# None Of The Above:

# Protest Voting in the World's Largest Democracy*

Gergely Ujhelyi, Somdeep Chatterjee, and Andrea Szabó†

July 3, 2020

**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. Using structural estimation, 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 the Editor, the anonymous referees, 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 mainstream parties. 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 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 difficult to know what fraction of these are simply 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). 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.

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

[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).

election results. Under the current system of voluntary voting, NOTA may be an effective tool to increase political participation. In counterfactual simulations, we estimate that NOTA would have a sizeable impact on election outcomes under compulsory voting.

Our project requires addressing three methodological challenges. (i) *Ecological inference.* We wish to estimate individual voter behavior in a counterfactual scenario without the NOTA option. However, because ballots are secret, individual voter behavior is observed neither with nor without NOTA. Instead, it must be inferred from aggregate data.[2] (ii) *Multiparty system.* Analyzing political outcomes in such systems faces well known difficulties due to the fact that outcomes are interrelated (e.g., vote share regressions have to account for the fact that vote shares sum to 1) and that parties may decide not to contest all races.[3] (iii) *Causality and mechanism.* Our goal is to compare voters' behavior with and without NOTA holding all else (e.g., the set of candidates running) constant. Each of these challenges requires us to make specific "structural" assumptions, but our setting also allows us to provide some reduced form evidence.

We begin with preliminary evidence on the aggregate impact of NOTA from a difference-in-differences exercise. 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. We find no evidence that NOTA affected the competitiveness of elections or the number of viable candidates.

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

[2] The ecological inference problem is central to much of political science. It is typically solved through statistical modeling, i.e., models not relying on utility maximization - see King (1997) and Cho and Manski (2008) for reviews.

[3] See Katz and King (1999) and Tomz et al. (2002) for a description of these simultaneity and selection issues and statistical procedures to deal with them. See Rekkas (2007) for a discussion.

voter demand for candidates. We adapt the BLP approach to consumer demand (Berry et al., 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 among general caste voters in reserved constituencies, 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 % 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 a NOTA option 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. Under compulsory voting, NOTA would gain added significance by giving voters the option to participate without influencing the electoral outcome (Ambrus et

al., 2017). 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.

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

## 2 Related literature

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 or against a particular candidate. 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 an option-specific utility component 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.[4]

[4] 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.

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 the Conclusion.

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). In our context, the structural approach is particularly useful to address the methodological challenges described in the introduction.

Finally, our paper relates to previous studies of NOTA-like votes. Most observational studies in this literature analyze "residual" (blank or invalid) votes, but this makes it difficult to separate intentional behavior from voting mistakes (McAllister and Makkai, 1993; Herron and Sekhon, 2005; Power and Garand, 2007; Superti, 2015). 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.[5] In a follow-up paper, Ambrus et al. (2018) conduct a lab-in-the-field experiment 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,

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

about the impacts of having a NOTA option on the ballot.

# 3 Background

## 3.1 The NOTA policy

In elections that use a paper ballot, 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 for any of the candidates.[6] In its 2013 decision, the Court agreed:

> "Democracy is all about choice. This choice can be better expressed by giving the voters an opportunity to verbalize themselves unreservedly and by imposing least restrictions on their ability to make such a choice. By providing NOTA button in the Electronic Voting Machines, it will accelerate the effective political participation in the present state of democratic system and the voters in fact will be empowered." (PUCL vs. Union of India, 2013, 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 would be the candidate who received the most votes among the non-NOTA votes.

These institutional features of the Indian NOTA make this policy ideal for the study of intentional voter behavior guided by a non-consequentialist motive. We review the difficulties that arise in interpreting null votes in other settings in Appendix 1.2.

[6] With electronic voting machines, the only way for a voter to have her non-vote recorded was to inform the clerk at the voting booth of her 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. See Appendix 1.1 for more details.

### 3.2 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.

Anecdotal evidence on how voters viewed NOTA suggests 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."[7] Based on this, we define protest voting simply as a revealed preference for voting "None Of The Above." In the rest of the paper, we will use the terms protest voter and NOTA voter interchangeably. To study protest voting in "normal"elections that do not have an explicit protest option, we will 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. 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

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

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 (Appendix 3.3.6).[8]

A second 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 significantly higher than among non-experienced states (1.28%, compared to 1.09%, $p = 0.027$).

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, 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 simply shows up to vote because she derives utility from participating. Such a voter, however, should still turn out in the absence of NOTA (and choose one of the candi-

[8] Note also that, based on our structural model, we estimate that illiterate voters (who may be more susceptible to mistakes) are *less* likely to choose NOTA.

dates). Our result that NOTA increased turnout are therefore inconsistent with this interpretation. Some of the substitution between *candidates* and NOTA could reflect indifferent voters who show up both with and without NOTA and vote randomly. However, as we show below, substitution to candidates is not uniform, which indicates that random voting cannot be responsible for our findings below.

### 3.3 Assembly elections and turnout in India

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 ("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 the 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 are typically held every 5 years but the electoral calendar varies 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. Since in these cases voters are faced with multiple choice sets, we exclude these states from the main analysis below.

All state and national elections are conducted by the Election Commission of India under the supervision of the chief election commissioner.[10] Election dates are set well in advance and declared as local holidays to reduce the cost of participation. Voting takes place using electronic voting machines (EVMs). Each candidate running in an election has a separate button assigned to

[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).

[10] Since independence, the Commission has emerged as a highly regarded institution with a large degree of autonomy (McMillan, 2010).

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 her 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. The reserved status of SC and ST constituencies is set at the same time as the electoral boundaries are drawn. In contrast to local governments, state elections have no political reservation for women.

The current electoral boundaries were set in April 2008 by a commission working under the Election Commission (see Table 1). This was the first time in over 30 years that electoral redistricting ("delimitation") took place in India. All constituency boundaries as well as the reservation status of the constituencies 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

[11] EVMs were introduced gradually beginning in 1999, and since 2004 all general and state elections have been conducted using these machines. See, e.g., https://pib.gov.in/newsite/mbErel.aspx?relid=104463. 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.

