"None Of The Above" Votes in India and the Consumption Utility of Voting*

Gergely Ujhelyi,† Somdeep Chatterjee,‡ and Andrea Szabó§

September 1, 2016

Abstract

Why do people vote? One possibility is that they derive consumption utility from doing so, but isolating this has proven empirically challenging. In this paper we study a recent natural experiment in India, where legislative elections have to provide a "None Of The Above" (NOTA) option to voters. Using the fact that NOTA cannot affect the electoral outcome we show that studying individual voters' behavior with and without NOTA provides a way to identify various components of the consumption utility of voting. To address the challenge that individual votes are not observable, we borrow techniques from the Industrial Organization literature to estimate a structural model of voter demand for candidates and perform counterfactual simulations removing the NOTA option. We complement this with a reduced-form analysis of NOTA in a difference-in-differences framework, exploiting variation in the timing of the reform created by the electoral calendar. Using both methods, we find that NOTA increased turnout. We find minimal substitution between candidates and NOTA, indicating that NOTA votes are cast by new voters who turn out to vote specifically for this option. This indicates the presence of an option-specific consumption utility of voting.

*We thank Francisco Cantú, Alessandra Casella, Suresh Naidu, Mike Ting, and seminar participants at Oxford, Columbia, and the 2016 Banff Workshop in Empirical Microeconomics for useful comments and suggestions. We also thank seminar participants at the Indian Statistical Institute, Indian Institute of Technology Kanpur, and Public Choice Society 2015 for useful comments on an earlier version.

†Department of Economics, University of Houston, Houston, TX, gujhelyi@uh.edu

‡Department of Economics, Flame University, Pune, India, somdeep.chatterjee@flame.edu.in

§Department of Economics, University of Houston, Houston, TX, aszabo2@uh.edu

1 Introduction

One possible solution to the "paradox" of why people bother to vote in large elections is that voting yields consumption utility. Such consumption utility could be derived from performing one's civic duty, expressing one's political views, or participating in a democracy (Downs, 1957; Riker and Ordeshook, 1968; Brennan and Lomasky, 1993). Empirically distinguishing these consumption motives from each-other and from other possible goals, such as a desire to affect the electoral outcome, is notoriously difficult. In this paper we propose to do this by using data from a natural experiment in Indian elections and estimating a structural model of voter turnout using techniques from the consumer demand literature.

To identify different components of the consumption utility of voting, we exploit a natural experiment in the world's largest democracy. Following a decision by the Indian Supreme Court, since September 2013 all state and national elections in the country must offer a "None Of The Above" (NOTA) option to voters. Votes cast for NOTA are counted (rather than simply discarded as invalid) but do not affect the outcome of the election (the winner is still the candidate with a plurality of votes among votes cast for candidates). In the five states that held local elections after the Supreme Court ruling 1.7 million voters chose NOTA, and in the 2014 national election 6 million voters voted for this option (representing 1.1% of all votes cast). While elections with a NOTA-type option have been used elsewhere, none of them came close to the scale of the Indian experiment.[1]

Because in the Indian system NOTA votes cannot affect the outcome of the election, voters who choose this option must be motivated by a consumption utility to vote. Such consumption utility can arise from two broad sources. It can be a general utility obtained from showing up at the polls (such as complying with a social norm to participate in the election), or it can be a utility specific to the option chosen by the voter (such as utility derived from expressing one's views).

Intuitively, we can distinguish between these two types of consumption utility by asking how a NOTA-voter would have behaved in the absence of the NOTA option. If without the NOTA option this voter would have voted for one of the candidates, this is consistent with both a general and an option-specific utility of voting. By contrast if without NOTA this voter would have abstained, then the NOTA vote cannot be explained by a general utility derived from showing up at the polls. Instead, the voter must be voting for NOTA in order to obtain a utility specific to this option. Thus, studying the counterfactual behavior of

[1] In most elections the only way for a voter to participate without voting for a candidate is to cast an invalid vote and these are difficult to distinguish from voting mistakes. In systems where a NOTA-type option is explicitly available to voters, it typically has electoral consequences, affecting who gets elected or whether the election has to be repeated. We review these different systems in section 3 below.

NOTA voters can test apart these two components of the consumption utility from voting.

Our empirical work seeks to test whether, following the introduction of NOTA, new voters showed up at the polls in order to vote for this option. This question is challenging because it requires making statements about *individual* voter behavior in the counterfactual no-NOTA scenario. Because ballots are secret, individual voter behavior is observed neither with nor without the NOTA option. Instead, it must be inferred from aggregate data.

To begin, we first ignore individual behavior and study the impact of NOTA on aggregate turnout in a reduced form framework. This exercise exploits variation in the effective timing of the NOTA reform created by the Indian electoral calendar: elections to the states' legislative assemblies occur at different times in different states. This allows us to study the impact of NOTA in a difference-in-differences framework by comparing the change in voter turnout in states not yet affected by the policy to changes in states that were already affected. From this analysis we estimate that, in the average electoral district, the introduction of the NOTA policy significantly increased turnout. This finding survives a variety of robustness checks and the magnitude of the effect (2-3 percentage points) is similar to the vote share of NOTA observed in the data.

While suggestive of new voters turning out to vote for NOTA, these aggregate patterns do not provide conclusive evidence because they mask the substitution between abstention, candidates, and NOTA at the individual level. In particular, we do not know whether NOTA voters would have abstained or voted for a candidate if NOTA was not available. To study this question, we relate the aggregate voting returns to individual voter behavior using a structural model of voter demand for candidates. We adapt the BLP model of Berry, Levinsohn and Pakes (1995) from the consumer demand literature, where consumers (voters) choose between the products (candidates) of firms (parties) in various markets (electoral districts). Voters have preferences over observed and unobserved candidate characteristics (including NOTA) and abstention. The model explicitly allows for heterogeneity in these preferences and links them to the aggregate vote shares we observe in the data. Estimating the model allows us to recover the parameters of individual voters' utility functions from this aggregate data. Using the estimates, we study how voters substitute between choosing NOTA, one of the candidates, and abstention in counterfactual simulations where the NOTA option is removed.

The results of this analysis indicate that NOTA increased turnout, which is in line with the aggregate patterns observed in the reduced form exercise. Furthermore we find that the magnitude of this increase explains virtually all the NOTA votes observed in the data. We find negligible substitution towards NOTA away from the candidates running for election. These results indicate that most voters who voted for NOTA would normally abstain. In turn, this provides evidence for the existence of consumption utility specifically from voting for NOTA. In this context, models that do not include an option-specific utility of voting would have a hard time explaining the data.

To the extent that participation in a democracy is valuable, our finding that having a NOTA option on the ballot can increase voter turnout is relevant in its own right, and provides support for the arguments of the Indian Supreme Court in introducing this policy.

Our paper is related to the vast literature on voter turnout, some of which we review in section 2 below. While several studies propose interesting models that are sufficient to explain a particular set of observations, there is much less work testing these models apart (i.e., asking whether a particular model is *necessary* to explain the data). Instead of proposing a new model of turnout, in this paper we focus on testing models apart by asking whether the concept of (option-specific) consumption utility is necessary to explain our data. This approach is similar in spirit to Coate and Conlin (2004) and Coate et al. (2008) who estimate and compare competing structural models of turnout on data from Texas liquor referenda.

While the idea of a consumption utility from voting is an old one, identifying it empirically has proved challenging. In the real world, it is difficult to find a situation where voting occurs while the probability of being pivotal is known to be zero. A number of field experiments have found that social pressure can increase turnout (e.g., Gerber et al., 2008; DellaVigna et al., 2015), which indicates the existence of a general consumption utility from voting (the utility derived from complying with social pressure). Identifying the utility obtained from voting for a specific option (as opposed to the general utility from showing up) is difficult even in laboratory settings. Lab experiments studying whether people vote for morally superior alternatives have found mixed results: Feddersen et al. (2009) and Shayo and Harel (2012) find evidence of consumption utility while Tyran (2004) and Kamenica and Egan Brad (2014) do not. Our paper adds to this literature by identifying consumption utility in real-world elections.

Methodologically, our paper offers a novel way to estimate the correlates of vote returns in multiparty elections. Some earlier approaches to this problem (e.g., Glasgow and Alvarez, 2005) have used discrete choice models with individual-level survey data, but such data is subject to well-known biases in voters' self-reported behavior (see, e.g., Selb and Munzert (2013) and the literature cited therein). Other studies use aggregate administrative data and purely statistical models to deal with the problem of conducting "ecological inference" regarding voter preferences (see Cho and Manski (2008) for a review). By contrast our BLP-based approach combines the advantages of a micro-founded discrete choice model with those of aggregate administrative data. It allows for rich heterogeneity in voter tastes for candidate characteristics and, because it is micro-founded, offers the possibility of conducting counterfactual experiments. In a different context, Rekkas (2007) also exploits some of these advantages of the BLP model in her study of campaign expenditures in the 1997 Canadian election. Our paper goes further by using panel data, allowing for heterogeneity in voters' preferences driven by demographics as in Nevo (2001), allowing for endogenous candidate choice by the competing parties, and by conducting counterfactual experiments using the estimated model.

Finally, our paper relates to previous studies of NOTA-type votes in the political science literature (reviewed in section 3 below). We differ from this literature by using NOTA votes to isolate the consumption utility from voting and by estimating a structural model that can be used to answer normative questions about the desirability of having this option on the ballot.

In the rest of the paper, section 2 explains how we propose to use NOTA votes to identify various components of the consumption utility from voting. Section 3 describes the NOTA policy, explains how it differs from similar options available to voters in other countries, and describes the Indian electoral setting we analyze. Section 4 describes our data and section 5 documents the pattern of NOTA votes and presents a difference-in-differences analysis of the effect of NOTA on turnout. Section 6 estimates the structural model and presents the counterfactual results. Section 7 concludes.

2 The consumption utility of voting

Why people vote is one of the classical questions of economics and political science. In the "calculus of voting" model (Downs, 1957; Riker and Ordeshook, 1968; Fiorina, 1976), voters consider both instrumental and consumption benefits. They vote for candidate j if

$$P_j B_j + (U_j + U_0) > c \tag{1}$$

and abstain otherwise (where $j = \arg\max_{j'}(P_{j'}B_{j'} + U_{j'})$ is the voter's preferred candidate). The first term on the left-hand side of (1) is the expected instrumental benefit, where P_j is an individual's probability of being pivotal in the election of candidate j and B_j is the benefit of the candidate winning. The second term is the consumption utility of voting, which captures a wide range of factors sometimes referred to as "expressive utility" or "civic duty": "1. the satisfaction from compliance with the ethic of voting [...] 2. the satisfaction from affirming allegiance to the political system [...] 3. the satisfaction from affirming a partisan preference [...] 4. the satisfaction of deciding, going to the polls, etc. [...] 5. the satisfaction of affirming one's efficacy in the political system" (Riker and Ordeshook, 1968, p28). We separate this consumption utility into two components to highlight that part of the utility (U_j) may depend on voting for a specific candidate j (e.g., the satisfaction from expressing partisan support), while part of it (U_0) only depends on showing up at the polls regardless of who one votes for (e.g., satisfaction from compliance with an ethical norm to vote). Finally, on the right-hand side of (1) c represents any direct or opportunity costs from voting. Without loss of generality, all terms in (1) are assumed to be non-negative.

Observing that in large elections the probability P_j of being pivotal is close to 0, the recent literature seeking to explain turnout within the framework of the calculus of voting equation (1) has followed various routes.[2] First, voters could overestimate P_j. Lab experiments show that, indeed, voters often overestimate the probability that their vote will matter and suggest that this can explain turnout decisions (Duffy and Tavits, 2008; Dittman et al., 2014). Relatedly, Ortoleva and Snowberg (2015) show that turnout is higher in populations with more overconfident voters. Under these conditions, turnout can be explained even if $(U_j + U_0) = 0$.

A second set of studies present models that feature an option-specific utility U_j. In Coate and Conlin (2004) and Feddersen and Sandorini (2006), this utility represents ethical considerations regarding what would be best for everyone in one's group. In other models, such as Shachar and Nalebuff (1999), U_j is created by the mobilization efforts of political leaders. In Degan and Merlo (2011), U_j includes a psychological disutility from the possibility of voting for the "wrong" candidate.

Another set of papers focus on the general utility U_0 from showing up to vote. For example, members of a group may observe turnout and draw inferences about whether an individual is an "ethical type" (Bénabou and Tirole, 2006; Gerber et al., 2008; Ali and Lin, 2013). Similarly, a voter may vote to avoid a feeling of shame from not having done his duty, especially if others will ask whether one has voted (Harbaugh, 1996; Blais, 2000; DellaVigna et al., 2015).

