Technology, Choice, and Fragmentation: The Political Effects Electronic Voting in India

Zuheir Desai* Alexander Lee†

March 30, 2017

Abstract

Electronic voting technology is often proposed as translating voter intent to vote totals better than alternative systems such as paper ballots. We suggest that EVMs can also alter vote choice, and in particular the way in which voters register anti-system sentiment. This paper examines the effects of the introduction of electronic voting machines in India, the world's largest democracy, using a difference in difference methodology that takes advantage of the technology's gradual introduction. EVMs are associated with dramatic declines in the incidence of invalid votes, and corresponding increases in vote for minor candidates. There is ambiguous evidence for EVMs decreasing turnout, no evidence for increases in voter error or fraud, and no evidence that machines with an auditable paper trail perform differently from other EVMs. The results highlight the interaction between voter technology and voter protest.

*Department of Political Science, University of Rochester, Harkness Hall, Rochester, NY 14627. Email: zuheir.desai@rochester.edu.

†Department of Political Science, University of Rochester, Harkness Hall, Rochester, NY 14627. Email: alexander.mark.lee@rochester.edu.

1 Introduction

Social scientists have long been aware that there is a gap between the preferences of voters and recorded vote totals, and that voting technology may have important effects on elections. Electoral outcomes may be influenced by fraudulent voting (whether individual or collective), or the process of voting may be so confusing that it causes voters to cast invalid votes, vote randomly, or abstain from voting entirely.

The relative merits of different types of voting technologies has thus been an area of active research. In particular, many studies have sought to study the impact of the fastest-growing voting technology in the 21st century, electronic voting machines, or EVMs (sometimes called DRE machines because they directly record votes within the machine itself). EVMs are often contrasted with paper ballots, the oldest and simplest voting technology.[1] The past decade has seen a variety of studies of effect of EVMs, mostly in the United States (Ansolabehere and Stewart, 2005; Stewart, 2006; Card and Moretti, 2007; Knack and Kropf, 2003; Lott Jr, 2009; Hanmer et al., 2010) but also internationally (Allers and Kooreman, 2008; Fujiwara, 2015), and in laboratory settings (Niemi et al., 2009). In general, these studies have found that EVMs are associated with small reductions in invalid or residual voting relative paper ballots, varying effects on turnout, and no effect on partisan outcomes, though exceptions can be found to all these generalizations. A separate literature in computer science has focused on the security flaws of EVMs (Kohno et al., 2004; Feldman et al., 2006; Kumar and Walia, 2011; Wolchok et al., 2010), though this does not directly address the question of whether the technology is more or less secure than alternatives in practice.

We make two primary contributions to this literature. Firstly, we examine the effect of EVM introduction on vote choice, and in particular on the ways in which voters express discontent. If EVMs reduce the number of invalid votes, what do the voters who would previously have cast invalid ballots do after their introduction? We suggest that many of the voters who spoiled their ballots did so deliberately, to protest the political system or the set of candidates presented to them. After the introduction of EVMs, these voters should begin to vote for candidates of non-viable and/or non-establishment parties, leading to an overall increase in the level of political fragmentation.

Secondly, this is the first study to our knowledge to examine the effect of EVMs in India, or

[1] Two other technologies, analog voting machines and punch cards, are sometimes discussed in the American contest.

in any developing country. With the exception of Fujiwara's (2015) study of middle-income Brazil, existing studies have generally focused on fairly wealthy and well-institutionalized democracies, in general and the United States in particular, even though EVMs are now very common in all types of democracies. Not only is the voter fraud problem that EVMs are designed to address severe in poor democracies (Callen and Long, 2015; Crost et al., 2013) but poor countries possibly might find the use of EVMs to be more intimidating or confusing than those in rich countries. The effects of Indian EVMs are also of considerable direct significance: The inexpensive Indian model is the world's second most used voting technology of any kind (after paper ballots), and is currently in use for elections in Kenya, Namibia, Nepal and Bhutan.

India presents formidable challenges to any voting technology, due to its many isolated rural polling booths, large number of illiterate voters, and robust tradition of electoral fraud (Verma, 2005, 2009). India is the world's largest user of EVMs, which were introduced on a very limited pilot basis in 1998, in a subset of constituencies in 1999, and in all constituencies in 2004.

To identify the causal effect of EVMs in India, we take advantage of their gradual introduction, using a difference-in-difference design with both constituency and year fixed effects to accounts for confounding spatial and temporal effects. While the subset of constituencies in which EVMs were introduced early was not representative of the country, we show that the two types of constituencies have fairly similar trends in electoral behavior before the introduction of EVMs. This strategy is supplemented by the use of an set of time-varying controls, and the use of an alternative sample with less variance between the treatment and control conditions.

We find that the introduction of EVMs in India had substantial effects on voter behavior. Most dramatically, the rate of invalid voting was slashed by more than 90%. This stems directly from the design of Indian EVMs, which make it difficult or impossible to cast spoilt ballots, either by accident or as an act of protest.

This raises the question of what these voters did when they were no longer permitted to spoil their ballots. We find that the introduction of EVMs is associated in a sharp rise in voting for minor parties, leading to an increase in the overall fragmentation of the vote. This effect holds even within parties: vote for the left parties and the Bahujan Samaj Party rise or stay constant with EVMs in states where they are insignificant electorally, but fall with EVM introduction in states where these parties are contenders for office. Suggestive evidence indicates that this is a result of a conscious choice by voters, rather than a result of random interactions with the machines.

By contrast, we find little evidence for other possible second order effects of EVM introduction. EVMs are associated with reductions in voter turnout, but these results are not robust, and may stem from unobserved differences between constituencies. Similarly, there is no evidence that EVMs have altered the number of confused voters, as measured by the tendency of voters to choose based on ballot position.

While EVMs are often thought to effect fraud, we find no systematic pattern in either direction. There is no evidence for the introduction of EVMs favoring specific parties, including state incumbents, or the major national parties. Similarly, there is no evidence that turnout was reduced in constituencies thought to be prone to ballot box stuffing. These results focus on the introduction of EVMs that did not have any form of paper receipt or auditable record, which many security specialists have proposed as a necessary additional security measure to prevent fraud and promote public trust in electronic voting (Kohno et al., 2004). In 2014, India began to introduce voter-verified paper audit trail (VVPAT) EVMs in a small number of constituencies, following a supreme court order. We find that VVPAT machines have no effect on any major electoral outcome.

Section Two reviews the literature on the effects of EVMs on invalid voting, and develops a hypothesis for how this trend should affect voting for minor parties. Section Three discusses several additional hypotheses drawn from the literature. Section Four describes electronic voting technology and the way is was introduced in India, while Section Five discusses the data, models and identification strategy. Section Six describes the main results for invalid voting and minor party voting, while Section Seven discusses the potential consequences of this reduction for other election outcomes. Section Eight concludes with a discussion of the implications of the Indian experience.

2 Voting Technology and Vote Choice

A variety of hypotheses have been advanced about the effect of electronic voting machines, each with strong normative implications. Below, we examine existing findings on the rate of invalid voting, the dependent variable most common in the literature. We then develop a hypothesis for how changes in invalid voting could alter the distribution of valid votes.

2.1 Invalid Votes and Undervotes

In most elections, the number of votes recorded for candidates for a specific office is less than the number of voters who turned out to vote, leaving a number of "residual" or "invalid" vote. This is usually the result of voters casting ballots that are blank, spoiled, or do not express a clear preference among the candidates. In a multi-office election, it can also result from voters casting votes in some races, (such as those at the top of the ballot) while not bothering with other races. Invalid voting may reflect voter confusion with the voting technology, which leaves them to deviate from the proper voting procedure even though they wished to caste a valid ballot. Finally, invalid voting may reflect a deliberate choice of a voter who dislikes the candidates or is indifferent between them, but who has come to vote because of legal compulsion, social pressure, or interest in another office on the ballot.

From a normative perspective, high residual votes are usually though to be undesirable, both because they mean that some voters do not influence the election, and because they may create controversy over the result. McAllister and Makkai (1993) finds that the poor and immigrants cast more invalid votes. Fujiwara (2015) also argues that the voters casting residual votes are disproportionately poor.

Electronic voting might well be expected to reduce the rates of invalid or uninterpretable votes. Unlike paper ballots, on which a voter potentially might make a wide variety of marks, or punch cards, with their "hanging chads" electronic voting allows only a finite set of inputs. Indeed, the only way for invalid votes to be recorded in most systems is if there is a technical problem with the machine, or if the voter leaves the polling booth midway through the process.

Many discussions of EVMs have focused on changes in residual or invalid votes. Ansolabehere and Stewart (2005) compare voting technologies in the US. They find that punch card systems have higher residual vote rates for top-ballot contests, though DRE machines perform worse than paper ballots. Stewart (2006) finds that counties that shifted from punch cards to DRE machines in 2004 had fewer residual votes than in 2000. Allers and Kooreman (2008) examine municipal and national elections in the Netherlands. They find that residual vote rates decreased in national elections after the introduction of DRE machines. Fujiwara (2015) finds a significant decrease in residual votes in Brazilian elections after the introduction of EVMs. Knack and Kropf (2003) and Hanmer et al. (2010) found that American EVMs have lower residual voting rates than paper ballots and (especially) punch cards, though lower than mechanical lever machines. Lott Jr (2009) found that EVMs reduced residual voting for prominent races but raised it for other races.

As the brief review above should make clear, EVMs have often (though not always) been found to reduce levels of residual voting. However, these effects are usually quite modest in size, a product of the relative modesty of invalid voting in developed democracies (Ansolabehere and Stewart, 2005; Stewart, 2006; Allers and Kooreman, 2008). Sample effect sizes in studies with fixed effect designs are generally in the range of one percentage point.[2] To the extent that that invalid voting is associated with poverty, lack of political sophistication, and illiteracy, we should expect poor countries to have higher base rates of invalid votes (keeping ballot design constant), and thus greater scope for improvement. The rate of invalid voting was .3% at the 1997 UK general election (Great Britain, 2008, 112) and 1.9% in the 1998 Indian general election, though both were conducted with paper ballots.

2.2 Protest Voting and Fragmentation

Who benefits from the decreasing number of invalid votes? We consider several possibilities, each of which represents a distinct theory of what motivates voters casting invalid ballots. The first is that invalid voting is a result of *voter protest*: That the voters dislike all the candidates or the political system in general, but choose to go to the polling place (either because of social pressure, legal compulsion, or moral conviction) and spoil their ballots. This is in accord with the limited literature on spoiled ballots, which finds that they are associated not just with low voter sophistication but with anti-system sentiment (Power and Roberts, 1995; Power and Garand, 2007; Rosenthal and Sen, 1973). If voters spoil their ballots to protest, EVMs ended that option: in the absence of a "none of the above", voters using EVMs are forced to choose a candidate or not vote at all.

Faced with a "forced choice," voters who dislike the set of candidates or political institutions may respond by deliberately voting for candidates who are unviable or not affiliated with the major parties, in a form of a protest or expressive vote. Deliberate voting for non-viable candidates is a well-known phenomena. Though Duverger's Law predicts that a plurality system will result in the domination by two parties within the electoral district (Duverger, 1954), many voters in such systems vote for minor parties. Subsequent authors have claimed that voters may choose these parties for a variety of reasons, including for expressive reasons

[2] Fujiwara (2015) finds a much higher value (12% percentage points), since he studies a context where mandatory voting meant that many unwilling voters caste invalid ballots.

