None Of The Above*

Gergely Ujhelyi, Somdeep Chatterjee, and Andrea Szabó†

March 12, 2019

Abstract

Who are "protest voters" and do they affect elections? We study this question using the introduction of a pure protest option ("None Of The Above") on Indian ballots. To infer individual behavior from administrative data, we borrow a model from the consumer demand literature in Industrial Organization. We find that in elections without NOTA, most protest voters simply abstain. Protest voters who turn out scatter their votes among many candidates and consequently have little impact on election results. From a policy perspective, NOTA may be an effective tool to increase political participation, and can attenuate the electoral impact of compulsory voting.

*We thank Sourav Bhattacharya, Francisco Cantú, Alessandra Casella, Aimee Chin, Julien Labonne, Arvind Magesan, Eric Mbakop, Suresh Naidu, Mike Ting, and especially Thomas Fujiwara for useful comments and suggestions. We also thank seminar participants at Oxford, Columbia, WUSTL, Calgary, the 2016 Wallis Institute Conference, the 2016 Banff Workshop in Empirical Microeconomics, NEUDC 2016, and the 2016 STATA Texas Empirical Microeconomics conference for comments. Thanks to seminar participants at the Indian Statistical Institute Kolkata, Indian Institute of Technology Kanpur, and Public Choice Society 2015 for feedback on an earlier version. We gratefully acknowledge use of the Maxwell/Opuntia Cluster and support from the Center for Advanced Computing and Data Systems at the University of Houston. A previous version of the paper circulated under "'None Of The Above' Votes in India and the Consumption Utility of Voting" (first version: November 1, 2015).

†Author affiliations and contact: Department of Economics, University of Houston, gujhelyi@uh.edu; Business Environment (Economics) Group, Indian Institute of Management Lucknow, somdeep@iiml.ac.in; Department of Economics, University of Houston, aszabo2@uh.edu. Ordering of authors: Ujhelyi (lead author), Chatterjee, Szabó (alphabetical).

1 Introduction

In many elections, some voters are thought to be motivated at least in part by a desire to protest. For example, voters may choose to abstain to express their dislike of all candidates running, or they may choose a fringe candidate to signal their dissatisfaction with the leading candidates. In recent years, protest voting has been proposed as partial explanation for the presidential victory of Donald Trump, the result of the Brexit referendum, and the surge in electoral strength of extremist parties in various countries.[1]

How many of these "protest voters" are there, and what choices do they make? How many of them abstain? Which candidates benefit from their votes, and do they affect electoral outcomes? These basic questions are difficult to answer empirically because voters' motivation is unobserved. Abstention may be an indication of protest, or it could just be a rational response to the low likelihood of being pivotal in the election. Votes for fringe or extremist candidates may reflect protest, or may indicate that voters actually agree with those candidates' views. One option to measure the protest motive is to conduct voter surveys, but there is skepticism regarding the veracity of responses to these types of questions. Another option is to study blank or spoiled ballots under the assumption that these reflect protest votes, but it is typically difficult to know what fraction of these simply reflect voting mistakes.

In this paper, we use a unique policy in India to identify protest voters and the choices they make using actual voter behavior. In 2013 the Indian Supreme Court mandated that all national and state elections must offer a "None Of The Above" (NOTA) option to voters. Votes cast for NOTA are counted and reported separately but do not directly affect the outcome of the election (in the sense that the winner is still the candidate with a plurality of votes among votes cast for candidates).[2] Thus, NOTA is quite literally a protest option. We define "protest voters" simply as the voters who chose NOTA once this option became available, and ask what choices these voters would have made without NOTA on the ballot. This will give us an estimate for the behavior of protest voters in "normal" elections that do not have the explicit protest option.

Studying this question is challenging because it requires inferring *individual* voter behavior in a counterfactual scenario without the NOTA option. However, because ballots are

[1] Descriptive headlines in recent years include "Protest Voters, What the Hell Have You Done?" (*Huffngton Post* 11/9/2016), "Martin Schulz: Trump win is Brexit-like protest vote" (*Politico,* 11/9/2016), and "'I thought I'd put in a protest vote': the people who regret voting leave" (*The Guardian* 11/25/2017).

[2] In some countries, options similar to NOTA have direct electoral consequences, e.g., by affecting whether the election has to be repeated (see section 2). While in India NOTA votes do not directly affect the outcome, having NOTA on the ballot may matter for outcomes indirectly, by changing voters' behavior. Whether this is the case is one of the central questions we analyze below.

secret, individual voter behavior is observed neither with nor without NOTA. Instead, it must be inferred from aggregate data.

We begin with some preliminary evidence on the aggregate impact of NOTA from a reduced form exercise. If protest voters normally abstain, then all else equal introducing NOTA should be associated with an increase in voter turnout. To test for this, we exploit 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 by comparing the change in voter turnout in states not yet affected by the policy to changes in states that were already affected. This analysis yields suggestive evidence that, in the average electoral district, the introduction of the NOTA policy significantly increased turnout. The magnitude of the estimate (2-3 percentage points) is statistically similar to the vote share of NOTA observed in the data, which is consistent with the idea that most protest voters normally choose abstention.

While suggestive, these aggregate findings mask the heterogeneity in substitution patterns at the individual level. Which candidates are chosen by NOTA voters who do turn out to vote when NOTA is not available? How does the behavior of protest voters differ across constituencies? To study these questions, we relate the aggregate voting returns to individual voter behavior using a model of voter demand for candidates. We adapt the BLP approach to consumer demand (Berry, Levinsohn and Pakes, 1995), 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 based on voter demographics, 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, the individual candidates, and abstention in counterfactual simulations where the NOTA option is removed.

The results of this analysis confirm that NOTA increased turnout. We estimate that, in the average constituency, approximately 2/3 of NOTA voters would normally abstain. The share of NOTA voters who normally abstain is larger in constituencies with more minority voters and fewer minority candidates, which is consistent with the NOTA option being valued by voters who feel more disenfranchised. Our estimates also show that the 1/3 of NOTA voters who normally turn out scatter their votes among several candidates. As a result, protest voting affects the winner in less than 0.5 percent of the elections in our sample.

Our exercise allows us to identify parties that benefit relatively more from the protest vote. We find that these parties are "fringe" parties that either represent very specific voter groups (such as a disenfranchised minority, or "the youth"), or have radical platforms (such as introducing the death penalty for corruption and rape). Although we estimate that these parties' supporters include relatively more protest voters, our findings indicate that the overwhelming majority of these supporters are *not* protest voters, in the sense that they chose them for other reasons than simply as a substitute for NOTA.

To the extent that all participation in a democracy is valuable, our finding that adding a NOTA option to the ballot can increase voter turnout is relevant in its own right - particularly since this policy is relatively simple to implement. This provides support for the arguments of the Indian Supreme Court, who stated increased turnout as its main goal when introducing NOTA.

In an effort to further boost turnout, Indian lawmakers and the courts are currently considering introducing compulsory voting. As recently emphasized by Ambrus et al. (2017), NOTA gains added significance in a compulsory voting system by giving voters the option to participate without influencing the electoral outcome. To study this issue, we perform a counterfactual analysis on the role of NOTA under compulsory vs. voluntary voting. We estimate that a quarter of currently abstaining voters would vote for NOTA under compulsory voting. Moreover, the availability of NOTA is an important mediator of the impact of compulsory voting on election results. Without NOTA, the number of electoral districts where compulsory voting changes the winner would be 50% higher.

Our paper contributes to several strands of the vast literature on voter behavior and turnout. In the Downsian "calculus of voting" paradigm (Downs, 1957; Riker and Ordeshook, 1968), voters' motivation is composed of the instrumental benefit from the possibility of being pivotal, and a consumption (or expressive) utility. The latter may be a general consumption utility from showing up to vote, or an option-specific utility associated with voting for a particular option (such as the satisfaction from casting a protest vote). Empirically identifying these consumption utility components of voting has proved challenging, since it is difficult to find real world settings where pivotality considerations can be completely shut down. A number of field experiments have found that social pressure can increase turnout (e.g., Gerber et al., 2008; DellaVigna et al., 2017), which indicates the existence of a general consumption utility (the utility from complying with a social norm to vote). 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 (a particular type of) option-specific consumption utility in real-world elections.[3]

As an alternative to the Downsian view, another strand of the literature on turnout emphasizes the signaling role of elections. In these models, voters are motivated in part by a desire to strategically signal their preferences or private information to politicians (Lohmann, 1993; Razin, 2003; Castanheira, 2003; McMurray, 2017) or to other voters (Lohmann, 1994; Piketty, 2000). Protest voting has a natural interpretation in these models as reflecting voters' attempt to communicate their dissatisfaction (Myatt, 2017), and our results can be viewed as providing an empirical test of this phenomenon. To be clear, we do not take a stand on whether the Downsian or the signaling view is correct: protest voting could reflect either consumption utility or a strategic signaling motivation. Our focus is on empirically identifying protest voting and some of its consequences. We briefly discuss implications of our findings for theories of voting in section 9 below.

Methodologically, our paper follows a growing literature using structural models to link aggregate administrative data on electoral outcomes to individual voter preferences (Coate and Conlin, 2004; Rekkas, 2007; Coate et al., 2008; Grypari, 2011; Kawai and Watanabe, 2013; Gordon and Hartmann, 2013; Gillen et al., 2015; Merlo and De Paula, 2017). This approach offers a useful way to estimate the correlates of vote returns in multiparty elections. Some earlier approaches to this problem used discrete choice models with individual-level survey data, but such data is subject to biases in voters' self-reported behavior (see, e.g., Selb and Munzert (2013) and the literature cited therein). Other studies used 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, the structural approach combines the advantages of a micro-founded discrete choice model with those of 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 simulations.

Finally, our paper relates to previous studies of NOTA-like votes. Most observational studies in this literature (reviewed in section 2 below) analyze "residual" (blank or invalid) votes, but this makes it difficult to separate intentional behavior from voting mistakes. A recent paper by Ambrus et al. (2017) studies the impact of NOTA in small experimental elections in a "Swing Voter's Curse" setting. There, NOTA allows uninformed voters to participate without adversely affecting the election result, particularly under compulsory voting.[4] In a follow-up paper, Ambrus et al. (2018) conduct a lab-in-the-field experiment

[3]Related studies by Pons and Tricaud (2018) and Kapoor and Magesan (2018) focus on the *implications* of option-specific consumption utility by showing that exogenous increases in the number of candidates can raise voter turnout.

[4]In Indian elections abstention is unlikely to be driven by the swing voter's curse. First, voters have many

before the 2016 US and Austrian presidential elections and show that voter behavior supports the interpretation of NOTA as a protest vote. We complement this literature by studying explicit NOTA votes in real-world elections, and by estimating a structural model that can be used to answer questions about the counterfactual behavior of these protest voters, and, more generally, about the impacts of having a NOTA option on the ballot.

Our study is subject to the usual tradeoffs between internal and external validity. While several features of the Indian context (including the rules of the NOTA policy and the relatively small share of voters choosing NOTA) has several advantages for identification, these features also mean that one must be careful in generalizing the findings to other contexts. We return to this issue at the end of the paper after presenting our findings.

In the rest of the paper, section 2 presents background information on NOTA and Indian elections. Section 3 describes the construction of our dataset. Section 4 documents the pattern of NOTA votes and presents a regression analysis of the effect of NOTA on turnout. Section 5 presents the structural model and section 6 explains the estimation. Section 7 contains the estimation results and section 8 the counterfactual analysis. Section 9 discusses additional implications of our findings, and section 10 concludes. An online Appendix contains further details.

2 Background

2.1 The 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:

> "Democracy is all about choice. This choice can be better expressed by giving

fringe candidates they can vote for without risking tipping the election for the wrong candidate. Second, since elections are large, the probability of tipping the election in any direction is small. On the other hand, in the lab the protest motivation for casting a NOTA vote is likely to be limited since there are no real candidates that one can protest against.

[5] With 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.

> 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, p44).

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.

2.2 Interpreting the NOTA vote

Although the popular discourse suggests that protest voting matters in elections, direct empirical evidence on this phenomenon is scarce. Part of the reason seems to be that protest voting is difficult to define (see Kselman and Niou (2011) and Myatt (2017) for discussions). It seems particularly difficult to establish to what extent a vote is *for* a candidate or *against* other candidates. NOTA allows us to sidestep these conceptual issues and focus on voters who have a revealed preference for a pure protest option. These are the voters we refer to as protest voters in this paper, and we will use the terms protest voter and NOTA voter interchangeably below. To study protest voting in "normal" elections that do not have an explicit protest option, we ask how NOTA voters would behave in the absence of NOTA, all else equal.

