Online Appendix to "None Of The Above: Protest Voting in the World's Largest Democracy"

Gergely Ujhelyi, Somdeep Chatterjee, and Andrea Szabó

July 3, 2020

Contents

1 Background

1.1 Voting machines and the NOTA policy

The Indian electoral rules (The Conduct of Election Rules, 1961 - CER)[1] explicitly recognize a voter's right to participate in the election by intentionally casting an invalid vote. The section of the CER relevant to electronic voting machines is known as Rule 49-O. According to this rule, a voter wishing to cast an invalid vote must give up the secrecy of the ballot: she must inform the election officer, who will record the non-vote along with the voter's signature or thumbprint. It is important to note that, with the type of voting machines used in India, other forms of intentionally casting an invalid vote (such as punching in an invalid candidate name or code) were not available. These voting machines simply involve pushing a button next to a candidate (see Figure A.1); in order for a vote to be recorded, one of the buttons must be pushed. Indeed, the government of India views the elimination of invalid voting as one of the salient features of electronic voting machines.[2]

In 2004, the People's Union for Civil Liberties (PUCL) challenged this process, arguing that asking voters to give up secrecy of the ballot when exercising their right to cast an invalid vote was unconstitutional. In its 2013 decision the Supreme Court agreed, declared Rule 49-O as unconstitutional, and mandated the introduction of the NOTA button on the voting machines.[3]

It is interesting to note that, with the paper ballots used before the voting machines, the CER also required a voter wishing to cast an invalid vote to give up anonymity. According to these rules, known as Rule 41(2)&(3), here the voter would have to return the invalid ballot personally to the election officer, thus revealing her identity. However, in practice, a voter could also leave the ballot blank and simply drop it in the ballot box. Thus, with paper ballots, anonymous invalid voting was still possible in practice. This difference relative to voting machines explains why these rules did not come under constitutional scrutiny until 2004, when all elections were conducted using voting machines.

1.2 NOTA-like options in other countries

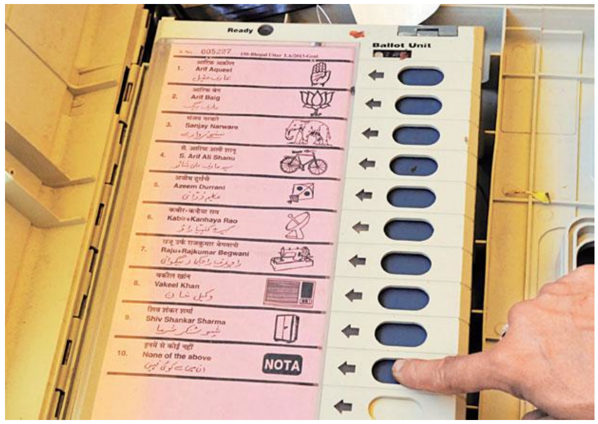
NOTA is an explicit option on the voting machine, and this makes it fundamentally different from simply casting an invalid vote as can be done in many countries. In the case of the latter,

[1] http://legislative.gov.in/sites/default/files/%282%29%20THE%20CONDUCT%20OF%20ELECTION %20RULES%2C%201961.pdf

[2] https://pib.gov.in/newsite/mbErel.aspx?relid=104463

[3] The Supreme Court's judgement in PUCL v Union of India is available at http://judis.nic.in/supremecourt/imgs1.aspx?filename=40835. For a summary of the case, see http://lawtimesjournal.in/peoples-union-for-civil-liberties-vs-union-of-india-anr-nota-case/

Figure A.1: Electronic voting machine with NOTA option



it is typically impossible to know whether such votes occur intentionally or by mistake, hence it is difficult to use them to draw conclusions regarding voters' intentional behavior (see, e.g., McAllister and Makkai, 1993; Herron and Sekhon, 2005; Power and Garand, 2007; Uggla, 2008; Driscoll and Nelson, 2014). For some applications, the fact that invalid votes also include voting mistakes will simply add measurement error to the "true" measure intended to capture negative votes. In other cases, however, this will have an important impact on the interpretation of the results. For example, more invalid votes among the less educated can mean either that these voters are more likely to make mistakes when filling out the ballot, or that they are particularly dissatisfied and intentionally cast invalid votes to express this.

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

Another feature that makes the Indian NOTA useful for the analysis of voters' motivations is the lack of electoral impact of the NOTA vote. Recall that NOTA vote can never "win." In addition, due to the first-past-the-post system, it has no impact on the allocation of legislative seats. By contrast in Colombia if the "blank vote" wins, new elections must be called with the rejected candidates prohibited from running again. In Spain, while the blank vote can never win, seats are allocated in a proportional system and a minimum 3% threshold must be reached for a party to enter parliament. Both of these systems could give voters an incentive to choose the blank vote strategically in order to affect the mix of candidates elected for office in the current election. Such incentives are not present in the Indian system.[4]

2 Data

2.1 Summary statistics

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

Table A.1: Summary statistics of the panel dataset

Variable	Obs	Mean	Std. Dev.	10%	90%
A. Candidate characteristics					
Female	9831	0.073	0.245	0.000	0.125
Age / 100	9831	0.433	0.145	0.300	0.600
Minority	9831	0.381	0.470	0.000	1.000
Ran in previous election	9831	0.619	0.443	0.000	1.000
Won in previous election	9831	0.142	0.253	0.000	0.571
Education	8989	0.588	0.457	0.000	1.000
Criminal history	8989	0.144	0.333	0.000	1.000
Assets (log/10)	9384	1.368	0.418	1.023	1.730
NOTA (0/1)	9831	0.053			
Broadcast allowance (100 minutes)	9831	0.673	0.773	0.000	1.800
INC (0/1)	9831	0.146			
BJP (0/1)	9831	0.144			
Independent (0/1)	9831	0.138			
Small party (0/1)	9831	0.120			
B. Constituency characteristics					
Eligible voters (1000)	1446	180.797	36.173	148.363	216.153
Turnout	1446	0.720	0.081	0.611	0.816
NOTA votes / total votes	520	0.021	0.013	0.007	0.037
NOTA votes / eligible voters	520	0.016	0.009	0.005	0.028
N. of candidates (before aggregation)	1446	11.370	4.909	6	17
N. of candidates (after aggregation)	1446	6.439	1.403	5	8
Reserved constituency (0/1)	1446	0.335			
Rainfall (cm/day)	1446	0.066	0.387	0.000	0.011
Minority population (%)	1446	0.353	0.182	0.186	0.617
Literate population (%)	1446	0.584	0.092	0.474	0.693
Rural workers (%)	1446	0.662	0.174	0.442	0.845