(Rosenthal and Sen, 1973) or for instrumental reasons to advance policy goals that voters could not achieve through major parties (Rosenstone et al., 1996). Voters could also be voting for minor parties as a signaling device in hopes of inducing downstream improvements in their preferred candidates performance (Kselman and Niou, 2011). Bélanger (2004) finds specifically that "third parties benefit from specific anti-party sentiment at the mass level"

Such a protest vote could take two forms with slightly different implications. Firstly, voters could choose to vote for a large party ideologically opposed to the political system as it currently exists, or rhetorically and organizationally distinct from the major parties, such as the National Front in France or the Australian Democrats in Australia. This existing work on such parties emphasizes the close association between their support and anti-system sentiment (Schedler, 1996; Pop-Eleches, 2010). While votes caste for these parties clearly express opposition to the political system as it currently exists, they are not always wasted, since they have some chance of winning the election.

Protest voters can only choose such national protest parties parties if elites choose to form them, their anti-system stance is credible, and can communicate that stance to voters. However, in many circumstances protest voters might not have a such a choice: either such a party does not exist, voters might be suspicious of the bone fides of any politician, or the voter may have no information about the non-major candidates on the ballot. Below, we suggest that our empirical case, India, may fit these criteria.

If no large protest party exists, protest voters will still avoid choosing from among candidates they regard as associated with the political system they dislike, while choosing (either arbitrarily or deliberately) among the parties not closely associated with the system. These parties may have genuine anti-system messages, simply be untainted with power and responsibility, or not engage in clientelist or dishonest voting practices of which the voter disapproves. This type of voting would increase the vote shares of minor parties or candidates relative to the front-runners or the establishment candidates, since successful candidates and parties will quickly be perceived as part of the the system.

If this version of the protest voting hypothesis is true, then we should expect to see an increase in the total vote-share of non-dominant, or minor candidates, after the introduction of EVMs. Since it increases the votes shares of the weakest candidates, EVM introduction should also be associated with increased fractionalization of the vote. This gives us Hypothesis One:

Hypothesis 1. *The introduction of EVMs should be associated with reductions in electoral fragmentation and increases in the vote for minor parties.*

3 Alternative Hypotheses: Error, Turnout and Fraud

3.1 Voter Error

The second explanation for invalid voting is that invalid voting is the result of *voter confusion*: That the voter intended to vote for a candidate, but was unable to mark the ballot in the legally prescribed manner—the equivalent of the "hanging chads" found in punch card systems. With paper ballots, a voter may easily cast an invalid vote–for instance by marking two candidates, or marking the wrong part of the ballot paper. With most EVMs, this possibility is removed: A voter must choose a candidate, and a "confused" voter would be one who presses the buttons incorrectly or arbitrarily. Voters, confused by EVM technology, might press buttons randomly or unskillfully, entering votes for candidates other than their preferred one. While such votes would be recorded as valid, they would not represent a normative improvement over spoiled ballots.

This effect would be intensified if EVMs are in fact *more* confusing than paper ballots, either because of some feature of the technology or the inherent difficulties of adopting a new system. While this is a common critique of EVMs among journalists, the empirical evidence is mixed. The literature on voter error seems to suggest that certain voting technologies are more prone to voter error than others, but to our knowledge none of these studies have focused on EVMs specifically. Dee (2007) looks at gubernatorial elections in California in 2003, where the ballot was constructed in a very complex fashion which served to confuse voters who used punch cards. Dee (2007) found that the users of punch cards were more likely to vote for bookend candidates compared with users of other technologies. Ansolabehere (2002) found that punch cards increase Wand et al. (2001) found that the so-called "butterfly" ballot (a punchcard format) led Florida voters to vote in error in the 2000 presidential election. In a series of laboratory experiments and expert interviews, Niemi et al. (2009) find that voters find some design features of different electronic voting systems "annoying, perplexing or disconcerting." However, they find that voters are not intimidated by EVMs.

How would a confused voter deal with an EVM? If they merely press buttons at random, the introduction of EVMs would lead to a corresponding increase in "random" voting." This may lead to votes being distributed to candidates evenly across the board, thereby increasing vote fragmentation within the constituency. Alternatively, buttons in certain positions may be more likely to be chosen by confused voters. Such patterns would accord with the large literature that shows that voters often cast ballots using arbitrary heuristics, such as the ballot position of the candidate (Krosnick et al., 2003; Ho and Imai, 2008; Alvarez et al., 2004). This gives us Hypothesis Two:

Hypothesis 2. *The introduction of EVMs should be associated with increases in vote for candidates in "favored" ballot positions.*

It is possible, of course, that EVMs would *reduce* voter confusion, in which case their introduction would lead to increases in valid voting, perhaps for parties favored by the poor and illiterate voters most likely to have been confused by paper ballots. This theory, developed by Fujiwara (2015), predicts that the introduction of EVMs should be associated with increases in vote for candidates in from parties favored by the poor. Note, however, that it seems reasonable that EVMs would be at least as confusing as paper ballots to voters who have no previous exposure to electronic voting (and previous exposure to paper ballots), even if we ignore the inherent simplicity of paper ballot technology. However, this effect might not hold if the requirements for valid paper ballots were previously complicated or involved.[3]

3.2 Turnout

The preceding discussion assumed that former protest voters would still go to the polls and make a choice after the introduction of EVMs. However, is certainly possible that voters unable to cast protest ballots would choose not to turn out at all. Note that there is an important objection to this theory: Voters who previously went to the polls to make a choice they knew will not influence the election may have strong social or personal reasons for casting that may be unrelated to the actual act of protest voting.

The actual evidence on the technology-turnout relationship is mixed. Card and Moretti (2007) find a negative correlation between turnout rates of Hispanics and EVM introduction in the 2004 election and claim this as a potential mechanism for the increase in Bush's vote share. Allers and Kooreman (2008) find that turnout decreases in Dutch municipal

[3] This appears to have been the case in Brazil before the introduction of EVMs (Fujiwara, 2015). There is also a strong relationship between poverty, ballot complexity and "informal voting" in Australia (McAllister and Makkai, 1993).

elections after the introduction of EVMs. Fujiwara (2015) finds no change in turnout after the introduction of EVMs in Brazil.

3.3 Electoral Manipulation

While increases in invalid or random votes or decreases in turnout are all problems for a democracy, they appear minor relative to the possibility of outright fraud. While some types of fraud, such as vote buying, fraudulent registration, voter impersonation, and voter intimidation, may operate under any kind of voting technology, others, such changes in recorded vote totals or the mass insertion of false votes, may be easier under some technologies than others. The manipulation of vote totals has been a part of political practice for many years in many countries, including Afghanistan (Callen and Long, 2015) and the Philippines (Crost et al., 2013).[4]

A large literature in computer science has found a startlingly large number of security issues that render EVMs susceptible to fraud. Kumar and Walia (2011) provide a thorough comparative overview of technological and security features EVMs around the world. According to them, most machines that are currently being used face significant security shortfalls. Wolchok et al. (2010) specifically look at Indian EVMs and point out the many security shortcomings in the machine, leaving it vulnerable to electoral manipulation. Kohno et al. (2004) and (Feldman et al., 2006) find that several brands of voting machines in use in the United States are susceptible to manipulation by both election official ("insiders") and ordinary voters ("outsiders"). However, empirical studies of EVM introduction have found little evidence of fraud (Fujiwara, 2015; Card and Moretti, 2007; Herron and Wand, 2007).

To counter the possibility of mass electoral fraud, security experts have often recommended the introduction of some sort of paper record of vote, to allow the totals reported by the machines to be crosschecked. Many US states now ban the use of EVMs without some sort form of paper trail. Theoretically, we should expect the introduction of this feature to enhance any security advantage of EVMs, and any disadvantage to be lessened.

Proponents of EVMs, by contrast, have declared that EVMs are less susceptible to fraud than other forms of election technology. In particular, the complexity of the machines, or their built-in security features, may make it more difficult to suddenly insert votes in large numbers (ballot box stuffing), which is easy with paper ballots. Indian EVMs, for instance, are built

[4] For general reviews of the electoral fraud literature see Lehoucq (2003) or Alvarez et al. (2009).

with an constraint of a maximum of five votes being cast a minute. Since the probability of law enforcement intervention increases over time, this feature makes brief "booth captures" more difficult. If this is correct, EVMs should be associated with reductions in turnout, not because voters are not turning out, but because the number of fraudulently cast ballots has been reduced. Unlike the the general decline in turnout discussed in the last section, this decline should be concentrated in regions that were more corrupt that others before (since the number of fraudulently cast ballots to be eliminated is larger). This gives us Hypothesis Three, that EVMs change the incidence of different types of electoral fraud:

Hypothesis 3. *A) The introduction of EVMs should be associated with changes in vote share for the candidates or parties best positioned to commit fraud. B) Any such effect will be attenuated by the introduction of voter verification technology. C) The introduction of EVMs should be associated with decreases in turnout in regions prone to electoral fraud.*

4 EVMs in India

4.1 Background

India is a federal parliamentary republic, and the world's largest democracy by population. Its electoral system closely mirrors that of Britain, with single-member districts whose members are elected using a first-past-the-post system. The directly elected lower house of national parliament, the Lok Sabha, contains 543 single member districts, each with a population of approximately two million, while state Vidhan Sabha districts are smaller. Since only the lower house of the legislature is directly elected and national, state and local elections are on different cycles, in most cases Indian voters only vote in one race in any given election.

State and national elections in India are administered by an independent national body, the Election Commission of India (ECI), which is granted wide powers over the bureaucracy and police during the election period. The ECI also supervises the creation of a register of eligible voters, enrollment in which is automatic. The commission is widely regarded as politically neutral and relatively efficient (McMillan, 2012), and takes extensive measures (not all of which can be discussed here) to guarantee the security of voters and the neutrality of the process.

Prior to 1998, all elections in India used paper ballots, with the names of the candidates and the symbols of their parties printed on the ballots. The voters marked the square next to the symbol of their preferred candidate and folded the ballot first vertically and then horizontally before putting it in the ballot box. The ballots were then counted in the presence of ECI officials and the parties, with "invalid votes" being those where no candidate preference could be assigned. Voters were not allowed to write in candidate names, or vote for a "none of the above" option.[5]

There were two main concerns with the use of traditional paper ballots in India. Firstly, a large portion of the electorate in India are not literate–48% of adults in 1991, and 35% in 2001. Despite extensive information campaigns by the ECI and the parties, India's very simple ballot structure, and the heavy use of party symbols, illiterate voters may sometimes have found it difficult to navigate and mark written ballots. This may plausibly have increased the likelihood to invalid or residual votes, especially among the very poor.

Secondly, despite the ECI's best efforts, post-independence India was no stranger to electoral fraud. While registration of nonexistent voters and vote buying have also been prominent, the most flagrant technique was the "capture" of polling booths. Armed operatives from one of the political parties would seize control of a polling station, intimidating officials, poll watchers and police, and stuff the ballot box with ballots favoring their candidate, sometimes modifying the officials' copy of the electoral register so that the turnout figures matched the number of ballots cast. This capture might last all day, but the stuffing process only required a few minutes, and the party operatives would often cast all possible votes and leave before the ECI and police had time to respond (Verma, 2005)

4.2 Electronic Voting in India

EVMs were first introduced in India in 1998 in a small number of constituencies in state assembly elections. Because of the success of this pilot program the ECI decided to implement the use of EVMs in the 1999 national elections. Forty five constituencies were selected in 17 states and 3 union territories. Importantly, these constituencies were *not* randomly selected. In general, they appear to have been more urban and wealthy than the country as a whole. The treated constituencies included all constituencies in Delhi, all but one constituency in Mumbai, and the larger cities in many other states, while not including any constituencies in the often unstable North-Eastern region of the country. All other constituencies continued to use paper ballots.

[5] "None of the above" (NOTA) became an option on Indian ballots in 2013.