This interpretation of NOTA votes rests on the assumption that votes cast on this option were a deliberate choice by voters who understood what they were doing. This assumption seems weaker than those typically made in interpreting voters' choices of particular candidates. Here, the act of voting for NOTA involved pushing a button that literally said "None Of The Above." The NOTA policy received wide news coverage in both national and local media. It was also featured explicitly in the regular voter education campaigns undertaken by the Election Commission. As a result we expect that most voters would be well-informed about the policy. Anecdotal evidence on how voters viewed NOTA also supports the idea that it was understood to be a protest option. One voter declared: "I am proud to take part in the democratic process. But the selection of candidates by the political parties has disappointed me. That's why I will be opting for NOTA" Another stated: "I am fed up of corruption. I will use NOTA as a weapon to register my protest against the political system in the country."[6]

Nevertheless, there could be potential concerns regarding voters' information that would affect the interpretation of our results. First, 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 that the majority of NOTA voters would have abstained without NOTA 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 turnout. Still, some of the substitution *from candidates* to NOTA could be due to voter confusion. To alleviate this concern, we checked whether voters disproportionately substituted to NOTA from candidates listed adjacent to it on the voting machine and did not find this to be the case (see the Appendix).[7]

Second, another way that voters may 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. While in this case NOTA votes may still reflect protest, they may also involve subtle strategic considerations about what will happen if NOTA wins. We find this 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 chose NOTA due to some confusion about the rules, 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 in 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 the 2014 general election the average 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 ($p = 0.027$), indicating that NOTA votes cannot be attributed simply to voter confusion.

A third potential concern arises if voters are *well*-informed. It is possible that the NOTA policy focused popular attention on elections, resulted in additional coverage and discussions in the media, etc. This additional attention could itself lead to an increase in turnout, confounding any effect that NOTA might have. While directly testing for this is difficult,

[6] http://www.deccanchronicle.com/140326/nation-politics/article/ignore-none-above-option-your-own-risk

[7] Note also that, according to our findings below, voters are more likely to substitute from candidates to NOTA in more literate constituencies, where voting errors may be less likely.

our findings below seem hard to reconcile with this interpretation. We find that new voters mostly voted for NOTA. If turnout was driven by an increased attention to elections in general, it seems more likely that the new voters would have shown up to vote for one of the candidates rather than NOTA.

Finally, it is theoretically possible that someone who votes for NOTA may be indifferent between all options on the ballot and would simply show up to vote because he derives utility from participating. Such a voter, however, should still turn out in the absence of NOTA (and choose one of the candidates), i.e., NOTA should not affect his turnout decision. Similarly, such a voter would have no reason to change his vote from a candidate to NOTA once NOTA is introduced. This interpretation would therefore imply that NOTA should have no impact on voter behavior, and this is not what we find.

2.3 NOTA-like options in other countries

NOTA is an explicit option on the voting machine, and this makes it fundamentally different from simply casting an invalid vote as can be done in many countries. In the case of the latter, it is typically impossible to know whether such votes occur intentionally or by mistake, hence 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.

In some countries, while there is no NOTA option on the ballot, blank votes are reported separately from invalid votes and are generally 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 when paper ballots were used. 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 through an increase in invalid rather than blank votes.

Another feature that makes the Indian NOTA useful for the analysis of voters' motivations is the lack of electoral impact of the NOTA vote. Recall that NOTA vote can never "win." In addition, 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. Both of these systems could give voters an incentive to choose the blank vote strategically in order to affect the mix of candidates elected for office in the current election. Such incentives are not present in the Indian system.[8]

2.4 Assembly elections in India

Various features of Indian state elections make this an ideal setting to study the NOTA policy.

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] The average constituency has approximately 180 thousand eligible voters and 11 candidates running. Many of these candidates get very few votes: on average the candidates finishing first and second receive 45% and 35% of the votes, respectively.

Table 1 shows the timing of state assembly elections in our study period. Elections

[8] Countries with a NOTA option similar to the Indian system include France and Brazil. In the US, the state of Nevada has a NOTA option in statewide races. On the latter, see Brown (2011) and Damore et al. (2012) who present correlations of NOTA votes with various election characteristics but do not discuss identification.

[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).

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 exclude these states from the analysis below.

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]

Voting in India takes place using electronic voting machines (EVMs). EVMs were introduced gradually beginning in 1999, and since 2004 all general and state elections have been conducted using these machines. Each candidate running in an election has a separate button assigned to them 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 and the machine beeps to confirm that the vote was successful.[11] 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

[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. The voting age is 18.

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

reservation status of the constituencies were 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.

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%.) This large turnout, and in particular large turnout among the poor, is the subject of an extensive literature in political science and anthropology. One prominent theme in this research is Indian voters' view of elections as being much more than a means to elect a government. In her extensive ethnographic study (conducted before the introduction of NOTA), Banerjee (2014) documents that

> "... the very act of voting is seen by them as meaningful, as an end in itself, which expresses the virtues of citizenship, accountability and civility that they wish to see in ordinary life, but rarely can. For these voters, Election Day creates a time out of time, a carnival space, where the everyday reality of inequality and injustice is suspended, and popular sovereignty asserted for a day." (p3)

Villagers describe everyday inequality of wealth and caste being set aside on election day as rich and poor alike must stand in line to cast their single vote ("I enjoy the identity of being equal with everyone at least for one day." p159). Participating in an election is a way for individuals to affirm their identity as belonging to the Indian state, and feel that the state is paying attention ("I vote to establish my identity and let the government know that there is someone with so-and-so name living in so-and-so village." p163). Voters also describe how they use elections to communicate their views on the different parties. One voter chose the BJP in order "to teach the Congress [party] a much needed lesson." (p153). Another voter voted against the BJP, his usual choice, to express his dissatisfaction with them and "to ensure that the party did not get the impression that most of the voters had been in their favor." (p154)

There appears to be clear anecdotal evidence to support the idea that various factors other than pivotality have an important role in the motivation of Indian voters.

3 Data

3.1 Samples used for analysis

Our dataset contains constituencies in 25 Indian states conducting assembly elections between 2006 and 2014. Each of these states held two elections over this period (see Table

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, Rajasthan,** NCT of Delhi	
2014	4	Andhra Pradesh*, Arunachal Pradesh*, Odisha*, Sikkim*	*National elections*
	10	Haryana, Maharashtra	
	12	Jammu & Kashmir, Jharkhand	

Notes: Elections in bold are included in both the panel and the repeated cross-section. Other elections listed are included only in the repeated cross-section, except as follows: * These state elections held simultaneously with national elections are excluded from the dataset. ** Bihar is excluded from the dataset because it is the only state holding elections in 2005 and 2010.

1). In the structural analysis, which requires panel data, we use a subset of this dataset as a panel of constituencies. In the reduced form exercise, we use the extended dataset as a repeated cross-section of constituencies.

Panel. The structural analysis below uses a panel of 723 constituencies in 5 states that conducted assembly elections in both 2008 and 2013 under the new electoral boundaries: Karnataka, Chhattisgarh, Rajasthan, Madhya Pradesh, and Mizoram. Together these five states represent over 140 million eligible voters, or about one fifth of eligible voters in India. One of these states, Karnataka (with 223 constituencies), held elections in both 2008 and 2013 without a NOTA option, while the remaining 4 states (520 constituencies) had a NOTA option in 2013 but not in 2008.[12]

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 2009 and 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 the 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 potentially confound the impact of NOTA in these states. We therefore decided to exclude these states from the panel.

Repeated cross-section. To use more observations and exploit policy variation across all available states, we use the extended dataset of the 25 states with elections between 2006 and 2014 as a repeated cross section of constituencies. This sample contains a total of 6685 constituency-year observations, and 1176 of these observations have the NOTA option. Like the panel, this data excludes the states that held assembly elections simultaneously with

[12] One other state, the National Capital Territory of Delhi, also held elections in 2008 and 2013 after delimitation. It is not included in the panel because its demographic data is not sufficiently disaggregated to permit matching with the constituency boundaries (see below).

[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.

national elections (Andhra Pradesh, Arunachal Pradesh, Odisha and Sikkim) since turnout considerations in these states are likely to be very different.[14]

Although less readily available than data for later years, in principle it may be possible to obtain electoral and demographic data for earlier elections in the repeated cross-section. We did not pursue this because of changes in voting technology in India between 1999-2004. Electronic voting machines were gradually introduced over this period, and we are not aware of any records on which constituencies had them and which ones did not. The introduction of voting machines could confound any patterns that we would observe in the data for this period (Fujiwara, 2015).

We next describe the information available in the panel and the extended dataset. The Appendix provides further details, including comparisons of states affected and unaffected by NOTA.

3.2 Election and candidate data

The basis for the electoral data is 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 this 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).[15] Voters' response to such candidate characteristics has been documented by a variety of studies in economics, psychology and political science.[16] We also use the Election Commission data to create two additional candidate characteristics: whether the candidate's party fielded a candidate in the previous election and whether that candidate won. Because for 2008 the "previous election" occurred before redistricting, we cannot create this measure at the constituency level. Instead, we calculate for each administrative district the fraction of constituencies within the district where the given party ran, and the fraction where it won, in the previous election and use these two variables as additional candidate characteristics.

We add to the above data further candidate and election characteristics from a number of sources (see the Appendix for details). First, we include data on the time allocated to each party on public TV and radio stations in each state. These time allowances are allocated to national and state-recognized parties separately in every state election, and they are based

[14] To maximize the number of observations, we include in the repeated cross-section the states that held elections in 2014 but not simultaneously with the national election. As we show below, excluding these states makes little difference for the results.

[15] In the analysis below we use a single characteristic, "minority" to describe SC and ST candidates. We checked that using SC and ST separately does not change our counterfactual results on the impact of NOTA.

[16] See, e.g., Chattopadhyay and Duflo (2004) on gender and Sigelman and Sigelman (1982) on age.

on the party's performance in the previous national and state election in that state. This variable serves as a pre-determined (in the current election) proxy for campaign advertising. Second, we add data on rainfall on election day, which a sizeable literature indicates could affect turnout. The rainfall variable is created based on gridded daily rainfall data obtained from the India Meteorological Department in 0.25×0.25 degree cells (which is smaller than the typical constituency). Different constituencies within the same state typically go to the polls in groups over a period of 2-3 days. We match the relevant daily rainfall grid to constituency boundaries and take the area-weighted average of the cells covering each constituency.

Finally, for some of the exercises below we include information on candidates' education level, criminal history, and assets. This data comes from affidavits that the Election Commission requires all candidates to file. The information has been collected and made publicly available at www.myneta.info and we use it to study the robustness of our findings. We do not use this data in the main specification for two reasons. First, data coverage for these variables is only about 92% of our sample. This creates a difficulty for the structural exercise where observing the full choice set of voters is crucial and candidates with missing characteristics cannot simply be dropped. Second, the fact that this information is based on candidates' self-reports raises potential concerns regarding its veracity. Moreover, veracity could be correlated with vote shares (e.g., if the affidavits of front running candidates receive more scrutiny than others').

Panel A of Table 2 shows summary statistics of the candidate characteristics.[17] The overwhelming majority of candidates are male: the average constituency has less than one female candidate. Average age is 43, and 38% of candidates are from the "minority" group (SC or ST). The latter figure includes the reserved constituencies: in non-reserved constituencies the share of minority candidates is 12%. Fifty-nine percent of the candidates completed high school, and 14% have a criminal conviction. The two largest parties, INC and BJP, field approximately 14.5% of the candidates.

As shown in Table 2 (Panel B) and Table 3, the average constituency has approximately 180 thousand eligible voters and turnout is around 71%. Approximately 30% of the constituencies are reserved for SC or ST candidates. Summary statistics of the electoral data in the 5-state panel are broadly similar to those in the extended dataset.

[17] As described in detail in Section 6 below, in each constituency, independent candidates and small party candidates are each aggregated into one "Independent" and one "Small party" candidate, respectively.