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 | | *National elections* |
| | 10 | Haryana, Maharashtra | |
| | 12 | Jharkhand | |
| 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 | | *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.

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.

# 4 Data

## 4.1 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 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 as there is too little overlap between the old and new constituencies to make constituency-level matching mean-

[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).

ingful.[13] While states on the 2009/14 election cycle also have consistent boundaries, we decided to exclude them from the panel. 2014 was a national election year that made headlines around the world for its unusual outcome (the BJP won by a landslide, the first time in 30 years that a single party won a majority of the legislative seats). 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. Focusing on states that held elections in a narrow (+/- 4 month) window around the NOTA policy is less subject to possible confounding effects.

*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 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 (see section 3.3) above. Based on previous research on the impact of voting machines, this could confound any patterns that we would observe in the data (Fujiwara, 2015; Debnath et al., 2018).

We next describe the information available in the panel and the extended dataset. Details, including summary statistics, are in Appendix 2.

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

[14] 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. In robustness checks, we show that excluding these states makes little difference for the results. We also confirm that the states with simultaneous national elections behave differently than others.

### 4.2 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 Appendix 2 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 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 $0.25 \times 0.25$ degree cells (which is smaller than the typical constituency). 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 by the civil

[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.

group ADR 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').

Summary statistics of the candidate characteristics appear in Table A.1. The overwhelming majority of candidates are male: the average constituency has less than 1 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 (approximately 30% of all constituencies) - in non-reserved constituencies the share of minority candidates is 12%. 59% of the candidates completed high school, and 14% have a criminal history.[17] The two largest parties, INC and BJP, field approximately 14.5% of the candidates. The average constituency has approximately 180,000 eligible voters and turnout is around 71%. Summary statistics of states in the panel are broadly similar to those of other states in the extended dataset (Appendix 3.2.3).

### 4.3 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. Appendix 2.3 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,

[17] Criminal history refers to the sum of all cases where a judge has decided to move ahead with a charge beyond the initial police investigation and prosecutorial action (see Vaishnav (2017, p318) for a detailed description). These are the cases candidates are required to disclose.

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.

# 5 Patterns in the data

## 5.1 NOTA votes

The first noteworthy feature of the data is that voters actually voted for NOTA, despite the fact that this could not affect the results of the election. In each of the 9 states 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] In our data, a total of 2.51 million voters chose NOTA.

To gauge the importance of NOTA for voters, a 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.

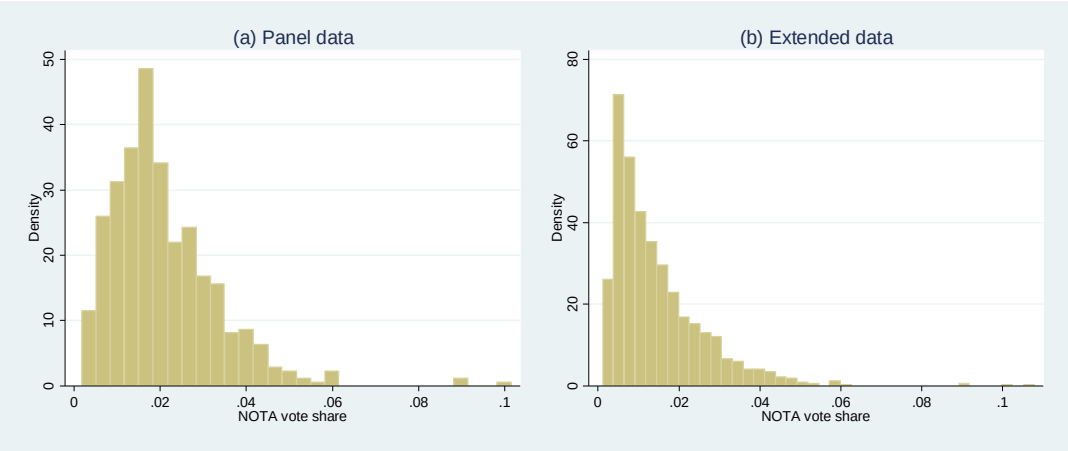
Figure 1 reveals some heterogeneity in NOTA votes across constituencies. To explore this further, in Appendix 3.1 we run cross-sectional regressions of the NOTA vote share on a variety of constituency characteristics. We find evidence of systematic heterogeneity: the NOTA vote share is significantly 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. Of course, it is impossible to know whether these aggregate correlations also hold at the individual level (e.g.,

[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) around 20% of Mexican voters exiting the election booth say that they have little or no confidence that the vote they just cast "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%).

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.

whether illiterate voters are more likely to vote for NOTA) without making structural assumptions. Our analysis below will relate heterogeneity in voter choices to voter demographics and explicitly model the choices of different groups of voters.

## 5.2 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 $\gamma_{cs}$ are constituency fixed effects (in the panel) or state fixed effects (in the repeated cross section), and the $\eta_t$ are year fixed effects. The parameter of interest, $\alpha_1$ 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 2 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-year 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. 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.

In Appendix 3.2 we provide suggestive evidence to support a causal interpretation of the turnout

[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 Appendix 3.3.1, we also report p-values obtained by randomization inference.

Table 2: 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] |
| Basic controls | x | x | |
| Extended controls | | | 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) and (3) are for the extended dataset (repeated cross-section). Basic controls are 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. Extended controls also add a reserved constituency indicator 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.

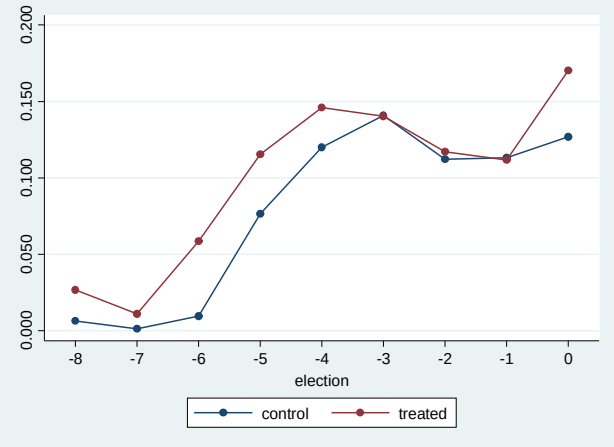
effect of NOTA. In particular, there do not appear to be large differences in turnout between the control and treatment states prior to the introduction of NOTA, either in levels or in trends (Figure 2). However, as we discuss in Appendix 3.2, institutional and political changes as well as data limitations require caution when comparing turnout figures over longer time periods.