Because of the perceived success of EVMs in this larger scale election, the ECI decided to use EVMs nationwide from 2004 onwards. In this paper will be focusing on three elections: 1998,1999 and 2004. To review, in 1998 none of the parliamentary constituencies (PCs) had EVMs, in 1999 only 45 PCs did, and in 2004 all PCs used EVMs.

The EVM adopted in India is manufactured by two government owned companies, Electronics Corporation of India (ECIL) and Bharat Electronics Limited (BEL). It differs considerably from the electronic voting machines in use in the United States, having a much simpler design, with only a basic set of programing instructions hardwired into the circuit board. The units are portable, can operate on battery power, and are (at approximately US $200 a unit) relatively cheap (Wolchok et al., 2010).

The basic design of the machine includes two main parts, a control unit and a ballot unit. By pressing a button on the control unit, the returning officer authorizes one vote from a particular ballot unit. The voter then presses the button on the ballot unit next to the symbol of their preferred candidate. This choice is then transmitted back to the control unit, where it is stored before the total votes for each booth are read out during the counting process.

There is space for 16 candidates on each ballot unit. If, for any constituency, there are more than 16 candidates, additional ballot units are linked together. Each polling booth can hold four ballot units and so up to 64 candidates can be accommodated. If there are more than 64 candidates, (a very rare event) then paper ballots are used.

Wolchok et al. (2010) suggest that Indian EVMs, like EVMs elsewhere in the world, suffer from security issues, and that a technically sophisticated group with access to the machines could modify the hardware in such a way as to produce desired results. These theoretical concerns parallel widespread rumors about attempts by the parties to modify the machines (Wolchok et al., 2010), and occasional reports of technical problems. Wolchok et al. (2010) are also critical of the ECI's procedures surrounding the storage of the machines, and skeptical that certain ECI security procedures (the random assignment of machines to booths, and the conduct of mock elections with machines before polling) address these concerns.

It should be noted, however that even if EVMs are vulnerable to a fraud, this does not mean that they are less vulnerable than alternative technologies. After all, a fraudster with access to stored boxes of paper ballots could produce a fraudulent result with considerably less effort and technical knowledge than that necessary to manipulate stored EVMs. For this reason, we will treat the security merits or demerits of EVMs as an open empirical question.

5 Data and Specifications

5.1 Data Structure and Sources

The primary data being used for this project is a three year panel (1998, 1999 and 2004) of Lok Sabha elections, though we will also examine some trends involving earlier and later elections.[6] There are 543 elected seats in the Lok Sabha, so our sample contains 1629 constituency-year observations. The data is taken from Kollman et al. (2011) supplemented by information taken directly from the Election Commission's reports. India had no constituency reapportionment between 1977 and 2004, so the unit of analysis remains constant.[7]

5.2 Basic Specification

There are obvious difficulties in interpreting even a strong association between presence of EVMs on political outcomes as causal since "treated" constituency-years are both later in time, and (within years) disproportionately urban and wealthy. For this reason, we use a difference-difference design to estimate the causal effect of electronic voting. The econometric specification is as follows:

$$y_{it} = \phi_i + \delta_t + \beta EVM_{it} + \epsilon_{it},$$

where y_{it} is the dependent variable, EVM_{it} is treatment assignment, δ_t is a vector of time fixed effects, ϕ_i is a vector of constituency fixed effects and ϵ_{it} is a noise term. Under the assumption that δ_t is the same for all $EVM_{it} = 1$ and $EVM_{it} = 0$, the treatment effect is $\mathbb{E}(y_{it}|EVM_{it} = 1) - \mathbb{E}(y_{it}|EVM_{it} = 0) = \beta$. In the next section, we will discuss the assumption that time trends in election dynamics are similar across the two groups.

5.3 Additional Tests and Models

To the further guard against confounding by differing time trends, some models include a set of time-varying controls. These include the proportion of the vote for each of the two largest national parties (the BJP and the Congress), the total number of candidates within

[6]The 1998 election was inconclusive, and resulted in a fragile coalition government led by the Hindu nationalist BJP. The snap 1999 elections resulted in a stronger BJP-led coalition government, which lost the 2004 elections to a coalition led by the Congress party.

[7]A reapportionment between the 2004 and 2009 elections affects comparability between these two periods.

the constituency, the Herfindahl index of party vote shares in the constituency, the margin of victory of the winning candidates, and turnout (in models where turnout is not the dependent variable).

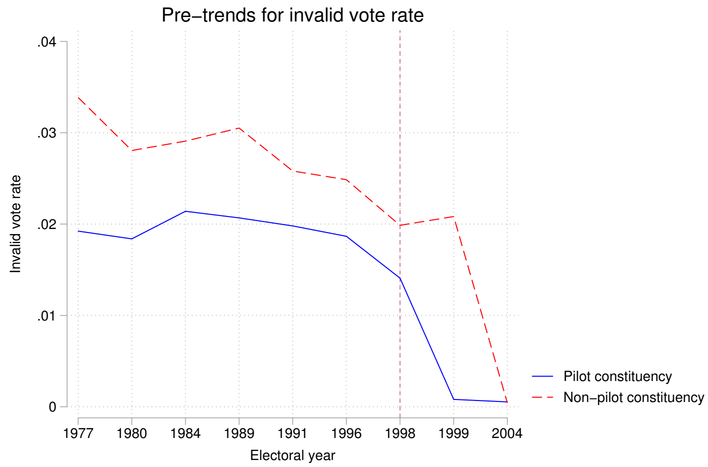
A second set of tests uses phase-year fixed effects. In each national election, the ECI divides the constituencies into 4-5 phases, each with its own polling day. This is done to make the task of the ECI easier and to ensure that adequate security measures can be implemented. Altogether, the data from the three elections in our study can be divided in 14 phase years. Disaggregating the data in this fashion allows us to account for confounders specific to particular elections days (such as rain, season or national news events).

A final set of tests limits the sample to only urban constituencies, defined as those where the the largest district in the constituency had an urban population of more than 40% at the 2001 census.[8] Overall, 27 of the 45 early treatment constituencies met this criteria, as did 57 of the 398 late treatment constituencies. While this approach does not create even quasi-random assignment, it does reduce the observed (and, possibly, unobserved) differences in political trends between the two groups.

Finally, standard errors, unless otherwise stated, have been clustered at the state-year level. The unit of observation in our econometric model is the constituency-year. However, constituencies are grouped at the state-level for elections, and have many common variables that affect them such as economic shocks, coordinated political campaigns, overall political context etc. Furthermore, in certain models, we also interact the EVM variable with data that is grouped at the state-level. Ignoring this grouped data problem can lead to inconsistent standard errors (Bertrand et al., 2004). Thus, whenever possible, we account for this problem by clustering standard errors at the state-year level. In general, clustering increases the standard errors, and we only report this more conservative estimate. In certain models, which we note, the clustered standard errors are smaller than OLS standard errors, and so we only report the more conservative OLS standard errors. In a few models (noted in the Table captions), there are very few observations and clusters and so we report OLS standard errors for these models.

[8] The census defines urban status in a relatively stringent fashion, coding many large and urbanized villages as rural.

Figure 1: Pre-trends for invalid vote rate



Notes. The blue solid line plots the average invalid rate in all pilot constituency across election years while the red dashed line plots the average invalid vote rate in non-pilot constituencies. The year 1998 marks the last election before the introduction of EVMs. Thus, 1999 is the first post-treatment year for the pilot constituencies. In the year 2004, the non-pilot constituencies also used EVMs.

6 Results

6.1 Pretrends

The identifying assumption in a difference-in-differences model is that the treatment group and the control group have parallel trends. This means that, absent the intervention, average changes would be the same across the treated and non-treated units. While this assumption cannot be directly tested, we provide evidence that pre-treatment trends were similar between the treatment and control groups.

Our identifying strategy relies on the fact that while all constituencies used paper ballots in 1998, 45 pilot constituencies used EVMs in 1999. Thus, these 45 pilot constituencies serve as our treatment group, and the rest are grouped as a control. Figure 1 shows over-time trends for invalid voting (calculated as the difference between turnout and the number of valid votes for candidates as a proportion of turnout) for the early-treatment and control groups.[9] The trends tend to move together over time, with the relatively wealthy early-treatment constituencies tending to have lower levels of invalid voting. This evolution is

[9] A small number control-group constituencies, mostly in Kashmir and Northeast India, either did not have elections or are missing data for certain years before 1996. These constituencies have been dropped from the pretrend analyses to make over-time comparisons valid.

interrupted in 1999, when invalid voting rates fell sharply in the early-treatment states, with not corresponding effect in the control states. The decline in invalid voting is distinctly noticeable. These findings are supported by Table A.15, which compares the effect of the intervention in the treatment districts to the lagged and led treatment, which should have no effect. The first column of that table compared the effect on invalid vote rates in the pilot constituencies in the treatment year (1999) versus all other election years. All but one of the placebo treatments (with 1989 being the exception) are statistically insignificant, and all are much smaller in sign. These results suggest that the parallel-trends assumption is justified for invalid votes.

In an online appendix, we show that pilot and non-pilot constituencies have similar trends not just in invalid voting, but in other outcomes as well. We show results for vote for state incumbents, vote for local incumbents, votes for the BJP and Congress, margin of victory and number of candidates.[10] In Section 7.2 we discuss the one outcome (turnout) for which the two groups do not appear to have parallel pretrends.

6.2 Invalid Voting

Table 1 shows results for the most common dependent variable in the literature, the rate of invalid votes. The introduction of EVMs is associated with a large, negative and statistically significant reduction in the rate of invalid voting. This holds in the naive regression (Column 1), the simple difference-in-differences model (Column 2) the model with controls (Column 3) and the urban subsample (Column 4). The substantive size of the coefficient is remarkable: The estimated effect of EVMs on invalid voting is higher than the mean level of invalid voting in 1998. Overall, the rate of invalid voting in India declined from 1.93% in 1998 to .04% in 2004.

These results indicate that EVMs succeed in reducing the rate of invalid voting—in fact, invalid voting appears to have been virtually eliminated. This result stems directly from the design of the machine: Indian EVMs, with their finite menu of buttons, make it almost impossible to cast an invalid ballot, whether deliberately or accidentally.

[10] Note that the massive year to year variation in the number of candidates seems to track very closely changes in the real value of the electoral deposit.

Table 1: Effects of EVM's on invalid vote rates

	(1)	(2)	(3)	(4)	(5)
EVM	-0.0196***	-0.0173***	-0.0185***	-0.0194***	-0.0169**
	(0.00190)	(0.00253)	(0.00277)	(0.00324)	(0.00491)
Year FE		Yes	Yes		Yes
Constituency FE		Yes	Yes	Yes	Yes
Controls			Yes	Yes	
Phase-year FE				Yes	
Constant	0.0201***	0.0243***	0.0143	0.0194	0.0255**
	(0.00192)	(0.00460)	(0.0133)	(0.0120)	(0.00796)
N	1629	1629	1628	1601	252
R^2	0.456	0.700	0.735	0.771	0.676

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on invalid vote rates in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and turnout rate, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

6.3 Effects on "enfranchisement"

The claim that this fall in invalid voting is normatively important hinges on the assumption that the voters who previously cast invalid votes now caste valid votes. If voters who previously caste invalid votes simply stoped turning out after the introduction of EVMs, the reform would have no political effect, and only a very doubtful normative value. This concern is especially valid because it appears that turnout may decrease with the introduction of EVMs (see below).

However, the results seem to suggest that EVMs had a net positive effect in terms of "enfranchisement", with the decline in invalid votes swamping the poorly estimated decline in turnout. These results are presented in Table A.1, which shows the results of a difference-in-differences model with the number of valid votes cast in a constituency as a dependent variable. The coefficient on EVMs is positive in all models, although it is only significant once accounting for other time-varying factors such as total number of electors and turnout rates. Thus, we can reasonably conclude that EVMs have an overall net positive effect on enfranchisement of Indian citizens since they resulted in less votes being disregarded as invalid. In the next section we analyze the consequences of this decline in invalid voting.