Table 2: Summary statistics of the panel dataset

Variable	Obs	Mean	Std. Dev.	10%	90%
A. Candidate characteristics					
Gender	9831	0.073	0.245	0.000	0.125
Age / 100	9831	0.433	0.145	0.300	0.600
Minority	9831	0.381	0.470	0.000	1.000
Ran in previous election	9831	0.619	0.443	0.000	1.000
Won in previous election	9831	0.142	0.253	0.000	0.571
Education	8989	0.588	0.457	0.000	1.000
Criminal conviction	8989	0.144	0.333	0.000	1.000
Assets (log/10)	9384	1.368	0.418	1.023	1.730
NOTA (0/1)	9831	0.053			
Broadcast allowance (100 minutes)	9831	0.673	0.773	0.000	1.800
INC (0/1)	9831	0.146			
BJP (0/1)	9831	0.144			
Independent (0/1)	9831	0.138			
Small party (0/1)	9831	0.120			
B. Constituency characteristics					
Eligible voters (1000)	1446	180.797	36.173	148.363	216.153
Turnout	1446	0.720	0.081	0.611	0.816
NOTA votes / total votes	520	0.021	0.013	0.007	0.037
NOTA votes / eligible voters	520	0.016	0.009	0.005	0.028
Reserved constituencies (0/1)	1446	0.335			
Rainfall (cm/day)	1446	0.066	0.387	0.000	0.011
Minority population (%)	1446	0.353	0.182	0.186	0.617
Literate population (%)	1446	0.584	0.092	0.474	0.693
Rural workers (%)	1446	0.662	0.174	0.442	0.845

Notes: The panel dataset contains the 2008 and 2013 state assembly elections in the states of Karnataka, Mizoram, Rajasthan, Madhya Pradesh, and Chhattisgarh. Gender is 0 if male and 1 otherwise (including one transgender candidate). Minority refers to SC or ST. Ran (Won) in previous election is the share of the constituencies within the district where the party's candidates ran (won) in the previous election. Education is 1 if completed high school. Criminal conviction is 1 if has criminal conviction. Assets is log(1+A)/10 where A is reported assets in Rp. Turnout is total votes divided by the number of eligible voters. Rural workers is the share of the rural workforce. In each constituency, independent candidates and small party candidates are each aggregated into one "Independent" and one "Small party" candidate, respectively, resulting in fractional values for these candidates (see Section 6.3 for details). Variables that only take the values 0 or 1 are marked (0/1). Data sources are described in the text.

Table 3: Summary statistics of constituencies in the extended dataset (repeated cross-section)

Variable	Obs	Mean	Std. Dev.	10%	90%
Eligible voters (1000)	6685	180.754	88.232	41.203	292.898
Turnout	6685	0.707	0.129	0.533	0.866
NOTA votes / total votes	1176	0.015	0.012	0.004	0.030
NOTA votes / eligible voters	1176	0.010	0.009	0.003	0.022
Reserved constituency (0/1)	6685	0.276			
Labor force participation	6685	0.572	0.065	0.503	0.665
Unemployment rate	6685	0.032	0.033	0.011	0.050
Household earnings (real Rp/week)	6685	1553.309	540.820	936.419	2135.973
Fraction illiterate	6685	0.301	0.118	0.116	0.479
Fraction primary school or less	6685	0.227	0.071	0.150	0.325
Sex ratio (females / 1000 males)	6685	986.086	63.851	928.189	1084.414
Fraction urban	6685	0.313	0.137	0.199	0.447
State NDP growth rate	6685	5.784	3.720	1.597	11.284
Election on weekend (0/1)	6685	0.258			
Rainfall (cm/day)	6684	0.083	0.216	0.000	0.194
Polling station density	6676	0.001	0.000	0.001	0.001
Redistricting (max overlap)	6084	0.898	0.187	0.551	1.000
Redistricting (fractionalization)	6084	0.864	0.239	0.421	1.000

Notes: 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. Polling station density is the number of voting stations per eligible voter. Redistricting (max overlap) is the largest area of a current constituency that was part of a single constituency before delimitation; (fractionalization) is a measure of territorial fractionalization as a result of redistricting. The construction of these measures is described in the Appendix.

3.3 Voter demographics

For the panel dataset used in 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. Because Census administrative areas do not correspond to the constituencies, this matching must be done using GIS boundary files. The Appendix contains the details of our matching procedure.

For the repeated cross-section exercise, demographic data comes from 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.[18] We obtained the individual level data and use it to create characteristics of the voting age population at the state-year or the district-year level for the reduced-form analysis. We complement this with data on the growth rate of per capita state domestic product from the Reserve Bank of India. Summary statistics of the demographic data appear in Table 2 (Panel B) and Table 3.

4 Patterns in the data

4.1 NOTA votes

The first noteworthy feature of the data is that voters actually voted for NOTA, despite the fact that voting for NOTA could not affect the results of the election. In the 9 states out of 25 in our data where NOTA was available, NOTA was chosen by a positive number of voters in every constituency.[19] The distribution of the NOTA vote share is shown on Figure 1. The average vote share is 1.5%, with a range of 0.1-11%. As a fraction of all eligible voters (including abstainers) 1% voted for this option.[20]

While 1-2% is a small fraction of votes, it should be noted that given the size of the Indian electorate, small fractions translate into large numbers. In our data, a total of 2.51

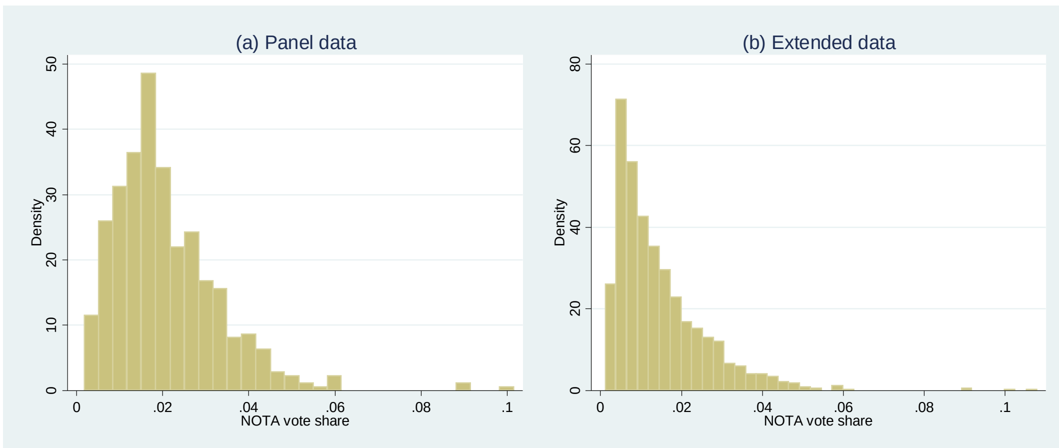
[18] Compared to the Census, the NSS data has less coverage in the cross section but is available at more frequent intervals, which makes it better suited for the aggregate reduced form analysis.

[19] While the fact that people voted for 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 20% of Mexican voters exiting the election booth say that they have little or no confidence that "the vote you cast for president will be respected and counted for the final result."

[20] In the panel data, in the 4 states affected by NOTA the average vote share of NOTA among total votes cast (eligible voters) was 2.1% (1.6%).

million voters chose NOTA. Another useful benchmark is the vote share of actual candidates. In the average constituency, NOTA received more votes than 7 of the candidates running for election. Furthermore, 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). In these elections, votes cast on NOTA would have been enough to flip the outcome.

Figure 1: Distribution of NOTA vote shares across constituencies



Notes: NOTA vote share is measured as a fraction of total votes cast in the constituencies where NOTA was available (elections held after September 2013). N = 520 for the panel and 1176 for the extended dataset.

Figure 1 reveals some heterogeneity in NOTA votes across constituencies. To explore this further, 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 minorities, a lower share of rural workers, and fewer candidates on the ballot. Our structural analysis below will relate heterogeneity in voter choices to voter demographics and model how different groups of voters choose between the different options on the ballot.

4.2 Preliminary evidence on the effect of NOTA on turnout

In this section we present a preliminary analysis of the aggregate behavior of NOTA voters. If NOTA voters normally abstain, then all else equal we expect the introduction of NOTA to be associated with an increase in turnout. Conversely, if NOTA voters normally vote for one of the candidates, then the introduction of NOTA should not affect turnout, but simply change the behavior of existing voters.

We use the fact that elections to the state assemblies are held at different times in different states (see Table 1) to conduct a reduced form analysis. We measure the effect of NOTA using the change in turnout in states exposed to NOTA relative to states that were not exposed. This complements our analysis below where we measure the effect of NOTA using counterfactual simulations on an estimated model.

We estimate the following specification:

$$Y_{cst} = \alpha_0 + \alpha_1 NOTA_{st} + \boldsymbol{\alpha}_2 \mathbf{X}_{cst} + \gamma_{cs} + \eta_t + \varepsilon_{cst}, \tag{1}$$

where Y_{cst} is turnout in constituency c of state s in year t, $NOTA_{st}$ equals 1 if the election features a NOTA option and 0 otherwise, $\mathbf{X}_{cst}$ are control variables, the γ_{cs} are constituency fixed effects (in the panel) or state fixed effects (in the repeated cross section), and the η_t are year fixed effects. The parameter of interest, α_1 is identified by comparing the *change* in turnout in the constituencies that held elections in both 2008-09 and 2013-14 without NOTA to the change in turnout in the constituencies that had NOTA in 2013-14 (but not in 2008-09).[21] Because the NOTA policy varies at the state level, inference needs to account for the clustering of constituencies by state.

Table 4 shows the results from estimating equation (1). In column (1), we use the panel dataset and control for the log number of eligible voters in a constituency and its square,[22] state labor force participation, weekly household earnings, and education, as well as constituency and year fixed effects. The coefficient indicates a turnout effect of +1.2 percentage points, but with only 5 states, this estimate is imprecise with a p-value of 0.171.[23]

In column (2) we repeat the regression using the extended dataset as a repeated cross section. The coefficient estimate on NOTA indicates a positive turnout effect of 3 percentage points that is statistically significant (and not statistically different from the 1.2 in the panel regression).[24] 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. These findings are suggestive that the presence of the NOTA option on the election ballot increased turnout in the average constituency.

[21] Recall that the data contains 25 states (6685 constituencies), 9 (1176) of which were affected by NOTA: 5 states (630 constituencies) in 2013 and 4 states (546 constituencies) in 2014.

[22] Omitting the squared term does not affect our findings. In most regressions below both of these variables are highly significant.

[23] With only 5 clusters, relying on the asymptotic standard error would be misleading. 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) designed for a very small number of clusters.

[24] We obtain similar inference using the bootstrap procedure as from the asymptotic standard errors. In the Appendix, we also report p-values obtained by randomization inference.

Moreover, in each case the 95% confidence interval around the point estimates includes the fraction of eligible voters who voted for NOTA in the data (1.6% in the panel or 1% in the repeated cross-section). This is consistent with the view that most NOTA voters would abstain in a "normal" election when this option was not present on the ballot.

Table 4: The impact of NOTA on turnout, regression estimates

	(1)	(2)	(3)
NOTA	0.012	0.029**	0.030*
asymptotic s.e.	(0.005)	(0.013)	(0.016)
bootstrap p-value	[0.171]	[0.026]	[0.066]
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	1446	6685	6685
States	5	25	25

Notes: Estimates of the effect of the NOTA policy on turnout from Eqn. (1). Column (1) is for the panel, columns (2-3) are for the extended dataset (repeated cross-section). Regressions control for 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. Asymptotic standard errors clustered by state in parentheses. To correct for the small number of clusters, 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.

While the number of available states and time periods impose fundamental limitations on this exercise, we attempt to rule out several natural confounds in the Appendix. First, we consider national election years. While elections held simultaneously with national elections are already excluded from the analysis, national elections held in the same year as a state election could also confound our estimates so we present results where such elections are also excluded. Second, we control for electoral redistricting. This is relevant because due to the timing of redistricting, none of the states affected by NOTA were redistricted while most of the states in the control group were. To account for this, we use GIS boundaries before and after redistricting to create measures of a constituency's exposure to redistricting and use these as additional controls in the regressions. Third, we searched news reports for state-specific events (such as terrorist attacks) that may act as a confound and drop the affected states from the analysis. Fourth, we include three additional controls for voting costs. We obtained information on the exact date of voting and control for elections held on weekend. We also obtained daily rainfall data which we matched to constituency boundaries and control for rainfall on election day. Finally, we obtained data on the number of voting stations in a constituency and control for the number of stations per voters as a measure of "congestion" at the voting booth. The results presented in the Appendix indicate that our estimates in Table 4 are generally robust to these alternative samples and controls.[25]

Our findings provide at least some evidence that most protest voters normally abstain. At the same time, without a model our ability to infer NOTA voters' counterfactual behavior using this aggregate analysis is fundamentally limited and makes it difficult to study heterogeneity in protest voters' behavior. To do this, we turn to a structural analysis that makes inference regarding individual voters' voting patterns possible.

5 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 BLP methodology for consumer demand estimation proposed by Berry, Levinsohn, and Pakes (1995). Although we are not the first to use this approach in the voting context, it is worth spelling out two factors that make this method particularly useful. 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 BLP's 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 discrete choice between voting for a candidate, voting for NOTA, or abstaining.[26] 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.[27] Finally, abstention (the "outside good" chosen by somebody

[25] The Appendix also presents regressions that check whether some of the turnout effects could be through changes in voter registration. We do not find this to be the case.

[26] 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.

[27] By contrast, studies of consumer choice have to rely on proxying the true market in which a set of

who does not choose any of the alternatives on the ballot) is well defined, and administrative data on its prevalence is readily available.[28]

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 administrative individual-level 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 surveys are susceptible to well-known biases (see, e.g., Selb and Munzert (2013) and the literature cited therein), being able to use the available administrative data for estimating a micro-founded model is 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 behavior 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.