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

Table A.2: Summary statistics of constituencies in the extended dataset (repeated cross-section)

Variable	Obs	Mean	Std. Dev.	10%	90%
Eligible voters (1000)	6685	180.754	88.232	41.203	292.898
Turnout	6685	0.707	0.129	0.533	0.866
NOTA votes / total votes	1176	0.015	0.012	0.004	0.030
NOTA votes / eligible voters	1176	0.010	0.009	0.003	0.022
Reserved constituency (0/1)	6685	0.276			
Labor force participation	6685	0.572	0.065	0.503	0.665
Unemployment rate	6685	0.032	0.033	0.011	0.050
Household earnings (real Rp/week)	6685	1553.309	540.820	936.419	2135.973
Fraction illiterate	6685	0.301	0.118	0.116	0.479
Fraction primary school or less	6685	0.227	0.071	0.150	0.325
Sex ratio (females / 1000 males)	6685	986.086	63.851	928.189	1084.414
Fraction urban	6685	0.313	0.137	0.199	0.447
State NDP growth rate	6685	5.784	3.720	1.597	11.284
Election on weekend (0/1)	6685	0.258			
Rainfall (cm/day)	6684	0.083	0.216	0.000	0.194
Polling station density	6676	0.001	0.000	0.001	0.001
Redistricting (max overlap)	6173	0.899	0.187	0.553	1.000
Redistricting (fractionalization)	6173	0.865	0.239	0.421	1.000

Notes: The repeated cross-section contains all assembly elections between 2006 and 2014 in 25 states. Turnout is total votes divided by the number of eligible voters. Polling station density is the number of voting stations per eligible voter. Redistricting (max overlap) is the largest area of a current constituency that was part of a single constituency before delimitation; (fractionalization) is a measure of territorial fractionalization as a result of redistricting. The construction of these measures is described in Section 3.3.3.

2.2 Rainfall and Broadcast allowance

The rainfall variable is created based on gridded daily rainfall data obtained from the India Meteorological Department in 0.25×0.25 degree cells. We match this grid to constituency boundaries and take the area-weighted average of the cells covering each administrative area on the relevant day (different constituencies within the same state typically go to the polls in groups over a period of 2-3 days). Figure A.2 illustrates the size of the rainfall grid relative to the constituencies.

Source: India Meteorological Department: New High Spatial Resolution (0.25X0.25 degree) Long Period (1901-2015) Daily Gridded Rainfall Data Set Over India (CD-ROM).

The Broadcast Allowance is total time allotted in minutes for Broadcast and Telecast in an election cycle. Political parties are provided free access to State owned Television and Radio for an allotted amount time. A base time is given to each National Party and Recognised State Party (recognized in the State) uniformly. Additional time is allotted to the parties on the basis of the poll performance of the parties in the last Lok Sabha and State Assembly election.

Source: Election Commission of India, http://eci.nic.in/eci_main1/Press_Release2013.aspx, http://eci.nic.in/eci_main1/press_release2008.aspx

Figure A.2: Example of daily rainfall grid and constituency boundaries (November 29, 2008)