Note that this estimated effect may represent an underestimate of the enfranchising effects of EVMs. To the extent that EVMs successfully prevent ballot box stuffing (their primary intended purpose) they should lead to a reduction in the number of legally valid but fraudulent votes, which would lead to EVM introduction having a spurious "disenfranchising" effect. The fact that the number of valid voters increases regardless is strong evidence that EVM introduction led many voters who would previously have cast invalid votes to caste valid ones.

6.4 H1: Increasing Fragmentation

If many voters previously casting spoilt ballots now cannot do so, they many now wish to cast valid votes in such a way as to protest the political system as a whole. The Indian state, with its widespread corruption and inefficiency, might plausibly stimulate such feelings. However, at least in the late 1990s, there existed no major party that could unambiguously be described as an anti-system party. At least at the national level, there existed a high degree of ideological agreement between the parties on many issues, with a general tendency towards a "moderate pluralism" that supports the legitimacy of the regime (Sridharan and Varshney, 2001). Moreover, parties have difficulty becoming electorally viable without becoming involved in the very system of funding irregularities, violence, over-centralization, and clientelism that might repel a protest voter (Vaishnav, 2011).[11]

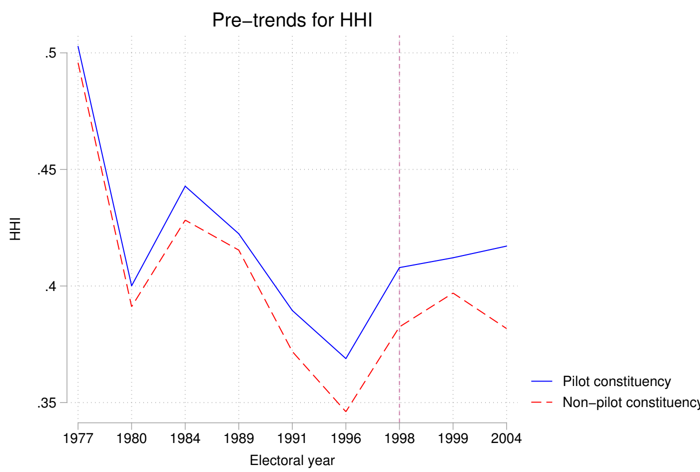
Indian, protest voters, then, must vote for small parties. Since these candidates have very small vote shares, an even increase would lead to an increase in HHI. We investigate these hypotheses in this section. We first use the Herfindahl-Hirschman Index (HHI) as a dependent variable and show that EVMs do lead to an increase in vote fragmentation. We then investigate whether this could be possible because of an increase in votes cast by error, or whether this is due to increased votes cast for unviable candidates. We suggest that the latter is the case, given that EVMs had a positive effect on minor party and independent vote shares, and that there is little evidence for random voting.

The HHI is simply the sum of the squared vote-shares of all candidates in a constituency.

[11] The recent history of the Aam Aadmi Party presents an interesting case study of the difficulties in creating a successful anti-system party in India.

It is an indicator of the relative fractionalization in the electoral district. Thus, in the case of a small number of parties dominating the election, the HHI is close to 1, and if there are many parties with similar vote shares, then it is close to 0. As before, we first present evidence on the pre-trends of HHI in Figure 2. The trends seem to move together, with the relatively urban and richer pilot constituencies showing more concentration of vote-shares and the non-pilot, rural constituencies being more competitive (in terms of the fragmentation of the vote). The trends are relatively parallel for pilot and non-pilot constituencies, and the effect of lags and leads of the treatment is statistically insignificant (Table A.14).

Figure 2: Pre-trends for HHI



Notes. The blue solid line plots the average HHI in all pilot constituency across election years while the red dashed line plots the average HHI in non-pilot constituencies. The year 1998 marks the last election before the introduction of EVMs. Thus, 1999 is the first post-treatment year for the pilot constituencies. In the year 2004, the non-pilot constituencies also used EVMs.

Table 2 shows the effect of EVMs on the HHI. EVMs have a negative effect on HHI in all models considered. The smaller the HHI, the more the fragmentation within the electoral district (as votes are divided among more candidates). Thus, a negative coefficient indicates that EVMs increase fractionalization. However, the coefficient in column 2, the standard difference-in-differences model, is not significant. This is because HHI is affected by other time varying variables such as the number of candidates in the district. After having controlled for these variables, columns 3 and 4 indicate that EVMs had a significant negative effect on the HHI. This effect is robust to phase-year controls.

This increase in fragmentation is paralleled by increases in the vote for minor candidates.

Table 2: Effects of EVM's on HHI

	(1)	(2)	(3)	(4)	
EVM	-0.00331	-0.0148	-0.0274**	-0.0252**	-0.0179
	(0.0213)	(0.0110)	(0.00819)	(0.00819)	(0.0215)
Year FE		Yes	Yes		Yes
Constituency FE		Yes	Yes	Yes	Yes
Controls			Yes	Yes	
Phase-year FE				Yes	
	0.390***	0.399***	0.384***	0.395***	0.457***
	(0.0127)	(0.0366)	(0.0819)	(0.0813)	(0.0172)
N	1629	1629	1628	1601	252
R^2	0.000	0.698	0.763	0.769	0.568

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on the HH index in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory and turnout rate, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

We use several measures of a "minor candidate", all based on a certain threshold of vote share that the candidate managed to gain in the election.[12] As such, it is an *ex post* measure of minor candidates.[13] We ran 5 different models where we designated candidates that received less than 1%, 2.5%, 5%, 7.5% and 10% of votes in the district as minor candidates. We looked at the effects of EVMs on the sum of the vote shares of such candidates. These candidates included independents and candidates of minor local or regional parties. However, the many strong regional party candidates were effectively excluded from the analysis. Note that the average effective number of parties in the 1998, 1999 and 2004 Lok Sabha elections was

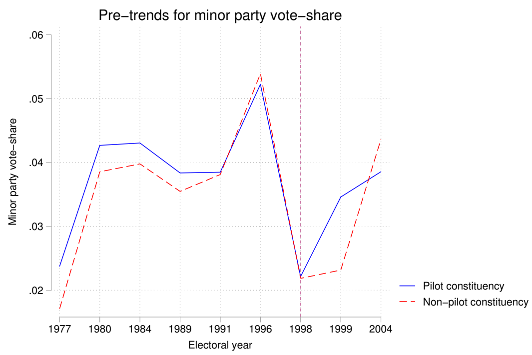
[12] While India does classify candidates (into national parties, state parties, unrecognized parties, and independents) these classifications are at best an imperfect guide to the viability and reputation of the candidates: Many national party candidates win tiny vote shares (and receive no help from the party organization), while many independent come in first or second, and have strong links to one or other of the parties. The boundary between independents and unrecognized parties is especially hazy.

[13] An ex-ante measure of minor party status is difficult because of the high level of variation in party viability from constituency to constituency and from election to election in India.

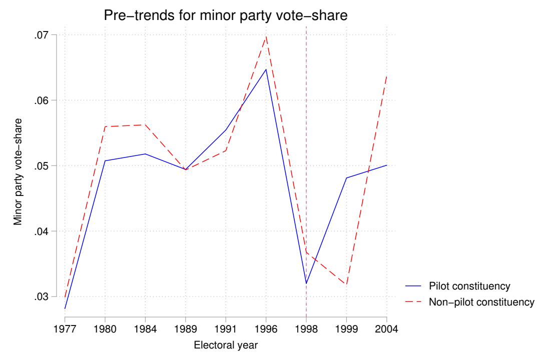
2.69, while the average vote share of the candidate coming third in the election was 10.94 %. Thus, all our measures of minor candidates captures the vote shares of candidates that placed on average worse than 3rd in the election, and most include candidates who are even less relevant.

Figure 3: Pre-trends for vote-share of "minor parties"

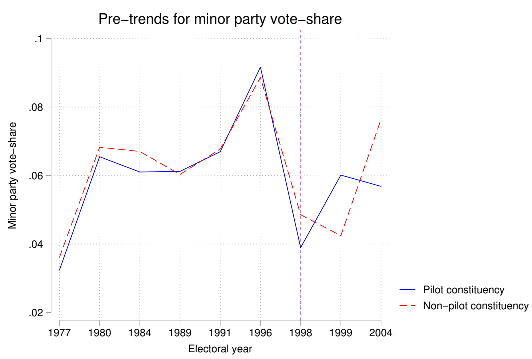
(a) < 2.5%



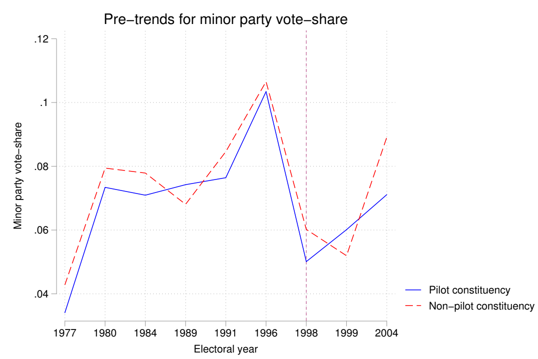
(b) < 5%



(c) < 7.5%



(d) < 10%



Notes. The blue solid line plots the average summed minor party/candidate vote share (with different thresholds) in all pilot constituency across election years while the red dashed line plots the average summed minor party/candidate vote shares in non-pilot constituencies. The year 1998 marks the last election before the introduction of EVMs. Thus, 1999 is the first post-treatment year for the pilot constituencies. In the year 2004, the non-pilot constituencies also used EVMs.

As with other dependent variables, we first present the pre-trends of the early-treatment and control groups. Figure 3 shows how these trends evolved before and after the treatment year in 1999 for all 5 measures of minor party vote-share. All pre-trends appear parallel. The jump in the average vote-share of minor candidates in pilot constituencies in the year 1999 is perceptible, as is the jump in the late-treatment constituencies (or the non-pilot constituencies) in 2004, when all constituencies used EVMs. The claim that pilot and non-pilot constituencies had parallel trends with regards to minor-candidate vote-shares is further bolstered by the results in Table A.15, where the coefficients on the interaction of pilot constituency and treatment leads are insignificant in all models.

The results of the difference-in-difference analysis is presented in Table 3. Overall, we estimate that EVM introduction was associated with a 2-3 percentage point increase in the vote share of minor candidates in the Lok Sabha elections. There are four separate panels in the table. The first panel presents the result of the base difference-in-difference analysis of all 5 models. Every model has a significant positive coefficient on EVM. This positive effect of EVMs on minor party vote-shares is robust and remains significant after the addition of control variables, the disaggregating of the data by phase-year, and running the vanilla regression on the urban sub-sample.

One potential interpretation of these results is that they reflect random voting: that is, confused vote for each candidate with an equal probability. Since there are many more electorally unviable candidates than there are major candidates in a single member electoral system, these minor candidates gain more vote share, thereby increasing fractionalization as well as votes for minor candidates (and all candidates in general). However, we provide evidence purely random voting claim. In Table A.2 and in Table A.3 we show that EVMs have no effect on the summed vote share of candidates receiving less than 0.5% of votes, or on the summed vote share of candidates receiving less than 20% of votes. Thus, the effect of EVMs is not uniform across the candidates, suggesting that voters may be focusing on the set of minor candidates with some level of name recognition, rather than all names on the ballot.