5.1 Specification: demand

Consider a constituency $c \in \{1, ..., C\}$ where voters can choose to vote for the candidates of parties $j = 1, ..., J$.[29] 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, caste, education, criminal convictions, assets, and the party's past electoral performance in the constituency (whether it fielded a candidate and whether he won). Unobserved characteristics include, for example, the candidate's physical appearance, or his expected probability of winning the election. Assume

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.

[28] 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, or no transportation?

[29] To simplify the notation, we do not index J with c, i.e., the variation in the set of parties across constituencies. In the data, voters' choice sets vary across constituencies and this is exploited in the estimation.

that the utility that voter i derives from voting for candidate $j \in \{1, ..., J\}$ can be specified as

$$U_{ijc} = \boldsymbol{\beta}_i \mathbf{x}_{jc} + \xi_{jc} + \varepsilon_{ijc}, \tag{2}$$

where $\mathbf{x}_{jc} = (x_{jc}^1, ..., x_{jc}^K)'$ is a vector of the observed characteristics of party j's candidate, including a full set of party indicators (i.e., the $\boldsymbol{\beta}_i$ coefficients include a set of party fixed effects). The term ξ_{jc} captures voters' valuation of unobserved candidate characteristics, and ε_{ijc} 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). As described below, the model allows for considerable flexibility in the distribution of the unobserved candidate characteristics ξ_{jc}. No explicit distribution for ξ_{jc} is required, and we also allow for correlation of these characteristics within a constituency. For example, if the election in constituency c is expected to be close, voters may be more motivated to turn out, and this could result in an increase in ξ_{jc} for some or all candidates j. Candidate platforms, which we have no data on, are implicit in the specification (2). Party platforms common across all candidates of a party are captured by the party fixed effects. Constituency-specific deviations from this platform could be captured by candidate characteristics, the heterogeneity in voter valuations $\boldsymbol{\beta}_i$ of the characteristics (including the party indicators), and ultimately the unobserved component ξ_{jc}.

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 voter characteristics:

$$\boldsymbol{\beta}_i' = \boldsymbol{\beta} + \boldsymbol{\Pi} \mathbf{d}_i + \boldsymbol{\Sigma} \mathbf{v}_i, \tag{3}$$

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 difficult to obtain reliable data on voters at more disaggregated levels than the constituency. The difficulty of directly matching voter demographics to candidate vote returns also makes more detailed treatments of preference heterogeneity (e.g., as in Petrin (2002)) impractical in this context.

In elections where the NOTA option is available, we simply include it in the list of candidates $\{1, ..., J\}$. NOTA's only characteristic is a NOTA indicator, which is equivalent to setting all other characteristics in $\mathbf{x}_{NOTA,c}$ equal to 0. We allow for individual heterogeneity in the utility of NOTA (from both observed and unobserved voter characteristics) through the corresponding coefficients in $\boldsymbol{\beta}_i$.

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 of voting. In consumer demand applications constructing the outside choice typically involves two sets of assumptions: assumptions about what consumers do when they don't 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. Let $j = 0$ indicate the abstention option and

$$U_{i0c} = \boldsymbol{\pi}_0 \mathbf{d}_i + \sigma_0 v_i^0 + \varepsilon_{i0c}. \tag{4}$$

This 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 (2) 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 (2) is equivalent to including them in the specification of the utility of abstention in (4). Thus, we are also allowing for further heterogeneity in voting costs as captured by these variables.

Denote the parameters of the model by $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\theta}_2)$, where $\boldsymbol{\theta}_2 = (\boldsymbol{\Pi}, \boldsymbol{\Sigma})$. Substituting (3) into (2), we can write

$$U_{ijc} = \delta_{jc} + \mu_{ijc} + \varepsilon_{ijc},$$

where $\delta_{jc} \equiv \boldsymbol{\beta} \mathbf{x}_{jc} + \xi_{jc}$ and $\mu_{ijc} \equiv (\boldsymbol{\Pi} \mathbf{d}_i + \boldsymbol{\Sigma} \mathbf{v}_i) \mathbf{x}_{jc}$. We assume that voter i chooses option j (vote for one of the candidates, vote for NOTA, or abstention) if $U_{ijc} > U_{ilc}$ for $l = 0, 1, ..., J$. Thus, voters choose between their options based on the observed and unobserved candidate characteristics, the observed and unobserved benefits of abstention, and their idiosyncratic shocks. This implicitly defines the set of demographics and unobserved variables for which voter i will choose option j:

$$A_{jc}(\mathbf{x}, \boldsymbol{\delta}_c(\boldsymbol{\beta}), \boldsymbol{\theta}_2) = \{(\mathbf{d}_i, \mathbf{v}_i, \boldsymbol{\varepsilon}_{ic}) | U_{ijc} > U_{ilc} \; for \; l = 0, 1, ..., J\},$$

where $\mathbf{x}$ are all the candidate characteristics, $\boldsymbol{\delta}_c = (\delta_{1c}, ..., \delta_{Jc})$, and $\boldsymbol{\varepsilon}_{ic} = (\varepsilon_{i1c}, ..., \varepsilon_{iJc})$.[30]

[30] This formulation assumes that the voter's choice regarding whether to vote is made simultaneously with the choice of who to vote for conditional on voting. Alternatively, one could imagine a two-stage model, where the voter first decides whether to vote, and then if he does vote, decides who to vote for. If the

Given the distribution of $(\mathbf{d}_i, \mathbf{v}_i, \boldsymbol{\varepsilon}_{ic})$, we can integrate over A_{jc} to obtain the vote shares $s_{jc}(\mathbf{x}, \boldsymbol{\delta}_c(\boldsymbol{\beta}), \boldsymbol{\theta}_2)$ predicted by the model. Under the assumed Type-I extreme value distribution for ε_{ijc}, these are given by

$$s_{jc}(\mathbf{x}, \boldsymbol{\delta}_c(\boldsymbol{\beta}), \boldsymbol{\theta}_2) = \int \frac{\exp\left[\delta_{jc} + \mu_{ijc} - \mu_{i0c}\right]}{1 + \sum_{q \geq 1} \exp\left[\delta_{qc} + \mu_{iqc} - \mu_{i0c}\right]} dF\left(\mathbf{d}_i, \mathbf{v}_i\right), \tag{5}$$

where $\mu_{i0c} \equiv \boldsymbol{\pi}_0 \mathbf{d}_i + \sigma_0 v_i^0$ and $F\left(\mathbf{d}_i, \mathbf{v}_i\right)$ denotes the distribution of the voter characteristics.[31] These predicted vote shares are a function of the data ($\mathbf{x}$), the parameters ($\boldsymbol{\theta}$), and the unobserved candidate characteristics ξ_{jc}. In the estimation, these predicted vote shares will be set equal to the vote shares observed in the data.

While the above model assumes sincere rather than strategic voting, it does control for some factors that would be important under strategic voting (i.e., if voters were driven in part by the expectation that their vote could be pivotal). First, under strategic voting candidates' unobserved characteristics are likely to be correlated because the expected electoral success of one candidate will be correlated with the expected electoral success of other candidates. In the above model, this would show up as correlation in the ξ_{jc} terms, and our estimation method corrects for this by using a cluster-robust procedure that allows for arbitrary correlation of ξ_{jc} within a constituency. Second, strategic considerations are likely to be stronger in smaller constituencies. By weighting observations with the number of eligible voters in a constituency we are giving lower weights to these constituencies.[32]

In principle, one could extend the model to explicitly include strategic considerations, but estimating such a model would be difficult. Kawai and Watanabe (2013) estimate a model of strategic voting that does not allow for abstention and note that including the abstention decision would make their procedure computationally challenging (p649-650). Because in our case understanding substitution between abstention and NOTA is crucial, we explicitly model abstention, but we do not model voters' strategic motivations. Two additional considerations offer some justification for this approach. First, as discussed in section 2, the descriptive literature argues that considerations other than pivotality play a major role in motivating Indian voters. Second, our focus is ultimately on the counterfactual results from adding or removing the NOTA option. Explicitly modeling pivotality considerations would only affect the results of this comparative static exercise if adding or removing NOTA changed the

voter knows all the available options on the ballot when deciding whether to turn out to vote, these two formulations are equivalent. Following most papers on voting, we assume that this is the case here.

[31] Note that s_{jc} is measured as a share of all eligible voters.

[32] Voters may also have subjective assessments that their vote might be pivotal. If these are independent across voters, they would be part of the ε_{ijc} term in (2) and simply integrated out from the market shares. If they are correlated across voters, they would be in ξ_{jc}, as above.

expected closeness of elections. To obtain some suggestive evidence on this, we investigate the impact of NOTA on different measures of closeness in the reduced form. These results, reported in the Appendix, suggest that whether or not a NOTA option was present did not affect the closeness of an election.[33]

5.2 Specification: supply

In this section we describe a simple model of party behavior and the supply of candidates in order to justify the instrumental variables used in the estimation below. 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 model tractable, we adopt a simple simultaneous-moves specification of the supply of candidates.

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 political parties that care about winning as well as the policy implemented by the winner. In constituency c, let party j's payoff be given by $V_{jc}(\mathbf{x}_c, \mathbf{s}_c)$, where $\mathbf{x}_c = (\mathbf{x}_{1c}, ..., \mathbf{x}_{Jc})$ are the characteristics of all candidates running in the election and $\mathbf{s}_c = (s_{1c}, ..., s_{Jc})$ are the vote shares that determine the winner. Vote shares are determined by candidates' observed characteristics as well as the voter valuations ξ_{jc}, as in equation (5). Thus, $s_{jc} = s_{jc}(\mathbf{x}_c, \boldsymbol{\xi}_c)$ where $\boldsymbol{\xi}_c = (\xi_{1c}, ..., \xi_{Jc})$ (to simplify the exposition, we suppress the parameters, including the party fixed effects in this section).[34]

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 "prices" q_{jc}^k of increasing a candidate's characteristic k in constituency c and must allocate its resources subject to a budget constraint $m = \sum_k q_{jc}^k x_{jc}^k \equiv \mathbf{q}_{jc}\mathbf{x}_{jc}$. For example, if $x^k = 1$ denotes a general caste candidate, q_{jc}^k may be the extra cost of finding such a candidate and convincing

[33] A more general question is whether strategic considerations would actually result in different voter behavior. Kawai and Watanabe (2013) show that, in the Japanese elections they consider, even though most voters are strategic, very few of them (between 1.4 and 4.2 percent) behave differently than they would if they were sincere.

[34] For simplicity in this section we also assume that the ξ_{jc} terms represent valuation shocks that the parties have no control over and that only affect their payoff through the vote shares s_{jc}. 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.

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 and their budget m as given.[35]

Suppose that parties choose the characteristics of their candidates simultaneously, after voter valuations $\boldsymbol{\xi}_c$ have been realized. In a Nash equilibrium, the characteristics of party j's candidate will satisfy

$$\mathbf{x}_{jc}^* \in \arg\max_{\mathbf{x}_{jc}} (V_{jc}(\mathbf{x}_c, \mathbf{s}_c(\mathbf{x}_c, \boldsymbol{\xi}_c)) | m = \mathbf{q}_{jc}\mathbf{x}_{jc})$$

or

$$\mathbf{x}_{jc}^* = \mathbf{x}_{jc}^*(\mathbf{x}_c, \boldsymbol{\xi}_c, \mathbf{q}_{jc}). \tag{6}$$

In words, candidate j's characteristics depend on the characteristics of all candidates running, voters' valuation shocks for all candidates, and party j's cost of increasing the various characteristics in the given constituency. This has two implications. First, the dependence of observed characteristics $\mathbf{x}_{jc}$ on voter valuations $\boldsymbol{\xi}_c$ creates an endogeneity problem for the estimation of the utility functions (2). Second, suppose that the prices $\mathbf{q}_{jc}$ for a given party are correlated across constituencies c. 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. Then (2) implies that the characteristics of a given party's candidates will be correlated across constituencies. As explained in section 6.2 below, this opens the possibility of using candidate characteristics in neighboring constituencies as instrumental variables in the estimation.

6 Estimation

6.1 Estimation algorithm

Estimation follows the algorithm proposed by BLP. The idea is to treat the unobserved characteristics ξ_{jc} 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 $\mathbf{x}_c$ and actual vote shares S_{jc}. BLP show that, for given $\boldsymbol{\theta}_2$, it is possible to numerically solve for $\boldsymbol{\delta}_c$ from the equations

[35] Since prices are allowed to differ by party, assuming that m is fixed across parties is without loss of generality.