While these studies convincingly demonstrate that the proposed models have explanatory power, it is not always clear to what extent these models are *necessary* to explain the data.[3]

[2] Missing from equation (1) are instrumental motivations other than those related to winning. For example, it is possible that a voter votes in order to signal his preferences to affect the policies chosen after the election. Or he may vote in order to encourage a candidate to run again in the future. It is possible to treat such motivations in a strategic setting but the likelihood that a voter's vote will be pivotal in affecting policy or encouraging a candidate is likely to be small (see, e.g., Razin (2003)). Here we follow most of the literature in assuming that if such motivations exist, they are sources of consumption utility and hence part of U_j.

[3] Examples of papers that highlight the value of testing models apart include the pair of studies by Coate and Conlin (2004) and Coate, Conlin, and Moro (2008) which explicitly compare several different models of turnout on the same dataset.

In particular, can we rule out that one or both components of the consumption utility $U_j + U_0$ is 0?

As we explain below, the Indian NOTA policy allows us to study this question by creating a "None Of The Above" option that voters can vote for but that, by design, cannot affect the electoral outcome. First, we document that voters actually choose NOTA. Because $P_{NOTA}B_{NOTA} = 0$, from equation (1) a voter who chooses NOTA must have

$$U_{NOTA} + U_0 > c \tag{2}$$

i.e., there has to be a positive consumption utility of voting.

Second, we ask how a voter who chose NOTA would have voted in the absence of the NOTA option. If the voter would have abstained, then $c \geq P_jB_j + (U_j + U_0)$. Combining with (2), we have

$$U_{NOTA} > P_jB_j + U_j,$$

i.e., there has to be a positive option-specific utility from voting for NOTA. For example, a voter may derive utility from expressing his disapproval of all the candidates. Conversely, if in the absence of NOTA the voter would have voted for one of the candidates, then it is possible that there is no option-specific utility but $U_0 > 0$ (for example, a voter may vote to satisfy social pressure while deriving no specific utility from voting for any of the options on the ballot.). Thus, studying voters' behavior with and without NOTA offers a test for the existence of an option-specific utility of voting.[4]

In this way, although not an actual candidate, studying NOTA votes gives us an opportunity to test both for the existence of a consumption utility from voting and for the existence of a consumption utility from voting for this particular option. The second exercise is empirically challenging because it requires making statements about individual behavior both with and without NOTA while, due to the secret ballot, this behavior is never observed.

[4]It is difficult to imagine a similar experiment where an actual candidate's probability of wining is administratively set to 0. (Having small-party candidates on the ballot who have little chance of winning is not the same experiment since voters' believing that $P_j > 0$ cannot be ruled out.) Feddersen et al. (2009) and Shayo and Harel (2012) create variation in P in lab experiments and find evidence that the "moral superiority" of an alternative affects voters' behavior when P is small.

3 Background

3.1 The Indian NOTA policy

In elections where a paper ballot is used, voters can participate without voting for any of the candidates: they can hand in an empty ballot or otherwise intentionally invalidate their vote. With the introduction of electronic voting machines Indian voters lost this possibility. In 2004, the citizen's group People's Union for Civil Liberties (PUCL) filed a petition with the Supreme Court to rectify this and give voters the ability to have their participation recorded without forcing them to vote on any of the candidates.[5] In its 2013 decision, the Supreme Court agreed:

> "For democracy to survive, it is essential that the best available men should be chosen as people's representatives for proper governance of the country. This can be best achieved through men of high moral and ethical values, who win the elections on a positive vote. Thus in a vibrant democracy, the voter must be given an opportunity to choose none of the above [...] Democracy is all about choice. This choice can be better expressed by giving the voters an opportunity to verbalize themselves unreservedly and by imposing least restrictions on their ability to make such a choice. By providing NOTA button in the Electronic Voting Machines, it will accelerate the effective political participation in the present state of democratic system and the voters in fact will be empowered." (PUCL vs. Union of India, 2013, p43-44).

Following the Supreme Court's decision, since September 2013, all state and national elections in India give voters the option of recording a "None Of The Above" vote on the voting machine. These votes are counted and reported separately but have no role in the outcome of the election. In particular, votes cast on NOTA affect neither the validity nor the winner of an election. Even if NOTA were to receive a majority of the votes, the winner of the election would be the candidate who received the most votes among the non-NOTA votes.

The NOTA policy received wide news coverage in both national and local media. In its decision the Supreme Court directed the Election Commission to undertake awareness programs to inform the electorate of the new policy, and voter education programs explicitly

[5] Under the electronic voting machines, the only way for a voter to have his non-vote recorded was to inform the clerk at the voting booth of his desire to do so. The clerk would then record this on the voter ledger together with the voter's thumbprint for identification. The PUCL argued that this was unconstitutional, violating the secret ballot.

focused on explaining this new option to voters. As a result we expect that most voters would be well-informed about the NOTA policy, including the fact that NOTA votes would not affect the electoral outcome.[6]

3.2 NOTA-like options in other countries

In most countries voters can effectively cast a "none of the above" vote by intentionally returning an invalid vote (e.g., leaving the ballot blank, writing on the ballot, or marking more than one candidate). Because it is typically impossible to know whether such votes occur intentionally or by mistake, it is difficult to use them to draw conclusions regarding voters' intentional behavior (see, e.g., McAllister and Makkai, 1993; Herron and Sekhon, 2005; Power and Garand, 2007; Uggla, 2008; Driscoll and Nelson, 2014). For some applications, the fact that invalid votes also include voting mistakes will simply add measurement error to the "true" measure intended to capture negative votes. In other cases, however, this will have an important impact on the interpretation of the results. For example, more invalid votes among the less educated can mean either that these voters are more likely to make mistakes when filling out the ballot, or that they are particularly dissatisfied and intentionally cast invalid votes to express this.[7]

In some countries, while there is no NOTA option on the ballot, blank votes are counted separately from invalid votes and are believed to represent a negative vote. In principle, this system could be equivalent to the Indian NOTA, but in practice the equivalence is unlikely to be perfect. First, blank votes could still represent voting mistakes, especially if there is a judgement call to be made about whether a vote is truly blank when it is being counted (for example, there could be markings on the side of the ballot, a small dot inside the checkbox, etc.). Fujiwara (2015) finds that the introduction of voting machines in Brazil reduced both blank and invalid votes among the less educated, which is consistent with both of these containing voting mistakes. Second, using the blank vote as an expression of dissatisfaction requires a shared understanding among voters regarding what the vote represents. Whether this social norm is operative in a given election is difficult to know with certainty. This is illustrated by the findings of Superti (2015) who studies a set of municipal elections in Spain - a country where the blank vote is generally understood to mean "none of the above." She shows that despite this common understanding, voter dissatisfaction following a ban which prevented the Basque nationalist party from contesting an election was likely expressed

[6] As we discuss in section 5, several patterns in the data also support this.

[7] These two interpretations also have different welfare implications regarding the desirability of having a NOTA option on the ballot. In the first case, NOTA only serves to confuse the less educated; in the second case, it gives disadvantaged segments of the population a voice.

through an increase in invalid rather than blank votes.

Another feature that makes India a cleaner case study than other systems for the analysis of voters' motivations is the lack of electoral impact of the NOTA vote. Recall that in India the NOTA vote can never "win," and due to the first-past-the-post system it has no impact on the allocation of legislative seats. By contrast in Colombia if the "blank vote" wins, new elections must be called with the rejected candidates prohibited from running again. In Spain, while the blank vote can never win, seats are allocated in a proportional system and a minimum 3% threshold must be reached for a party to enter parliament. In both of these systems choosing the blank vote as opposed to choosing one of the parties has immediate electoral consequences, affecting the mix of candidates eventually elected for office. In the Indian case, NOTA votes cannot be driven by electoral motivations in the current election.[8]

3.3 Assembly elections in India

We study voters' behavior under NOTA in the context of Indian state elections. In the Indian federal system, state governments are responsible for most areas of local significance, including health care, education, public works, police and security, and disaster management. State legislative assemblies are elected in single-member electoral districts (called "constituencies") in a first-past-the-post system. The party or coalition that wins the most number of seats in an assembly forms the state government headed by a Chief Minister and his council of ministers.[9]

Table 1 shows the timing of state assembly elections in our study period. Elections are typically held every 5 years but the electoral calendar varies widely across states. For example, some states held assembly elections in 2007 and 2012 while others in 2008 and 2013; some states always go to the polls in March while others always do so in November. This variation in the timing of elections creates an important source of identification for the analysis below.

In most states assembly elections are conducted separately from other elections. Four states, Andhra Pradesh, Arunachal Pradesh, Odisha and Sikkim, hold elections simultaneously with national elections. We will exclude these states from the analysis below.

[8] A voter's motivation (with any vote under any system) can always include a desire to affect long-run outcomes, e.g., by signaling his political preferences to the eventual winner in order to affect policy, or by encouraging a candidate to run in future elections. Because a single vote is just as unlikely to be pivotal in affecting these outcomes as it is in affecting who wins, we think that these motivations are best viewed as alternative sources of the consumption utility derived from voting.

[9] In states that have a bicameral legislature, the system just described applies to the lower house. Members of the upper house are either elected by the lower house or appointed by the Chief Minister or the Governor (the representative of the federal government in the states).

Table 1: Timeline of events in the study period

Year	Month	State assembly elections	Other events
2006	4	Assam	
	5	Kerala, Puducherry, Tamil Nadu, West Bengal	
2007	2	Manipur, Punjab, Uttarakhand	
	5	Uttar Pradesh	
	6	Goa	
	12	Gujarat, Himachal Pradesh	
2008	2	Tripura	
	3	Meghalaya, Nagaland	
	4		Delimitation
	5	Karnataka	
	11	Madhya Pradesh, NCT of Delhi	
	12	Chhattisgarh, Jammu & Kashmir, Mizoram, Rajasthan	
2009	4	Andhra Pradesh*, Arunachal Pradesh*, Odisha*, Sikkim*	National elections
	10	Haryana, Maharashtra	
	12	Jharkhand	
2010	10	Bihar*	
2011	4	Assam, Kerala, Puducherry, Tamil Nadu	
	5	West Bengal	
2012	1	Manipur, Punjab, Uttarakhand	
	3	Goa, Uttar Pradesh	
	11	Himachal Pradesh	
	12	Gujarat	
2013	2	Meghalaya, Nagaland, Tripura	
	5	Karnataka	
	9		NOTA policy introduced
	11	Chhattisgarh, Madhya Pradesh	
	12	Mizoram, NCT of Delhi, Rajasthan	
2014	4	Andhra Pradesh*, Arunachal Pradesh*, Odisha*, Sikkim*	National elections
	10	Haryana, Maharashtra	
	12	Jammu & Kashmir, Jharkhand	

Notes: * excluded from the dataset.

All state and national elections in India are conducted by the Election Commission of India under the supervision of the chief election commissioner. Since independence, the Commission has emerged as a highly regarded institution with a large degree of autonomy (McMillan, 2010). Election dates are set well in advance and declared as local holidays to reduce the cost of participation. Polling stations ("booths") are spread out throughout each constituency and enlisted voters are assigned to specific booths. Voters go to their designated booth to cast their vote with their Elector's Photo Identification Card.[10] Generally these booths are set up in neighboring schools or public buildings within a very small radius of one's residence. Participation rates in Indian elections tend to be high. In our state election data, average turnout is 71% and only 7% of the constituencies had turnout lower than 50%. (By comparison, turnout in US midterm elections is typically around 40%.) The voting age is 18, and the average constituency has approximately 180 thousand eligible voters.

Since 2004 all voting in India has taken place using electronic voting machines (EVMs).[11] Each candidate running in an election has a separate button assigned to him on the machine. Next to the button is the symbol identifying the candidate (to accommodate illiterate voters) and the voter pushes the button to record his vote. A light illuminates confirming that the vote was successful.[12] Under the NOTA policy, one of the buttons on the machine is assigned to the NOTA option.

In the Indian system of political reservation, some constituencies are designated Scheduled Caste (SC) and some Scheduled Tribe (ST). In these, only candidates from the given caste can run (to win, they must still obtain a plurality of all votes regardless of voters' caste). The reserved status of SC and ST constituencies is set at the same time as the electoral boundaries are drawn. In contrast to local (village) governments, state elections have no political reservation for women.

The current electoral boundaries were set in April 2008 by a commission working under the Election Commission (see Table 1). This was the first time in over 30 years that electoral redistricting ("delimitation") took place in India. All constituency boundaries as well as the reservation status of the constituencies was fixed by the delimitation commission in order to reflect population figures of the 2001 census. As described below, this redistricting poses challenges for the construction of our dataset and our empirical strategy.

[10] Voter Registration is a one time procedure. Except in special cases (such as for convicted criminals), once registered as a voter, a person can vote in all subsequent elections without having to go through any further registration process. Once registered the voter's name is on the voters' list and he or she gets the identification card which needs to be produced at the polling station before being allowed to vote.

[11] Electronic voting machines in India were introduced gradually beginning in 1999. Since 2004 all general and state elections are conducted using these machines.

[12] These machines are simpler to operate than some of the EVMs used in other countries that sometimes require a voter to follow written instructions, enter a candidate's number on a keypad, etc.