6.5 Within Party Results

The consequences of the increase in protest voting can be illustrated for two political parties or groups of parties that might appeal to some groups of protest voters. The Bahujan Samaj Party was founded upon the grievances of voters from the formerly "untouchable" Scheduled Castes, and is vocal in its belief that other political parties cannot properly represent these voters (Chandra, 2007). Similarly, India's major left wing parties[14] are all vocal in their condemnation of the other parties, who they see as tools of global capitalism and rural

[14] Communist Party of India, Communist Party of India (Marxist) and Communist Party of India (Marxist-Leninist) (Liberation), Revolutionary Socialist Party and Forward Bloc

Table 3: Effects of EVM's on Minor party vote shares

Panel A: Basic diff-in-diff

	(1) $< 2.5\%$	(2) $< 5\%$	(3) $< 7.5\%$	(4) $< 10\%$
EVM	0.0140^{***}	0.0257^{***}	0.0321^{***}	0.0224^{**}
	(0.00362)	(0.00527)	(0.00607)	(0.00775)
Constant	0.0109^{+}	0.0464	0.0466	0.0468
	(0.00563)	(0.0285)	(0.0288)	(0.0296)
N	1629	1629	1629	1629
R^2	0.622	0.608	0.570	0.548

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel B: Includes controls

	(1) $< 2.5\%$	(2) $< 5\%$	(3) $< 7.5\%$	(4) $< 10\%$
EVM	0.0202^{***}	0.0314^{***}	0.0373^{***}	0.0253^{***}
	(0.00313)	(0.00402)	(0.00583)	(0.00686)
Constant	0.00101	0.0716^{+}	0.0841^{+}	0.138^{**}
	(0.00977)	(0.0403)	(0.0476)	(0.0503)
N	1628	1628	1628	1628
R^2	0.713	0.666	0.610	0.592

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel C: Includes phase-year FE

	(1) $< 2.5\%$	(2) $< 5\%$	(3) $< 7.5\%$	(4) $< 10\%$
EVM	0.0208^{***}	0.0317^{***}	0.0366^{***}	0.0247^{**}
	(0.00299)	(0.00404)	(0.00655)	(0.00769)
Constant	0.00823	0.0714^{+}	0.0956^{*}	0.144^{**}
	(0.00956)	(0.0383)	(0.0466)	(0.0507)
N	1601	1601	1601	1601
R^2	0.727	0.680	0.625	0.600

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel D: Urban subsample

	(1) $< 2.5\%$	(2) $< 5\%$	(3) $< 7.5\%$	(4) $< 10\%$
EVM	0.0159^{***}	0.0250^{***}	0.0310^{***}	0.0212^{+}
	(0.00450)	(0.00609)	(0.00871)	(0.0107)
Constant	0.0207^{***}	0.0180^{***}	0.0175^{***}	0.0493^{+}
	(0.00254)	(0.00255)	(0.00442)	(0.0282)
N	252	252	252	252
R^2	0.579	0.560	0.469	0.456

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. Panel A conducts a basic diff-in-diff regression for all 5 measurements of minor candidate vote share on EVM, Panel B includes controls, Panel C replaces electoral year fixed effects with phase-year fixed effects, and Panel D focuses on the urban subsample (without controls). All standard errors have been clustered at the state-year level.

"feudalism" (Communist Party of India , Marxist).

However, the nature of these parties differs from state to state. In many states, both the BSP and the left resemble classic protest parties, winning few elections, dispensing no patronage, and depending upon the support of the poor. In the states where they are strongest, however, (Uttar Pradesh for the BSP, West Bengal, Kerala and Tripura for the left) they are the establishment, frequently winning elections and forming the state government—in 1999 the Left Front had governed West Bengal for 22 years. As a consequence, the parties in these areas include many more "normal" office-seeking politicians, win support from a wide variety of social groups, and form mutually profitable relationships with the states' business communities. For these reasons, voters in these states with a grudge against the party system would be unlikely to choose these parties.

Table A.4 and Table A.5 show that the effect of EVMs on support for the BSP depends on what state the BSP candidate is competing in. While their vote share decreases in their strongholds in UP (where they identified with the establishment) it increases in other (where their rhetoric may qualify the BSP as a protest party). Similarly, Table A.6 and Table 3 show that the vote share of a candidate affiliated to a Communist party is falling in states where these parties are well-established, and increasing in states where the Communist parties are minor parties. All of these results except the last are statistically significant in the core diff-in-diff model.

7 Results for Alternative Hypotheses

7.1 H2: Voter Confusion and Voter Error

Do EVMs lead to increases in valid votes cast in error, even as the number of invalid votes decreases? While an association between EVMs and voter error is certainly plausible, it is difficult *a priori* to know what forms this error would take. Two possibilities are discussed here, both based on the intuition that confused voters would be more likely to buttons in certain positions than in others, even if they are unfamiliar with the candidate at that position.

The first possibility is that EVMs would increase "donkey voting:" choosing the candidate first on the machine. A large literature has shown that voting based on ballot order is a common heuristic among voters (Krosnick et al., 2003; Ho and Imai, 2008), and we have some anecdotal evidence that voters in India use EVMs in this way (Banerjee, 2015). Ballot order in India is not random, with the candidates of nationally recognized parties being listed first (in alphabetical order in the state's official language) followed by the candidates of state recognized parties, and then all other candidates. Using this structure, we were able to reconstruct the ballot order for all the Hindi-speaking states during our time period. Table A.8 shows the relationship between EVMs and voting for the first placed candidate. There is little or no relationship between EVMs and such voting, in the basic difference-in-difference model, and the effect remains statistically insignificant in all models.

Another common voting error is to caste votes for candidates immediately above or below their actual preferred choice (Alvarez et al., 2004). Table A.10 and Table A.11 examine whether such "proximity effects" are exacerbated by EVMs. There is no evidence that candidates immediately above or below the winner on the ballot benefit from the introduction of EVMs.

Another variation of the ballot-order effect hypothesis is specific to the Indian context. Recall that the ballot units of Indian EVMs can only accommodate 16 candidates, and that additional candidates require the use of multiple ballot units. Presented with multiple ballot units, a voter might try to select candidates from both (thus recording a vote on only one), or choose only from a single unit. This would tend to favor the candidates in prominent positions on the second or subsequent units, usually independent or minor party candidates who would ordinarily gain very few votes.

Table A.9 has as its dependent variable the vote for the 17th candidate on the ballot—the top candidate on the second unit. Since elections with such a large number of candidates are rare, these models have very few observations. The estimated effect of being places 17th is large and positive: About 2 percentage points, substantially higher than the mean vote share of these candidates. However, the effect is badly estimated (understandably, given the small number of observations and the number of covariants in the difference-in-difference model) and is nowhere near statistically significant at traditional levels. However, Table A.12 shows the large positive effect does not appear to extend to other high ballot positions on the second unit and low ballot positions on the first unit (e.g. the 15th, 16th, 18th and 19th place candidates).

Another variant of the voter confusion hypothesis is that EVMs *reduce* confusion. This should increase the vote share of parties favored by the poor, since poor voters are most likely to become confused by ballot structure and caste invalid ballots. Fujiwara (2015) finds that the introduction of EVMs, and the consequent decline in invalid ballots, increased votes for parties favored by the poor. There are several reasons to be skeptical that this finding would apply to India, not least because poor voters India can be found among the supporters of all political parties due to clientelist distribution (Thachil, 2014). In addition, India's ballot structure (in contrast to Brazil's) was extremely simple under the paper ballot system, making procedural errors by illiterate voters less likely than they would have been in Brazil.

Table A.17 examines whether the introduction of EVMs impacted the support of parties favored by the poor. These "poor voted parties" were determined by looking at data on party voters from the Indian National Election Study, taken from Shastri et al. (2009) and Palshikar et al. (2014). Parties that received a plurality of votes from voters classified as "very poor" were coded as a "poor party".[15] It appears EVMs have no substantial association with vote for these parties, with the estimated effects being inconsistent in sign and statistically insignificant.

7.2 Turnout

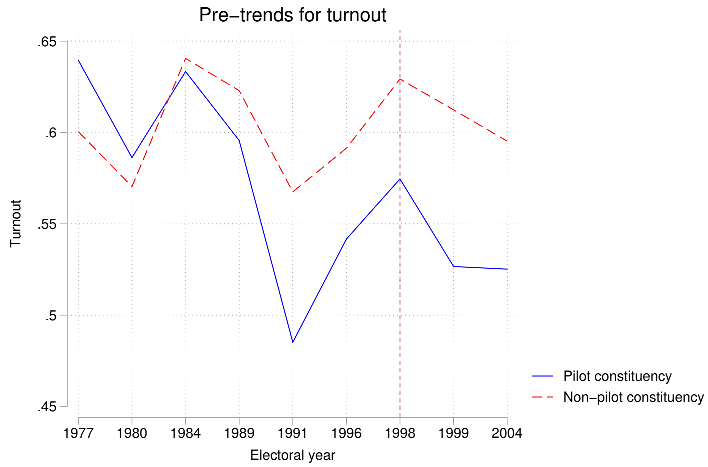
Do voters who previously caste invalid ballots still turn out? In the Indian case, since EVMs make it impossible to cast an invalid ballot, voters who intentionally casted spoiled ballots could now loose their incentive to go to the polls. As before, we examine the pre-trends of the early treatment pilot constituencies and the non-pilot constituencies. Compared to the pre-trends of invalid vote rates, the pre-trends for turnout do not show evidence of parallel trends.

In particular, figure 4 shows that there was a perceptible negative trend in turnout in the treatment districts relative to the control districts in the early 1990s,[16] though the gap did not appear to be increasing in the two elections before 1999. This trend may reflect the growing turnout gap in India between poor and rich voters, with rich voters tending to be less involved (Ahuja and Chhibber, 2012). The second column in Table A.15 conducts a placebo analysis by comparing the effects of the EVM treatment on turnout in the year 1999 versus other electoral years. According to the results, while the pilot constituencies have consistently smaller turnout rates compared to the non-pilot constituencies across all electoral years, though the difference in the treatment year is larger than any other year.

[15] In a few states where breakdowns by class are not available, poor parties were those preferred by a plurality of scheduled caste voters.

[16] The perceptible drop in turnout in pilot constituencies in 1991 could be because of the assassination of Rajiv Gandhi midway through the elections.

Figure 4: Pre-trends for turnout



Note: The blue solid line plots the average turnout rate in all pilot constituency across election years while the red dashed line plots the average turnout vote rate in non-pilot constituencies. The year 1998 marks the last election before the introduction of EVMs. Thus, 1999 is the first post-treatment year for the pilot constituencies. In the year 2004, the non-pilot constituencies also used EVMs.

Table A.13 examines the effect of EVM introduction on voter turnout. The results suggest that EVMs have a slight negative effect on turnout. Substantively, the effect is a little over two percentage points, a little smaller than the overall observed decline in turnout during this period, from 62.5% in 1998 to 59% in 2004. However, these results are imprecisely estimated as only the second model (Column 2) delivers a coefficient that is significant at the 10% level.[17]

Thus, we are cautious about whether EVMs affect turnout. Firstly, the pre-trends seem to suggest that the parallel trends assumption is not valid in the case of turnout. Secondly, an analysis of lags and leads of treatment in Table A.15 shows that the pilot constituencies consistently were different from non-pilot constituencies in terms of turnout. Finally, since the OLS standard errors are much smaller than the clustered standard errors, it seems likely that turnout rates within each state-year dyad are not independent. However, if true these results would provide additional evidence that EVMs change how voter protest is expressed.

.

[17] The results in Table A.13 report standard errors that are clustered at the state-year level. The OLS standard errors of the first three models in fact render the coefficients on EVM to be significant at the 5% level.