$s_{jc}(\mathbf{x}, \boldsymbol{\delta}_c, \boldsymbol{\theta}_2) = S_{jc}$, i.e., equating the model-predicted vote shares in (5) to those observed in the data. Using the resulting values of $\delta_{jc}(\boldsymbol{\theta}_2)$, one can express the unobserved candidate characteristics as $\xi_{jc}(\boldsymbol{\theta}) = \delta_{jc}(\boldsymbol{\theta}_2) - \boldsymbol{\beta}\mathbf{x}_{jc}$. Given the data and with δ_{jc} computed, this is a standard econometric error, which depends nonlinearly on the parameters of the model. While $\xi_{jc}(\boldsymbol{\theta})$ may not be independent of $\mathbf{x}_{jc}$, one can find a suitable set of instruments $\mathbf{Z}_{jc}$ and use the moment conditions $E[\xi_{jc}(\boldsymbol{\theta})|\mathbf{Z}_{jc}] = 0$ to estimate the parameters using GMM.

When forming the sample analog of the moment conditions, we weight observations by the number of eligible voters in a constituency. Specifically, for each moment condition $E[\xi_{jc}(\boldsymbol{\theta})|z_{jc}] = 0$, we use $\frac{1}{N}\sum\limits_{j,c} n_c \xi_{jc}(\boldsymbol{\theta}) z_{jc} = 0$, where N is the number of observations (candidates), and n_c is the number of eligible voters in constituency c divided by the average number of eligible voters in all constituencies. Weighting the moment conditions in this way ensures that our estimates are not too sensitive to the politics of a few small constituencies which may be very different from those of larger constituencies.

Letting $\tilde{\boldsymbol{\xi}}(\boldsymbol{\theta}) = [n_1 \boldsymbol{\xi}_1(\boldsymbol{\theta}), ..., n_C \boldsymbol{\xi}_C(\boldsymbol{\theta})]'$ denote the vector of errors, we find

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \ \tilde{\boldsymbol{\xi}}(\boldsymbol{\theta})' \mathbf{Z} \mathbf{W}^{-1} \mathbf{Z}' \tilde{\boldsymbol{\xi}}(\boldsymbol{\theta}), \tag{7}$$

where $\mathbf{Z}$ is the matrix of instruments, and $\mathbf{W}^{-1}$ is the weighting matrix.[36]

To compute the estimate in (7), we use the standard two-step GMM procedure. We first set $\mathbf{W} = \mathbf{Z}'\mathbf{Z}$ and compute an initial estimate of the parameters, $\boldsymbol{\theta}^{Step1}$. We then use this initial estimate to compute a robust weight matrix and use this updated weight matrix to compute the final parameter estimates. For the robust weight matrix we use the cluster-robust formula $\mathbf{W} = \sum\limits_{c}^{C} \mathbf{Z}_c' \tilde{\boldsymbol{\xi}}_c(\boldsymbol{\theta}^{Step1}) \tilde{\boldsymbol{\xi}}_c(\boldsymbol{\theta}^{Step1})' \mathbf{Z}_c$, i.e. we allow for both heteroskedasticity and correlation of the errors ξ_{jc} across candidates within a constituency. As discussed above, this is important if the expected closeness of the race results in correlation between unobserved voter preferences for some of the candidates.

The covariance matrix of the estimated parameters is computed using the standard formulas. To test for the joint significance of the nonlinear parameters, we report Newey-West D tests (Newey and West, 1987). Further computational details are provided in the Appendix.

[36] For given $\boldsymbol{\theta}_2$, the linear coefficients $\boldsymbol{\beta}$ (which include the party fixed effects) can be obtained analytically from (7). Unlike Nevo (2001), we are able to estimate the party fixed effects and the other coefficients in $\boldsymbol{\beta}$ in the same step because we have variation in candidate characteristics for a given party across constituencies.

6.2 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_{jc} and $\mathbf{x}_{jc}$ are uncorrelated. Second, instruments are needed because some of the candidate characteristics could be chosen by the parties as in the supply model above (see equation (6)).[37] In the estimation we treat candidates' gender, age, and minority status as endogenous.

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 instrument by taking the fraction of female candidates of the given party in other constituencies in the state, and another instrument 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).

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 5.2. For a given party, the "prices" $\mathbf{q}_{jc}$ of increasing particular candidate characteristics are likely to be correlated across constituencies c 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 so that, for a given party, the characteristics of its candidates $\mathbf{x}_{jc'}$ in constituencies $c' \neq c$ are relevant instruments for the characteristics of the candidate running in constituency c. Similarly,

[37] The endogeneity problem is likely to be present even if candidate characteristics are assumed to be fixed at the time of the election. Although we control for a large number of characteristics $\mathbf{x}_{jc}$, there will always be characteristics (for example, physical appearance) that we do not have data on, but that could be correlated with both $\mathbf{x}_{jc}$ and voters' valuation ξ_{jc}.

every other party's characteristics $\mathbf{x}_{j'c'}$ in constituencies $c' \neq c$ are correlated with its characteristics $\mathbf{x}_{j'c}$ in constituency c due to the correlation in the prices faced by that party. Thus, under (6) $\mathbf{x}_{j'c'}$ are also relevant instruments for $\mathbf{x}_{jc}$.

In the analysis below, we compare the results using 3 different sets of instruments, summarized in Table 5. The first set includes the average of the gender and age characteristics. As described above, we compute the average characteristic of a party's candidates in all other constituencies in the state. We also take the interaction of these averages with the state dummies to allow for more flexibility in the instruments' impact by state. Similarly, we compute the average characteristics of all candidates in other constituencies, and we do this for both elections in the data (2008 and 2013). We increase this set with further instruments in order to help identify the nonlinear parameters of the model. The second set of instruments augments the first set by adding interactions of the NOTA indicator with the constituency-level average of the demographics used in the estimation (minority population, literacy rate, share of rural workers). These variables should help identify the corresponding preference heterogeneity parameters for NOTA (especially the parameters on the interaction of NOTA with the individual demographics $\mathbf{d}_i$). For the third set, we augment the first set with instruments for the minority candidate characteristic, constructed similarly to those for gender and age. Summary statistics for the instruments appear in the Appendix.

Table 5: Description of the Instrumental Variables

	Number of IVs	First set	Second set	Third set
Average gender of candidates per state	7	x	x	x
Average age of candidates per state	7	x	x	x
Average minority status of candidates per state	6			x
NOTA x demographics	3		x	

Notes: Average gender per state instruments include (i) the average gender of a party's candidates in other constituencies within the state in the given election, interacted with state indicators; and (ii) the average gender of all candidates in other constituencies in the state, for both the 2008 and the 2013 elections separately. Average age and minority instruments constructed similarly. For minority, the interaction with one of the state indicators (Mizoram) is dropped because all constituencies in this state are reserved. NOTA instruments are the interaction of the NOTA indicator with the following average demographics in the constituency: minority population, literate population, rural workers.

To get a sense of the strength of these instruments we report "first stage" F statistics in the Appendix. Because the full model we estimate below is nonlinear, these tests are merely suggestive.[38] The first instrument set appears to be relatively stronger than the second and

[38] Weak identification in nonlinear GMM is an active area of current research and we know of no test

third sets.[39] The first two sets pass the conventional weak IV test ($F > 10$) while the third one does not. For reasons discussed further below the preferred specifications that we use for the counterfactual analysis will not use the third instrument set.

The identifying assumption for estimation, expressed in the moment conditions, is that unobserved voter valuations for a particular candidate in a constituency are conditionally independent of the instruments. One case in which this assumption will hold is if parties do not condition their choice of candidates on the popularity shocks ξ_{jc} in (6). 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. A second case in which the identifying assumption will hold is if, controlling for party-specific means and demographics, constituency-specific voter valuations ξ_{jc} 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 of a party would be captured by the party dummies). See Hausman (1996) and Nevo (2001) for analogous assumptions in the consumer demand literature.

In our setting, a natural possibility is the correlation of preference shocks across constituencies within a particular state or a particular year, as would be the case, e.g., if a party conducted a particularly effective campaign in that state or that year only. Two features of the exercise mitigate this possibility. First, in our sample 65% of the parties field candidates in only one of the states and 60% field candidates in only one year. For these parties, controlling for a party fixed effect captures any correlation between the voter valuations ξ_{jc} across constituencies in a given state or a given election. Second, the included Broadcast allowance control (which varies at the party/state/year level) should capture some of the local advertising effects described above. As we show below, our preferred specifications always pass the standard overidentification test, which increases our confidence in the validity of these instruments in our setting.[40]

directly applicable to our setting. Some recent studies tackle this issue by using Chamberlain-type optimal instruments but this would require specifying the exact form of endogeneity, i.e., parametrizing the supply model of candidate characteristics.

[39] Recall that the second instrument set augments the first with instruments designed to help identify *nonlinear* coefficients on NOTA in the full model. These extra instruments only vary across NOTA observations so naturally they are weak instruments in the *linear* model where there are no coefficients on NOTA that require an instrument.

[40] We also experimented with specifications that included state $\times$ party or year $\times$ party fixed effects but found that including this many fixed effects (46 and 30, respectively) made it impossible to identify the nonlinear model.

6.3 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). In IO applications, the solution to the analogous problem of many small products is typically to assume that these are part of the broadly defined "outside option." We cannot follow this route here since the outside option is a well-defined choice (abstention) and the focus of our study. Instead, we deal with this problem by aggregating small parties and independent candidates. Specifically, 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).[41] We also create an "Independent Party" containing all independent candidates, and aggregate them within constituencies in the same way. After this aggregation, we are left with a total of 22 parties (see the Appendix for summary statistics before and after aggregation).

We include in the analysis the full list of candidate characteristics available in the data: gender, age, caste, party, the fraction of constituencies in the district where the party's candidates ran or won in the previous election, and in some specifications education and criminal convictions. We select the voter demographics to be included based on OLS regressions of the share of NOTA on demographics, and include those that indicated significant heterogeneity in voter preferences for NOTA (see the discussion in section 4 and the Appendix). We include minority population, literacy, and the share of rural workers.[42]

The BLP algorithm requires numerically solving the integral in (5) 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

[41] The Appendix shows that our results do not change if we change the aggregation threshold to a party fielding candidates in either 1/2 or 1/4 of the constituencies in a given state.

[42] We also experimented with including additional demographic variables but found that the resulting model was difficult to identify. In some cases (e.g., fraction of male population) it is clear that there is not enough variation in the data to identify nonlinear coefficients.

probabilities of voting for each candidate, and averaging across simulations to obtain the simulator for the integral. Further details are discussed in the Appendix.

7 Estimation results

7.1 Parameter estimates

We begin by modeling heterogeneity in voter preferences through Normally distributed random coefficients (i.e., setting $\mathbf{\Pi} = 0$, and capturing heterogeneity through $\mathbf{\Sigma v}_i$ in equation (3) and $\sigma_0 v_i^0$ in equation (4)). Estimation results for these specifications (reported in the Appendix) yield the following patterns. First, most of the linear parameters on the candidate characteristics are statistically significant, indicating that these variables are relevant determinants of average voter utility in a constituency over and above the party labels (since the mean utility always includes party fixed effects). Second, the random coefficients on gender and age tend to be large and statistically significant while for the other characteristics they are usually small and insignificant. This suggests the presence of significant heterogeneity in voter preferences for gender and age but not for other characteristics (including the party dummies). For these other characteristics, controlling for their mean valuation (together with all other candidate and constituency characteristics) appears to leave little individual heterogeneity for the model to explain. Third, we find that these Normal random coefficients specifications of the model are generally inadequate: the J-test always rejects these specifications.

In Table 6 we relax the assumption of Normally distributed random coefficients and instead model voter heterogeneity using the distribution of voter demographics from the Census ($\mathbf{\Sigma} = \mathbf{0}, \mathbf{\Pi} \neq \mathbf{0}$). We allow for heterogeneity in voter preferences for candidate gender and age, since this is where the random coefficients specifications suggested the presence of significant heterogeneity, as well as for NOTA. The first three specifications in Table 6 differ in the instrument sets used for identification.

Columns (1) and (2) of Table 6 paint a consistent picture. Many of the characteristics continue to exhibit significant correlation with mean voter valuations: the linear parameters indicate that voters have a preference for non-minority candidates and candidates fielded by parties that were more successful in the last election. Based on the nonlinear parameters, we see a statistically significant preference for female candidates among less literate voters, for male candidates among the more literate, and for older candidates among minority voters and rural workers.

We also find that voters have a lower utility from abstaining in reserved constituencies and in elections held on days with less rain (though the latter is not statistically significant). The former may reflect a positive utility associated with participating in elections in reserved constituencies, while the latter is consistent with the literature on increased voting costs due to bad weather.

We find that these specifications of voter heterogeneity perform much better than the Normal random coefficients specifications. They pass the J-test for the validity of the moment conditions, and the Newey-West D-test always rejects the null that the nonlinear parameters included in the specifications are jointly 0. The third instrument set in column (3) yields somewhat different patterns for mean voter valuations and heterogeneity, but the J-test rejects this specification. In the counterfactual exercise below we will use column (2) as our main specification.