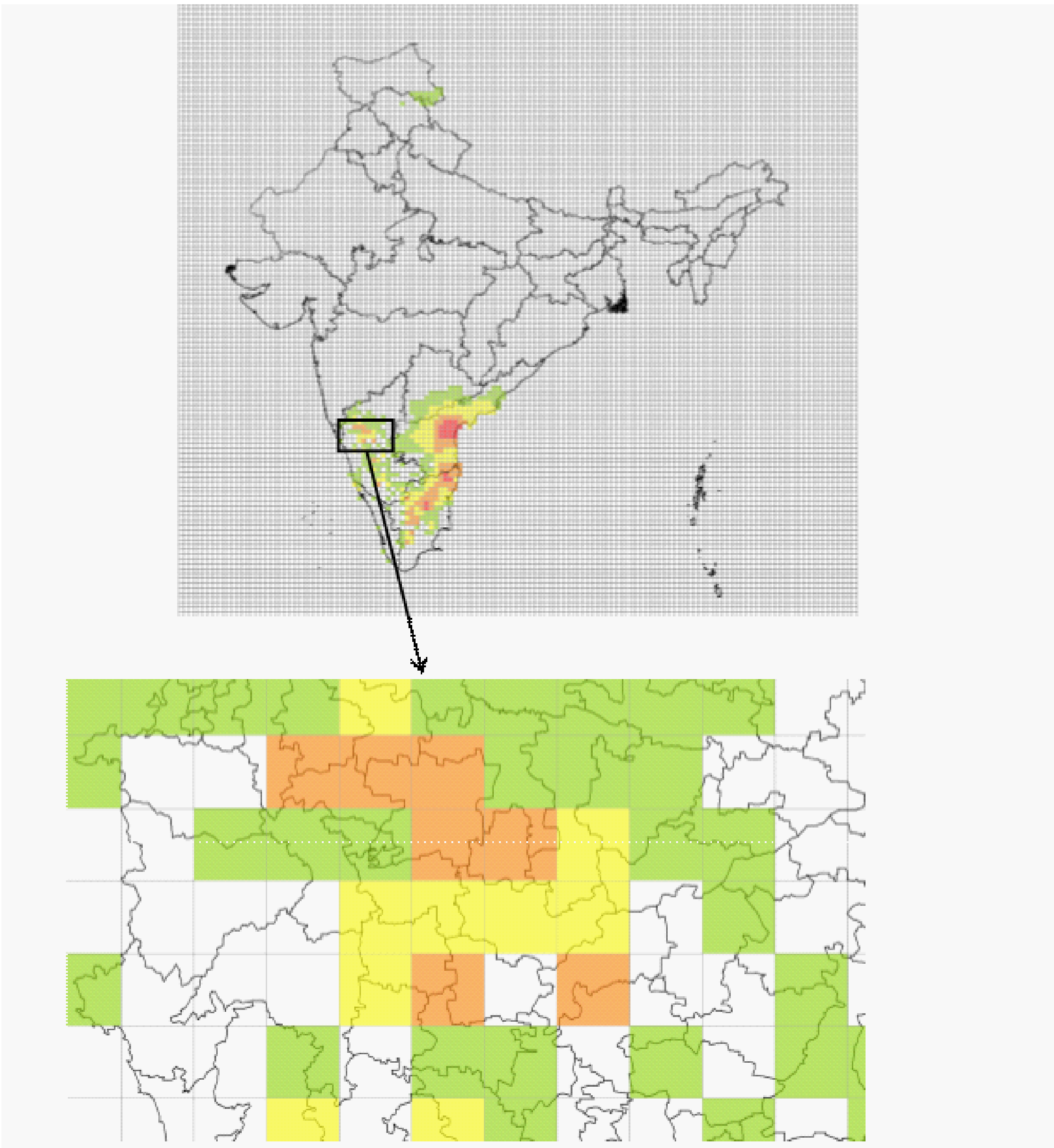


Table A.3: Detailed results of the matching procedure

State	Karnataka	Chhattisgarh	Madhya Pradesh	Mizoram	Rajasthan
Step 1:					
Villages in 2001 census (a)	27,783	19,860	52,548	729	39,993
Duplicate village names 2001 (b)	1,387	1,634	3,513	0	1,150
Share (b/a)	5%	8%	7%	0%	3%
Villages 2001 without duplicates (a-b)	26,396	18,226	49,035	729	38,843
Villages in 2011 census (c)	27,710	19,637	52,352	726	43,497
Duplicate village names 2011 (d)	1,434	1,030	2,720	0	1,423
Share (d/c)	5%	5%	5%	0%	3%
Villages 2011 without duplicates (c-d)	26,276	18,607	49,632	726	42,074
Step 2: Simple matches (e)	25,676	11,791	38,943	593	34,049
Step 3: Matched without sub-district name (f)	112	5,175	8,911	73	428
Step 4: Matched with alternative spelling (g)	15	382	324	0	565
Total matched (e+f+g)	25,803	17,348	48,178	666	35,042
Not matched, share of 2001 ((a-b-e-f-g)/(a-b))	2%	4%	2%	9%	10%

Notes: Results of Steps 1-4 of the procedure for matching the 2001 and the 2011 Census using village names. See the text for descriptions of each step.

2.3 GIS matching of Census data to electoral data at the constituency level

GIS matching of the Census and electoral data is necessary because in India the Census areas and the constituencies do not coincide. Boundary files for the 2013 electoral constituencies are publicly available. In order to match the electoral data to the most recent (2011) Census data, we need to overcome the difficulty that the 2011 Census boundary files are not publicly available. We do this using boundary files from the previous (2001) Census. We first match villages in the 2011 and the 2001 Census using village names. Next, we match the 2001 sub-districts to each 2013 electoral constituency using GIS boundary files.[5] Details of the matching are described below.

I. Matching villages in the 2001 and the 2011 census. Administrative boundaries in India change over time, with sub-districts, districts, and even states splitting up into new units. Our matching procedure is based on the smallest administrative unit available in the Census, the village. To match village names in the 2001 and 2011 census, we proceeded through the following steps. The detailed results for each step are described in Table A.3.

1. Eliminate duplicate village names in every sub-district in both the 2001 and the 2011 dataset. Across the 5 states, this results in 5.5% of the villages being dropped in 2001 and 4.6% in 2011.

2. Match the two datasets by (state, district name, sub-district name, village name). One state, Rajasthan, had a new district created in 2011 (Pratapgarh) which was carved out from 3 other districts (Chittaurgarh, Udaipur, and Banswara). For this state, we repeated this step three times, replacing the new district name with each of the three parent districts.

3. For the villages not yet matched, repeat the match by (state, district name, village name). This results in additional matches, reflecting changes in the boundaries of sub-districts within districts

4. For the villages not yet matched, allow for variations in spelling. Specifically, for villages not yet matched we repeat the match by (state, district name, village name), allowing for the following variations in both the 2001 and 2011 datasets:[6]

(i) Double letters (e.g., two r instead of one) for each letter in a village name.

(ii) One of the following extra letters anywhere in the village name: a, h, e, n, i; or an extra u after o.

(iii) A one-letter change in the village name: a to e, r to d, t to r, h to n, d to g, n to g,

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

[6] We established these rules by running the match and inspecting the unmatched names, and then including any reasonable new match as a new rule.

Table A.4: Comparing matched and unmatched constituencies

	Matched	Unmatched	p-value for equality
Number of eligible voters (1000)	167.393	121.174	0.44
	(6.881)	(68.307)	
Turnout	0.690	0.659	0.73
	(0.010)	(0.109)	
Election closeness	0.095	0.112	0.40
	(0.003)	(0.025)	
Reserved constituency	0.335	0.459	0.65
	(0.039)	(0.313)	
N	723	61	
States	5	4	

Notes: 2008 characteristics of the constituencies matched to the Census data and those that are lost during the matching, in the states of Chhattisgarh, Karnataka, Madhya Pradesh, Mizoram, and Rajasthan. The p-value for the equality of means test in the last column is from OLS regressions of each variable on a "matched" indicator, computed using a bootstrap clustered by state.

o to u.

These resulted in a small number of additional matches (see Table A.3).

II. Matching 2001 sub-districts to electoral constituencies. Of the 854 constituencies that were not redistricted and held elections in both 2008 and 2013, we have constituency boundary files for 850. The 2001 Census boundary files allowed us to match 723 of these to sub-districts in the Census. Delhi is responsible for most of the attrition during the matching: we lose all 70 constituencies in this state. In the 5 remaining states, we lose 21 constituencies in Karnataka, 12 in Madhya Pradesh, 27 in Mizoram, and 1 in Rajasthan (no constituencies are lost in Chhattisgarh). Table A.4 compares the electoral characteristics of constituencies in these states that were successfully matched to those that were not and shows no statistically significant differences. As shown below, our estimates and counterfactual results are robust to dropping Mizoram altogether.