7.3 H3a: Partisan Effects

If EVMs were in fact altering the chances of successful electoral fraud, we should expect their introduction to increase the vote for specific parties or types of parties, especially those likely to be able to fraudulently manipulate the machines. Note that while study of whether different voting technologies favor or disfavor particular political parties has been the topic of discussion in the literature, but there is little proof of systematic effects (as opposed to analyses of particular races) (Stewart, 2011).

We do not find any systematic effects on the vote-shares of specific political parties, such as the INC or the BJP, or on the vote-shares of electoral alliances such as the NDA, UPA[18] or the Third Front. These results are not reported for reasons of space, but are available on request from the authors.

Table Six analyses the effect of EVMs on the vote share of the incumbent party of the state government. In the Indian context, the state government is the agency with effective control over the police and the district administration, which they might use for electoral advantage. Despite the Election Commission's careful attempts to limit such influence, the state incumbents clearly have a much greater opportunity to engage in fraud than any other party, and a decline of the vote for this party in areas with EVMs would be strong evidence for fraud. Conversely, if EVMs had a positive effect on state incumbent vote share, we might suspect the sort of systematic machine tampering feared by Kumar and Walia (2011) .

However, Table A.16 shows that there is little evidence for such an effect. EVMs have a small positive relationship with state incumbent vote share, but this effect is statistically insignificant at conventional levels.[19] State incumbents thus appear not to be affected by the introduction of EVMs, either because of the quality of the Election Commission's precautions or because EVMs in ineffective in preventing the types of fraud they use.

7.4 H3b: Voter Verification

One of the defining features of "direct recording" EVMs is that votes are recorded on the the memory unit of the machine, rather than on paper. This makes it impossible for voters to directly verify that their vote has been cast in the way that they wish, and theoretically

[18]While the UPA was formally created after the 2004 election, the INC was allied with several regional parties during the 1998 and 1999 elections. We also examined whether EVMs had any effect on the vote-shares of the INC+allies, and found no systematic effects.

[19]While EVMs seem to have a significant negative effect on state incumbent vote share after the introduction of our standard controls, once we control for minor party vote shares, this effect disappears.

possible to alter vote totals within the machine in ways that would be difficult to detect. The most commonly recommended solution to this problem is a voter-verified paper audit trail (VVPAT) (Kohno et al., 2004). VVPAT machines differ from other EVMs in that the voter receives a paper "receipt" for her vote, which can than be compared to the machine-reported totals in a post-election audit.

In 2013, the Indian supreme court ordered the election commission to introduce VVPAT technology in all elections. In the 2014 national election, eight constituencies had VVPAT. This makes possible a difference in difference-analysis-similar to that in Section Five, using two years (2009 and 2014). Since the announced goal of VVPAT is the reduction in fraud, we will focus on the results for two outcomes that might plausibly be correlated with fraud: The level of voting for the state incumbent party and the turnout rate.[20]

Tables A.18 and A.19 show the results this analysis. Relative to ordinary machines, the introduction of VVPAT machines appears to have no negative effect on turnout or vote for incumbents: If anything, turnout appears to increase very slightly in treated constituencies. In the appendix, we show that the parallel trend assumption holds for the eight constituencies assigned to VVPAT in 2014.[21]

The fact that the effect of VVPAT machines is indistinguishable from that of non-auditable EVMs does not mean that these innovations are useless, since this technology may possibly prevent election fraud in the future. It does however, indicate that these machines are not associated with changes in political outcomes relative to 2004 and 2009, either because no large-scale fraud occurred during this period or because VVPAT has not decreased the types of fraud that did take place.

7.5 H3c: Turnout and Fraud

Perhaps the effect of EVMs can be found not in the vote totals but in the turnout figures? We especially focus on region variational in the turnout effect, given that we expect to see decreases in turnout in constituencies that are more prone to booth-capturing. If booth capturing was common before 1999, some portion of the turnout recorded by the EC represents fraudulent votes, entered into the voter register and ballot box by armed goons. If the introduction of EVMs reduced the incidence of booth capture (as it was designed to do) we should

[20] Results showing VVPAT has no association with invalid voting are not reported for reasons of space.

[21] Because of the redistricting between the 2004 and 2009 election, the constituencies in the pretend analysis are those covering a majority of the territory of the 2014 districts, rather than the same units.

expect turnout to decline with their introduction in areas where this practice was common.

Interestingly, the effect of EVMs on turnout is not larger in areas that would intuitively be identified as more corrupt. One commonly used measure of corruption in Indian public life is the tendency of many candidates to face serious criminal charges (Vaishnav, 2011; Aidt et al., 2011). Using Aidt et al.'s (2011) data on the criminal status of candidates in the 2004 and 2009 elections, we defined a constituency as "criminal" if there was at least one criminal who ran for election. Table A.14 results show that there is no estimated effect of EVMs on turnout in these constituencies. Similar results (not reported for reasons of space) could be obtained by interacting EVM introduction with state-level poverty, insurgency, or location in the Hindi belt).

These weak results are consistent with design of the machines, since EVMs do not make it impossible for political parties to capture polling booths, but only increase the time it takes to do so. While party goons can still take control of polling booths, the delay built into the machines means they must maintain control for a longer time if they wish to cast all the booth's ballots. Anecdotal evidence suggests that political parties still indulge in mass fraudulent voting, even with the presence of EVMs (Rohde, 2004)

8 Conclusion

These results show that the switch from paper to electronic voting in India was associated with substantial political effects. Invalid voting was virtually eliminated, with this decline also being associated with an increase in the vote for smaller political parties, often from outside the traditional party system. At the same time, EVMs had modest or null effects on voter error and voter turnout. There is also little evidence that EVMs had a impact on fraud, either for better or for worse.

In several respects, these results, particularly the decline of invalid voting and the null result for fraud, echo the findings of the existing literature. They show, however, that EVMs can have some normatively desirable effects (reducing spoiled ballots) even in poor and fraud prone democracies that might seem unpromising ground for any sort of voting technology. In poor democracies, Indian style electronic voting, with its fixed menu of choices, may represent a way of increasing the number of valid votes, to say nothing of its advantages in streamlining the counting process. This is not to imply, of course, that EVMs, particularly non-auditable models, represent a perfectly secure voting technology, or that they have completely eliminated voter confusion. It simply means that in an environment with many invalid votes, even the modest interface improvements of first generation voting machines may represent a valuable improvement on the status quo.

While these results are in line with existing findings, the increase in voting for minor parties associated with EVMs represents a fascinating area for future research. At a minimum, they suggest a partial alternative perspective on the much commented on increase in the fragmentation of the Indian party system in the 1990s (Chhibber and Nooruddin, 1998) and the persistence of "non-duvergerian" outcomes at the district level in India (Diwakar, 2007). Taken more broadly, they suggest that voting technology, can have a substantial, and consequential, effects on how anti-system sentiments are expressed within the electoral system.

References

Ahuja, Amit and Pradeep Chhibber, "Why the poor vote in India:"If I don't vote, I am dead to the state"," *Studies in comparative international development*, 2012, *47* (4), 389–410.

Aidt, Toke, Miriam A Golden, and Devesh Tiwari, "Incumbents and criminals in the indian national legislature," 2011. https://www.repository.cam.ac.uk/bitstream/handle/1810/242058/cwpe1157.pdf?sequence=1&isAllowed=y.

Allers, Maarten A and Peter Kooreman, "More evidence of the effects of voting technology on election outcomes," *Public Choice*, 2008, *139* (1), 159–170.

Alvarez, R Michael, Melanie Goodrich, Thad E Hall, D Roderick Kiewiet, and Sarah M Sled, "The complexity of the California recall election," *Political Science and Politics*, 2004, *37* (01), 23–26.

_ **, Thad E Hall, and Susan D Hyde**, *Election fraud: detecting and deterring electoral manipulation*, Brookings Institution Press, 2009.

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

Ansolabehere, Stephen, "Voting machines, race, and equal protection," *Election Law Journal*, 2002, *1* (1), 61–70.

Banerjee, Mukulika, *Why India Votes?*, Routledge, 2015.

Bélanger, Éric, "Antipartyism and Third-Party Vote Choice A Comparison of Canada, Britain, and Australia," *Comparative Political Studies*, 2004, *37* (9), 1054–1078.

Bertrand, M, E Duflo, and S Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates," *The Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Callen, Michael and James D Long, "Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan," *American Economic Review*, jan 2015, *105* (1), 354–381.

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

Chandra, Kanchan, *Why ethnic parties succeed: Patronage and ethnic head counts in India*, Cambridge University Press, 2007.

Chhibber, Pradeep and Irfan Nooruddin, "Party competition and fragmentation in Indian national elections: 1957-1998," in P. Wallace and R. Ray, eds., *Indian Politics and the 1998 Election*, Sage, 1998, pp. 36–54.

Communist Party of India (Marxist), "Communist Party of India (Marxist): Party Programme," 2000.

Crost, Benjamin, Joseph Felter, Hani Mansour, and Daniel I Rees, "Election Fraud and Post-Election Conflict: Evidence from the Philippines," *IZA Discussion Paper*, 2013.

Dee, Thomas S, "Technology and voter intent: Evidence from the California recall election," *The Review of Economics and Statistics*, 2007, *89* (4), 674–683.

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

Duverger, Maurice, *Political Parties*, New York: Wiley, 1954.

Feldman, Ariel J, J Alex Halderman, and Edward W Felten, "Security analysis of the Diebold AccuVote-TS voting machine," 2006. Center for Information Technology Policy, Princeton University.

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

Great Britain, *The governance of Britain: review of voting systems, the experience of new voting systems in the United Kingdom since 1997*, The Stationery Office, 2008.

Hanmer, Michael J, Won-Ho Park, Michael W Traugott, Richard G Niemi, Paul S Herrnson, Benjamin B Bederson, and Frederick C Conrad, "Losing Fewer Votes The Impact of Changing Voting Systems on Residual Votes," *Political Research Quarterly*, 2010, *63* (1), 129–142.

Herron, Michael C and Jonathan Wand, "Assessing partisan bias in voting technology: The case of the 2004 New Hampshire recount," *Electoral Studies*, 2007, *26* (2), 247–261.

Ho, Daniel E and Kosuke Imai, "Estimating causal effects of ballot order from a randomized natural experiment the California alphabet lottery, 1978–2002," *Public Opinion Quarterly*, 2008, *72* (2), 216–240.

Jr, John R Lott, "Non-voted ballots, the cost of voting, and race," *Public Choice*, 2009, *138* (1-2), 171–197.

Knack, Stephen and Martha Kropf, "Voided ballots in the 1996 presidential election: a county-level analysis," *Journal of Politics*, 2003, *65* (3), 881–897.

Kohno, Tadayoshi, Adam Stubblefield, Aviel D Rubin, and Dan S Wallach, "Analysis of an electronic voting system," in "Security and Privacy, 2004. Proceedings. 2004 IEEE Symposium on" IEEE 2004, pp. 27–40.

Kollman, Ken, Allen Hicken, Daniele Caramani, and David Backer, "Constituency-level elections archive," \url{ http://www. electiondataarchive. org} 2011.

Krosnick, Jon A, Joanne M Miller, and Michael P Tichy, "An unrecognized need for ballot reform: The effects of candidate name order on election outcomes," *Rethinking the vote: The politics and prospects of American election reform*, 2003, pp. 51–73.

Kselman, Daniel and Emerson Niou, "Protest voting in plurality elections: a theory of voter signaling," *Public Choice*, 2011, *148* (3/4), 395–418.

Kumar, Sanjay and Ekta Walia, "Analysis of electronic voting system in various countries," *International Journal on Computer Science and Engineering*, 2011, *3* (5), 1825–1830.

Lehoucq, Fabrice, "Electoral fraud: Causes, types, and consequences," *Annual review of political science*, 2003, *6* (1), 233–256.