In columns (4) and (5) we estimate the model including as additional candidate characteristics education, criminal history and assets from the candidate affidavits filed with the Election Commission. This requires dealing with the difficulty that data on these characteristics is missing for approximately 8% of our sample. Since here the BLP procedure requires observing the full set of available options in a constituency, candidates with an incomplete vector of characteristics cannot simply be dropped from the estimation. Including these candidates in the "outside option," as is sometimes done in consumer demand applications, is also not feasible since in our case the outside option is a well-defined choice (abstention). Instead, we replace missing values with 0 and use as additional candidate characteristics an indicator taking a value of 1 for missing education or criminal history, and another indicator taking the value of 1 for missing asset information. In all specifications we always use these characteristics together with their corresponding missing-indicators.

The findings from column (4) and (5) show that including education, criminal history, and assets in mean voter valuations has little effect on the estimates reported above. Interestingly, we find consistently positive and significant parameters on education and crime in mean voter utility, indicating that voters have a preference for both of these characteristics.[43] For the counterfactual exercises below we also present results using the specification in column (5) as a robustness check.

In the Appendix, we report further specifications that allow for different sources of heterogeneity for abstention and NOTA, and that exclude the state of Mizoram, which is a potential outlier in several dimensions. We show that the counterfactual simulations below yield similar results using these alternative specifications.

[43] The latter is consistent with Vaishnav's (2017) discussion on the electoral success of criminally convicted politicians in India.

Table 6: Parameter estimates of the full model using voter demographics

	(1)	(2)	(3)	(4)	(5)
Linear parameters					
Gender	6.951**	6.846***	4.382***	7.331***	7.711***
	(2.765)	(2.295)	(1.683)	(2.426)	(2.626)
Age	-1.432	-0.734	4.668***	-0.195	-1.475
	(3.956)	(3.241)	(1.596)	(3.294)	(3.405)
Minority	-7.517***	-5.404***	-0.531*	-5.130***	-5.311***
	(1.903)	(0.794)	(0.307)	(0.775)	(0.844)
Ran	0.287	0.214	0.069	0.199	0.215
	(0.207)	(0.178)	(0.105)	(0.176)	(0.179)
Won	0.548***	0.548***	0.497***	0.527***	0.624***
	(0.184)	(0.173)	(0.130)	(0.172)	(0.165)
NOTA	-4.374***	-3.981***	-3.762***	-3.957***	-4.058***
	(0.534)	(0.207)	(0.117)	(0.201)	(0.213)
Reserved SC	5.792***	3.955***	0.205	3.797***	3.902***
	(1.635)	(0.568)	(0.273)	(0.550)	(0.585)
Reserved ST	3.305**	1.704***	0.952***	1.669***	1.723***
	(1.388)	(0.292)	(0.118)	(0.277)	(0.284)
Rainfall	-0.133	-0.136	-0.218*	-0.129	-0.111
	(0.086)	(0.088)	(0.120)	(0.085)	(0.083)
Broadcast	-0.273	-0.248	-0.090	-0.247	-0.276
	(0.289)	(0.268)	(0.166)	(0.264)	(0.269)
Education				0.459***	0.490***
				(0.114)	(0.124)
Crime				0.269**	0.283**
				(0.117)	(0.115)
Missing educ/crime				-0.003	0.039
				(0.129)	(0.171)
Assets					-0.888*
					(0.465)
Missing assets					-0.145
					(0.206)
Nonlinear parameters (Π)					
Gender x Minority	-2.125	-4.719	-6.539	-6.000	-7.563
	(4.869)	(4.082)	(5.359)	(4.643)	(5.257)
Gender x Literate	-13.750	-13.375**	-4.184	-14.750**	-15.625**
	(8.537)	(6.795)	(4.167)	(6.893)	(6.822)
Age x Minority	20.000***	18.188***	-8.100***	17.375***	17.375***
	(3.524)	(2.786)	(3.052)	(2.695)	(2.835)
Age x Rural worker	7.625*	9.500***	12.525***	8.813***	8.500***
	(4.511)	(2.623)	(2.615)	(2.628)	(2.760)
NOTA x Rural worker	-0.906	-0.064	0.295*	-0.045	-0.094
	(2.451)	(0.469)	(0.173)	(0.447)	(0.464)
J	8.083	10.092	110.844	9.182	13.229
df	6	9	12	9	9
p-value	0.232	0.270	0.000	0.421	0.153
Newey-West D	36.966	37.377	48.302	33.41	40.40
p-value	0.000	0.000	0.000	0.000	0.000

Notes: Parameter estimates from the BLP model using voter demographics ($\Pi \neq 0$). Columns 1-3 are for the first, second, and third instrument set, respectively. Columns 4 and 5 use the second instrument set. The linear parameters also include indicators for parties, states, and years. Standard errors robust to heteroskedasticity and intra-constituency correlation in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively. J is the overidentification test statistic with corresponding degrees of freedom and p-value. Newey-West D is a likelihood ratio test for the null hypothesis that the nonlinear parameters are jointly 0 with the corresponding p-value. N = 9831.

8 Counterfactual simulations

In this section we use the estimated model to evaluate the impact of NOTA. First, we perform a counterfactual experiment where NOTA is removed. This allows us to study how many new voters turned out in order to vote for this option, and how many voters switched from candidates to NOTA. It also allows us to investigate from what type of parties / candidates voters switched to NOTA. This in turn helps us understand the behavior of protest voters in elections where the NOTA option is not available. In addition, this exercise helps us quantify whether the Indian Supreme Court's argument that NOTA would boost political participation was justified. As a second policy experiment, we ask how the presence of NOTA would mediate the impact of introducing compulsory voting, which is an issue of current policy relevance in India and elsewhere. Throughout we focus on the preferred specification in column (2) of Table 6, and present results obtained from additional specifications in the Appendix.

8.1 The impact of NOTA

To evaluate the impact of NOTA on voter behavior, we restrict attention to those constituencies in our data that had the NOTA option available in 2013 (520 constituencies, comprising 101.4 million eligible voters, with a total of 3073 candidates). We perform a counterfactual experiment where the NOTA option is removed and study voters' resulting choices, holding everything else constant. 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. In constituency c the impact of NOTA on candidate j's vote share (as a fraction of eligible voters) is given by $\Delta s_{jc} \equiv s_{jc} - s_{jc}^*$ and its impact on turnout is $\Delta(1 - s_{0c}) \equiv s_{0c}^* - s_{0c}$, where stars denote the counterfactual value. Relative to the reduced form exercise, this analysis makes it possible to study the impact of NOTA on individual voter behavior, holding all else constant.[44]

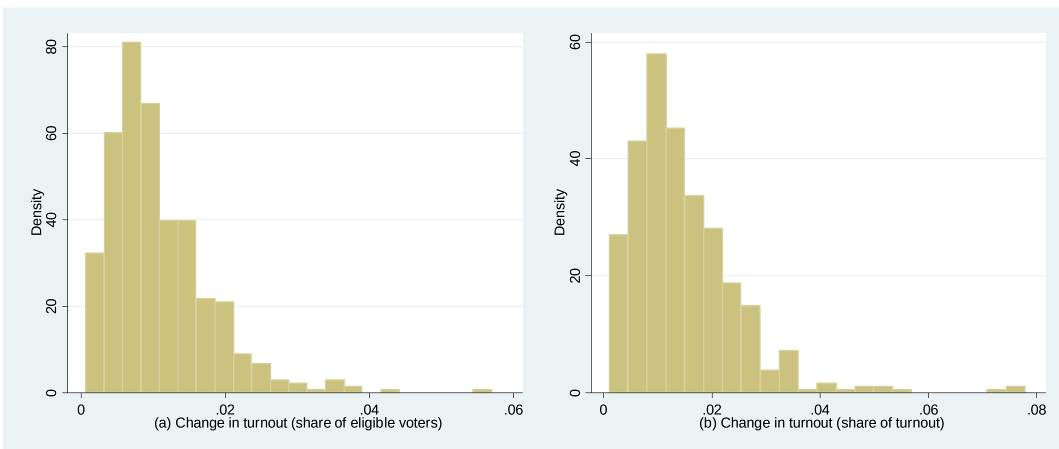
8.1.1 Turnout vs candidate votes

The simulated impact of introducing the NOTA option on turnout is shown in panel (a) of Figure 2. This figure shows the distribution of change in turnout, $\Delta(1 - s_{0c})$, across constituencies. The average increase in turnout is 1.08 percentage points, which is close to

[44]Note that by holding all else constant (including the potentially endogenous variables) we are studying how NOTA voters would behave if the only change was the presence/absence of the NOTA option. This is exactly the question of interest for understanding the behavior of protest voters with/without NOTA. Studying changes in other endogenous variables would require an explicit model of how those variables are determined.

the 1.57 percent of eligible voters who voted for NOTA in this sample. The simulated increase in turnout due to NOTA is at least 0.5 percentage points in 82% of the constituencies. Panel (b) presents the change in turnout as a share of the counterfactual without-NOTA turnout ($\Delta(1 - s_{0c})/s^*_{0c}$) and shows similar patterns. The simulation indicates that NOTA increased turnout by 1.44 *percent*, or, equivalently, reduced abstention by 4.7 *percent*. These findings are in line with those of the reduced form analysis in Table 4. They indicate that NOTA raised turnout, and that, on average, most of the protest vote is accounted for by new voters who would have abstained without a NOTA option on the ballot.

Figure 2: Impact of NOTA on turnout

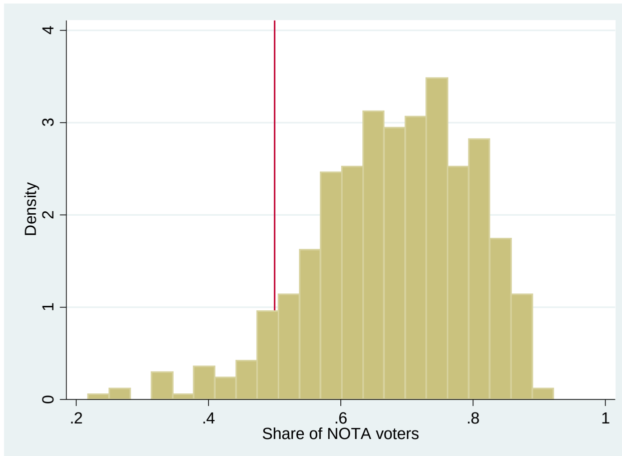


Notes: Distribution of the changes in turnout across constituencies as a share of eligible voters (panel (a)) and as a share of the (counterfactual) without-NOTA turnout (panel (b)). N = 520.

What fraction of protest voters abstain or vote for candidates in each constituency? To answer this question, we aggregate all votes cast on candidates at the constituency level and compute $\sum_{j \neq 0, NOTA} \Delta s_{jc}$, the reduction in the vote share of all candidates in constituency c due to NOTA. Writing $s_{NOTA,c} = \Delta(1 - s_{0c}) + \sum_{j \neq 0, NOTA} |\Delta s_{jc}|$, we compute the fraction of NOTA votes explained by new voters (as opposed to substitution away from candidates) as $\frac{\Delta(1-s_{0c})}{s_{NOTA,c}}$. This represents the fraction of protest voters who would choose to abstain if NOTA was not available. Figure 3 shows the distribution of this measure across constituencies. The figure reveals that in almost all constituencies (93%) the majority of NOTA votes is due to new turnout, i.e., most protest voters would abstain in the absence of NOTA. In the average constituency, 68% of NOTA voters would abstain in an election where everything else was held constant but the NOTA option was removed from the ballot.

Figure 4 shows the estimated impact of NOTA on individual candidates' vote shares, Δs_{jc}. Panel (a) contains all candidates. Among these, the average reduction in vote shares

Figure 3: Share of NOTA voters abstaining without NOTA



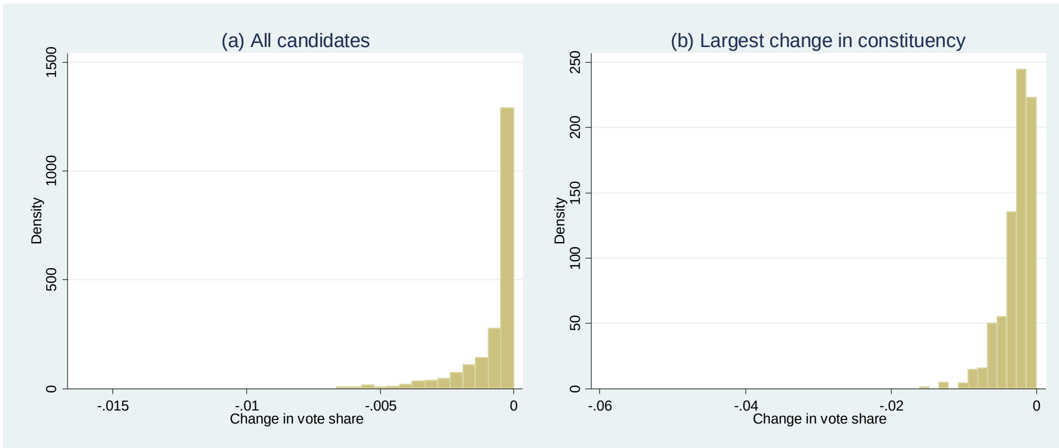
Notes: Share of NOTA voters in a constituency who choose to abstain when NOTA is removed. The vertical line represents 0.5. N = 520.