### 5.3 Robustness

The number of available states and time periods impose fundamental limitations on our ability to control for all possible confounds in both the cross-section and time series. With this caveat, we now probe the robustness of the findings in several ways.

*National elections.* In the main specification we excluded the four states holding elections simultaneously with the national election in 2009 and 2014. In columns (1) and (2) of Table 3 we add these states to the sample, and include an interaction between an indicator for these states (National) and NOTA. An additional difficulty in these regressions is that a new state, Telangana, was carved out of Andhra Pradesh in 2014 and held its own elections. In columns (3) and (4) we repeat the regressions without Andhra Pradesh. In these regressions, the estimated turnout

Figure 2: The evolution of turnout in the control and treated states




*Notes:* Evolution of turnout controlling for state and year fixed effects and log(eligible voters). Election counts are for each state over the period 1967-2014 relative to the last election observed in the data. Election -9 is the excluded category. NOTA was introduced in treated states between election -1 and 0. See Appendix 3.2.3 for details.

effect of NOTA for the non-simultaneous election states is $2.5 - 3.5$ percentage points, which is similar to what we obtained in Table 2. The coefficients on the interaction terms indicate that the simultaneous election states behave differently from the others. A likely explanation for this difference is that turnout for these states measures a combination of voter turnout in the state and the national elections. Overall national election turnout in India increased in 2014 compared to 2009 - for the simultaneous states, this may confound the impact of the 2013 NOTA policy on state election turnout, and result in the generally positive interaction effects we observe.

Table 3: The impact of NOTA on turnout, with and without national elections

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| NOTA | 0.035** | 0.025 | 0.027** | 0.029* | 0.030* | 0.031* |
| | (0.016) | (0.017) | (0.013) | (0.016) | (0.015) | (0.015) |
| NOTA x National | 0.071** | 0.032 | 0.064** | -0.041 | | |
| | (0.031) | (0.041) | (0.031) | (0.035) | | |
| Basic controls | x | | x | | x | |
| Extended controls | | x | | x | | x |
| R2 | 0.18 | 0.20 | 0.18 | 0.19 | 0.19 | 0.21 |
| *N* | 7737 | 7737 | 7149 | 7149 | 5680 | 5680 |
| States | 29 | 29 | 28 | 28 | 22 | 22 |

*Notes:* Estimates of the effect of the NOTA policy on turnout using the repeated cross section. Columns (1) - (4) also include the states of Arunachal Pradesh, Odisha, and Sikkim, and columns (1) and (2) also include Andhra Pradesh. National is 1 for these states and 0 otherwise. Columns (5) and (6) exclude all elections held in 2009 and 2014. All regressions control for state and year fixed effects. Control variables are listed in Table 2. Standard errors clustered by state in parentheses. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

In columns (5) and (6) we ask what happens if we exclude from the sample *all* states holding elections in 2009/2014 (not just those that hold elections simultaneously with the national elections). States holding elections in a national election year could potentially also see an impact from national events in that year, like the wave of support for the BJP in the 2014 national elections. This has the potential to confound our estimates of the NOTA policy.[25] In fact, we find that our estimates remain very similar.

*Systematic Voters' Education and Electoral Participation (SVEEP).* In Appendix 3.3.2, we investigate whether SVEEP, a major voter education program undertaken by the Indian government, could confound our findings. The Election Commission launched SVEEP beginning with the December 2009 state election in Jharkhand, and the program was implemented in every subsequent state and national election. Given the timing of the program, its simple presence/absence is unlikely to confound our estimates, which rely on a policy change four years later, in 2013. This is confirmed by controlling for an indicator for the presence of SVEEP which leaves our results unchanged. A more challenging concern is that local features of the program could be correlated with the NOTA policy. We discuss two strategies to control for this in Appendix 3.3.2. One uses the length of states' SVEEP Action Plans as a proxy for the intensity of SVEEP activities. The other uses turnout levels in the previous national election - this is based on the fact that the Election Commission directed states to focus education efforts on areas with low turnout. We find that our estimates of the impact of NOTA are robust to controlling for SVEEP intensity using these proxies.

*Electoral redistricting.* Due to the timing of electoral redistricting in 2008, none of the states affected by NOTA were redistricted while most of the states in the control group were. To account for this, in Appendix 3.3.3 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. We find similar results to those above.

*State-specific events.* We searched news reports for state-specific events (such as terrorist attacks) that may act as a confound and dropped the affected states from the analysis. We also ran regression dropping each state, one at a time. These results, shown in Appendix 3.3.4, indicate

[25] Because split-ticket voting (constituencies voting for different parties at the state and national levels) is common in India, it is a priori not obvious that events affecting national turnout would affect non-simultaneous assembly elections. Note also that increased support for the BJP would presumably lead to more BJP votes rather than NOTA votes, so this is unlikely to explain the turnout effects from NOTA.

that our results are robust.

*Voting costs.* In Appendix 3.3.5, we include three additional controls for voting costs. We obtained information on the exact date of voting and control for elections held on a 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 Appendix 3.3.5 indicate that our estimates in Table 2 are robust.[26]

### 5.4 NOTA and political competition

In this section we explore some of the possible effects of NOTA on the number of candidates and the competitiveness of elections. A candidate may be dissuaded from entering the race if he believes that most people who would vote for him are likely to choose NOTA instead.[27] Given the size of the NOTA vote, this type of consideration is likely to be more relevant for candidates who have little chance of winning.