McAllister, Ian and Toni Makkai, "Institutions, society or protest? Explaining invalid votes in Australian elections," *Electoral Studies*, 1993, *12* (1), 23–40.

McMillan, Alistair, "The Election Commission of India and the regulation and administration of electoral politics," *Election Law Journal*, 2012, *11* (2), 187–201.

Niemi, Richard G, Michael J Hanmer, Benjamin B Bederson, Frederick G Conrad, and Michael W Traugott, *Voting technology: The not-so-simple act of casting a ballot*, Brookings Institution Press, 2009.

Palshikar, S, K C Suri, and Y Yadav, *Party Competition in Indian States: Electoral Politics in Post-Congress Polity*, Oxford University Press, 2014.

Pop-Eleches, Grigore, "Throwing out the bums: Protest voting and unorthodox parties after communism," *World Politics*, 2010, *62* (02), 221–260.

Power, Timothy J and J Timmons Roberts, "Compulsory voting, invalid ballots, and abstention in Brazil," *Political Research Quarterly*, 1995, *48* (4), 795–826.

_ **and James C Garand**, "Determinants of invalid voting in Latin America," *Electoral Studies*, 2007, *26* (2), 432–444.

Rohde, David, "Maujempur Journal; On a Newfangled Voting Machine, the Same Old Fraud," *New York Times*, 2004.

Rosenstone, S.J., L.R. Behr, and E.H. Lazarus, *Third Parties in America: citizen response to major party failure*, Princeton: Princeton University Press, 1996.

Rosenthal, Howard and Subrata Sen, "Electoral Participation in the French Fifth Republic," *The American Political Science Review*, 1973, *67* (1), 29–54.

Schedler, Andreas, "Anti-political-establishment parties," *Party politics*, 1996, *2* (3), 291–312.

Shastri, S, K C Suri, and Y Yadav, *Electoral Politics in Indian States: Lok Sabha Elections in 2004 and Beyond*, Oxford University Press, 2009.

Sridharan, Eswaran and Ashutosh Varshney, “Toward moderate pluralism: Political parties in India,” *Political Parties and Democracy, Baltimore und London*, 2001, pp. 206–237.

Stewart, Charles, “Residual vote in the 2004 election,” *Election Law Journal*, 2006, *5* (2), 158–169.

_ , “Voting technologies,” *Annual Review of Political Science*, 2011, *14*, 353–378.

Thachil, Tariq, “Elite Parties and Poor Voters: Theory and Evidence from India,” *American Political Science Review*, 2014, *108* (02), 454–477.

Vaishnav, Milan, “The market for criminality: Money, muscle and elections in India,” 2011. Working Paper.

Verma, Arvind, “Policing elections in India,” *India Review*, 2005, *4* (3-4), 354–376.

_ , “Situational prevention and elections in India,” *International Journal of Criminal Justice Sciences*, 2009, *4* (2), 83.

Wand, Jonathan N, Kenneth W Shotts, Jasjeet S Sekhon, Walter R Mebane Jr, Michael C Herron, and Henry E Brady, “The butterfly did it: The aberrant vote for Buchanan in Palm Beach County, Florida,” *American Political Science Review*, 2001, pp. 793–810.

Wolchok, Scott, Eric Wustrow, J Alex Halderman, Hari K Prasad, Arun Kankipati, Sai Krishna Sakhamuri, Vasavya Yagati, and Rop Gonggrijp, “Security analysis of India’s electronic voting machines,” in “Proceedings of the 17th ACM conference on Computer and communications security” ACM 2010, pp. 1–14.

Online Appendix

Table A.1: Effect of EVM's on valid votes

	(1)	(2)	(3)	(4)	(5)
EVM	37587.5	13693.7	19611.9	21811.9	21793.0
	(26410.9)	(12612.0)	(14394.4)	(14053.4)	(18490.7)
Year FE		Yes	Yes		Yes
Constituency FE		Yes	Yes	Yes	Yes
Controls			Yes	Yes	
Phase-Year FE				Yes	
Constant	675191.1***	760420.1***	564393.8***	560744.5***	567489.2***
	(14469.8)	(18519.1)	(78041.1)	(75481.9)	(36239.8)
N	1629	1629	1628	1601	252
R^2	0.012	0.926	0.932	0.941	0.968

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on valid votes in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and total number of electors, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

Table A.2: Effect of EVM's on vote share of candidates receiving less that 0.5% of votes

	(1)	(2)	(3)	(4)	(5)
EVM	0.000887	-0.00274**	0.000608	0.000724	-0.00247
	(0.00161)	(0.000965)	(0.00103)	(0.00104)	(0.00176)
Year FE		Yes	Yes		Yes
Constituency FE		Yes	Yes	Yes	Yes
Controls			Yes	Yes	
Phase-year FE				Yes	
Constant	0.00754***	0.000908	-0.0130***	-0.0115**	0.0117***
	(0.000857)	(0.00103)	(0.00365)	(0.00371)	(0.00222)
N	1629	1629	1628	1601	252
R^2	0.003	0.646	0.823	0.828	0.634

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on the vote share of candidates receiving less than 0.5% of vote share in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and total number of electors, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

Table A.3: Effect of EVM's on vote share of candidates receiving less that 20% of votes

	(1)	(2)	(3)	(4)	(5)
EVM	0.0231	0.0237^{+}	0.00537	0.00445	0.0317
	(0.0189)	(0.0140)	(0.0103)	(0.0105)	(0.0196)
Year FE		Yes	Yes		Yes
Constituency FE		Yes	Yes	Yes	Yes
Controls			Yes	Yes	
Phase-year FE				Yes	
Constant	0.111^{***}	0.0507^{+}	0.456^{***}	0.440^{***}	0.0580^{*}
	(0.0108)	(0.0299)	(0.0734)	(0.0645)	(0.0249)
N	1629	1629	1628	1601	252
R^2	0.012	0.583	0.716	0.726	0.555

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on the vote share of candidates receiving less than 20% of vote share in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and total number of electors, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

Table A.4: Effect of EVM's on BSP vote share in UP

	(1)	(2)	(3)	(4)
EVM	0.0153	-0.0281^{+}	-0.0101	-0.00241
	(0.0118)	(0.0156)	(0.0246)	(0.0256)
Year FE		Yes	Yes	
Constituency FE		Yes	Yes	Yes
Controls			Yes	Yes
Phase-year FE				Yes
Constant	0.216***	0.234***	0.402***	0.277***
	(0.00633)	(0.0155)	(0.0722)	(0.0766)
N	255	255	255	236
R^2	0.007	0.804	0.834	0.887

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on BSP vote share in Lok Sabha electoral constituencies in Uttar Pradesh. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and total number of electors, Column (4) replaces electoral year fixed effects by phase-year fixed effects. The table reports robust standard errors for all models.

Table A.5: Effect of EVM's on BSP vote share in states other than UP

	(1)	(2)	(3)	(4)
EVM	0.00769	0.00730^{+}	0.00752^{+}	0.00793
	(0.00586)	(0.00388)	(0.00418)	(0.00507)
Year FE		Yes	Yes	
Constituency FE		Yes	Yes	Yes
Controls			Yes	Yes
Phase-year FE				Yes
Constant	0.0141^{**}	0.000597	0.00178	0.00241
	(0.00456)	(0.00474)	(0.0322)	(0.0282)
N	1374	1374	1373	1365
R^2	0.007	0.756	0.766	0.773

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on BSP vote share in Lok Sabha electoral constituencies outside Uttar Pradesh. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and total number of electors, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Standard errors have been clustered by state-year for all models.

Table A.6: Effect of EVM's on Left vote share in West Bengal, Kerala, and Tripura

	(1)	(2)	(3)	(4)
EVM	0.0380	-0.0400^{+}	-0.0444^{+}	-0.0508*
	(0.0301)	(0.0229)	(0.0255)	(0.0255)
Year FE		Yes	Yes	
Constituency FE		Yes	Yes	Yes
Controls			Yes	Yes
Phase-year FE				Yes
Constant	0.377***	0.445***	0.411^{+}	0.496*
	(0.0175)	(0.0223)	(0.236)	(0.245)
N	192	192	192	192
R^2	0.009	0.918	0.926	0.929

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on Left vote share in Lok Sabha electoral constituencies in West Bengal, Kerala and Tripura. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and total number of electors, Column (4) replaces electoral year fixed effects by phase-year fixed effects. The table reports robust standard errors for all models.

Table A.7: Effect of EVM's on Left vote share in states other than West Bengal, Kerala and Tripura

	(1)	(2)	(3)	(4)
EVM	-0.00171	0.00635	0.00616	0.00964
	(0.00728)	(0.0117)	(0.0118)	(0.0116)
Year FE		Yes	Yes	
Constituency FE		Yes	Yes	Yes
Controls			Yes	Yes
Phase-year FE				Yes
Constant	0.0203***	0.0176	0.0511	0.0685*
	(0.00459)	(0.0158)	(0.0374)	(0.0331)
N	1437	1437	1436	1409
R^2	0.000	0.675	0.680	0.688

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on Left vote share in Lok Sabha electoral constituencies outside West Bengal, Kerala and Tripura. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and total number of electors, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Standard errors have been clustered by state-year for all models.

Table A.8: Effect of EVM on vote share of candidate placed 1st on the ballot list in the Hindi belt

	(1)	(2)	(3)
EVM	-0.0102	0.00472	-0.0436
	(0.0418)	(0.0477)	(0.0406)
Year FE		Yes	Yes
Constituency FE		Yes	Yes
Controls			Yes
Constant	0.246***	0.138	0.283*
	(0.0195)	(0.0852)	(0.121)
N	675	675	616
R^2	0.001	0.529	0.524

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on the vote share of the 1st placed candidate in the Hindi belt. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and the turnout rate. Standard errors are clustered at the state-year level.

Table A.9: Effect of EVM's on voting for candidate placed 17th on the ballot list in the Hindi belt

	(1)	(2)	(3)
EVM	0.0273**	0.0229	0.00720
	(0.00906)	(0.0563)	(0.0721)
Year FE		Yes	Yes
Constituency FE		Yes	Yes
Controls			Yes
Constant	0.00227	0.00106	0.290
	(0.00486)	(0.0438)	(0.203)
N	66	66	66
R^2	0.124	0.552	0.688

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on the vote share of the 17th placed candidate in the Hindi belt. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and the turnout rate. Robust standard errors have been reported.

Table A.10: Effect of EVM on vote share of candidate placed below the eventual winner on the ballot list in the Hindi belt

	(1)	(2)	(3)
EVM	-0.00876	-0.0314	-0.0382
	(0.0189)	(0.0457)	(0.0428)
Year FE		Yes	Yes
Constituency FE		Yes	Yes
Controls			Yes
Constant	0.109***	0.0259	-0.102
	(0.0133)	(0.0228)	(0.0989)
N	651	651	650
R^2	0.001	0.453	0.479

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on the vote share of the candidate placed below the winner in the Hindi belt. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and the turnout rate. Standard errors have been clustered at the state-year level.

Table A.11: Effect of EVM on vote share of candidate placed above the eventual winner on the ballot list in the Hindi belt

	(1)	(2)	(3)
EVM	-0.0266	0.0167	0.0354
	(0.0294)	(0.0763)	(0.0732)
Year FE		Yes	Yes
Constituency FE		Yes	Yes
Controls			Yes
Constant	0.136***	0.0881	-0.0265
	(0.0176)	(0.0995)	(0.158)
N	500	500	499
R^2	0.007	0.587	0.629

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on the vote share of the candidate placed below the winner in the Hindi belt. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and the turnout rate. Standard errors have been clustered at the state-year level.