(as a fraction of eligible voters) is 0.08 percentage points, and the change is smaller than 0.5 percentage points in absolute value for 98% of the candidates. Small changes for the average individual candidate could be due in part to the large number of candidates. Therefore panel (b) shows, for each constituency, the candidate who experienced the largest change in vote share ($\max_j |\Delta s_{jc}|$ for each c). The largest change is 0.28 percentage points in the average constituency, and it is below 0.5 percentage points in absolute value for 85% of the constituencies.

The pattern on Figure 4 suggests that, in an election without the NOTA option, protest votes that are cast on candidates tend to be scattered across many candidates. This in turn implies that, in this setting, the impact of protest votes on the outcome of the election will be limited. Indeed, according to our counterfactual exercise, the introduction of NOTA changed the identity of the winner in only 2 out of our 520 elections (0.4%). (In both of these elections the INC and the BJP switched places, see Table 7 below.) The small impact of NOTA on election outcomes is in spite of the fact that in the data, the number of NOTA votes would have been enough to change the winner in 43/520 elections (8.3%) if protest voters had coordinated on the right candidate.

In the Appendix, we repeat this counterfactual experiment with specifications that use the first instrument set, that include education, criminal history and assets as additional candidate characteristics, or that include small variations in the demographic interactions used in the model. We find similar results to those reported above: depending on the specification, the average increase in turnout is between 0.98 and 1.10 percentage points and

Figure 4: Impact of NOTA on candidates' vote shares



Notes: Distribution of the change in vote shares (as a fraction of eligible voters) across candidates. Panel (a) is for all 3073 candidates. Panel (b) is for the candidate with the largest change in each of the 520 constituencies.

accounts for 63 to 69% of NOTA votes in the average constituency. The average change in candidate vote shares is small, with values between 0.08 and 0.10 percentage points, and the election outcome is always affected in only 2 constituencies.

8.1.2 Which parties are substitutes for NOTA?

We now focus on the protest voters who do substitute between candidates and NOTA, and ask which parties they would choose if the NOTA option was removed. To do this, we first aggregate candidates by party across constituencies in the data. The first three columns of Table 7 present, for each party, the number of candidates fielded (in the 520 constituencies), the number of constituencies won, and the fraction of all eligible voters who voted for them. We then do a similar aggregation for the counterfactual results, and present the difference in columns (4) (in absolute terms) and (5) (relative to the actual vote shares).

Which parties' supporters include relatively more protest voters? Column (5) of Table 7 indicates that the four parties that lose the largest share of their voters to NOTA are BYS, CSM, GGP, and JGP. Two common characteristics of these parties are that (i) none of them won a single constituency in the data, and (ii) internet searches on their platform and history indicate that they are quite radical, and/or represent a very specific voter group. For example, the JGP's platform includes the end of the system of political reservation, terminating all subsidies, giving cash transfers to all voters, and introducing the death penalty for rape and corruption.[45] The GGP was formed in 1991 to represent the Gondi people and to establish a separate Indian state. The CSM broke away from the BJP in 2008 in order to provide a "third alternative."[46] The BYS's platform appears to be to "support the role of Indian youth."[47] Based on our counterfactual simulations, supporters of these fringe parties include relatively larger shares of protest voters.[48]

Although the results indicate that these parties' supporters include relatively more protest voters, we find that the majority of their voters are *not* protest voters. Most of these parties' voters appear to support them for reasons other than simply as a substitute for NOTA.

It is worth emphasizing that allowing for voter heterogeneity in the estimation was crucial to obtain these substitution patterns between the various choices. A restricted model with $\Pi = \Sigma = 0$ would have ruled out these patterns since, as is well known, this "Logit" specification *assumes* that substitution patterns only depend on observed choice shares. In our case, this would mean that the change caused by NOTA would, by construction, be similar for options with similar choice shares in a constituency. This is illustrated in the last two columns of Table 7 which show the counterfactual implications that would be obtained from the Logit specification. Since abstention has a similar choice share as the BJP and the INC in the data, the Logit specification would imply that substitution towards NOTA is mechanically similar for these three options, with changes of 0.487, 0.465, and 0.501 percentage points (column 6). By contrast, in the full model substitution towards NOTA is much larger for abstention (1.085 percentage points vs. 0.201 and 0.169 in column 4).

[45] https://en.wikipedia.org/wiki/Jago_Party

[46] http://economictimes.indiatimes.com/news/politics-and-nation/chhattisgarh-regional-party-merges-with-bjp/articleshow/30958876.cms

[47] www.byspindia.org

[48] In the Appendix, we repeat the entire exercise using a lower threshold for aggregating small parties. We find a fifth party, the SHS, whose supporters include a relatively larger (2%) share of protest voters. This is one of the most radical parties in India, associated with fascism and a number of violent ethnic riots.

Table 7: The impact of NOTA on parties

Choice	Data			Change due to NOTA: Full model				Change due to NOTA: Logit	
	N. of candidates	Elections won	Percent of all voters	Votes (ppoints)	Votes (percent)	Extra wins	Extra losses	Votes (ppoints)	Votes (percent)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
BJP	507	361	32.910	-0.201	-0.609	1	1	-0.465	-1.413
BSP	499	8	3.623	-0.030	-0.818	0	0	-0.053	-1.467
BYS	102	0	0.102	-0.002	-1.506	0	0	-0.002	-1.793
CSM	54	0	0.223	-0.004	-1.869	0	0	-0.003	-1.188
GGP	44	0	0.203	-0.003	-1.711	0	0	-0.003	-1.525
INC	519	127	26.732	-0.169	-0.632	1	1	-0.501	-1.875
Independents	469	15	4.963	-0.045	-0.909	0	0	-0.075	-1.513
JGP	85	0	0.060	-0.001	-1.235	0	0	-0.001	-1.490
MNF	10	1	0.057	0.000	-0.225	0	0	-0.001	-1.813
NPEP	133	4	1.294	-0.009	-0.712	0	0	-0.008	-0.648
NOTA	520	0	1.578	1.578	100	0	0	1.578	100
SP	194	0	0.432	-0.003	-0.714	0	0	-0.006	-1.340
Small parties	446	4	2.692	-0.026	-0.976	0	0	-0.030	-1.110
ZNP	11	0	0.016	0.000	-0.256	0	0	-0.001	-3.462
Abstention			25.114	-1.085	-4.319	0	0	-0.487	-1.938
Total	3593	520	100	0		2	2	0	

Notes: Tabulation of all the choices available in the data used for the counterfactual exercise. For each party column (1) shows the total number of candidates and (2) the number of constituencies won. Column (3) is the share of all voters (out of 101.384 million eligible voters) choosing each option in the data. Column (4) gives the simulated effect of introducing NOTA in the full model, (5) is (4) divided by (3), times 100. Columns (6) and (7) show the number of additional constituencies won and lost by each party as a result of NOTA. Columns (8)-(9) show the impact on vote shares from the Logit model ($\Pi = \Sigma = 0$).

8.1.3 Heterogeneity across constituencies

In which constituencies are protest voters more likely to abstain, vote for the major parties, or vote for other parties? Table 8 explores the heterogeneity across constituencies in the simulated behavior of protest voters. "Major parties" refer to the INC and the BJP. Reported coefficients are from a 3-equation SUR system where the fraction of NOTA voters who choose each option is regressed on various constituency observables (since shares sum to 1, coefficients in the third column are obtained as linear combinations of the first two). Here, all covariates are entered simultaneously; the Appendix shows regressions with one covariate at a time with substantively similar results.

According to these correlations, protest voters are more likely to abstain in constituencies with more minority voters and fewer minority candidates. One possible interpretation of these patterns is that NOTA is an attractive choice for minority voters to register their protest, particularly in constituencies where there are no minority candidates who could play this role. This is also consistent with the case studies of Banerjee (2014) discussed in section 2.4 where voters view elections as an opportunity to express their political relevance. Minority voters with little political representation may feel particularly invisible to the state, and for them voting for NOTA may be a particularly valuable form of political expression.

We also find that protest voters are more likely to choose non-major parties in smaller and more literate constituencies, and (not surprisingly) in constituencies with more candidates running. They are also more likely to choose non-major parties in constituencies that have female candidates (who often run as independents).

As we show in the Appendix, there are no obvious patterns in the geographic distribution of the simulated behavior of protest voters.

8.2 NOTA and compulsory voting

Having a NOTA option on the ballot gains added significance in a compulsory voting system. In this case, NOTA is the only option available to a voter who wishes not to influence the outcome of the election without incurring the fine associated with abstention. In a series of lab experiments, Ambrus et al. (2017) find that the share of NOTA votes almost doubles under compulsory voting, and this can have consequences for who wins the election.

Compulsory voting has been a recurring policy question in India for many years. The Indian parliament has considered bills proposing to introduce compulsory voting several times in the past, but these were never passed. The advent of NOTA appears to have re-energized proponents of compulsory voting, a group that includes the currently ruling BJP party and prime minister Modi. In 2015, the state of Gujarat introduced compulsory voting

Table 8: Heterogeneity in the behavior of NOTA voters

Dep. Var.:	Share of NOTA voters switching to		
	Abstention	Major parties	Other parties
Reserved	0.010	-0.007	-0.003
	(0.013)	(0.013)	(0.008)
Literacy	-0.020	-0.110	0.130***
	(0.078)	(0.078)	(0.050)
Size	0.072*	-0.010	-0.062**
	(0.042)	(0.042)	(0.027)
Minority pop.	0.211***	-0.208***	-0.003
	(0.040)	(0.040)	(0.025)
Rural workers	0.045	-0.114***	0.069***
	(0.039)	(0.039)	(0.025)
N.candidates	-0.000	-0.005***	0.005***
	(0.001)	(0.001)	(0.001)
Female cand.	-0.000	-0.012	0.012*
	(0.010)	(0.010)	(0.007)
Minority cand.	-0.041***	0.046***	-0.005
	(0.015)	(0.015)	(0.010)

Notes: System of seemingly unrelated regressions of the share of NOTA voters switching to abstention, major parties (INC or BJP) or other parties in each constituency in the no-NOTA counterfactual. Size is log(number eligible voters). Female cand. is 1 if there is at least one female candidate running, Minority cand. is 1 if there is at least one minority candidate. Each regression includes state fixed effects. Standard errors in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively. N=520.

in local elections but this was struck down by the state's High Court. In early 2017 the Indian Supreme Court held hearings on the question of nation-wide compulsory voting.[49]

In this section we use our model to simulate the impact of introducing compulsory voting with and without the NOTA option. We use the same 520 constituencies as for our no-NOTA counterfactuals above (the constituencies where NOTA was available in 2013). For both the no-NOTA and with-NOTA scenario, we simulate the impact of compulsory voting by removing the abstention option. Thus, we assume full compliance with compulsory voting.[50] We predict voter behavior in the compulsory system and study the resulting vote shares.

The results are summarized in Tables 9 and 10, which have a similar structure to Table 7 above. The Appendix provides further details. When NOTA is available (Table 9), com-

[49] https://www.telegraphindia.com/1170124/jsp/nation/story_132101.jsp#.WQER4I-cGUk

[50] In practice the expected fine associated with abstention is unlikely to be prohibitive so some voters will still choose to abstain, and abstention could be heterogenous across voter-types. In this case, details of the compulsory voting policy, such as the magnitude of any fines, will change the impact of the policy both with and without NOTA. Our results below should be interpreted as providing a benchmark under the assumption that compliance with the policy is complete. Future work exploring the implications of relaxing this assumption would be useful.