4 Data

4.1 Samples used for analysis

Our analysis uses two samples of constituencies: a panel serves as our primary dataset, and we use a repeated cross-section as a secondary sample.

The instrumental variables used in the structural analysis require panel data, and our main sample is a panel of 854 constituencies in the 6 states that conducted assembly elections in both 2008 and 2013 under the new electoral boundaries: Karnataka, Chhattisgarh, Rajasthan, Madhya Pradesh, Delhi, and Mizoram (see Table 1). One of these states, the 223 constituencies in Karnataka held elections in both years without a NOTA option, while the 630 constituencies in the remaining 5 states had a NOTA option in 2013 but not in 2008.

The main obstacle to extending the panel data to more constituencies is the delimitation (electoral redistricting). This makes it impossible to include elections before April 2008 in the panel as there is too little overlap between the old and new constituencies to make constituency-level matching meaningful.[13] For example, although 3 other states also held elections in both 2008 and 2013, they did so in February-March and had their constituency boundaries redrawn between the two elections in April 2008 so we cannot include these states in the panel. Other states with consistent electoral boundaries in our study period are those holding elections in 2014. However, 2014 was a national election year that made headlines around the world for its unusual outcome (the BJP led by Narendra Modi won by a landslide, the first time in 30 years that a single party won a majority of the legislative seats). Because 2014 state assembly elections took place either simultaneously with or after the national election (and in the latter case more than a year after the NOTA policy was introduced), the national election could confound the impact of NOTA in these states. We therefore decided to exclude these states from the panel analysis.

To obtain more power for a reduced form analysis, we use as a secondary dataset a repeated cross section of constituencies in 25 states that conducted elections between 2006 and 2014. Like the panel, this dataset excludes the states that held assembly elections simultaneously with national elections (Andhra Pradesh, Arunachal Pradesh, Odisha and Sikkim) since turnout considerations in these states are likely to be very different.[14] It

[13] Using GIS software we have computed the maximum overlap of each current constituency's area with an old constituency. For example, a maximum overlap of 80% indicates that 80% of the current constituency's area came from one constituency, while 20% came from one or more other constituencies. We find that half of the current constituencies have a maximum overlap of 62% or less and a quarter of the constituencies have a maximum overlap of 50% or less. This makes it impossible to match electoral data across constituencies in a meaningful way.

[14] Since our main goal with the repeated cross section is to increase power, we include the states that held elections in 2014 but not simultaneously with the national election. Excluding these states makes little

also excludes the state of Bihar because its unique election calendar (2005 and 2010) would require earlier data on voter demographics than we have access to. We have a total of 6685 constituency-year observations in this repeated cross section, and 1176 of these observations were affected by NOTA.

We next describe the information available in the primary (panel) and secondary (repeated cross-section) datasets.

4.2 Election and candidate data

The electoral data comes from the Election Commission of India, which provides information on assembly elections at the candidate level. Apart from standard electoral variables (candidate's party and vote share; number of eligible voters in the constituency) a key feature of the data is the presence of several candidate characteristics. The administrative data includes information on each candidate's age, gender, and caste (General, ST or SC).

Table 2 shows summary statistics of the electoral data at the constituency level for the panel and the repeated cross section. The average constituency has approximately 180 thousand eligible voters and 11 candidates competing. The overwhelming majority of candidates are male: the average constituency has less than one female candidate. The median age of candidates in a constituency is typically between 38 and 53. Approximately 13% of the constituencies are reserved for SC and 15% for ST. The average non-reserved constituency has 1.3 SC candidates and less than 0.5 ST candidates. Average turnout is 71% and the average vote share of the winning candidate is 45%. Summary statistics of the electoral data in the 6-state panel and the 25-state repeated cross section are generally similar.

Table 3 shows a more detailed distribution of candidate characteristics in the panel dataset. Each year we have approximately 10,000 candidates. The share of female candidates is 7% and the share of general caste candidates 62%. The largest four parties (INC, BJP, BSP, SP) field, respectively, 9, 8, 8, and 4% of all candidates. About 40% of the candidates run as independents not affiliated with any party.

4.3 Voter demographics

Our first source of demographic information is various waves of the National Sample Survey, conducted by the Indian Ministry of Statistics and Program Implementation since 1950. Each wave contains close to half a million individual surveys covering all Indian states, and is designed to be representative of the population at the subdistrict level. We obtained the individual level data and use it to create characteristics of the voting age population at the

difference for the results.

Table 2: Summary statistics of the electoral data at the constituency level

Variable	Obs	Mean	Std. Dev.	10%	90%
A. Panel					
Number of candidates	1708	11.32	4.96	6	18
Female candidates	1708	0.80	1.00	0	2
Median candidate age	1708	44.07	5.45	38	51
Eligible voters (1000)	1708	175.54	47.15	139.76	218.02
Turnout	1708	0.71	0.09	0.58	0.82
Winning vote share	1708	0.44	0.09	0.33	0.55
NOTA votes / total votes	630	0.019	0.013	0.006	0.035
NOTA votes / eligible voters	630	0.014	0.009	0.004	0.026
Non-reserved constituencies:					
Number of SC candidates	1144	1.31	1.43	0	3
Number of ST candidates	1144	0.48	1.04	0	2
B. Repeated cross section					
Number of candidates	6685	10.54	5.35	5	17
Female candidates	6685	0.73	0.97	0	2
Median candidate age	6685	44.99	5.78	38	53
Eligible voters (1000)	6685	180.75	88.23	41.20	292.90
Turnout	6685	0.71	0.13	0.53	0.87
Winning vote share	6685	0.45	0.10	0.32	0.56
NOTA votes / total votes	1176	0.015	0.116	0.004	0.030
NOTA votes / eligible voters	1176	0.010	0.009	0.003	0.022
Non-reserved constituencies:					
Number of SC candidates	4842	1.24	1.56	0	3
Number of ST candidates	4842	0.26	0.79	0	1

Notes: The panel dataset contains the 2008 and 2013 state assembly elections in the states of Karnataka, NCT Delhi, Mizoram, Rajasthan, Madhya Pradesh, and Chhattisgarh. The repeated cross-section contains all assembly elections between 2006 and 2014 in 25 states. Turnout is total votes divided by the number of eligible voters. Winning vote share is the winner's share of all non-NOTA votes. Source: Election Commission of India.

Table 3: Candidate characteristics in the panel data

Variable	All	2008	2013
Age	44.42	43.98	44.87
Female	0.07	0.07	0.07
General caste	0.61	0.62	0.61
SC	0.21	0.21	0.20
ST	0.15	0.16	0.13
Selected parties:			
INC	0.09	0.09	0.08
BJP	0.08	0.08	0.08
BSP	0.08	0.08	0.07
SP	0.04	0.05	0.03
Independent	0.39	0.42	0.36
NOTA	0.03	-	0.06
N	19957	9762	10195

Notes: Average age and fraction of candidates with different characteristics in the 2008 and 2013 state assembly elections in the states of Karnataka, NCT Delhi, Mizoram, Rajasthan, Madhya Pradesh, and Chhattisgarh. Source: Election Commission of India.

state-year or the district-year level for the reduced-form analysis. Table 4 summarizes these variables for the 25 states in the repeated cross section. We complement this with data on the growth rate of per capita state domestic product from the Reserve Bank of India.

For the structural exercise, demographic characteristics are needed at the constituency level. We are not aware of any existing dataset with appropriate coverage. We create the necessary dataset using the 2011 Indian Census by aggregating village-level information and matching it to constituencies using GIS coordinates. Specifically, we obtained GIS boundary files for the 2013 electoral constituencies and the 2001 census. To use data from the 2011 census, we proceed in two steps. First, we match sub-districts ("tehsil") in the 2001 census to the 2011 census using village names.[15] Administrative boundaries in India change over time, with tehsils, districts, and even states splitting up into new units. This step of our matching procedure is based on the smallest administrative unit available in the census, the village. Second, we match the 2011 census data to each 2013 electoral constituency using the 2001 sub-district boundaries. We use area-weighted averages to compute values for constituencies that overlap several sub-districts.

Of the 854 constituencies, we have constituency boundary files for 850. We were able to match 723 of these to the sub-district data from the census. The location of these con-

[15] Sub-districts, called tehsils in most states, are administrative units above the villages and below the districts and the states.

Table 4: Voter demographics at the state level (repeated cross-section)

	Obs	Mean	Std. Dev.	Min	Max
Labor force participation	50	0.58	0.07	0.46	0.75
Unemployment rate	50	0.02	0.03	0.00	0.15
Real household earnings (Rp per week)	50	1708	665	864	4335
Fraction illiterate	50	0.25	0.13	0.04	0.52
Fraction primary school or less	50	0.23	0.09	0.13	0.53
Female per 1000 male	50	987	72	790	1172
Fraction urban	50	0.32	0.18	0.09	0.97
State NDP growth rate	50	6.11	4.57	-5.38	24.31

Notes: Source for all variables except NDP growth rate: National Sample Survey, rounds 62, 64, 66, 68, 71. Individual surveys for respondents above 18 were aggregated to the state level. Household earnings deflated to 2001 prices using the CPI from the Reserve Bank of India. Source for NDP growth rate: Reserve Bank of India.

stituencies is shown on Figure 1. Most of the constituencies we lose during the matching (70) are in NCT Delhi. We lose this entire state because the census data is not sufficiently disaggregated. Of the matched constituencies, 520 are affected by NOTA in 2013 and 203 are not (in the full panel, these numbers are respectively 630 and 224).

Table 5 shows the summary statistics of the census data at the constituency level. The variables include basic demographic characteristics such as gender, caste, literacy, and employment as well as economic characteristics of the households (infrastructure and asset ownership).

5 Patterns in the data

5.1 NOTA votes

The first noteworthy feature of the data is that a positive number of voters voted for NOTA. Despite the fact that voting for NOTA could not affect the results of the election, in the 9/25 states in our data affected by the policy a total of 2.51 million voters chose this option.[16] The distribution of the NOTA vote share is shown on Figure 2. NOTA was chosen by a positive number of voters in every constituency, receiving an average vote share of 1.5% with a range

[16] While the fact that people voted on an option that could not affect the election might seem surprising, this behavior is not qualitatively different from votes cast on small extra-parliamentary parties, or from voting in an election where voters have no trust in the integrity of the election and that their vote will actually be counted. For example, in Cantú and García-Ponce (2015), despite having just voted, around 5% of Mexican voters exiting the election booth say that they have no confidence that "the vote you cast for president will be respected and counted for the final result," and another 15% state that they have little confidence.

Figure 1: Constituencies in the merged dataset

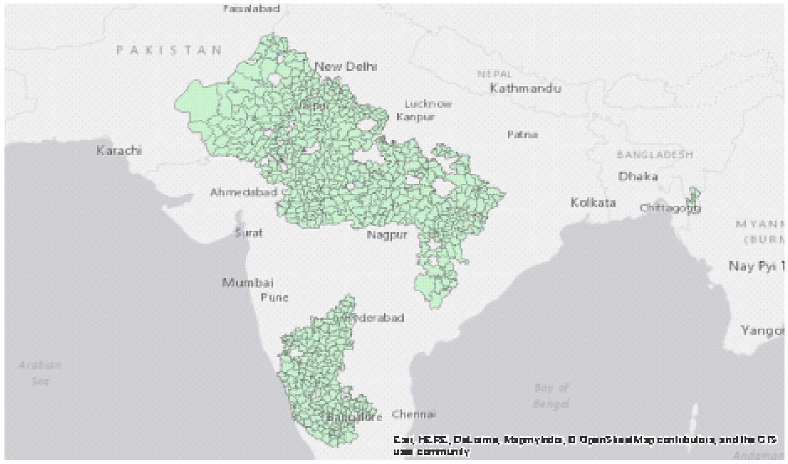
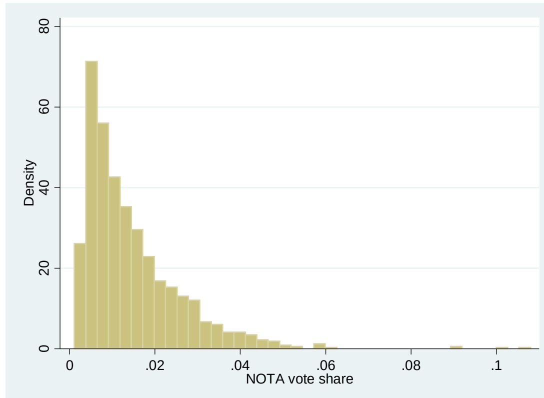


Table 5: Demographic characteristics of the constituencies from the Indian Census (panel dataset)

Variable	Mean	Std. Dev.	10%	90%
Literate	0.58	0.09	0.47	0.69
Fraction male	0.51	0.01	0.50	0.53
Fraction SC	0.18	0.07	0.08	0.26
Fraction ST	0.29	0.18	0.16	0.58
Fraction employed	0.46	0.05	0.40	0.52
Fraction rural workers	0.66	0.17	0.44	0.84
Fraction of households with infrastructure:				
No latrine	0.71	0.22	0.37	0.91
Water near premises	0.47	0.10	0.35	0.59
Water on premises	0.26	0.16	0.10	0.47
Fraction of households owning asset:				
Car	0.03	0.02	0.01	0.06
Computer	0.01	0.01	0.002	0.02
Phone	0.55	0.20	0.25	0.78
TV	0.37	0.17	0.15	0.61

Notes: Source: Indian Census, 2011. Village level census data was merged to assembly constituency GIS boundary files as described in the text. N = 723.