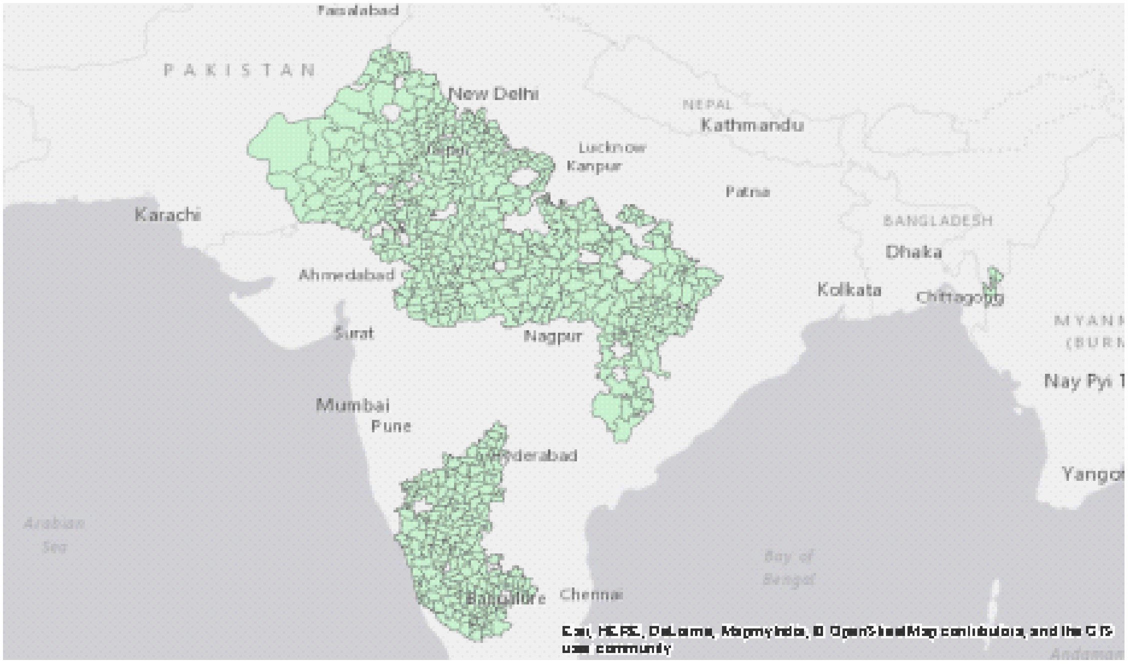
Of the 723 matched constituencies, 71.9% (520) are affected by NOTA in 2013. In the full 854-constituency panel the corresponding figure is 73.8% (630 constituencies). The location of the matched constituencies is shown on Figure A.3.

3 Patterns in the data

3.1 The correlates of NOTA votes

In this section we investigate the correlation between NOTA votes and constituency characteristics. We use the same dataset as in the structural analysis and run simple cross-sectional

Figure A.3: Constituencies in the merged dataset



regressions on the 520 constituencies that are affected by the NOTA policy in 2013. We include state fixed effects and, to avoid confounding our estimates by differential turnout across constituencies, we measure NOTA vote shares as a fraction of total votes cast.[7]

The results in Table A.5 indicate substantial heterogeneity in NOTA votes across constituencies. For example, the NOTA vote share is significantly higher in reserved constituencies and in constituencies with more illiterate voters, more women, more ST, and a lower share of rural workers. Each of these patterns is consistent with a variety of possible explanations. One possible interpretation is that NOTA votes are higher in more economically disadvantaged constituencies, reflecting a general dissatisfaction with elected leaders in these constituencies. Note however that the coefficients remain unchanged if we add controls for various indicators of infrastructure and economic activity in column (2). Another possible interpretation is that NOTA votes come from politically underrepresented voters, such as women, and non-SC or ST voters in reserved constituencies. Without further analysis it is impossible to know what these constituency-level correlations imply about the determinants of *individual* choices.

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

Table A.5: The correlates of NOTA votes

	(1)	(2)	(3)	(4)
Constituency characteristics:				
Literacy	-0.034***	-0.035***	-0.029***	-0.026**
	(0.008)	(0.010)	(0.008)	(0.010)
Percent rural worker	-0.012***	-0.019***	-0.013***	-0.015***
	(0.003)	(0.005)	(0.003)	(0.005)
Percent SC	0.014*	0.013	0.006	0.008
	(0.008)	(0.009)	(0.007)	(0.008)
Percent ST	0.011***	0.007*	0.010***	0.006*
	(0.003)	(0.004)	(0.003)	(0.004)
Percent male	-0.246***	-0.228***	-0.164***	-0.189***
	(0.031)	(0.046)	(0.033)	(0.044)
Reserved SC	0.005***	0.005***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Reserved ST	0.011***	0.011***	0.008***	0.008***
	(0.002)	(0.001)	(0.002)	(0.002)
Size	-0.008**	-0.008**	-0.005	-0.005
	(0.003)	(0.004)	(0.003)	(0.004)
No latrine		0.001		0.004
		(0.004)		(0.004)
Water nearby		0.016**		0.014**
		(0.006)		(0.006)
Water at home		0.010*		0.013**
		(0.006)		(0.005)
Percent employed		0.015		-0.004
		(0.011)		(0.011)
Car ownership		0.022		0.019
		(0.037)		(0.034)
Computer ownership		-0.030		0.034
		(0.057)		(0.052)
Phone ownership		-0.009		-0.010*
		(0.006)		(0.005)
TV ownership		-0.003		-0.007
		(0.006)		(0.006)
Candidate characteristics:				
Number of candidates			-0.001***	-0.001***
			(0.000)	(0.000)
No female			-0.001	-0.001
			(0.001)	(0.001)
<15% female			-0.000	-0.000
			(0.001)	(0.001)
Median age			-0.000	-0.000
			(0.000)	(0.000)
No SC			0.000	-0.000
			(0.001)	(0.001)
<15% SC			0.000	0.000
			(0.001)	(0.001)
No ST			-0.001	-0.000
			(0.001)	(0.001)
<10% ST			0.001	0.001
			(0.001)	(0.002)
R^2	0.57	0.60	0.63	0.66
N	520	520	520	520

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

In columns (3) and (4) we add candidate characteristics to the regression. We find that constituencies with more candidates running have lower NOTA vote shares, which is consistent with NOTA reflecting dissatisfaction with the menu of candidates being offered. We do not find evidence that the presence of female, SC or ST candidates is correlated with NOTA votes.

In Table A.6 we explore the correlation between caste and NOTA votes further by including interactions of reservation status and the share of minority (SC or ST) population. We find that a larger SC/ST population is correlated with more NOTA votes, and particularly so in reserved constituencies. This pattern, too, has several possible interpretations. As we show below, our estimated model suggests that it is the non-SC/ST voters who are more likely to vote NOTA in the reserved constituencies.

Table A.6: The correlates of NOTA votes: NOTA and caste

	(1)	(2)	(3)
Literacy	-0.030***	-0.031***	-0.027***
	(0.008)	(0.010)	(0.008)
Percent rural worker	-0.012***	-0.019***	-0.013***
	(0.003)	(0.005)	(0.003)
Percent SC/ST	0.010**	0.006	0.004
	(0.004)	(0.004)	(0.004)
Percent male	-0.246***	-0.223***	-0.172***
	(0.028)	(0.042)	(0.030)
Reserved	0.002	0.002	-0.001
	(0.002)	(0.002)	(0.002)
Reserved x Percent SC/ST	0.013**	0.013**	0.014***
	(0.006)	(0.006)	(0.005)
Size	-0.008**	-0.008**	-0.005
	(0.004)	(0.004)	(0.003)
Economic indicators		x	
Candidate characteristics			x
R^2	0.58	0.59	0.65
N	520	520	520

Notes: The dependent variable is the share of NOTA votes among all votes cast. Regressions at the constituency level for the cross-section of constituencies affected by the NOTA policy in 2013 in the panel dataset. All regressions include state fixed effects. Economic indicators and candidate characteristics refer to the additional variables included in Table A.5. Robust standard errors in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively.