Table 9: The impact of compulsory voting with NOTA available

Choice	Voluntary voting (data)			Change due to compulsory voting			
	N. of candidates	Elections won	Percent of all voters	Votes (ppoints)	Votes (percent)	Extra wins	Extra losses
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
BJP	507	361	32.910	6.674	20.280	18	50
BSP	499	8	3.623	1.786	49.315	7	3
BYS	102	0	0.102	0.077	75.356	0	0
CSM	54	0	0.223	0.097	43.666	0	0
GGP	44	0	0.203	0.064	31.471	0	0
INC	519	127	26.732	5.745	21.490	45	21
Independents	469	15	4.963	2.071	41.716	6	2
JGP	85	0	0.060	0.050	83.076	0	0
MNF	10	1	0.057	0.007	12.791	0	0
NPEP	133	4	1.294	0.359	27.741	0	2
NOTA	520	0	1.578	6.928	439.037	1	0
SP	194	0	0.432	0.184	42.499	0	0
Small parties	446	4	2.692	1.070	39.745	1	0
ZNP	11	0	0.016	0.002	14.321	0	0
Abstention			25.114	-25.114	-100		
Total	3593	520	100	0		78	78

Notes: For each party column (1) shows the total number of candidates and (2) the number of constituencies won in the data (when voting is voluntary). Column (3) is the share of all voters (out of 101.384 million eligible voters) choosing each option in the data. Column (4) gives the simulated effect of introducing compulsory voting, (5) is (4) divided by (3), times 100. Columns (6) and (7) show the number of additional constituencies won and lost by each party as a result of compulsory voting.

pulsory voting increases the share of NOTA votes among eligible voters by a factor of five (6.9 percentage points). Over one third of the voters "forced" to turn out by compulsory voting (1.578 + 6.928 = 8.506 percent of eligible voters) choose NOTA, and we find that NOTA actually "wins" in one of the constituencies. In terms of relative popularity (column (5)), the two biggest winners after NOTA are the fringe parties BYS and JGP discussed above. We find that, when NOTA is available, compulsory voting changes the winner in 78 (15%) of elections (columns (6)-(7)). In terms of election outcomes, the INC is a net winner ($+45-21=+24$ seats) and the BJP a net loser ($+18-50=-32$ seats) from compulsory voting according to our results. Among smaller parties, the BSP gains while the NPEP looses a few seats.[51]

The impact of compulsory voting without NOTA is shown in Table 10. In this case, all new voters are forced to vote for actual candidates. In terms of relative popularity, the

[51] In constituencies where independent and small party candidates represent aggregates of multiple candidates, counterfactual wins by these candidates may not be meaningful. Ignoring these cases does not change the patterns described here (see the Appendix).

Table 10: The impact of compulsory voting when NOTA is not available

Choice	Voluntary voting			Change due to compulsory voting			
	N. of candidates	Elections won	Percent of all voters	Votes (ppoints)	Votes (percent)	Extra wins	Extra losses
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
BJP	507	361	33.111	9.290	28.056	28	77
BSP	499	8	3.652	2.656	72.722	13	3
BYS	102	0	0.104	0.145	139.488	0	0
CSM	54	0	0.227	0.132	58.208	0	0
GGP	44	0	0.206	0.116	56.549	0	0
INC	519	127	26.901	8.109	30.144	64	33
Independents	469	15	5.009	3.059	61.070	9	2
JGP	85	0	0.061	0.083	135.775	0	0
MNF	10	1	0.057	0.010	16.914	1	0
NPEP	133	4	1.303	0.607	46.580	1	2
NOTA	0	0	0	0			
SP	194	0	0.436	0.298	68.437	0	0
Small parties	446	4	2.719	1.690	62.168	1	0
ZNP	11	0	0.016	0.004	26.697	0	0
Abstention			26.199	-26.199	-100		
Total	3073	520	100	0		117	117

Notes: For each party, columns (1-3) describe the counterfactual scenario with NOTA removed. Column (1) shows the total number of candidates, (2) the number of constituencies won, and column (3) the share of all voters (out of 101.384 million eligible voters) choosing each option. Column (4) gives the simulated effect of introducing compulsory voting without NOTA. (5) is (4) divided by (3), times 100. Columns (6) and (7) show the number of additional constituencies won and lost by each party as a result of compulsory voting.

biggest winners from compulsory voting are now the BYS and the JGP, which more than double their number of voters. Election outcomes now change in 50% more constituencies than they did when NOTA was available: there is a different winner in 117 or 22.5% of constituencies. Compared to the with-NOTA case, we generally find that the same parties are net winners and losers, but the magnitude of both the net wins and losses is now larger. For example, the INC now has a net gain of 31 seats, while the BJP a net loss of 49 seats. In this sense, the presence of NOTA attenuates the impact of compulsory voting on election outcomes.

9 Discussion

9.1 Implications for theories of voting

Our estimation framework is most naturally interpreted based on the Downsian view of voting, where consumers derive consumption utility from turning out as well as from choosing a particular option. Empirically identifying these consumption motives for voting in real-world settings is often challenging. Results from field experiments that create a social pressure to vote indicate the existence of a general consumption utility from turning out (Gerber et al., 2008; DellaVigna et al., 2017). An option-specific consumption utility (utility from choosing a particular option) has been difficult to identify even in laboratory experiments (see Tyran (2004), Feddersen et al. (2009), Shayo and Harel (2012), Kamenica and Egan Brad (2014)). In our setting, all else equal, a NOTA voter who abstains when this option is not available must derive an option-specific utility from choosing NOTA. Thus, our results indicate the presence of an option-specific consumption utility in real-world elections.

Another possible interpretation of NOTA votes is in terms of models that focus on the messages voters send to politicians (or other voters) through their behavior at the polls (Lohmann, 1993; Piketty, 2000). In these models, there is typically more than one way to send the same signal in equilibrium. For example, to express their preference regarding a particular policy, voters could coordinate on voting for one of the small parties, or they could coordinate on abstaining (see Castanheira, 2003, Section 5 for a nice discussion of this point). Voters signal through abstention in Shotts (2006) and McMurray (2017), and by voting for specific candidates in Piketty (2000), Castanheira (2003), Razin (2003), and Myatt (2016). Thus, how exactly voters signal to politicians is an empirical question. From this perspective, our results can be interpreted as showing that, for one particular signal (signaling their protest), most voters appear to use abstention rather than a vote for one of the candidates.

9.2 Generalizability of the findings

As discussed above, several features of the Indian context are advantageous for our analysis. Main among these are the rules of the NOTA policy precluding a direct electoral impact of protest votes, and the relatively small fraction of voters choosing NOTA. The latter in particular means that ignoring strategic responses to NOTA (by voters or by parties) seems plausible. Of course, these same features also mean that care must be taken in generalizing the findings to contexts where the fraction of protest voters may be larger.

One relevant consideration in generalizing the findings is the fact that Indian turnout is so high. For example, our finding that only a minority of NOTA voters substituted away from candidates means that *even in an environment where many voters were already voting for candidates*, few of them appeared to be doing so purely to express protest.

Similarly, we found that introducing NOTA can increase turnout even when abstention is already low. Our main counterfactual exercise implies that NOTA reduced abstention by 4.7 *percent* in the average constituency. According to estimates available in the literature, even highly personalized interventions like phone calls to voters typically reduce abstention by under 10 percent (Nickerson, 2006). In Gerber et al. (2008), informing voters that their voting behavior will be scrutinized by researchers reduced abstention by 3.6 percent; threatening to publicly reveal whether or not they voted reduced abstention by an additional 8.2 percent. Compared to these more costly interventions, the impact of NOTA seems remarkably large in a setting where abstention rates were low to begin with.

10 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. We use the behavior of NOTA voters in the absence of NOTA to study protest 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 in the absence of NOTA most protest voters would have abstained. Hence, our results indicate that the NOTA policy increased turnout. We also find that protest voters who normally turn out scatter their votes across many candidates. As a result, protest voting (and therefore the NOTA policy) affects the outcome of very few elections.

Our results also allow identifying specific parties whose supporters include relatively more protest voters. These tend to be small parties representing specific voter groups and / or advocating extreme policies. However, we find that the overwhelming majority of these parties' supporters are not protest voters, in the sense that they choose them for reasons other than merely a substitute for NOTA. This type of finding may help interpret the rise of extremist parties in various settings.

Was the NOTA policy worth it? Introducing a NOTA option on the voting machines is a fairly simple, low-cost policy: a back-of-the-envelope calculation in the Appendix puts the direct cost at a few dollars for every 1 million eligible voters. The above results indicate that NOTA fulfilled the Supreme Court's stated goal of increased voter participation. In this sense, the policy may have been cost-effective.[52] At the same time, NOTA may have externalities, both negative and positive. On the negative side, the several millions of extra voters showing up at the polls may have created extra costs in terms of election administration, or increased the waiting time for other voters. On the positive side, increased civic participation in elections may spill over to improved civic participation in other areas. In the long run, it may help hold politicians accountable, or crowd out less desirable forms of political expression (e.g., violence). Quantifying these effects is beyond the scope of this paper, but some of these possibilities would be interesting to explore in future research.

References

[1] Ambrus, A., B. Greiner, and A. Sastro (2017): "The Case for Nil Votes: Voter Behavior under Asymmetric Information in Compulsory and Voluntary Voting Systems," *Journal of Public Economics* 154, 34-48.

[2] Ambrus, A., B. Greiner, and A. Zednik (2018): "The effects of a 'None of the above' ballot paper option on voting behavior and election outcomes," working paper.

[3] Banerjee, M. (2014): *Why India Votes?* Routledge, New Delhi, India

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

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

[6] Brown, A.R. (2011): "Losing to nobody? Nevada's "none of these candidates" ballot reform," *The Social Science Journal* 48, 364–370.

[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.

[52] According to our results, in most cases the policy generated consumption utility for voters without affecting the winner of the election. Of course, the welfare calculation would be very different if NOTA had overturned more election results.

[9] Castanheira, M. (2003): "Why Vote for Losers?" *Journal of the European Economic Association* 1(5), 1207-1238.

[10] Chattopadhyay, R., and E. Duflo (2004): "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," *Econometrica* 72(5), 1409–1443.

[11] 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.

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

[13] 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.

[14] Damore, D.F., M.M. Waters, and S. Bowler (2012): "Unhappy, Uninformed, or Uninterested? Understanding "None of the Above" Voting," *Political Research Quarterly* 65(4), 895-907.

[15] DellaVigna, S., J.A. List, U. Malmendier, and G. Rao (2017): "Voting to Tell Others," *Review of Economic Studies* 84(1), 143-181.

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

[17] 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.

[18] 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.

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

[20] 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.

[21] Gillen, B.J., S. Montero, H.R. Moon, and M. Shum (2015): "BLP-Lasso for Aggregate Discrete-Choice Models with Rich Demographic Covariates," working paper.

[22] Gordon, B.R., and W.R. Hartmann (2013): "Advertising Effects in Presidential Elections," *Marketing Science* 32(1), 19-35.

[23] Grypari, I. (2011): "Issue Emphasis in U.S. Presidential Elections," working paper.

[24] 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.

[25] 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.

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

[27] Kapoor, S., and A. Magesan (2018): "Independent Candidates and Political Representation in India," *American Political Science Review* 112(3), 678-97.

[28] Kawai, K., and Y. Watanabe (2013): "Inferring Strategic Voting," *American Economic Review* 103(2), 624–662.

[29] Kselman, D., and E. Niou (2011): "Protest voting in plurality elections: a theory of voter signaling, " *Public Choice* 148, 395-418.

[30] Lohmann, S. (1993): "A Signaling Model of Informative and Manipulative Political Action," *American Political Science Review* 87(2), 319-333.

[31] Lohmann, S. (1994): "Information Aggregation Through Costly Political Action," *American Economic Review* 84(3), 518-530.

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

[33] 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.

[34] McMurray, J. (2017): "Voting as communicating: Mandates, multiple candidates, and the signaling voter's curse," *Games and Economic Behavior* 102, 199-223.

[35] Merlo, A., and A. de Paula (2017): "Identification and Estimation of Preference Distributions When Voters Are Ideological," *Review of Economic Studies* 84, 1238-1263.

[36] Myatt, D.P. (2017): "A Theory of Protest Voting," *Economic Journal* 127(603), 1527-1567.

[37] 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.

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

[39] Newey, W.K., and K.D. West (1987): "Hypothesis Testing with Efficient Method of Moments Estimation," *International Economic Review* 28(3), 777-787.

[40] Nickerson, D.W. (2006): "Volunteer Phone Calls Can Increase Turnout," *American Politics Research* 34(3), 271-292.

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

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

[43] Piketty, T. (2000): "Voting as communicating," *Review of Economic Studies* 67, 169-191.

[44] Pons, V., and C. Tricaud (2018): "Expressive Voting and Its Cost: Evidence from Runoffs with Two or Three Candidates," *Econometrica* 86(5), 1621-1649.

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

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

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

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

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

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

[51] Shotts, K.W. (2006): "A signaling model of repeated elections," *Social Choice and Welfare* 27(2), 251-261.

[52] Sigelman, L., and C. Sigelman (1982). "Sexism, Racism, and Ageism in Voting Behavior: An Experimental Analysis," *Social Psychology Quarterly* 45(4), 263-269.

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

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

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

[56] Vaishnav, M. (2017): *When Crime Pays: Money and Muscle in Indian Politics*, Yale University Press, New Haven, CT.

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