Figure 2: Distribution of NOTA vote shares across constituencies



Notes: NOTA vote share is measured as a fraction of total votes cast. N = 1176.

of 0.1-11%. As a fraction of all eligible voters (including abstainers) 1% voted for NOTA.[17] In the average constituency, NOTA received more votes than 7 of the candidates running for election. In 97 constituencies out of 1176, the vote share of NOTA was larger than the winning margin (the difference between the vote share of the winner and the runner up).

One consequence of the introduction of NOTA is simply the appearance of another option on the ballot. A potential concern is that this new option confused some voters who chose it by mistake. Our findings below on increased turnout are difficult to reconcile with this interpretation. If NOTA had simply confused voters at the voting booth, we would not expect to find a positive impact on voter turnout. An alternative way that voters might be confused is if they mistakenly thought that voting for NOTA would somehow affect the electoral result (for example, that the election would be invalid if NOTA obtained a majority). We find this interpretation implausible for two reasons. First, given the 1.5% actual vote share on NOTA, voting for NOTA to invalidate the election would have required not just confusion about electoral rules but also extremely unrealistic expectations about the number of voters planning to vote for NOTA. Second, if voters mistakenly thought that NOTA would affect elections, we would expect them to be less likely to vote for NOTA as they gain more experience. To check for this, we looked at the 2014 general elections, held at the same time in all states. Some of these states already had experience with NOTA at the assembly elections in 2013, while others did not. If the use of NOTA in 2013 was due to voter confusion, we would expect the experienced states to vote for NOTA less than the inexperienced states. In fact, the opposite is true: in this general election the average

[17] In the panel data, in the 5 states affected by NOTA the average vote share of NOTA among total votes cast (eligible voters) was 1.9% (1.4%).

NOTA vote share among the experienced states was 1.28%, compared to 1.09% among non-experienced states. Voters in states that had more experience with NOTA were significantly *more* likely to use it ($p = 0.027$).

Figure 2 reveals some heterogeneity in NOTA votes across constituencies. Our structural analysis below will relate this heterogeneity to voter demographics and model how different groups of voters choose between the different options on the ballot. As a precursor to this analysis, in the Appendix we run cross-sectional regressions of the NOTA vote share on a variety of constituency characteristics. We find evidence of systematic heterogeneity: the NOTA vote share is significantly higher in reserved constituencies and in constituencies with more illiterate voters, more women, more ST, and a lower share of rural workers. One possible interpretation of these correlations is that economically disadvantaged and / or politically disenfranchized voters obtain more utility from expressing themselves by voting for NOTA.

5.2 The effect of NOTA on turnout

As discussed in section 2, testing for a consumption utility of voting requires inferring the changes in individual voters' behavior following the introduction of NOTA. Specifically, did the NOTA policy lead to some voters choosing to vote for NOTA instead of abstaining? While this question is difficult to answer using reduced form methods, in the Indian case we can use a simple difference-in-differences approach to answer a related question: did the NOTA policy lead to some voters choosing to vote instead of abstaining?

While identifying the impact of NOTA on turnout is challenging because the introduction of the policy took place at the same time across India, we can exploit variation in the Indian electoral calendar for a difference-in-differences analysis. Specifically, we use the fact that elections to the state assemblies are held at different times in different states (see Table 1). For the panel dataset, our specification is the following:

$$Y_{cst} = \alpha_0 + \alpha_1 NOTA_{st} + \boldsymbol{\alpha}_2 \mathbf{X}_{cst} + \gamma_c + \eta_t + \varepsilon_{cst}, \tag{3}$$

where Y_{cst} is turnout in constituency c of state s in year t, $NOTA_{st}$ equals 1 if the NOTA policy is in place and 0 otherwise, $\mathbf{X}_{cst}$ are control variables, and γ_c and η_t are constituency and year fixed effects, respectively. Using states that held elections in 2008 and 2013, the parameter of interest, α_1 is identified by comparing the change in turnout in the states that held elections in both years without NOTA to the change in turnout in the states that were affected by NOTA in 2013 (but not in 2008). For the repeated cross-section sample, the specification is identical to (3) except that we replace the constituency fixed effects γ_c with state fixed effects γ_s.[18]

Table 6 shows the results from estimating equation (3). In column (1), we use the panel dataset and control for the number of eligible voters in a constituency, state labor force participation, weekly household earnings, and education (as well as constituency and year fixed effects). The coefficient estimate on NOTA indicates a positive turnout effect but with the small number of states and state-level variation in the policy, the estimate is highly imprecise.[19] In column (2) we repeat the same specification for the repeated cross-section, replacing the constituency fixed effects with state fixed effects. The point estimate on NOTA remains similar but the precision improves drastically, indicating a statistically significant turnout effect of 3 percentage points.[20] In column (3) we add as additional controls a dummy for reserved constituencies as well as the following state-level variables: unemployment, sex ratio, urbanization, and the growth rate of state per capita net domestic product. The estimated effect of NOTA remains robust to these additional controls.

The main threat to identification in the regressions presented in Table 6 is other events or policies that may affect changes in turnout between assembly elections held before and after the introduction of NOTA. In the Appendix, we present a number of robustness checks: we exclude national election years, control for the extent to which constituencies were affected by electoral redistricting, drop specific states where political events during our period of study might potentially confound the effect of NOTA, and finally include additional controls for voting costs, such as weekend elections, weather, and the density of voting stations in a constituency. We find that our estimates are robust to these alternative samples and specifications.

Overall, these results indicate that the presence of the NOTA option on the election ballot increased turnout in the average constituency. It is interesting to note that in each case the 95% confidence interval around the estimates includes the fraction of eligible voters who voted for NOTA in the data (1% in the repeated cross-section). This may indicate that NOTA voters turn out to vote specifically for this option and would abstain if this option was not present on the ballot. At the same time, without a model our ability to infer NOTA voters' counterfactual behavior using this aggregate analysis is fundamentally limited. Did the NOTA policy lead to some voters choosing to vote *for NOTA* instead of abstaining?

[18] Recall that the panel includes constituencies from 6 states, 5 of which were affected by NOTA. The repeated cross section contains constituencies from 25 states, 9 of which were affected by NOTA (5 in 2013 and 4 in 2014).

[19] Because the NOTA policy varies at the state level, inference needs to account for clustering. Given the small number of clusters, we obtain the p-value for the effect of NOTA by using a wild bootstrap procedure as recommended by Cameron and Miller (2015) with the 6-point weight distribution of Webb (2013).

[20] We obtain similar inference using either standard errors clustered by state or the wild bootstrap procedure.

Table 6: The impact of NOTA on turnout, DD estimates

	(1)	(2)	(3)
NOTA	0.035	0.029**	0.030*
s.e.		(0.013)	(0.016)
bootstrap p-value	[0.587]	[0.036]	[0.060]
Eligible voters, labor force participation, hh earnings, education	x	x	x
Political reservation, unemployment, sex ratio, urbanization, NDP growth rate			x
Constituency FE	x		
State FE		x	x
R^2	0.78	0.18	0.19
N	1708	6685	6685
States	6	25	25

Notes: Estimates of the effect of the NOTA policy on turnout from Eqn. (2). All regressions control for time fixed effects, the log number of eligible voters in a constituency and its square, and the following state-level variables: labor force participation, real weekly household earnings, fraction of illiterates, fraction with primary school or less as highest education. Column (3) also controls for reserved constituencies and the following state level variables: unemployment, sex ratio, fraction urban, and the growth rate of net domestic state product. Standard errors clustered by state in parentheses. The bootstrap p-value was computed using a wild bootstrap procedure with a 6-point weight distribution. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

Our structural analysis below will make such inference regarding individual voters' voting patterns possible.

6 Estimating the effect of NOTA from a demand system for candidates

In this section we estimate a model of voter choice among candidates, NOTA and abstention, adapting the consumer demand model of Berry, Levinsohn and Pakes (1995) (BLP).[21] Two factors make this approach particularly useful in the voting context. First, the rules governing elections imply that several assumptions of the model naturally hold. Second, some of the estimation challenges addressed by the method - notably the need to make inferences regarding individual behavior from aggregate data - are central to any voting application. We discuss each of these factors in turn.

While the typical Industrial Organization applications view the static discrete choice framework as an approximation, the rules governing elections actually make this model quite realistic in the electoral context. By the nature of elections, voters are restricted to a

[21] See also Nevo (2001) and Petrin (2002).

discrete choice between voting for a candidate, voting for NOTA, or abstaining.[22] Choices are made simultaneously by all voters in a given race - unlike consumers, voters cannot adjust the timing of their choice. When making a choice, the voter has before him a complete list of all available options on the ballot, in contrast to a consumer who may not be aware of all available brands of the product he is considering buying. Electoral competition takes place in markets (electoral districts) that are administratively defined and where vote shares are fully observed by the researcher.[23] Finally, abstention (the "outside good" chosen by somebody who does not choose any of the alternatives on the ballot) is well defined, and administrative data on its prevalence is readily available.[24]

What makes this estimation framework particularly attractive in the electoral context is that it addresses two estimation challenges central to most of these applications. First, due to the secret ballot, the need to infer individual behavior from aggregate data is of major importance in this literature. While in IO applications one could in principle obtain individual data (e.g., from consumer loyalty programs), in the voting context this is virtually impossible as in most cases administrative data on individual choices simply does not exist. Because of this, most existing research either relies exclusively on aggregate analysis, or uses voter survey data to analyze individual behavior. Since voter survey data is susceptible to well-known biases, being able to use the available administrative data for estimating a micro-founded model is highly valuable.

Second, legislative elections often create a problem of dimensionality due to the presence of many candidates. Even in two-party systems, the same party will have different candidates in different electoral districts. This makes it important to study voter choices in "characteristic space," i.e., projecting the heterogeneity between candidates onto a small number of characteristics such as gender, age, race, etc. This results in similar gains as IO researchers have noted in going from products to product characteristics. But because there are typically more candidates running for election than any given industry has products, the gains in the electoral application are likely to be even larger.

[22] By contrast, when choosing which product to buy, consumers may purchase a mixture of products even if they only consume one at any given time.

[23] By contrast, studies of consumer choice have to rely on proxying the true market in which a set of products compete, e.g., based on geographic areas. It is also common in these studies to rely on a sample of products and stores, while in the electoral context complete data on the vote shares of all candidates (including abstention) is readily available.

[24] By contrast, defining the relevant outside good for, e.g., the market for new cars, requires making assumptions: is it a used car, public transportation, nothing?

6.1 Specification: demand

Consider a constituency $t \in \{1, ..., T\}$ where voters can vote for candidates $j \in \{1, ..., J\}$. If available, we include the NOTA option in the list of candidates, and we let $j = 0$ indicate abstention (the outside option). Each candidate is described by a set of characteristics observed by the researcher and a set of unobserved characteristics. Besides the candidate's party, characteristics observed in our data include gender, age, and caste. Unobserved characteristics include, for example, the candidate's experience. Assume that the utility that voter i derives from voting for candidate $j \in \{1, ..., J\}$ can be specified as

$$U_{ijt} = \boldsymbol{\beta}_i \mathbf{x}_{jt} + \bar{\xi}_j + \xi_{jt} + \varepsilon_{ijt}, \tag{4}$$

where $\mathbf{x}_{jt} = (x_{jt}^1, ..., x_{jt}^K)'$ is a vector of the observed characteristics of party j's candidate, $\bar{\xi}_j$ is the average popularity of the party, ξ_{jt} captures voters' valuation of unobserved candidate characteristics, and ε_{ijt} is a stochastic term with mean zero drawn from a Type-I extreme value distribution (the role of this assumption will be made clear below). Unless stated otherwise we treat the NOTA option as another candidate, including a NOTA indicator to identify characteristics (e.g., gender) which are only defined for actual candidates.[25]

Voter preferences for the various candidate characteristics are represented by the coefficients $\boldsymbol{\beta}_i = (\beta_i^1, ..., \beta_i^K)$. These vary across individuals based on demographic variables and unobserved characteristics:

$$\boldsymbol{\beta}_i' = \boldsymbol{\beta} + \boldsymbol{\Pi}\mathbf{d}_i + \boldsymbol{\Sigma}\mathbf{v}_i, \tag{5}$$

where $\mathbf{d}_i = (d_i^1, .., d_i^D)'$ is a vector of "observed" demographic variables, $\mathbf{v}_i = (v_i^1, ..., v_i^K)'$ are "unobserved" voter characteristics, and the parameters are in the $(K \times 1)$ vector $\boldsymbol{\beta}$, the $(K \times D)$ matrix $\boldsymbol{\Pi}$, and the $(K \times K)$ scaling matrix $\boldsymbol{\Sigma}$. We assume that the $\mathbf{v}_i$ are drawn from independent Normal distributions with mean 0. As in most consumer demand applications, "observed" variables are individual characteristics whose empirical distribution is known (from census data), while the distribution of "unobserved" characteristics has to be assumed. While individual level consumption data is sometimes available, in the voting context, given the secrecy of the ballot, it is generally impossible to directly match individual characteristics to votes.