3.2 Comparison of states before NOTA

3.2.1 Trends

For the reduced-form analysis to identify the causal effect of NOTA, turnout in the treatment and the control states must have parallel trends. The usual way to provide suggestive evidence on this assumption is to compare trends and levels in the two groups before the policy.

Data limitations and the changing nature of Indian politics over time make it difficult to provide a fully convincing analysis of pre-trends.[8] To provide some suggestive comparisons, we obtained historical state-level turnout data for all states in the sample. On Figure A.4, we first plot average turnout separately for the control and the treatment states, both for the 25 states in the extended dataset and the 5-state panel, from 1990 until the introduction of NOTA in 2013. In the raw data there do not seem to obvious differences in trends before the introduction of NOTA in the two groups of states.

Next, we control for state and year fixed effects as well as the log number of eligible voters. Since elections take place in different years, we use elections as our time unit for the purpose of visualizing trends. Election 0 corresponds to the last election observed in the sample (these are the elections after NOTA for the treated states), election -1 is the previous election for each state, -2 the one before that, and so on. Elections -9 to 0 take place in years 1967-2014. The average year of each election if shown in Table A.7 separately for the control and treatment group.

Table A.7: Average year of elections -9,...,0 in treatment and control states

Election	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
Control	1971.9	1975.8	1979.9	1984.1	1988.5	1992.7	1997.4	2002.0	2006.9	2011.9
Treatment	1975.0	1977.6	1980.5	1984.1	1988.6	1993.8	1998.1	2003.4	2008.3	2013.4

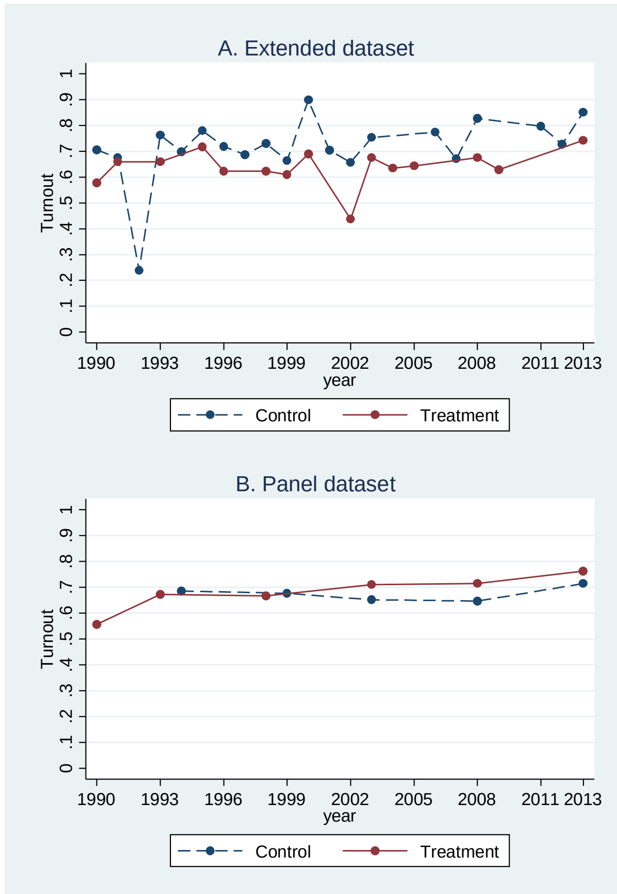
We run the regression

$$Turnout_{st} = \sum_{j=-8}^{0} (\lambda^j Election^j_{st} \times NotaState_s + \theta^j Election^j_{st}) + \rho \log(Voters_{st}) + \gamma_s + \eta_t + \varepsilon_{st}, \tag{1}$$

where $Turnout_{st}$ is turnout in state s in year t if the state had an election in t, $Election^j_{st}$ is 1 for election j, with $j = -9$ serving as the base category, $NotaState_s$ is 1 for states in the

[8] Difficulties in comparing turnout figures over time stem from changing political boundaries (redistricting as well as new states created in 1972, 1987, and 2000), the gradual introduction of electronic voting in 1999-2004, and periods of political instability (for example, the period 1989-1999 had 5 national elections as successive governments collapsed).

Figure A.4: The evolution of turnout in the control and treatment states before NOTA



Notes: Average turnout in control states and treatment states up until the introduction of NOTA in 2013. Panel A is the extended dataset with 9 treatment and 16 control states. Panel B is the 5-state panel with 4 treatment and 1 control state.

treatment group (states which have NOTA in election 0), $Voters_{st}$ is the number of eligible voters, and γ_s and η_t are state and year fixed effects. We estimate this regression for the 25 states in the extended sample. We have 8.4 elections for the average state (states created in 2000 have 3 elections), for a total of 210 observations. The estimated evolution of turnout for control states (θ^j) and treatment states ($\theta^j + \lambda^j$) is shown on Figure 2 in the paper. Except for two breaks, in elections -7 (around 1980) and -4 (around 1993), the evolution of turnout prior to NOTA is very similar in the two groups.

3.2.2 Levels

Turning to levels, Table A.8 and A.9 compare treatment and control states before the introduction of NOTA. Table A.8 presents summary statistics separately for each state in the panel dataset. Karnataka is the control state unaffected by NOTA, the other states are treatment states, where NOTA was available in 2013 but not in 2008. Summary statistics are for 2008 (before NOTA). In the last column, comparing the treated states to the control state does not reveal large differences before the introduction of NOTA: we find 1 significant p-value out of 13 variables. (In the structural exercise below, we will repeat the analysis excluding the state of Mizoram, which does appear to be an outlier on several dimensions.)

Table A.9 compares control and treatment states in the extended dataset. For each state, values are for the first election in the sample (which took place before the introduction of NOTA for every state). We again do not see a large difference between the two groups: 1 significant p-value at 5% and an additional 2 at 10%. None of the electoral variables are significantly different.

3.2.3 States in the panel and the extended dataset

Table A.10 investigates to what extent the 5 states in the panel are representative of the extended dataset. The table compares these 5 states to the other 20 states in the extended dataset. As above, for each state, values are for the first election in the sample (which took place before the introduction of NOTA). The two groups show few differences: 2 significant p-values at 1%. None of the electoral variables are significantly different.[9]

[9] States in the panel have more reserved constituencies, with a p-value of 12%. This is due in part to Mizoram, where all constituencies are reserved. Without Mizoram the p-value of the difference increases to 20%.

