

Sondheimer, R. M., & Green, D. P. (2010). Using experiments to estimate the effects of education on voter turnout. *American Journal of Political Science*, *54*(1), 174–189.

Southwell, P. L. (1988). The mobilization hypothesis and voter turnout in Congressional elections, 1974-1982. *Western Political Quarterly*, *41*(2), 273–287.

Stein, R. M., Vonnahme, G., Byrne, M., & Wallach, D. (2008). Voting technology, election administration, and voter performance. *Election Law Journal*, *7*(2), 123–135.

Sundaram, A., & Vanneman, R. (2008). Gender differentials in literacy in India: The intriguing relationship with womens labor force participation. *World Development*, *36*(1), 128–143.

Suzuki, M., & Chappell Jr, H. W. (1996). The rationality of economic voting revisited. *The Journal of Politics*, *58*(1), 224–236.

Thachil, T. (2014a). Elite parties and poor voters: Theory and evidence from India. *American Political Science Review*, *108*(2), 454–477.

Thachil, T. (2014b). *Elite parties, poor voters: How social services win votes in India.* Cambridge University Press.

Thomas, S. (1991). The impact of women on state legislative policies. *The Journal of Politics*, *53*(4), 958–976.

Uhlaner, C. J., Cain, B. E., & Kiewiet, D. R. (1989). Political participation of ethnic minorities in the 1980s. *Political Behavior*, *11*(3), 195–231.

Vaishnav, M. (2017). *When crime pays: Money and muscle in Indian politics.* Yale University Press.

Verba, S., Burns, N., & Schlozman, K. L. (1997). Knowing and caring about politics: Gender and political engagement. *The Journal of Politics*, *59*(4), 1051–1072.

Verba, S., & Nie, N. (1972). *Participation in America: Political democracy and social equality*. New York: Harper and Row.

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.

Wilkinson, S. I. (2006). The politics of infrastructural spending in India. *Department of Political Science, University of Chicago, mimeo*, *31*.

Witsoe, J. (2012). Everyday corruption and the political mediation of the Indian state: An ethnographic exploration of brokers in Bihar. *Economic and Political Weekly*, *47*(6), 47–54.