To complete the choice set, the utility of the "outside option" must be specified. In the voting context, this is the utility from abstention, which also includes any direct and indirect costs voting. In consumer demand applications constructing the outside choice typically involves two sets of assumptions: assumptions about what consumers do when they don't

[25] For example, if the only candidate characteristic is gender (g_j), then $x_j = [(1 - n_j)g_j, n_j]$ where n_j is equal to 1 for NOTA and 0 otherwise.

purchase a specific product, and assumptions about what constitutes a market. In the voting context neither of these is necessary, since electoral constituencies are exogenously given and voters who do not vote necessarily abstain. We let

$$U_{i0t} = \boldsymbol{\pi}_0 \mathbf{d}_i + \sigma_0 v_i^0 + \varepsilon_{i0t}, \tag{6}$$

which allows for the utility of abstention (hence the cost of voting) to vary by observed demographics and unobserved voter characteristics. As discussed below, we also include in (4) state and year fixed effects and indicators for whether the constituency is reserved for SC or ST candidates. Since voter choices will be determined by the differences in utilities, including these variables in (4) is equivalent to including them in the specification of the utility of abstention in (6). Thus, we are also allowing for further heterogeneity in voting costs as captured by these variables.

Let $\bar{\boldsymbol{\xi}} = (\bar{\xi}_1, ..., \bar{\xi}_J)$, $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}, \bar{\boldsymbol{\xi}})$, and $\boldsymbol{\theta}_2 = (\boldsymbol{\Pi}, \boldsymbol{\Sigma})$, and let $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ represent the parameters of the model. Substituting (5) into (4), we can write

$$U_{ijt} = \delta_{jt} + \mu_{ijt} + \varepsilon_{ijt},$$

where $\delta_{jt} \equiv \boldsymbol{\beta}\mathbf{x}_{jt} + \bar{\xi}_j + \xi_{jt}$ and $\mu_{ijt} \equiv (\boldsymbol{\Pi}\mathbf{d}_i + \boldsymbol{\Sigma}\mathbf{v}_i)\mathbf{x}_{jt}$. Voters choose to vote for one of the candidates (including NOTA) or abstain. This implicitly defines the set of demographics and unobserved variables for which voter i will choose candidate j:

$$A_{jt}(\mathbf{x}, \boldsymbol{\delta}_t(\boldsymbol{\theta}_1), \boldsymbol{\theta}_2) = \left\{ (\mathbf{d}_i, \mathbf{v}_i, \boldsymbol{\varepsilon}_{it}) | U_{ijt} \geq U_{ilt} \ for \ l = 0, 1, ..., J \right\},$$

where $\mathbf{x}$ are all the candidate characteristics, $\boldsymbol{\delta}_t = (\delta_{1t}, ..., \delta_{Jt})$, and $\boldsymbol{\varepsilon}_{it} = (\varepsilon_{i1t}, ..., \varepsilon_{iJt})$. Given the distribution of $(\mathbf{d}_i, \mathbf{v}_i, \boldsymbol{\varepsilon}_{it})$, we can integrate over A_{jt} to obtain the vote shares $s_{jt}(\mathbf{x}, \boldsymbol{\delta}_t(\boldsymbol{\theta}_1), \boldsymbol{\theta}_2)$ predicted by the model. Under the assumed Type-I extreme value distribution for ε_{ijt}, these are given by

$$s_{jt}(\mathbf{x}, \boldsymbol{\delta}_t(\boldsymbol{\theta}_1), \boldsymbol{\theta}_2) = \int \frac{\exp\left[\delta_{jt} + \mu_{ijt} - \mu_{i0t}\right]}{1 + \sum\limits_{q \geq 1} \exp\left[\delta_{qt} + \mu_{iqt} - \mu_{i0t}\right]} dF\left(\mathbf{d}_i, \mathbf{v}_i\right), \tag{7}$$

where $\mu_{i0t} \equiv \boldsymbol{\pi}_0 \mathbf{d}_i + \sigma_0 v_i^0$ and $F\left(\mathbf{d}_i, \mathbf{v}_i\right)$ denotes the distribution of the voter characteristics. These predicted vote shares are a function of the data ($\mathbf{x}$), the parameters ($\boldsymbol{\theta}$), and the unobserved candidate characteristics ξ.

6.2 Specification: supply

While some political economy models treat candidates as exogenously given, others, notably the citizen-candidate literature, emphasize that politician characteristics may emerge endogenously in the political process (Osborne and Slivinski, 1996; Besley and Coate, 1997). To allow for this possibility while keeping the problem tractable, we adopt a simple simultaneous-moves specification of the supply of candidates. We will use this framework to justify the instrumental variables we use in the estimation below.[26]

As in the citizen-candidate literature suppose that implemented policies depend on elected politician's characteristics and that candidate characteristics emerge endogenously in the political process. In particular, suppose that candidates are chosen by a political party that cares about winning as well as the policy implemented by the winner. In constituency t, party j's payoff is given by $v_{jt}(\mathbf{x}_t, \mathbf{s}_t)$, where $\mathbf{x}_t = (\mathbf{x}_{1t}, ..., \mathbf{x}_{Jt})$ are the characteristics of all candidates running in the election and $\mathbf{s}_t = (s_{1t}, ..., s_{Jt})$ are the vote shares that determine the winner. Vote shares are determined by candidates' observed characteristics as well as the voter valuations ξ_{jt}, as in equation (7). Thus, $s_{jt} = s_{jt}(\mathbf{x}_t, \boldsymbol{\xi}_t)$ where $\boldsymbol{\xi}_t = (\xi_{1t}, ..., \xi_{Jt})$ (to simplify the exposition, we set $\bar{\boldsymbol{\xi}} = 0$ in this section).[27]

Given a party's membership, fielding candidates with some characteristics may be easier than others. For example, a lower caste party may find it difficult to field general caste candidates. A simple way to capture this is by supposing that party j faces a budget constraint $m = \sum_k q_{jt}^k x_{jt}^k \equiv \mathbf{q}_{jt}\mathbf{x}_{jt}$, where m is the budget available to spend on candidates (assumed constant for simplicity) and q_{jt}^k is the "price" of increasing a candidate's characteristic k in constituency t. For example, if $x^k = 1$ denotes a general caste candidate, q_{jt}^k may be the extra cost of finding such a candidate and convincing him to run. Prices will generally depend on such factors as a party's membership, the economic and demographic characteristics of a constituency, the prestige associated with a political career in the local population, etc. We assume that parties take these prices as given.

Suppose that parties choose the characteristics of their candidates simultaneously, after voter valuations $\boldsymbol{\xi}_t$ have been realized. In a Nash equilibrium, the characteristics of party

[26] In the consumer demand literature, it is common to model firms that compete on prices but take all other product characteristics as exogenously given. In our context, there is no natural separation between endogenous and exogenous candidate characteristics so we will treat all characteristics except NOTA as potentially endogenous.

[27] For simplicity in this section we also assume that the ξ_{jt} terms represent valuation shocks that the parties have no control over and that only affect their payoff through the vote shares s. Allowing these to also capture unobserved (to the researcher) candidate characteristics that the parties can affect would change the exposition without affecting the main argument.

j's candidate will satisfy

$$\mathbf{x}_{jt}^* \in \arg\max_{\mathbf{x}_{jt}} (v_{jt}(\mathbf{x}_t, \mathbf{s}_t(\mathbf{x}_t, \boldsymbol{\xi}_t)) | m = \mathbf{q}_{jt}\mathbf{x}_{jt})$$

or

$$\mathbf{x}_{jt}^* = \mathbf{x}_{jt}^*(\mathbf{x}_t, \boldsymbol{\xi}_t, \mathbf{q}_{jt}). \tag{8}$$

In words, candidate j's characteristics depend on the characteristics of all candidates running, voters' valuation of the unobserved characteristics, and party j's cost of increasing the various characteristics in the given constituency. This has two implications. First, the dependence of observed characteristics $\mathbf{x}_{jt}$ on voter valuations $\boldsymbol{\xi}_t$ creates an endogeneity problem for the estimation of the utility functions (4). Second, it is plausible that the prices $\mathbf{q}_{jt}$ for a given party are correlated across constituencies t. For example, a lower caste party is likely to face a higher price to field a general caste candidate in all constituencies within a state. This implies that the characteristics of a given party's candidates will be correlated across constituencies. As explained below, this opens the possibility of using candidate characteristics in neighboring constituencies as instrumental variables in the estimation.

6.3 Estimation

6.3.1 Estimation algorithm

Estimation follows the algorithm proposed by BLP. The idea is to treat the unobserved characteristics ξ as the econometric error and derive moment conditions that can be used to estimate the parameters using Generalized Method of Moments (GMM). Detailed treatments of the procedure can be found in BLP and Nevo (2000, 2001) so we only provide a brief summary below.

Consider a dataset with information on candidate characteristics $\mathbf{x}$ and actual vote shares S_{jt}. BLP show that, for given $\boldsymbol{\theta}_2$, it is possible to numerically solve for $\boldsymbol{\delta}_t$ from the equations $s_{jt}(\mathbf{x}, \boldsymbol{\delta}_t, \boldsymbol{\theta}_2) = S_{jt}$, i.e., equating the model-predicted vote shares to those observed in the data. Using the resulting values of $\delta_{jt}(\boldsymbol{\theta}_2)$, we express the unobserved candidate characteristics as $\xi_{jt}(\boldsymbol{\theta}) = \delta_{jt}(\boldsymbol{\theta}_2) - \bar{\xi}_j - \boldsymbol{\beta}\mathbf{x}_{jt}$. Given our data and with δ_{jt} computed, this is a standard econometric error, which depends nonlinearly on the parameters of the model. While we do not expect $\xi_{jt}(\boldsymbol{\theta})$ to be independent of $\mathbf{x}_{jt}$, we can find a suitable set of instruments $\mathbf{Z}_{jt}$ and use the moment conditions $E[\xi_{jt}(\boldsymbol{\theta})|\mathbf{Z}_{jt}] = 0$ to estimate the parameters using GMM. Thus, we find

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \ \boldsymbol{\xi}(\boldsymbol{\theta})'\mathbf{Z}\mathbf{W}^{-1}\mathbf{Z}'\boldsymbol{\xi}(\boldsymbol{\theta}),$$

where $\boldsymbol{\xi}(\boldsymbol{\theta})$ is the vector of errors, $\mathbf{Z}$ is the matrix of instruments, and $\mathbf{W}$ is the weighting matrix.

To compute the estimate, we use the standard two-step GMM procedure (Greene, 2003, p206). We first set $\mathbf{W} = \mathbf{Z}'\mathbf{Z}$ and compute an initial estimate of the parameters, $\boldsymbol{\theta}^1$. We then use this initial estimate to recompute a robust weight matrix $\mathbf{W} = \frac{1}{n}\sum_{j,t}^{n}[\xi_{jt}(\boldsymbol{\theta}^1)]^2\mathbf{Z}'_{jt}\mathbf{Z}_{jt}$, and use this updated weight matrix to compute the final parameter estimates.

6.3.2 Identification

Identification of the model relies on moment conditions corresponding to included exogenous variables and excluded instruments. In this framework, the need for instrumental variables arises for two reasons. First, instruments are needed to generate enough moment conditions to identify the nonlinear parameters in voters' utility functions. Thus, instruments are necessary even if ξ_{jt} and $\mathbf{x}_{jt}$ are uncorrelated. Second, instruments are needed because some of the candidate characteristics could be endogenous, as suggested by equation (8).[28] In the context of consumer demand estimation, where "voters" are the consumers and "candidates" are the products, it is common to use instruments based on the characteristics of other products produced by the same firm and the characteristics of products produced by other firms (e.g., BLP; Nevo, 2001). A natural counterpart in our setting is to think of firms as the parties that field the candidates. Using this analogy, we use as instruments the average characteristics of a given party's candidates in other constituencies within the state as well as the average characteristics of all candidates in other constituencies within the state. For example, for candidate gender we create an IV by taking the fraction of female candidates of the given party in other constituencies in the state, and another IV by taking the fraction of female candidates among all candidates in other constituencies in the state. What makes these instruments possible is the variation in the characteristics of a given party's candidates across constituencies as each election is contested by a different set of individuals. This avoids the difficulties that sometimes arise in the consumer demand literature from insufficient variation in the characteristics of a firm's products across markets (e.g., Nevo, 2001).[29]

Beyond the analogy to product characteristics, a rationale for these instruments in our case may be given based on the supply of candidates available to each party, as in section

[28] The endogeneity problem is likely to be present even if candidate characteristics are assumed to be fixed at the time of the election. We only observe a partial list of characteristics $\mathbf{x}_{jt}$, and unobserved characteristics that influence ξ_{jt} (including experience, voting record, qualifications, physical appearance, etc.) are likely to be correlated with these.