Table A.8: Characteristics of states in the panel dataset

	Control	Treatment					p-value
	Karnataka	Chhattisgarh	Madhya Pradesh	Mizoram	Rajasthan	All	for equality
Constituency characteristics							
Number of eligible voters (1000)	174.993	169.080	156.154	15.148	181.134	164.426	0.75
Turnout	0.683	0.706	0.705	0.830	0.666	0.693	0.64
Election closeness	0.100	0.089	0.099	0.065	0.089	0.093	0.17
Reserved constituency	0.241	0.433	0.376	1.000	0.296	0.371	0.18
State characteristics							
Number of constituencies	203	90	218	13	199	130.000	0.27
Labor force participation	0.661	0.746	0.648	0.678	0.601	0.668	0.85
Unemployment rate	0.013	0.007	0.011	0.025	0.019	0.015	0.57
Household earnings (real Rp/week)	1547.857	863.851	914.498	1875.160	1274.700	1232.052	0.31
Fraction illiterate	0.356	0.398	0.398	0.039	0.523	0.340	0.90
Fraction primary school or less	0.206	0.307	0.284	0.361	0.174	0.282	0.18
Sex ratio	1001.416	990.720	936.186	1009.557	1002.068	984.633	0.43
Fraction urban	0.335	0.167	0.258	0.465	0.257	0.287	0.54
State NDP growth rate	11.284	6.251	2.923	8.207	2.718	5.025	0.02

Notes: Characteristics of each state in the panel dataset in 2008 (before NOTA was available). Karnataka is the "control" state that does not have NOTA in 2013. For constituency-level variables the values shown are averages within each state. The p-value for the equality of means test in the last column is from OLS regressions of each variable on a "treatment" indicator. For constituency-level variables we obtained the p-values through a bootstrap clustered by state.

Table A.9: Control and treatment states before NOTA in the extended dataset

	Control	Treatment	p-value for equality
Constituency characteristics			
Number of eligible voters (1000)	165.956	180.764	0.67
	(24.852)	(24.811)	
Turnout	0.691	0.655	0.51
	(0.049)	(0.020)	
Election closeness	0.104	0.103	0.91
	(0.010)	(0.005)	
Reserved constituency	0.262	0.293	0.58
	(0.033)	(0.048)	
State characteristics			
Number of constituencies	135.313	130.667	0.91
	(26.540)	(28.628)	
Labor force participation	0.591	0.592	0.97
	(0.016)	(0.030)	
Unemployment rate	0.049	0.022	0.02
	(0.010)	(0.003)	
Household earnings (real Rp/week)	1,410.241	1,610.451	0.36
	(108.754)	(188.210)	
Fraction illiterate	0.251	0.327	0.21
	(0.027)	(0.052)	
Fraction primary school or less	0.260	0.215	0.23
	(0.024)	(0.028)	
Sex ratio	997.418	942.847	0.06
	(15.913)	(22.701)	
Fraction urban	0.288	0.357	0.43
	(0.035)	(0.079)	
State NDP growth rate	7.663	3.832	0.08
	(1.456)	(1.487)	
Constituencies	2165	1176	
States	16	9	

Notes: Average characteristics with standard errors in parentheses of the control and treatment states before NOTA was available. For each state, values included are for the first election in the sample. The p-value for the equality of means test in the last column is from OLS regressions of each variable on a "treatment" indicator. For constituency-level variables we obtained the p-values allowing for clustering by state.

Table A.10: Comparison of states in the panel to others in the extended dataset before NOTA

	Panel	Other states	p-value for equality
Constituency characteristics			
Number of eligible voters (1000)	163.797	173.429	0.70
	(11.254)	(22.848)	
Turnout	0.688	0.675	0.77
	(0.013)	(0.041)	
Election closeness	0.096	0.106	0.29
	(0.004)	(0.009)	
Reserved constituency	0.344	0.251	0.12
	(0.055)	(0.030)	
State characteristics			
Number of constituencies	156.800	127.850	0.51
	(38.630)	(22.652)	
Labor force participation	0.667	0.572	0.00
	(0.024)	(0.015)	
Unemployment rate	0.015	0.046	0.00
	(0.003)	(0.008)	
Household earnings (real Rp/week)	1295.213	1529.093	0.28
	(191.260)	(110.852)	
Fraction illiterate	0.343	0.262	0.32
	(0.081)	(0.026)	
Fraction primary school or less	0.267	0.238	0.47
	(0.034)	(0.022)	
Sex ratio	987.990	975.218	0.56
	(13.294)	(17.079)	
Fraction urban	0.296	0.317	0.75
	(0.050)	(0.044)	
State NDP growth rate	6.276	6.286	1.00
	(1.623)	(1.357)	
Constituencies	784	2618	
States	5	20	

Notes: Average characteristics with standard errors in parentheses of the 5 states in the panel and the other 20 states in the extended dataset before NOTA. For each state, values included are for the first election in the sample. The p-value for the equality of means test in the last column is from OLS regressions of each variable on a "panel state" indicator. For constituency-level variables we obtained the p-values allowing for clustering by state.

3.3 Robustness of the regression estimates

3.3.1 Randomization inference

The reduced form results presented in Section 5.2 of the paper suggest that the NOTA policy increased turnout. As discussed there, our ability to conduct inference is fundamentally limited by the number of states and time periods. In the paper, we attempt to remedy this by reporting bootstrapped standard errors, and find that these lead to similar inference as standard errors obtained with the asymptotic cluster-robust formula. We now check if the inference that NOTA increased turnout can be supported further by conducting randomization inference tests (see MacKinnon and Webb (2018) for a recent analysis of randomization inference in the panel data context).

We replicate the "experiment" involving the timing of the Supreme Court's decision with respect to the states' electoral calendar by allowing every combination of states to serve as a potential treatment or control group. For every artificial treatment assignment, we select 9 of the 25 states as the treatment group (the same number as in the true treatment group). Each of these states is assigned the NOTA policy between its first and second election, while the other 16 states are assumed not to have NOTA in either election. For each treatment assignment, we run specification (2) in Table 3 in the paper and collect the t-statistics of the estimated treatment effects. We obtain the p-value of the true treatment effect based on the rank of the true t-statistic in this distribution.

First, we consider all possible treatment assignments with 9 treatment and 16 control states (2,042,975 possible combinations). To obtain the (approximate) p-value, we estimate treatment effects as described above for 1000 randomly chosen assignments from this set. This yields a p-value of 0.069 for the null of no treatment effect against the one-sided alternative of a positive treatment effect, and a p-value of 0.14 against the two-sided alternative.