To explore the scope for such "supply" effects, in Table 4 we estimate regressions using various measures of the number of candidates as dependent variable. In columns (1) and (2), we use the total number of candidates in each constituency. The results indicate a statistically significant reduction of around 3 candidates in the average constituency (relative to a mean of 11). The next two columns study the number of effective candidates (the inverse of the Herfindahl index in each constituency, computed without the NOTA option). We again find significant negative effects, but these are much smaller than in the first two columns: here, NOTA is associated with a reduction of around 0.4 candidates relative to a mean of 3. The change between columns (1) - (2) and (3) - (4) is consistent with the idea that some of the small candidates may have decided to stay out of the race once NOTA was introduced.

In columns (5) and (6), we repeat the regressions focusing on candidates who are *not* independent or small-party candidates. To do this we aggregate all independent candidates in a constituency, and do the same for small-party candidates (we will use this aggregation in the struc-

[26] Appendix 3.3.7 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.

[27] Deadlines for nominating candidates are typically 2-3 weeks before the election. The first state to hold elections after NOTA (Chhattisgarh) had its nomination deadline 1 month after the NOTA judgement was published on 9/27/2013.

tural analysis below, see Appendix 4.2 for details). In these regressions, we find that the effect of NOTA on the number of candidates is statistically insignificant and small in magnitude (around -0.7 relative to a mean of 6). This confirms that to the extent that entry effects are present, they occur within the group of independent and small-party candidates. Aggregating these appears to be an effective way to control for supply effects on the number of candidates in this setting.[28]

Table 4: NOTA and the number of candidates

| Dep. Var: | Number of candidates | | Number of effective candidates | | Number of candidates after aggregation | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| NOTA | -2.959** | -2.890*** | -0.441*** | -0.331** | -0.664 | -0.765 |
| | (1.208) | (0.664) | (0.128) | (0.119) | (0.878) | (0.462) |
| Basic controls | x | | x | | x | |
| Extended controls | | x | | x | | x |
| $R^2$ | 0.07 | 0.13 | 0.05 | 0.06 | 0.25 | 0.34 |
| $N$ | 6685 | 6685 | 6685 | 6685 | 6685 | 6685 |
| States | 25 | 25 | 25 | 25 | 25 | 25 |

*Notes:* Estimates of the effect of the NOTA policy on the number of candidates using the repeated cross section sample. In columns (5) and (6) we aggregate all independent candidates in a constituency and do the same for all small-party candidates (parties running in less than 1/3rd of the constituencies in a state). All regressions control for state and year fixed effects. Control variables are listed in Table 2. Standard errors clustered by state in parentheses. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

While NOTA does not seem to cause changes in the number of viable candidates, it may still affect election closeness. In Appendix 3.4, we run regressions using various measures of the distance between the vote shares of winners and runner-ups. There is no evidence that NOTA affected the competitiveness of elections using these measures.

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

Our findings above provide evidence consistent with the existence of protest voters who normally abstain. What fraction of protest voters would normally abstain, and what fraction would vote for candidates? For the latter, which parties do they choose? Do protest voters matter for election outcomes? As highlighted in the Introduction, these questions are impossible to answer without

[28] Previous research shows that fewer independent candidates tends to have a negative effect on turnout (Kapoor and Magesan, 2018). Taken together, our findings in Tables 2 and 4 thus indicate that the positive direct effect of NOTA on turnout was large enough to offset any negative effects.

making additional assumptions. Not only do we need to estimate simultaneous outcomes in a multiparty system, we wish to conduct ecological inference about individual voters' behavior, and study protest voters' choices with and without the NOTA option *holding all else constant*. To make explicit the assumptions we rely on, we turn to a structural analysis.

We estimate a model of voter choice among candidates, NOTA and abstention, by adapting the BLP methodology for consumer demand estimation proposed by Berry et al. (1995). Although we are not the first to use this approach in the voting context, it is worth spelling out the factors that make this method particularly useful.

First, the rules governing elections imply that several assumptions of the model, which in IO applications have to be assumed, naturally hold in this case. In elections, voters are restricted to a discrete choice between voting for a candidate, voting for NOTA, or abstaining, and choices are made simultaneously by all voters in a given race.[29] Electoral competition takes place in markets (electoral districts) that are administratively defined, and voters are explicitly presented with a complete list of all available options on the ballot.[30] Complete administrative data is typically available on both vote shares and abstention.[31]

Second, the method addresses several estimation challenges that are central to any voting application - most importantly the ecological inference problem of inferring individual behavior from aggregate data. 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. Finally, the BLP method also naturally allows for multiparty elections, including partially contested races.[32]

[29] By contrast, when choosing which product to buy, consumers may purchase a mixture of products, and can substitute intertemporally.

[30] By contrast, studies of consumer choice have to rely on proxying the true market in which a set of products compete (e.g., based on geographic areas), and assume that consumers are aware of all available brands. Unlike in the voting context, defining the relevant "outside good" also requires assumptions.

[31] By contrast, most consumer studies are forced to rely on a *sample* of products and stores,

[32] 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 useful to study voter behavior in "characteristic space," i.e., projecting the heterogeneity between candidates onto a small

### 6.1 Specification: demand

Consider a constituency $c \in \{1, ..., C\}$ where voters can choose to vote for the candidates of parties $j = 1, ..., J$.[33] Each candidate is described by a set of characteristics observed by the researcher (such as party or caste) and a set of unobserved characteristics (such as the candidate's physical appearance or expected likelihood of winning). Assume 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. The term $\xi_{jc}$ captures voters' valuation of unobserved candidate characteristics, and $\varepsilon_{ijc}$ 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 $\xi_{jc}$. No explicit distribution for $\xi_{jc}$ 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 $\xi_{jc}$ for some or all candidates $j$. Voter expectations regarding policy platforms or clientelistic transfers, which we have no data on, are implicit in the specification (2). For example, promises regarding policies or transfers that are common across all candidates of a party are captured by the party fixed effects. Constituency-specific deviations from the average 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 $\xi_{jc}$.

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:

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

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.

[33] 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.

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 data on individual consumers is sometimes available, in the voting context, given the secrecy of the ballot, it is generally impossible to directly match voter characteristics to administrative data on choices. 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 $\boldsymbol{\varepsilon}_{ic} = (\varepsilon_{i1c}, ..., \varepsilon_{iJc})$.