[29] As usual, variables that enter the utility function and are treated as exogenous serve as their own instruments. In our case, this includes state, year, and party fixed effects, as well as the NOTA indicator and its interaction with demographics.

6.2. For a given party, the "price" of increasing a particular candidate characteristic is likely to be correlated across constituencies due to the composition of the party's membership, demographic characteristics of the constituency, etc. This implies that the characteristics of a particular party's candidates are likely to be correlated across constituencies.

The identifying assumption, expressed in the moment conditions, is that unobserved voter valuations for a particular candidate in a constituency are conditionally independent of these instruments. One case in which this assumption will hold is if, controlling for party-specific means and demographics, constituency-specific voter valuations ξ_{jt} are independent across constituencies (but may be correlated for a given constituency over time). This rules out a popularity shock to *some* of a party's candidates as would be caused, e.g., by a regionally coordinated advertising campaign (a campaign raising the popularity of all candidates would be captured by the party dummies). See Hausman (1996) and Nevo (2001) for analogous assumptions in the consumer demand literature. A second case in which the identifying assumption will hold is if parties do not condition their choice of candidates on the popularity shocks ξ_{jt} in (8). For example a party with an SC base may find it impossible to respond to a popularity shock by finding a candidate from a different caste in time for the election. In this case, the mix of candidate characteristics offered by a party would reflect the supply of characteristics in the relevant population, rather than respond to popularity shocks among voters.

Under either interpretation, (8) implies that characteristics of candidates fielded by a given party in different constituencies will only be correlated due to correlation in the prices $(\mathbf{q}_{j1}, ...\mathbf{q}_{jT})$. Thus, for a given party, the characteristics of its candidates $\mathbf{x}_{jt'}$ in constituencies $t' \neq t$ are valid instruments for the characteristics of the candidate running in constituency t. Similarly, every other party's characteristics $\mathbf{x}_{j't'}$ in constituencies $t' \neq t$ are correlated with its characteristics $\mathbf{x}_{j't}$ in constituency t due to the correlation in the prices faced by that party. Thus, under (8) $\mathbf{x}_{j't'}$ are also valid instruments for $\mathbf{x}_{jt}$.

As described above, the main concern regarding the use of Hausman instruments is the correlation of ξ_{jt} across some (but not all) constituencies. In our setting, a natural possibility is correlation within a particular state, as would be the case if a party conducted a successful campaign within that state only. This concern is mitigated however by the fact that in our sample 73% of the parties field candidates in only one of the states.[30] For these parties, controlling for a party fixed effect captures any correlation between the voter valuations ξ_{jt} across constituencies. To control for such correlation for the other 27% of the parties, we also experiment with a specification that includes state $\times$ party fixed effects.

[30] Once we aggregate small parties and independents as described in the next section, this figure becomes 65%.

6.3.3 Practical issues

As described above, parties play an important role in our specification: we estimate party fixed effects and we define our IVs based on parties. One difficulty arises because of the presence of many small parties. There are a total of 200 parties in the data, but half of them field candidates in only 1 of every 40 constituency within a state. A second, related difficulty is the presence of independent candidates (candidates not affiliated with any party). There are 6751 of these candidates in the data, but 70% of them receive less than 1% of the votes in a constituency and only 3% receive more than 10%. Each of these parties and candidates adds a new parameter that is difficult to identify due to the small number of constituencies where the party is represented (in the extreme case of an independent candidate running in only one year, identifying the fixed effect is not possible). To circumvent these difficulties, we create a "Small Party" category comprising parties fielding candidates in less than a third of the constituencies in any given state and we average all small party candidates' characteristics within a constituency (we do this after constructing the instruments so that the individual IVs are aggregated also). We also create an "Independent Party" containing all independent candidates within a state, and aggregate them within constituencies in the same way. After this aggregation, we are left with a total of 22 parties.

We include in the analysis the full list of candidate characteristics available in the data: gender, caste, age and party. We select the constituency characteristics to be included based on the variables that indicated significant heterogeneity in voter preferences for NOTA in the regressions in section 5. We include fraction male, literacy, fraction SC, fraction ST, and the share of rural workers.

The BLP algorithm requires numerically solving the integral in (7) to obtain the predicted market shares. We do this in the standard way by drawing individual voters from the distribution of demographics in each constituency, computing the predicted individual probabilities of voting for each candidate, and averaging across simulations to obtain the simulator for the integral.

6.4 Results

6.4.1 Parameter estimates

We present parameter estimates for different specifications of the above model in Tables 7 and 8. First, we set $\mathbf{\Pi}$ and $\mathbf{\Sigma}$ equal to 0 so that voter heterogeneity only enters through the ε_{ijt} terms in equation (4). This is the conditional Logit model, and we estimate it with and without instruments in columns (1) and (2) of Table 7. We report coefficient estimates on the candidate characteristics as well as a subset of the control variables (reserved constituency indicators and dummies for 3 major parties).

In column (3), we keep $\mathbf{\Pi} = 0$ but allow for random coefficients on all the candidate characteristics as well as the utility of abstention (through $\mathbf{\Sigma v}_i$ in equation (5) and $\sigma_0 v_i^0$ in equation (6)). The estimates of $\mathbf{\Sigma}$ are in the first column of Table 8. Finally, we present the full model which allows for both observed and unobserved heterogeneity in voters' evaluation of the various candidate characteristics. These estimates are in column (4) of Table 7 ($\boldsymbol{\beta}$) and the remainder of Table 8 ($\mathbf{\Sigma}$ and $\mathbf{\Pi}$). Moving from column (1) to column (2) of Table 8, we kept those elements of $\mathbf{\Sigma}$ that were statistically significant.

Table 7: Estimates of the linear parameters of the demand system

Variable	OLS Logit (1)	Instrumental Variable Logit (2)	Random Coefficients (3)	Full model (4)
Female	-0.090**	-0.153	-1.547*	-3.725
	(0.044)	(0.268)	(0.928)	(2.992)
Age	1.029***	8.925***	8.013***	10.231*
	(0.116)	(1.423)	(2.097)	(5.767)
SC candidate	-0.370***	0.273	0.377	0.120
	(0.047)	(0.223)	(0.385)	(1.266)
ST candidate	-0.043	0.082	0.143	2.002
	(0.067)	(0.148)	(0.386)	(2.878)
NOTA	0.269***	-3.439***	-4.263***	-61.830***
	(0.074)	(0.076)	(0.890)	(11.052)
Reserved SC	0.302***	-0.202	-0.258	-0.140
	(0.048)	(0.189)	(0.360)	(0.954)
Reserved ST	0.290***	0.351***	0.734***	-1.923
	(0.064)	(0.115)	(0.194)	(2.688)
Party: BJP	2.664***	-5.135***	-4.677***	-1.994
	(0.042)	(0.728)	(0.990)	(3.104)
Party: INC	2.738***	-5.157***	-4.724***	-2.089
	(0.041)	(0.745)	(1.017)	(3.177)
Party: BSP	0.325***	-7.039***	-6.802***	-4.348
	(0.045)	(0.644)	(0.895)	(2.933)

Notes: The table reports estimates of the linear parameters of the model (β). All specifications include a full set of party dummies (three of which are reported in the table) as well as dummies for states, years, and constituency reservation status. Columns (2)-(4) include instrumental variables as described in the text. Robust standard errors in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively.

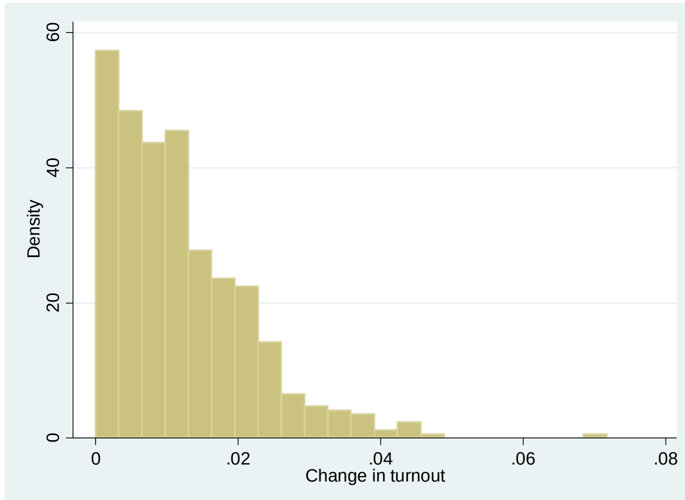
In Table 7, coefficients change substantially between the OLS and IV specifications, suggesting that instrumenting is important. Tables 7 and 8 reveal that voters value candidate characteristics other than party affiliation. For example, older male candidates tend to receive more votes. Allowing for taste heterogeneity among voters regarding these characteristics in the full model substantially reduces the estimated impact of party affiliation on votes (column (4) of Table 7). In Table 8, areas with a higher share of ST voters yield more votes for ST candidates (this effect is identified from variation across non-reserved constituencies). NOTA is a less popular option in constituencies with more literate voters and in rural areas.

Table 8: Estimates of the nonlinear parameters of the full model

Variable	Standard Deviations (1)	Standard Deviations (2)	Interactions with Demographic Variables: Fraction SC	Fraction ST	Fraction male	Literacy	Rural
Female	-3.873**	-7.081*	-	-	6.214	-	-
	(1.513)	(3.813)			(6.498)		
Age	-0.915	-	-	-	-	-	-
	(2.453)						
SC candidate	0.003	-	0.060	-	-	-	-
	(1.799)		(2.188)				
ST candidate	1.951**	4.971***	-	6.452*	-	-	-
	(0.859)	(1.708)		(3.556)			
NOTA	1.343	-	-	-43.704	-42.581	-45.274*	-24.270**
	(0.839)			(43.322)	(34.231)	(24.276)	(11.305)
Constant	-0.060	-	-	-	-	13.066***	6.082***
	(1.650)					(1.881)	(1.927)

Notes: The table reports estimates of the nonlinear parameters of the model (Π and Σ). 'Standard deviations' refer to the Σ parameters of the random coefficients. The specification in column (1) corresponds to column (3) in Table 7 and restricts $\Pi = 0$. The remaining columns of the table correspond to the specification in column (4) in Table 7. Robust standard errors in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively.

Figure 3: Impact of NOTA on turnout



Notes: Distribution of the changes in turnout across constituencies. Mean = 0.0119, median = 0.0100, N = 519.

6.4.2 Counterfactual analysis: The impact of NOTA

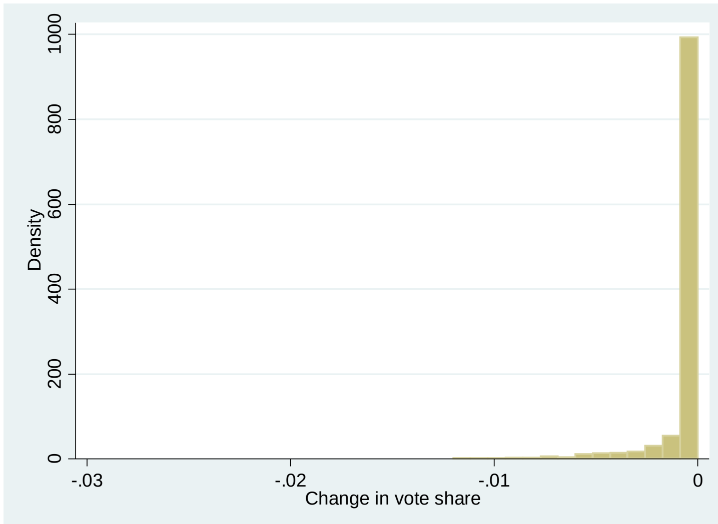
In this section we use the estimated model to evaluate the impact of introducing NOTA. We restrict attention to those constituencies in our data that had the NOTA option available in 2013 and perform a counterfactual experiment where the NOTA option is removed. We compute new vote shares and turnout rates under this counterfactual scenario, and calculate the impact of NOTA as the difference between the actual and the counterfactual outcomes.

The estimated impact of NOTA on turnout is shown in Figure 3. The average increase in turnout is above 0.5 percentage points in 74% of the constituencies, with an average of 1.2 percentage points. This is somewhat lower than the effect we obtained from the reduced form exercise in section 5.2, but very close to the 1.4 percent of eligible voters who voted for NOTA in the data.