One issue with this procedure is that, since states differ in the number of constituencies, the number of units (constituencies) in the artificial treatment assignments can differ considerably from the true treatment assignment. To achieve assignments more comparable to the actual one, we restrict the set of possible treatment assignments by considering only assignments where the total number of treated constituencies is within +/- 20% of the true number of treated constituencies (1176). This yields 1,374,186 possible assignments, and we again obtain an approximate p-value from 1000 randomly chosen assignments from this set. This yields a p-value of 0.027 (one-sided) and 0.074 (two-sided) for the estimated treatment effect. Sampling from assignments within 10% (5%) of the true number of treated constituencies yields a one-sided p-value of 0.031 (0.021) and a two-sided p-value of 0.080 (0.064).

3.3.2 Systematic Voters' Education and Electoral Participation (SVEEP)

In 2009, the Election Commission of India launched a major voter education program, SVEEP, with the twin goals of educating voters about the electoral process and increasing voter participation. The program was rolled out over several years following states' electoral calendars, beginning with the December 2009 assembly election of Jharkhand. It has been in place in every subsequent state and national election since. The program directs election officers at the state and district level to develop strategies for increasing voter awareness and participation. It emphasizes the importance of understanding "gaps" in turnout and developing activities in partnership with local organizations (civil groups, government organizations, the media, etc.) to address them. No systematic information appears to exist on these activities, but see Election Commission of India (2014) for a wide range of examples undertaken by the different states.

One concern is that the mere presence/absence of SVEEP could affect turnout. However, given the timing of the SVEEP program, the year fixed effects in our regressions will capture the presence/absence of SVEEP for most states. The exception is the group of states on the 2009/2014 electoral cycle: since in the 2009 election SVEEP was present in some of these states but not others, these states will be differentially affected by SVEEP between 2009 and 2014. Based on the robustness checks in Table 4 in the paper, where we leave out all elections held in 2009 and 2014, we already know that this does not explain away our results. We now confirm this also by controlling for a SVEEP indicator that takes a value of 1 for elections held after December 2009. The estimates, shown in columns (1) and (2) of Table A.11, are similar to those in the paper.

A second concern stems from the fact that the SVEEP program was decentralized to the state and district levels. The effect of NOTA could thus be confounded if the local features of the program were correlated with the NOTA policy. Ideally, one would measure the extent and type of SVEEP activities for each state or district and include it as a control. This is difficult to do, especially since no systematic documentation appears to exist on what specific activities were undertaken. As a second-best, we pursue two alternatives.

Table A.11: Controlling for SVEEP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NOTA	0.030**	0.035**	0.033**	0.044**	0.031**	0.029**	0.028*	0.032*
	(0.013)	(0.016)	(0.014)	(0.017)	(0.012)	(0.012)	(0.013)	(0.016)
SWEEP Y/N	-0.019	-0.049***	-0.020	-0.058***	-0.034*	-0.070***	-0.342***	-0.375***
	(0.017)	(0.017)	(0.017)	(0.015)	(0.019)	(0.023)	(0.037)	(0.028)
SWEEP intensity			-0.014***	-0.016***	0.029	0.060*		
			(0.004)	(0.002)	(0.018)	(0.030)		
Rel. turnout 2009							0.323***	0.327***
							(0.030)	(0.028)
Basic controls	x		x		x		x	
Extended controls		x		x		x		x
R2	0.18	0.19	0.19	0.20	0.18	0.20	0.28	0.30
N	6,685	6,685	6,685	6,685	6,685	6,685	6,685	6,685
States	25	25	25	25	25	25	25	25

Notes: SWEEP Y/N is an indicator equal to 1 beginning with the Jharkhand election in December 2009. SWEEP intensity is measured by the length of the state's 2014 SWEEP Action Plan divided by the number of constituencies in columns (3) and (4), and by an indicator equal to 1 if length per constituency is above-median in columns (5) and (6). Relative turnout 2009 is turnout in the 2009 parliamentary constituency divided by average turnout in the state.

First, to measure the intensity of SVEEP across states we use the length of states' SVEEP Action Plans for the 2014 national election. The Election Commission directed states to create such an action plan ahead of the national election, and issued a template for what the plan should include.[10] In spite of the template, the length and level of detail of the documents produced by the states varied widely. At one extreme, Tripura and Manipur have plans of less than 10 pages, filling in each section of the template with a few sentences. At the other extreme, the reports of Gujarat, Madhya Pradesh and Uttarakhand are above 100 pages. Punjab's report is 776 pages. The difference in length reflects a difference in detail. For example, a section entitled "C.1.1 Information and motivation" in the Tripura report contains two short bullet points, while the corresponding section in the Gujarat report is 9 pages long, with detailed examples of activities to be undertaken (see Figure A.5).

It seems plausible that the different length of the action plans are indicators of the different intensity of SVEEP in each state. First, the amount of work put into the document is likely to reflect the state's capacity for implementing SVEEP and state election officers' commitment to this program. Second, the extent of the state's past SVEEP efforts is likely to be reflected in the length of the plans (e.g., a state with more experience with SVEEP would find it easier to compile a longer, more detailed plan). Indeed, all action plan documents include explicit discussions of past SVEEP activities undertaken by the state (see the examples in Figure A.5). For these reasons the length of the action plans for 2014 may be a reasonable indicator of states' SVEEP efforts in the assembly elections that took place during the program period (after December 2009).

To create a proxy for the intensity of SVEEP, we divide the length of the action plans (number of pages) by the number of constituencies in the state (and use 0 for elections prior to SVEEP). Columns (3) and (4) of Table A.11 include this variable as a control in the turnout regressions, and we find that the effect of NOTA remains robust. In columns (5) and (6) we replace the linear proxy with an indicator equal to one if the length per constituency is above the median across states. Again we find similar results.

The intensity of SVEEP could be endogenous to actual or expected turnout. As a second attempt to control for the intensity of SVEEP activities, we use *past* turnout as a proxy. The Election Commission has explicitly called on states to identify areas with low turnout relative to the rest of the state and target interventions to those areas. Thus, areas with relatively low turnout may have more intense SVEEP activities. Because of the redistricting that took place in 2008, for most states we cannot use previous turnout in the same constituency (nor were the states able to target SVEEP based on this measure). Instead, we use turnout in

[10] The template is available at https://ecisveep.nic.in/files/category/2-publications. 2014 was the first time that the states were required to create action plans.