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 $\varepsilon_{ijc}$, 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.[34] These predicted vote shares are a function of the data ($\mathbf{x}$), the parameters ($\boldsymbol{\theta}$), and the unobserved candidate characteristics $\xi_{jc}$. 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 allows controlling 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 $\xi_{jc}$ terms, and our estimation method corrects for this by using a cluster-robust procedure that allows for arbitrary correlation of $\xi_{jc}$ within a constituency. Second, strategic considerations are likely to be stronger in smaller constituencies. In the estimation we will weight observations with the number of eligible voters in a constituency,

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

thus giving lower weights to these constituencies.[35]

In principle, one could extend the model to explicitly include strategic considerations, but estimating such a model would be difficult as it would have to deal with voters' beliefs that their vote is pivotal. Kawai and Watanabe (2013) describe a procedure to do this in a model that does not allow for abstention and discuss the computational difficulties that would arise if they were to include voters' choice to abstain (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 3, 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 expected closeness of elections. As discussed in Section 5.4, the reduced form results show no indication that the NOTA policy affected election closeness.[36]

### 6.2 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

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

[36] 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.

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 $\xi_{jc}$, 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).[37]

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 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.[38]

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

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

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 (6) implies that the characteristics of a given party's can-

[37] For simplicity in this section we also assume that the $\xi_{jc}$ 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.

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

didates will be correlated across constituencies. As explained in section 6.3 below, this opens the possibility of using candidate characteristics in neighboring constituencies as instrumental variables in the estimation.

### 6.3 Estimation and identification

Estimation follows the Generalized Method of Moments (GMM) algorithm proposed by Berry et al. (1995). Detailed treatments of the procedure can be found in Berry et al. (1995) and Nevo (2000, 2001). Briefly, consider a dataset with information on candidate characteristics $\mathbf{x}_c$ and actual vote shares $S_{jc}$. Berry et al. (1995) show that, for given $\boldsymbol{\theta}_2$, it is possible to numerically solve for $\boldsymbol{\delta}_c$ from the equations $s_{jc}(\mathbf{x}, \boldsymbol{\delta}_c, \boldsymbol{\theta}_2) = S_{jc}$, 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 $\xi_{jc}(\boldsymbol{\theta}) = \delta_{jc}(\boldsymbol{\theta}_2) - \boldsymbol{\beta}\mathbf{x}_{jc}$, a nonlinear function of the model parameters. Identification relies on moment conditions $E[\xi_{jc}(\boldsymbol{\theta})|\mathbf{Z}_{jc}] = 0$ where the $\mathbf{Z}_{jc}$ are suitable instruments, and estimation is via GMM (see Appendix 4 for further details).

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 $\xi_{jc}$ 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)).[39] 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., Berry et al., 1995; Nevo, 2001). Based on the analogy between firms selling products and parties fielding candidates, 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.[40]

[39] 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 are characteristics (e.g., physical appearance) that we do not have data on, but that could be correlated with both $\mathbf{x}_{jc}$ and voters' valuation $\xi_{jc}$.

[40] What makes these instruments possible is the variation in the characteristics of a given party's candidates across

Beyond the analogy with 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 6.2. 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, 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 Appendix, we compare three sets of instruments constructed based on the logic above. In our preferred specification, we use the average gender and age of a party's candidates in all other constituencies in the state, interacted with state dummies; the average gender and age of all candidates in other constituencies in the state for both elections in the data (2008 and 2013); and finally interactions of the NOTA indicator with the constituency demographics used in the estimation (minority population, literacy rate, share of rural workers).[41]

The identifying assumption for estimation, expressed in the moment conditions, is that voters' valuation for a candidate's unobserved characteristics in a constituency ($\xi_{jc}$) is conditionally independent of the average (observed) characteristics of candidates in *other* constituencies (i.e., the instruments). In Appendix 4.3 we discuss several detailed examples where this assumption is likely or unlikely to hold. As we show below, our preferred specifications always pass the standard overidentification test, which increases our confidence in the validity of our identifying assumption.

Details of our implementation of the BLP algorithm are described in Appendix 4. To deal with the presence of many small parties and independent candidates, we aggregate these into one "small party" and one "independent" party, respectively (see Appendix 4.2). We include in the analysis the full list of candidate characteristics available in the data described in Section 4.2 above. We select the voter demographics to be included based on OLS regressions of the share of NOTA on

constituencies as each election is contested by a different set of individuals. This avoids the difficulties that arise in the consumer demand literature when there is insufficient variation in the characteristics of a firm's products across markets (Nevo, 2001).

[41] The NOTA interactions should help identify the corresponding preference heterogeneity parameters for NOTA (especially the parameters on the interaction of NOTA with the individual demographics $\mathbf{d}_i$).

demographics. We include minority population, literacy, and the share of rural workers as these indicate significant heterogeneity in voter preferences for NOTA (see the discussion in Section 5 and Appendix 3.1).[42] Our standard errors allow for both heteroskedasticity and correlation of the errors $\xi_{jc}$ across candidates within a constituency (see Appendix 4.1). As discussed in section 6.1, this is important if, e.g., the expected closeness of the race results in correlation between unobserved voter preferences for some of the candidates. We pay close attention to the numerical stability of our estimates through our choice of optimizer, convergence criterion, and by eliminating a source of instability in the typical codes used to implement the procedure (see Appendix 4.5).

# 7 Estimation results

We begin by modeling heterogeneity in voter preferences through Normally distributed random coefficients (i.e., setting $\mathbf{\Pi} = 0$, and capturing heterogeneity through $\mathbf{\Sigma}\mathbf{v}_i$ in equation (3) and $\sigma_0 v_i^0$ in equation (4)). Estimation results for these specifications (reported in Appendix 5.1) 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 fixed effects. Second, the random coefficients on female 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.

Next, we relax the assumption of Normally distributed random coefficients and instead model voter heterogeneity using the distribution of voter demographics from the Census ($\mathbf{\Sigma} = \mathbf{0}$, $\mathbf{\Pi} \neq \mathbf{0}$) (Appendix 5.2). 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 het-

[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.

erogeneity, as well as for NOTA. Estimates using slightly different instrument sets 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.[43] 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).