The estimated impact of NOTA on individual candidates' vote shares is shown in Figure 4. The reduction in vote shares is smaller than 0.5 percentage points in absolute value for 96% of the candidates, with a mean of 0.06 percentage points. This indicates that substitution from voting for a candidate to voting for NOTA is minimal.

To gauge the impact of NOTA on parties, Table 9 aggregates voter choices by party. For each party in the data, the first column gives the number of candidates and the second column shows the fraction of the 101.2 million eligible voters who voted for that party. The third column shows the difference relative to the counterfactual without NOTA. For the two

Figure 4: Impact of NOTA on candidates' vote shares



Notes: Distribution of the changes in vote shares across candidates. Mean = -0.0006, median = -0.0000, N = 3068.

largest parties, BJP and INC, we estimate a loss in total votes of around 0.15 percentage points. The change in the votes cast on other parties is even smaller. By contrast the change in the overall abstention rate (and hence turnout) is a magnitude larger at 1.2 percentage points. This is similar to the estimated effect across all constituencies.

It should be emphasized that allowing for the flexible random coefficients specification above was crucial to obtain these results. The more restrictive Logit model could not possibly have resulted in these effects. As is well known, the Logit specification implies that substitution patterns only depend on observed choice shares. In our case, this would imply that adding the NOTA option would, by construction, cause the biggest change in the most popular candidate's vote share. This is illustrated by the last column of Table 9 which shows the counterfactual implications that would be obtained from the Logit specification. As can be seen, this would imply that substitution towards NOTA is similar for voters of the two largest parties and for abstaining voters.

The patterns emerging from the demand model are similar to those observed in the difference-in-differences analysis. The results indicate that NOTA votes are mostly cast by voters who would have chosen to abstain in the counterfactual without NOTA and turn out to vote specifically for NOTA. This provides strong evidence that voters derive positive consumption utility from voting for this option.

Table 9: Impact of NOTA on vote shares by party

			Change with NOTA	
Party	N. of candidates	Percent of all voters	Full model	Logit
BJP	506	32.90	-0.17	-0.63
BSP	498	3.63	-0.02	-0.09
BYS	102	0.10	0.00	-0.01
CSM	54	0.22	0.00	-0.01
GGP	44	0.20	0.00	-0.01
INC	518	26.74	-0.14	-0.58
Independents	468	4.96	-0.03	-0.11
JGP	85	0.06	0.00	0.00
MNF	10	0.06	0.00	0.00
NPEP	133	1.30	-0.01	-0.02
SP	194	0.43	0.00	-0.01
ZNP	11	0.02	0.00	0.00
Small Party	445	2.70	-0.01	-0.05
Abstention		25.11	-1.20	-0.76

Notes: The table shows, for each party, the total number of candidates and the corresponding share of all voters in the data (out of 101.168 million eligible voters). The last two columns are the simulated effects of introducing NOTA in the full random coefficients specification as well as in the more restrictive Logit model.

7 Conclusion

This paper analyzed India's NOTA policy which gives people the option to participate in elections and cast a valid "None of the Above" vote without the possibility of affecting the electoral outcome. Individuals who choose to vote for NOTA but would abstain otherwise must derive a consumption utility from voting that is specific to this option. Thus, the NOTA policy makes it possible to test apart various components of the consumption utility of voting. To address the challenge that, due to the secret ballot, individual choices are not observable, we estimate counterfactual voter behavior using a structural model and techniques borrowed from the consumer demand literature. The model allows for rich heterogeneity in voter preferences and relates these parameters to the aggregate vote returns. In counterfactual simulations, we find that the NOTA policy resulted in increased turnout. Based on the estimated model, virtually all the NOTA votes observed in the data represent new voters who showed up specifically to vote for NOTA and who would have abstained in the absence of this option. These patterns are also supported by the findings from a reduced-form analysis. Our results show that, in this context, the presence of an option-specific consumption utility from voting is necessary to explain the data. For example, voters may derive utility from expressing their protest against one or more of the candidates running for election. In our sample, models that do not incorporate an option-specific utility of voting would have a hard time explaining the data.

To the extent that voter participation is valuable in a democracy, our results suggest that having a NOTA option on the ballot may be a desirable policy. It creates both political participation and individual utility.

References

[1] Ali, N., and C. Lin (2013): "Why People Vote: Ethical Motives and Social Incentives," *American Economic Journal: Microeconomics* 5(2), 73–98.

[2] Bénabou, R. J., and J. Tirole. (2006). "Incentives and Prosocial Behavior," *American Economic Review* 96(5), 1652-1678.

[3] Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841-890.

[4] Besley, T., and S. Coate (1997): "An Economic Model of Representative Democracy," *Quarterly Journal of Economics* 112, 85-114.

[5] Blais, A. (2000): *To vote or not to vote*, University of Pittsburgh Press, Pittsburgh, PA.

[6] Brennan, G., and L. Lomasky (1993): *Democracy and Decision: The Pure Theory of Electoral Preference*. Cambridge University Press, Cambridge, UK.

[7] Cameron, A.C., and D.L. Miller (2015): "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources* 50(2), 317-372.

[8] Cantú, F., and O. García-Ponce (2015): "Partisan losers' effects: Perceptions of electoral integrity in Mexico," *Electoral Studies* 39, 1-14.

[9] Cho, W.K.T., and C.F. Manski (2008): "Cross-Level/Ecological Inference," in: J.M. Box-Steffensmeier, H.E. Brady, and D. Collier (eds.) *The Oxford Handbook of Political Methodology*, Oxford University Press, Oxford, UK.

[10] Coate, S. and M. Conlin (2004): "A Group Rule: Utilitarian Approach to Voter Turnout: Theory and Evidence," *American Economic Review* 94(5) 1476-1504.

[11] Coate, S., M. Conlin, and A. Moro (2008): "The performance of pivotal-voter models in small-scale elections: Evidence from Texas liquor referenda," *Journal of Public Economics* 92, 582–596.

[12] Degan, A. and A. Merlo (2011): "A Structural Model of Turnout and Voting in Multiple Elections," *Journal of the European Economic Association* 9(2), 209–245.

[13] DellaVigna, S., J.A. List, U. Malmendier, and G. Rao (2015): "Voting to Tell Others," working paper, UC Berkeley.

[14] Dittmann, I., D. Kubler, E. Maug, and L. Mechtenberg (2014): "Why Votes Have Value: Instrumental Voting with Overconfidence and Overestimation of Others' Errors," *Games and Economic Behavior* 84, 17–38.

[15] Downs, A. (1957): *An economic theory of democracy*, Harper and Row, New York, NY.

[16] Driscoll, A., and M.J. Nelson (2014): "Ignorance or Opposition? Blank and Spoiled Votes in Low-Information, Highly Politicized Environments," *Political Research Quarterly* 67(3), 547-561.

[17] Duffy, J. and M. Tavits (2008): "Beliefs and Voting Decisions: A Test of the Pivotal Voter Model," *American Journal of Political Science* 52(3), 603–618.

[18] Feddersen, T. and A. Sandorini (2006): "A Theory of Participation in Elections," *American Economic Review* 96(4), 1271-1282.

[19] Feddersen, T., S. Gailmard, and A. Sandroni, A. (2009): "Moral bias in large elections: theory and experimental evidence," *American Political Science Review* 103, 175–192.

[20] Fiorina, M. P. (1976): "The Voting Decision: Instrumental and Expressive Aspects," *Journal of Politics* 38(2), 390-413.

[21] Fujiwara, T. (2015): "Voting Technology, Political Responsiveness, and Infant Health: Evidence From Brazil," *Econometrica* 83(2), 423–464.

[22] Gerber, A.S., D.P. Green, and C.W. Larimer (2008). "Social Pressure and Voter Turnout: Evidence from a Large-scale Field Experiment," *American Political Science Review* 102(1), 33-48.

[23] Glasgow, G., and R.M. Alvarez (2005): "Voting behavior and the electoral context of government formation," *Electoral Studies* 24, 245–264.

[24] Greene, W.H. (2002): *Econometric Analysis*, Fifth Ed., Prentice Hall, Upper Saddle River, NJ.

[25] Harbaugh, W. T. (1996): "If People Vote Because They like to, then Why do so Many of Them Lie?," *Public Choice* 89(1/2), 63-76.

[26] Hausman, J. (1996): "Valuation of New Goods Under Perfect and Imperfect Competition," *in:* T. Bresnahan and R. Gordon (eds.): *The Economics of New Goods, Studies in Income and Wealth Vol. 58*, National Bureau of Economic Research, Chicago, IL.

[27] Kamenica, E., and L. Egan Brad (2014): "Voters, dictators, and peons: expressive voting and pivotality," *Public Choice* 159(1), 159-176.

[28] Herron, M. C. and J. S. Sekhon (2005): "Black Candidates and Black Voters: Assessing the Impact of Candidate Race on Uncounted Vote Rates," *Journal of Politics*, 67(1), 154–177.

[29] McAllister, I. and T. Makkai (1993): "Institutions, Society or Protest? Explaining Invalid Votes in Australian Elections," *Electoral Studies* 12(1), 23-40.

[30] McMillan, A. (2010): "The Election Commission," *in*: N.G. Jayal and P.B. Mehta (eds): *The Oxford Companion to Politics in India*, Oxford University Press, Oxford, UK.

[31] Nevo, A. (2000): "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy* 9, 513-548.

[32] Nevo, A. (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69, 307-342.

[33] Ortoleva, P. and E. Snowberg (2015): "Overconfidence in Political Behavior," *American Economic Review* 105(2), 504–535.

[34] Osborne, M., and A. Slivinski (1996): "A Model of Political Competition with Citizen-Candidates," *Quarterly Journal of Economics* 111, 65-96.

[35] Petrin A. (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy* 110, 705-729.

[36] Power, T. J. and J. C. Garand (2007): "Determinants of invalid voting in Latin America," *Electoral Studies* 26, 432-444.

[37] Razin, R. (2003): "Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates," *Econometrica* 71(4), 1083-1119.

[38] Rekkas, M. (2007): "The Impact of Campaign Spending on Votes in Multiparty Elections," *Review of Economics and Statistics* 89(3): 573–585.

[39] Riker, H. W. and P. C. Ordeshook (1968): "A Theory of the Calculus of Voting," *American Political Science Review* 62(1), 25-42.

[40] Selb, P., and S. Munzert (2013): "Voter overrepresentation, vote misreporting, and turnout bias in postelection surveys," *Electoral Studies* 32, 186–196.

[41] Shachar, R., and B. Nalebuff (1999): "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review* 89(3), 525-47.

[42] Shayo, M. and A. Harel (2012): "Non-consequentialist voting," *Journal of Economic Behavior and Organization* 81, 299– 313.

[43] Superti, C. (2015): "Vanguard of the Discontents: Blank and Null Voting as Sophisticated Protest," working paper, Harvard University.

[44] Tyran, J.-R. (2004): "Voting when money and morals conflict: an experimental test of expressive voting," *Journal of Public Economics* 88, 1645-1664.

[45] Uggla, F. (2008): "Incompetence, Alienation, or Calculation? Explaining Levels of Invalid Ballots and Extra-Parliamentary Votes," *Comparative Political Studies* 41(8), 1141–1164.

[46] Webb, M. (2013): "Reworking Wild Bootstrap Based Inference for Clustered Errors," working paper, University of Calgary.

A Appendix

A.1 The correlates of NOTA votes

In this section we investigate the correlation between NOTA votes and constituency characteristics. We use the same sample as in our structural exercise (the panel dataset matched to the census data) and run simple cross-sectional regressions on the 520 constituencies that are affected by the NOTA policy in 2013. We include state fixed effects and, to avoid confounding our estimates by differential turnout across constituencies, we measure NOTA vote shares as a fraction of total votes cast.[31]

The results are in Table 10. We find substantial heterogeneity in NOTA votes across constituencies. For example, the NOTA vote share is significantly higher in reserved constituencies and in constituencies with more illiterate voters, more women, more ST, and a lower share of rural workers. Each of these patterns is consistent with a variety of possible explanations. One possible interpretation is that NOTA votes are higher in more economically disadvantaged populations, reflecting a general dissatisfaction with elected leaders in these constituencies. Note however that the coefficients remain unchanged if we add controls for various indicators of infrastructure and economic activity in column (2). Another possible interpretation is that NOTA votes come from politically underrepresented voters, such as women and non-SC or ST voters in reserved constituencies.

In columns (3) and (4) we add candidate characteristics to the regression. These estimates are merely suggestive because candidate entry will be affected by their expected vote share, and this could be correlated with the fraction of voters choosing NOTA. We find that constituencies with more candidates running result in lower NOTA vote shares, which is consistent with NOTA reflecting dissatisfaction with the menu of candidates being offered. We do not find evidence that the presence of female, SC or ST candidates affects NOTA votes.

[31] Using NOTA votes as a share of eligible voters yields very similar results.