Table A.12: Effect of EVM on vote share of candidate placed around the 17th place on the ballot list in the Hindi belt

	(1)	(2)	(3)	(4)
	15th	16th	18th	19th
EVM	-0.000895	-0.00294	-0.0357	-0.00209
	(0.0151)	(0.00898)	(0.0263)	(0.00651)
Constant	0.0322**	0.0136*	0.0403*	0.00766+
	(0.0101)	(0.00559)	(0.0158)	(0.00414)
N	144	111	50	42
R^2	0.000	0.001	0.037	0.003

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on the vote share of candidates around 17th place in the Hindi belt. Columns (1)-(4) report the OLS estimates of the regression of the 15th placed candidate's, the 16th placed candidate's, the 18th placed candidate's and the 19th placed candidate's vote share on EVM respectively. Robust standard errors have been reported.

Table A.13: Effect of EVM's on turnout

	(1)	(2)	(3)	(4)	
EVM	-0.0346	-0.0238^{+}	-0.0219	-0.0177	-0.0140
	(0.0286)	(0.0122)	(0.0141)	(0.0150)	(0.0178)
Year FE		Yes	Yes		Yes
Constituency FE		Yes	Yes	Yes	Yes
Controls			Yes	Yes	
Phase-year FE				Yes	
Constant	0.619***	0.727***	0.741***	0.766***	0.465***
	(0.0142)	(0.0217)	(0.0497)	(0.0476)	(0.0269)
N	1629	1629	1628	1601	252
R^2	0.024	0.848	0.850	0.874	0.908

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on turnout rates in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, and the HH index score, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

Table A.14: Effect of EVM*criminal constituency on turnout

	(1)	(2)	(3)	(4)	(5)
EVM	-0.0112	-0.0282^{+}	-0.0266	-0.0185	-0.00870
	(0.0308)	(0.0153)	(0.0172)	(0.0184)	(0.0198)
EVM*criminal constituency	-0.0421*	0.00743	0.00837	0.00229	-0.0106
	(0.0175)	(0.0135)	(0.0135)	(0.0139)	(0.0105)
Year FE		Yes	Yes		Yes
Constituency FE		Yes	Yes	Yes	Yes
Year*criminal constituency FE		Yes	Yes	Yes	Yes
Controls			Yes	Yes	
Phase-year FE				Yes	
Constant	0.619***	0.731***	0.700***	0.706***	0.469***
	(0.0142)	(0.0233)	(0.0718)	(0.0576)	(0.0270)
N	1629	1629	1628	1601	252
R^2	0.037	0.849	0.852	0.876	0.908

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on turnout rates in constituencies that had a criminal candidate run in 2004. Column (1) runs a simple OLS model, Column (2) reports the results of a triple differences regression with constituency specific fixed effects, electoral year fixed effects, and an interaction of year dummies and criminal constituency dummy, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and turnout rate, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts the triple difference regression on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

Table A.15: Leads of the treatment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	Invalid votes	HHI	Turnout	2.5%	5%	7.5%	10%
Pilot*1989	-0.00159	0.00980	0.00220	-0.00410	-0.00308	-0.000225	0.00666
	(0.00496)	(0.0243)	(0.0251)	(0.00422)	(0.00869)	(0.0102)	(0.0141)
Pilot*1991	0.00110	0.0204	-0.0543*	-0.00445	0.00220	0.000334	-0.00485
	(0.00472)	(0.0264)	(0.0259)	(0.00529)	(0.00924)	(0.0103)	(0.0135)
Pilot*1996	-0.00415	0.00866	-0.0476*	-0.00409	-0.00178	0.00685	0.00308
	(0.00626)	(0.0229)	(0.0230)	(0.00583)	(0.0108)	(0.0125)	(0.0147)
Pilot*1998	-0.00369	0.0114	-0.0522*	-0.00213	-0.00124	-0.00509	-0.00327
	(0.00614)	(0.0221)	(0.0220)	(0.00415)	(0.00771)	(0.00928)	(0.0119)
Pilot*1999	***-0.0181*****	***0.00142***	***-0.0835*****	***0.00903****	***0.0198****	***0.0220****	***0.0150***
	(0.00622)	***(0.0224)***	***(0.0271)***	***(0.00441)***	***(0.00840)***	***(0.00981)***	***(0.0124)***
Pilot*2004	0.00216	0.0210	-0.0674**	-0.00790	-0.0105	-0.0152	-0.0116
	(0.00603)	(0.0238)	(0.0236)	(0.00527)	(0.00913)	(0.0102)	(0.0130)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.0382***	0.446***	0.688***	0.0287***	0.0457***	0.0540***	0.0645**
	(0.00837)	(0.0221)	(0.0382)	(0.00869)	(0.0124)	(0.0163)	(0.0203)
N	3235	3777	3234	3235	3235	3235	3235
R^2	0.342	0.357	0.490	0.421	0.390	0.364	0.342

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This table assigns placebo treatments to pilot constituencies in electoral years prior to 1999. Column (1) investigates the effect of placebo EVM treatment on invalid vote rates, Column (2) does the same for turnout rates and Columns (3)-(7) look at the leads of treatment for the different measures of minor party vote-share. The actual treatment year for pilot constituencies is 1999, and is marked by the bold and italic row. All errors have been clustered at the state-year level.

Table A.16: Effect of EVM's on state incumbent vote share

	(1)	(2)	(3)	(4)	(5)
EVM	0.00162	0.000623	-0.00332	-0.00709	-0.00454
	(0.00651)	(0.00396)	(0.00393)	(0.00437)	(0.00693)
Year FE		Yes	Yes		Yes
Constituency FE		Yes	Yes	Yes	Yes
Controls			Yes	Yes	
Phase-year FE				Yes	
Constant	0.0429***	0.0785***	0.104**	0.0922**	0.0322*
	(0.00385)	(0.0162)	(0.0357)	(0.0329)	(0.0144)
N	1617	1617	1616	1589	243
R^2	0.000	0.624	0.682	0.693	0.595

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on state incumbent vote share in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and turnout rate, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

Table A.17: Effect of EVM's on Poor Voted Party votes

	(1)	(2)	(3)	(4)	(5)
EVM	0.0359	-0.00801	-0.0184	-0.0137	0.000546
	(0.0603)	(0.0230)	(0.0183)	(0.0181)	(0.0412)
Year FE		Yes	Yes		Yes
Constituency FE		Yes	Yes	Yes	Yes
Controls			Yes	Yes	
Phase-Year FE				Yes	
Constant	0.358^{***}	0.361^{***}	0.149^{+}	0.144	0.369^{***}
	(0.0361)	(0.0191)	(0.0815)	(0.0948)	(0.0324)
N	831	831	831	811	166
R^2	0.016	0.811	0.898	0.908	0.681

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on the vote-share of parties voted by the poor in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and turnout rate, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

Table A.18: Effect of VVPAT on turnout

	(1)	(2)	(3)	(4)
VVPAT	-0.0235	0.0195	0.0233	0.0112
	(0.0342)	(0.0198)	(0.0175)	(0.0317)
Year FE		Yes	Yes	Yes
Constituency FE		Yes	Yes	Yes
Controls			Yes	
Constant	0.634***	0.720***	0.648***	0.532***
	(0.0225)	(0.0445)	(0.0436)	(0.0522)
N	1086	1086	1086	75
R^2	0.000	0.944	0.954	0.929

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. This Table shows the impact of EVMs on invalid vote rates in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and turnout rate, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

Table A.19: Effect of VVPAT on state incumbent vote share

	(1)	(2)	(3)	(4)
VVPAT	-0.00880*	-0.00270	-0.00388	-0.00269
	(0.00421)	(0.00308)	(0.00295)	(0.00349)
Year FE		Yes	Yes	Yes
Constituency FE		Yes	Yes	Yes
Controls			Yes	
Constant	0.0228***	0.0351*	0.0265	0.0315***
	(0.00228)	(0.0133)	(0.0229)	(0.00465)
N	1078	1078	1078	73
R^2	0.001	0.853	0.903	0.773

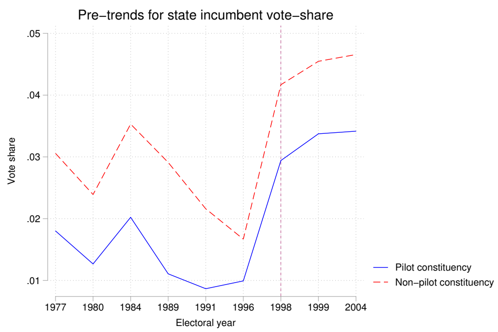
Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

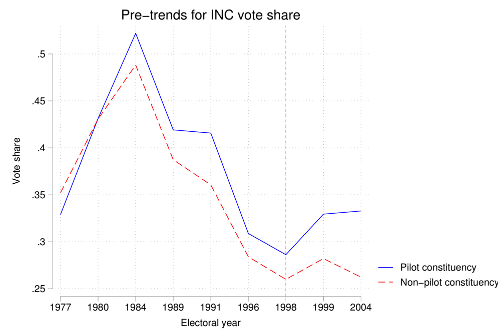
Notes. This Table shows the impact of EVMs on invalid vote rates in Lok Sabha electoral constituencies. Column (1) runs a simple OLS model, Column (2) reports the results of a basic diff-in-diff regression with constituency specific fixed effects and electoral year fixed effects, Column (3) includes time-varying control variables such as the INC vote share, BJP vote share, number of candidates in the constituency, the eventual margin of victory, the HH index score and turnout rate, Column (4) replaces electoral year fixed effects by phase-year fixed effects. Finally, Column (5) conducts a basic diff-in-diff regression with constituency and year fixed effects on constituencies with more than 40% of its population living in urban areas. Standard errors have been clustered by state-year for all models.

Figure A.1: Pre-trends for other control variables

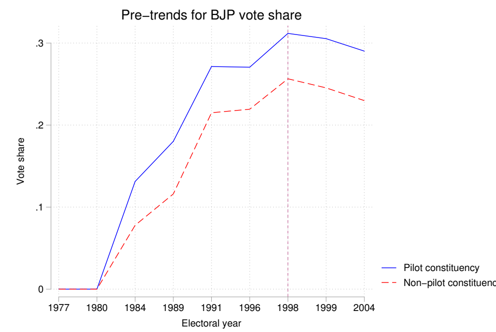
(a) State incumbent vote-share



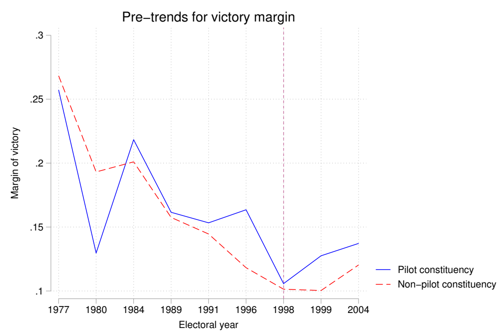
(b) INC vote share



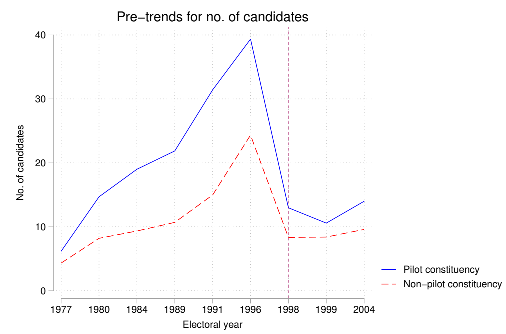
(c) BJP vote-share



(d) Margin of victory



(e) Number of candidates



Notes. The blue solid line plots the average values of the different variables in all pilot constituency across election years while the red dashed line plots the average summed minor party/candidate vote shares in non-pilot constituencies. The year 1998 marks the last election before the introduction of EVMs. Thus, 1999 is the first post-treatment year for the pilot constituencies. In the year 2004, the non-pilot constituencies also used EVMs.