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 and West (1987) D-test always rejects the null that the nonlinear parameters included in the specifications are jointly 0.

We also 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 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.[44] 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.

We find 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. Voter preference for candidates with a criminal history is consistent with Vaishnav (2017) who provides an extensive documentation of this phenomenon in India. In

[43] The heterogeneity in preference for female candidates based on literacy is broadly consistent with the Chattopadhyay and Duflo (2004) finding that female politicians (at the village level) tend to come from more disadvantaged socio-economic backgrounds.

[44] 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).

particular, Vaishnav argues that voters take a candidate's criminal history "serves as a signal of their enhanced capacity and willingness to do whatever it takes to protect their supporters' interests." (Vaishnav, 2017, p168). Interestingly, Vaishnav argues that this is particularly the case for a criminal history involving *serious* crimes (typically crimes involving jail time). In Appendix 5.3, we show that our results are entirely consistent with this: voters' preference for criminal history appears to be driven by serious crimes.

In Appendix 5.4, 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.

# 8 Counterfactual simulations

## 8.1 The impact of NOTA

In this section we use the estimated model to evaluate the impact of NOTA. First, we perform a counterfactual experiment where the NOTA option is removed, holding everything else constant. 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.

Note that the estimation above focused on describing voter choices (the demand side) and used instruments derived from a stylized supply model that allows for endogenous candidate choices. In the counterfactual analysis, we study voter behavior holding all else (including any supply effects) constant. This is the relevant exercise for understanding the behavior of protest voters with or without NOTA; it is not designed to provide a full analysis of the impact of the NOTA policy. The latter would require a more detailed supply side model to study candidate responses.[45]

[45] Our findings from section 5.4 indicate some of the difficulties that would be involved in studying supply side responses. These effects may be present among small-party candidates - but estimating voter demand for these candidates individually (without aggregation) may not be possible.

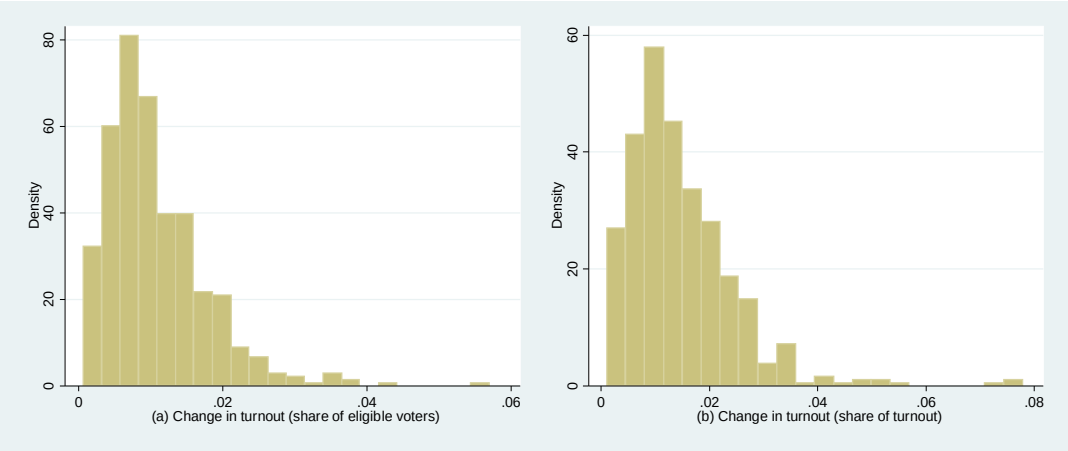
To evaluate the impact of NOTA on voter behavior, we restrict attention to those constituencies in our data that had the NOTA option available in 2013 (520 constituencies, comprising 101.4 million eligible voters, with a total of 3073 candidates). We compute new vote shares and turnout rates under the counterfactual no-NOTA 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. Throughout we focus on the preferred specification in column (2) of Table A.28, and present results obtained from additional specifications in Appendix 5.4.

### 8.1.1 Turnout vs candidate votes

The simulated impact of introducing the NOTA option on turnout is shown in panel (a) of Figure 3. 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 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 2. 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.

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 4 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

Figure 3: 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.

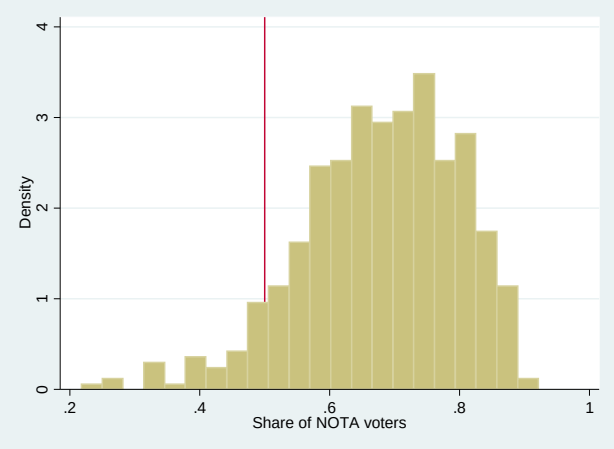
would abstain in the absence of NOTA. In the average constituency, 68% of NOTA voters would abstain in an election where everything else was held constant but the NOTA option was removed from the ballot.

Figure 5 shows the estimated impact of NOTA on individual candidates' vote shares, $\Delta s_{jc}$. Panel (a) contains all candidates. Among these, the average reduction in vote shares (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.[46]

The pattern on Figure 5 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

[46] The findings are inconsistent with the hypothesis that NOTA voters are indifferent between all candidates on the ballot and substitute to them randomly. In this case, NOTA voters should substitute to all candidates within a constituency to approximately the same extent. Pearson Chi-squared tests reject this null hypothesis for all constituencies but one.

Figure 4: 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.