Table 10: The correlates of NOTA votes

	(1)	(2)	(3)	(4)
Constituency characteristics:				
Reserved SC	0.005***	0.005***	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Reserved ST	0.011***	0.011***	0.009***	0.009***
	(0.002)	(0.001)	(0.002)	(0.002)
Literacy	-0.021***	-0.035***	-0.015**	-0.026**
	(0.006)	(0.010)	(0.006)	(0.010)
Size	-0.006*	-0.008**	-0.003	-0.005
	(0.003)	(0.004)	(0.003)	(0.004)
Fraction male	-0.215***	-0.230***	-0.133***	-0.190***
	(0.031)	(0.046)	(0.033)	(0.043)
Fraction SC	0.002	0.010	-0.006	0.006
	(0.008)	(0.009)	(0.007)	(0.008)
Fraction ST	0.010***	0.008*	0.010**	0.007
	(0.004)	(0.004)	(0.004)	(0.004)
No latrine		0.002		0.004
		(0.004)		(0.004)
Water nearby		0.016**		0.014**
		(0.006)		(0.006)
Water at home		0.011*		0.013**
		(0.006)		(0.005)
Fraction employed		0.015		-0.004
		(0.011)		(0.011)
Rural workers		-0.019***		-0.015***
		(0.005)		(0.005)
Car ownership		0.023		0.020
		(0.037)		(0.034)
Computer ownership		-0.029		0.036
		(0.057)		(0.052)
Phone ownership		-0.009		-0.010*
		(0.006)		(0.005)
TV ownership		-0.003		-0.007
		(0.006)		(0.006)
Candidate characteristics:				
Number of candidates			-0.001***	-0.001***
			(0.000)	(0.000)
No female			-0.001	-0.001
			(0.001)	(0.001)
<15% female			-0.000	-0.000
			(0.001)	(0.001)
Median age			-0.000	-0.000
			(0.000)	(0.000)
No SC			0.000	-0.000
			(0.001)	(0.001)
<15% SC			-0.000	0.000
			(0.001)	(0.001)
No ST			-0.000	-0.000
			(0.001)	(0.001)
<10% ST			0.002	0.001
			(0.001)	(0.002)
R^2	0.57	0.60	0.63	0.66
N	520	520	520	520

Notes: The dependent variable is the share of NOTA votes among all votes. Regressions at the constituency level for the cross-section of constituencies affected by the NOTA policy in 2013. Merged dataset: average demographic characteristics are from the census, average candidate characteristics are from the Election Commission. All regressions include state fixed effects. Robust standard errors in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively.

A.2 Robustness of the DD estimates

This section explores the robustness of the difference-in-difference estimates presented in section 5.2 of the paper to various political events affecting one or more states.

A.2.1 National elections

In our study period, Indian national elections took place in Spring 2009 and 2014. Recall that we do not include in the analysis the four states that hold their assembly elections simultaneously with the national election. In the remaining states, because split-ticket voting (constituencies voting for different parties at the state and national levels) is common in India, it is ex ante not obvious that events affecting national turnout, like the wave of support for the BJP in the 2014 national elections, would affect assembly elections.[32] If national elections did have an impact, we would expect this to be the strongest for states that held assembly elections in October and December of the national election years. If national elections impacted turnout in the assembly elections in these states, this has the potential to confound our estimates of the NOTA policy introduced between the two national election years in September 2013.

In Table 11 we exclude the national election years from the sample. Columns (1) and (2) exclude 2014 and columns (3) and (4) exclude both 2009 and 2014. Odd numbered columns correspond to the specification in column (2) of Table 6 with the basic controls and even numbered columns to column (3) with the extended controls. We find that all point estimates are similar to, and if anything slightly larger than the 3 percentage points effect we found in Table 6. National elections do not appear to confound the estimates reported in the main text.

A.2.2 Redistricting

Another potential confound is the electoral redistricting that took place in April 2008. Because elections are held every 5 years and NOTA was introduced in September 2013, none of the states that were affected by NOTA in our period of study were redistricted, while most states that were not affected by NOTA were redistricted. Thus, redistricting has the potential to confound our estimates of NOTA.[33]

To control for this, we create a constituency-level measure of redistricting by using GIS boundary files to compare constituencies before and after the delimitation. Our first measure

[32] Note also that increased support for the BJP would lead to more BJP votes rather than NOTA votes.

[33] For example, if redistricting lowered turnout, our estimate of NOTA's effect of turnout would likely be biased upward.

Table 11: Effect of NOTA on turnout, excluding national election years

	(1)	(2)	(3)	(4)
NOTA	0.033**	0.033**	0.030*	0.031*
	(0.015)	(0.016)	(0.015)	(0.015)
Basic controls	x		x	
Extended controls		x		x
Excluded years	2014	2014	2009, 2014	2009, 2014
R^2	0.18	0.20	0.19	0.21
N	6139	6139	5680	5680
States	25	25	22	22

Notes: Estimates of the effect of the NOTA policy on turnout from Eqn. (2) using the repeated cross section sample with specific years excluded. All regressions control for state and year fixed effects, the log number of eligible voters in a constituency and its square, and the following state-level variables: labor force participation, real weekly household earnings, fraction of illiterates, fraction with primary school or less as highest education. The Extended controls specification also controls for reserved constituencies and the following state level variables: unemployment, sex ratio, fraction urban, and the growth rate of net domestic state product. Standard errors clustered by state in parentheses. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

calculates for each current constituency that was redistricted in our study period the largest area that was part of a single constituency before the redistricting. For example, a value of 0.8 for this "maximum overlap" measure indicates that 80% of the current constituency's area was part of a single constituency pre-delimitation (while the remaining 20% was part of one or more different constituencies). The higher the maximum overlap, the less a constituency was affected by redistricting. Our second measure, rather than focus on the largest area of overlap, uses each overlapping area to create an index of "territorial fractionalization." If a constituency overlaps with n pre-delimitation constituencies with $s_1, ..., s_n$ denoting the share of its area falling in each of these, then the fractionalization index is $1 - \sum_{i=1}^{n} s_i^2$. The larger this value, the more the current constituency was affected by redistricting. Both of these measures are available for 22 states (constituency boundary files are not available for the states of Assam, Manipur, and Nagaland).

Table 12 presents regressions corresponding to specifications (2) and (3) in Table 6 controlling for these measures of redistricting. The first two columns repeat columns (2) and (3) in Table 6 on the 22 states with available redistricting measures. Columns (3) and (4) then add the maximum overlap measure and columns (5) and (6) the territorial fractionalization index. As can be seen, adding either measure of redistricting to the regressions causes little change in the estimated effect of NOTA. The estimates also retain their significance, except for column (6) where the standard error increases just enough to yield a p-value of 0.101.[34]

[34] The coefficients on the redistricting measures are never statistically significant.

Table 12: Effect of NOTA on turnout, controlling for redistricting

	(1)	(2)	(3)	(4)	(5)	(6)
NOTA	0.033**	0.022*	0.031**	0.020*	0.030**	0.020
	(0.016)	(0.011)	(0.015)	(0.011)	(0.014)	(0.012)
Basic controls	x		x		x	
Extended controls		x		x		x
R^2	0.21	0.23	0.21	0.23	0.21	0.23
N	6084	6084	6084	6084	6084	6084
States	22	22	22	22	22	22

Notes: Estimates of the effect of the NOTA policy on turnout from Eqn. (2) using the repeated cross section sample. Columns (1) and (2) are run on the states with available constituency boundary files. Columns (3) and (4) control for redistricting using the maximum overlap measure and columns (5) and (6) using the territorial fractionalization index. All regressions control for state and year fixed effects, the log number of eligible voters in a constituency and its square, and the following state-level variables: labor force participation, real weekly household earnings, fraction of illiterates, fraction with primary school or less as highest education. Even-numbered columns also control for reserved constituencies and the following state level variables: unemployment, sex ratio, fraction urban, and the growth rate of net domestic state product. Standard errors clustered by state in parentheses. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

A.2.3 State-specific events

Turning to state-specific events that may confound our estimates, we identified four states where various events may plausibly affect 2013 or 2014 turnout relative to the previous election (that is, turnout in the with-NOTA election relative to turnout in the without-NOTA election). In Chhattisgarh, Maoist insurgents conducted terrorist attacks in 2010 and May 2013, between the 2008 and 2013 elections in this state. In Jammu & Kashmir, various incidents occurred between its 2008 and 2014 elections, including a border skirmish in January 2013 between India and Pakistan described by observers as one of the worst in 10 years. In Delhi, a new anti-corruption party, Aam Aadmi entered politics in 2012, energized voters, and emerged as the second-largest party in the 2013 assembly election. Finally, Maharashtra held its 2009 election a year after the 2008 terrorist attacks in Mumbai on several hotels and public buildings, and security concerns may have depressed voter turnout there.

In Table 13, we repeat specifications (2) and (3) from Table 6 excluding each of these states one at a time and then all four of them. The results corresponding to the first specification are in column (1) and column (2) corresponds to the second specification with the extended set of controls. All these coefficients are close to the 3 percentage point effect found in Table 6. The events in these four states do not appear to drive the estimated effect of NOTA on turnout reported in the main text.

Table 13: Effect of NOTA on turnout, robustness to state-specific events

Excluded state	Effect of NOTA		N
	Basic controls	Extended controls	
Chhattisgarh	0.025*	0.035	6505
	(0.014)	(0.021)	
Maharashtra	0.031**	0.031*	6109
	(0.015)	(0.015)	
Delhi	0.029**	0.031*	6545
	(0.013)	(0.016)	
Jammu and Kashmir	0.030**	0.031*	6511
	(0.014)	(0.016)	
All four	0.028*	0.039*	5615
	(0.015)	(0.019)	

Notes: Estimates of the effect of the NOTA policy on turnout from Eqn. (2) using the repeated cross section sample with specific states excluded. All regressions control for state and year fixed effects, the log number of eligible voters in a constituency and its square, and the following state-level variables: labor force participation, real weekly household earnings, fraction of illiterates, fraction with primary school or less as highest education. The Extended controls specification also controls for reserved constituencies and the following state level variables: unemployment, sex ratio, fraction urban, and the growth rate of net domestic state product. Standard errors clustered by state in parentheses. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

A.2.4 Voting costs

In this section we include further controls in the difference-in-differences specification in an attempt to control for any time-varying differences in voting costs across constituencies. First we obtained data from the Election Commission on the day of the week that the elections in each constituency were held. We create a dummy for wether the election was held on a weekend, as this might affect the cost of turnout. Constituencies within a state typically go to the polls in groups over a period of 2-3 days, so this variable varies at the constituency level. In column (1) and (2) of Table 14 we find that controlling for the Weekend dummy has no impact on our results.

Second, we obtained data on rainfall on each election day. Several studies document that bad weather can raise the cost of turnout. We matched the geo-coded weather data given in 1x1 degree cells to state boundaries and took the area-weighted average of the cells covering each state to obtain an estimate for each state for each election day. Just as for the Weekend dummy, rainfall data also varies at the constituency level. Columns (3) and (4) in Table 14 show that controlling for rainfall does not affect our estimates of the impact of NOTA.

Third, we obtained data on the number of voting stations in each constituency. We divide this by the number of eligible voters in order to proxy for the convenience of voting. For example, a low number of voting stations per voters may lead to long wait times at the

Table 14: Effect of NOTA on turnout with controls for voting costs

	(1)	(2)	(3)	(4)	(5)	(6)
NOTA	0.028*	0.030*	0.029**	0.030*	0.051***	0.052**
	(0.015)	(0.016)	(0.013)	(0.015)	(0.015)	(0.019)
Weekend	-0.003	-0.008				
	(0.009)	(0.009)				
Rainfall			0.001	0.005		
			(0.005)	(0.006)		
Voting stations					79.791**	86.529***
					(30.391)	(30.277)
Basic controls	x		x		x	
Extended controls		x		x		x
R^2	0.18	0.19	0.18	0.19	0.2	0.22
N	6685	6685	6684	6684	6676	6676
States	25	25	25	25	25	25

Notes: Estimates of the effect of the NOTA policy on turnout from Eqn. (2) using the repeated cross section sample with additional controls. Weekend is a dummy equal to 1 for elections held on a weekend. Rainfall is rainfall on election day in cm. Voting stations is the number of voting stations per eligible voters. All regressions control for state and year fixed effects, the log number of eligible voters in a constituency and its square, and the following state-level variables: labor force participation, real weekly household earnings, fraction of illiterates, fraction with primary school or less as highest education. The Extended controls specification also controls for reserved constituencies and the following state level variables: unemployment, sex ratio, fraction urban, and the growth rate of net domestic state product. Standard errors clustered by state in parentheses. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

voting booth and discourage some people from voting. We include this variable as a control in columns (5) and (6) of Table 14. These estimates should be interpreted with care since the number of voting stations could be endogenous for a number of reasons (for example, areas with historically high turnout may receive more stations). Nevertheless, it is reassuring that controlling for differences in voting costs as proxied by the number of stations per voter actually reinforces our findings.