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 5 below.)[47]

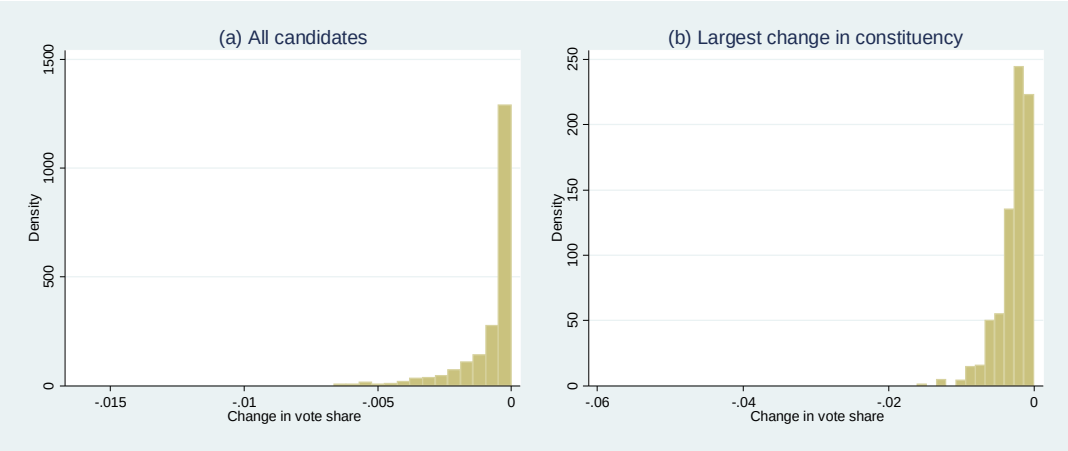
In Appendix 5.4, we repeat this counterfactual experiment with specifications that modify the instrument set, that include education, criminal history and assets as additional candidate characteristics, or that include small variations in the demographic interactions used in the model. We find similar results to those reported above: depending on the specification, the average increase in turnout is between 0.98 and 1.10 percentage points and 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 5 present, for each party, the number of candidates fielded (in the 520 constituencies), the number of con-

[47] The result that the NOTA option changed the winner in only 2/520 elections is not mechanical. 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 runner-up.

Figure 5: 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.

stituencies 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). Larger numbers in absolute value indicate that more voters substitute between that party and NOTA.

In column (4), we find that most NOTA voters who substitute to candidates vote for one of the two major parties (BJP or INC). One possible explanation is that these voters view choosing one major party as a powerful way to express their protest against the other. Consistent with this, we find that, across constituencies, the number of NOTA voters who substitute towards the BJP is positively correlated with the share of constituencies in the district won by the INC, the party in power at the national level throughout our period of study.[48]

Of course, since both the BJP and the INC have many voters (and many candidates), even gaining a sizeable share of NOTA voters causes little change in their overall political support. Which parties benefit *relatively* more from protest voters when NOTA is not available? Column (5) of Table 5 indicates that the four parties that owe the largest share of their support to protest voters 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

[48] The correlation between the number of NOTA voters choosing the BJP (INC) and the share of constituencies won by the INC (BJP) is 0.225 (0.015).

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.[49] 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."[50] The BYS's platform appears to be to "support the role of Indian youth."[51] Based on our counterfactual simulations, supporters of these fringe parties include relatively larger shares of protest voters.[52]

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 5 which show the counterfactuals 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 8). 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).

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

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

[51] www.byspindia.org

[52] In Appendix 5.4, 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 5: The impact of NOTA on parties

| | | 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 ($\Pi = \Sigma = 0$).

Table 6: The correlates of NOTA votes among simulated voters

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Literate | -0.029*** | 0.004*** | 0.005*** | 0.003*** | 0.005*** |
| | (0.008) | (0.000) | (0.000) | (0.000) | (0.000) |
| Rural worker | -0.009*** | -0.023*** | -0.026*** | -0.022*** | -0.026*** |
| | (0.003) | (0.001) | (0.001) | (0.001) | (0.001) |
| Minority | 0.015*** | -0.025*** | -0.027*** | -0.016*** | -0.018*** |
| | (0.003) | (0.001) | (0.001) | (0.001) | (0.001) |
| Reserved SC | 0.005*** | 0.004*** | | 0.005*** | |
| | (0.001) | (0.001) | | (0.001) | |
| Reserved ST | 0.012*** | 0.024*** | | 0.044*** | |
| | (0.001) | (0.001) | | (0.003) | |
| Size | -0.005 | -0.003 | | -0.002 | |
| | (0.004) | (0.004) | | (0.004) | |
| Minority x Reserved SC | | | | -0.004*** | -0.004** |
| | | | | (0.001) | (0.001) |
| Minority x Reserved ST | | | | -0.036*** | -0.042*** |
| | | | | (0.003) | (0.004) |
| State FE | x | x | | x | |
| Constituency FE | | | x | | x |
| $R^2$ | 0.56 | 0.45 | 0.51 | 0.51 | 0.60 |
| N | 520 | 520,000 | 520,000 | 520,000 | 520,000 |

*Notes:* Column (1) is a constituency level regression of the share of NOTA votes among all votes cast for the 520 constituencies affected by NOTA, using the simulated voters aggregated to the constituency level. Columns (2) - (5) are individual-level regressions of the probability of choosing NOTA for the simulated voters. Robust standard errors clustered by constituency in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively.

### 8.1.3 Heterogeneity among NOTA voters

We now explore what the estimated model implies about heterogeneity at the aggregate (constituency) and individual level. To do this, we use the 1000 voters simulated in each constituency and their predicted choices. Recall that, at the constituency level, the model matches vote shares perfectly. Thus, across constituencies, patterns predicted by the model only differ from the data due to the simulation error. Column (1) of Table 6 regresses the share of NOTA votes on constituency characteristics using the simulated choices aggregated to the constituency level. We see more NOTA votes in constituencies with fewer literate voters, more minorities (SC/ST), and fewer rural workers. In Column (2), we use the individual simulated voters and study the correlates of their probabilities of choosing the NOTA option. This regression reveals that NOTA is more likely to be chosen by voters who are *more* literate, are *not* SC/ST, and are not rural workers. In Column

(3) we show that replacing constituency characteristics with constituency fixed effects yields the same patterns.

Based on the predictions from our estimated model, the correlation between NOTA votes and literacy or minority status changes sign between the aggregate and the individual level. For example, this suggests that while less literate constituencies have more NOTA votes, it may be the more educated in these constituencies who choose this option. Similarly, while constituencies with larger shares of SC/ST have more NOTA votes, it may be the general caste voters who vote NOTA in these constituencies. These patterns highlight the ecological inference problem in this setting, and indicate that conclusions based on the aggregate data alone could be misleading.

To explore the above patterns further, columns (4) and (5) include interactions between constituency reservation and individual caste. We find that, according to the model, NOTA votes are especially likely among general caste voters in reserved constituencies. For example, in column (4), the probability of voting NOTA among general caste voters is 0.5 percentage points higher in SC-reserved constituencies and 4.4 percentage points higher in ST-reserved constituencies. The corresponding figures for minority voters are 0.1 and 0.8 percentage points ($0.005 - 0.004 = 0.001$ and $0.044 - 0.036 = 0.008$). This may indicate that general caste voters in reserved constituencies perceive a lack of representation and express this by choosing NOTA.

In Appendix 5.5, we also explore heterogeneity in NOTA voters' substitution patterns (from abstention or from major/non-major parties). We find that, without a NOTA option, general caste NOTA voters in reserved constituencies are particularly likely to abstain. By contrast, minority NOTA voters in these constituencies are relatively more likely to choose one of the candidates.

## 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 any fines 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 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.[53]

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.[54] We predict voter behavior in the compulsory system and study the resulting vote shares.

The results are summarized in Table 7, which has a similar structure to Table 5 above. Appendix 5.6 provides further details. When NOTA is available (Panel A), compulsory voting increases the share of NOTA votes among eligible voters by a factor of 5 (6.9 percentage points). Over 1/3 of the voters "forced" to turn out by compulsory voting ($1.578 + 6.928 = 8.506\%$ 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 loses a few seats.[55]

The impact of compulsory voting without NOTA is shown in Panel B. In this case, all new voters are forced to vote for actual candidates. In terms of relative popularity, the biggest winners from compulsory voting are now the BYS and the JGP, which more than double their number of

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

[54] 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.

[55] 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 Appendix 5.6).

Table 7: The impact of compulsory voting with and without NOTA

| | 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) |
| A. With NOTA | | | | | | | |
| 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 |
| B. Without NOTA | | | | | | | |
| 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 | - | - | - | - | - | - | - |
| 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:* Panels A and B show the impact of introducing compulsory voting when NOTA is or is not available, respectively. For each party, column (1) shows the total number of candidates, (2) the number of constituencies won, and (3) the share of all votes received (out of 101.384 million eligible voters) when voting is voluntary (in Panel A, this is the observed 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.

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 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 estimate the behavior of NOTA voters in the absence of NOTA to study protest voting. To infer individual choices from aggregate data in a multiparty system holding everything else constant, we use a structural model and techniques borrowed from the consumer demand literature.

We find that in the absence of NOTA, most protest voters abstain, while those who turn out scatter their votes across many candidates. As a result, protest voting (and therefore the NOTA policy) affects turnout, but has little impact on the outcome of elections. In counterfactual simulations, we also show that NOTA can significantly attenuate the impact of compulsory voting on election outcomes.

Our results 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.

One implication of our findings is that NOTA may be a low-cost way to increase voter participation, fulfilling the Supreme Court's stated goal of increased turnout at the cost of a few dollars for every 1 million eligible voters.[56]

What do these results imply about voters' motivations more broadly? Our findings are most naturally interpreted based on the Downsian view where, beyond the instrumental benefit from being

[56] Appendix 6 provides a back-of-the-envelope cost-benefit analysis and further discussion.

pivotal, voters also derive consumption ("expressive" or "non-consequentialist") utility. Empirically identifying these consumption motives for voting in real-world settings is challenging (Spenkuch, 2018; Pons and Tricaud, 2018). It is particularly difficult to separate a general utility obtained from turning out from an option-specific utility from choosing a particular option or candidate. In our setting, any general utility from turning out is available to voters both with and without NOTA on the ballot, while the option-specific utility from choosing NOTA is only available in the first case. Therefore, as we show formally in Appendix 7, our result that most voters substitute from abstention to NOTA indicate that these voters derive an option-specific consumption utility from voting.

Another possible interpretation of NOTA votes is in terms of models where voters send signals to politicians or other voters through their behavior at the polls (Lohmann, 1993; Piketty, 2000). Theoretically, there is typically more than one way to send the same signal. Voters are assumed to 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, which action serves as a signal in a particular setting is an empirical question. From this perspective, our results indicate that in the absence of NOTA, most voters appear to use abstention rather than a vote for one of the candidates to signal their protest.

Given that India is the largest democracy in the world with roughly 800 million registered voters, understanding this context is clearly important in its own right. At the same time, a natural question is to what extent the findings here might translate into other settings. One relevant consideration in generalizing the findings is the fact that Indian turnout is relatively high. Our results imply that introducing a pure protest option can increase political participation even in a setting where abstention was already low. At the same time, we found that even in an environment where many voters were already voting for candidates, few of these choices were made purely to express protest.

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

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

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

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

[10] 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.

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

[12] 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.

[13] Debnath, S., M. Kapoor, and S. Ravi (2018): "The Impact of Electronic Voting Machines on Electoral Frauds, Democracy, and Development," working paper.

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

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

[16] 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.

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

[18] 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.

[19] 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.

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

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

[22] 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.

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

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

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

[26] Katz, J.N., and G. King (1999): "A Statistical Model for Multiparty Electoral Data," *American Political Science Review* 93(1), 15-32.

[27] King, G. (1997): *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton University Press, Princeton, NJ.

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

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

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

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

[32] 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.

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

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

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

[36] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[51] Spenkuch, J. L (2018): "Expressive vs. strategic voters: An empirical assessment," *Journal of Public Economics* 165, 73-81.

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

[53] Tomz, M., J.A. Tucker, and J. Wittenberg (2002) : "An Easy and Accurate Regression Model for Multiparty Electoral Data," *Political Analysis* 10(1), 66-83.

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

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