Figure A.5: SVEEP Action Plans

[…]
b) SVEEP FOR MOTIVATION

The timeline for SVEEP motivational activities will broadly be divided into two:

- SVEEP prior to the announcement of elections
- SVEEP from the time of announcement of elections

SVEEP motivation in the period prior to the announcement of elections:

a) In this period we will mostly concentrate on educating voters about voter registration processes, and will also build up awareness about the reasons to vote and why every vote counts. Making sure that people understand the voters' list (electoral rolls, process of registration), decide to register, are able to register and can register correctly are the challenges of SVEEP.
[…]
d) For engaging Youth the "Young Voters festival 2013" was recently concluded with participation of 1951 schools and Colleges from across the state. Our Election Emissary of 2012 Shri Dhvanit Thaker, RJ of Radio Mirchi is proposed to be continued and we also propose to add 2-3 more talented RJs as our Emissaries. […]

SVEEP Motivation from the date of announcement of elections:

There will be several physical events and activities to motivate the voters to turnout. The district wise details of such events and activities are attached at the last part of this Plan. Some of the major motivational activities are as follows:
Ekrar Patra:

- This was an exercise similar to the "Sankalpa Patra" exercise, wherein school children got their parents to accept their responsibility to register as electors.
- This was carried out between 6th Sept and 15th Sept, 2013.

Sankalpa Patra

- This will be done on the lines of the same activity undertaken in 2012 general elections
- The Sankalpa Patra will have two pledges as follows:
  - Pledge to turnout to vote
  - Pledge to vote ethically and not vote for bribe/cash/liquor etc.

Other activities:

- Various programmes will be held viz. Street Plays, different competitions amongst the students for inter personal contact with youth.
- Listing and detailing of all Resident Welfare Associations in all urban local bodies falling in Assembly Constituency for tapping their resources to facilitate voters staying in these societies.

[…]

Excerpts from section C.1.1 Information and Motivation from the Gujarat 2014 SVEEP Action Plan (pp51-60)

Playing of Audio Jingles and telecasting of the messages of National Icons through DD and local cable channels.

As done during Last Assembly Election 2013, the message of ethical voting will be played through Local Media for awareness of common mass. Mobile Van will be introducing for massive campaign on active participation and ethical voting.

Full text of Section C.1.1 Information and Motivation from the Tripura 2014 SVEEP Action Plan (p3)

the 2009 national election. We merge turnout information from the 2009 national election for each parliamentary constituency (parliamentary constituencies contain multiple assembly constituencies in a state). For elections under SVEEP, we divide these values by turnout in the average parliamentary constituency in the state; for elections prior to SVEEP, we assign a value of 0. We introduce this variable as a control in Table A.11, columns (7) - (9). Here too, we find that adding the control leaves the NOTA coefficient unchanged.

Based on these proxies for SVEEP, we do not see evidence that our findings above are due to any confounding effects this policy may have had on the impact of NOTA.

3.3.3 Redistricting

Another potential confound is the electoral redistricting that took place in April 2008. Because elections are held every 5 years and NOTA was introduced in September 2013, none of the states that were affected by NOTA in our period of study were redistricted, while most states that were not affected by NOTA were redistricted. Thus, redistricting has the potential to confound our estimates of NOTA.[11]

To control for this, we create a constituency-level measure of redistricting by using GIS boundary files to compare constituencies before and after the delimitation. Our first measure calculates for each current constituency that was redistricted in our study period the largest area that was part of a single constituency before the redistricting. For example, a value of 0.8 for this "maximum overlap" measure indicates that 80% of the current constituency's area was part of a single constituency pre-delimitation (while the remaining 20% was part of one or more different constituencies). The higher the maximum overlap, the less a constituency was affected by redistricting. Our second measure, rather than focus on the largest area of overlap, uses each overlapping area to create an index of "territorial fractionalization." If a constituency overlaps with n pre-delimitation constituencies with $a_1, ..., a_n$ denoting the share of its area falling in each of these, then the fractionalization index is $1 - \sum_{i=1}^{n} a_i^2$. The larger this value, the more the current constituency was affected by redistricting. Both of these measures are available for 22 states (constituency boundary files are not available for the states of Assam, Manipur, and Nagaland).

Table A.12 presents regressions corresponding to Table 3 in the paper controlling for these measures of redistricting. The first two columns repeat columns (2) and (3) in Table 3 in the paper on the 22 states with available redistricting measures. Columns (3) and (4) then add the maximum overlap measure and columns (5) and (6) the territorial fractionalization index. As can be seen, adding either measure of redistricting to the regressions causes little change in the estimated effect of NOTA. The estimates also retain their significance, except for column (6) where the standard error increases just enough to yield a p-value of 0.106.[12]

3.3.4 State-specific events

Turning to state-specific events that may confound our estimates, we identified four states where various events may plausibly affect 2013 or 2014 turnout relative to the previous election (that is, turnout in the with-NOTA election relative to turnout in the without-NOTA election). In Chhattisgarh, Maoist insurgents conducted terrorist attacks in 2010

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

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

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

	(1)	(2)	(3)	(4)	(5)	(6)
NOTA	0.033**	0.021*	0.031**	0.019*	0.030**	0.020*
	(0.015)	(0.011)	(0.014)	(0.011)	(0.014)	(0.011)
Control for redistricting	none	none	maxo	maxo	fract	fract
Basic controls	x		x		x	
Extended controls		x		x		x
R^2	0.20	0.21	0.20	0.21	0.20	0.21
N	6173	6173	6173	6173	6173	6173
States	22	22	22	22	22	22

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

and May 2013, between the 2008 and 2013 elections in this state. In Jammu & Kashmir, various incidents occurred between its 2008 and 2014 elections, including a border skirmish in January 2013 between India and Pakistan described by observers as one of the worst in 10 years. In Delhi, a new anti-corruption party, Aam Aadmi entered politics in 2012, energized voters, and emerged as the second-largest party in the 2013 assembly election. Finally, Maharashtra held its 2009 election a year after the 2008 terrorist attacks in Mumbai on several hotels and public buildings, and security concerns may have depressed voter turnout there.

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

While we did not find specific events in other states that may have affected turnout and whose timing coincided with NOTA, we may not have found all such events. To allow for this, we ran regressions excluding each state one at a time. The distribution of the resulting parameter estimates and p-values is shown in Table A.14. The 3 percentage point effect found in the main text turns out to equal both the mean and the median of the distribution of coefficients in these regressions. Of these coefficients, 88% are statistically significant at the 10 percent level and 68% are significant at the 5 percent level.

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

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

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

3.3.5 Voting costs

In this section we include further controls in the difference-in-differences specification in an attempt to control for any time-varying differences in voting costs across constituencies. First we obtained data from the Election Commission on the day of the week that the elections in each constituency were held. We create a dummy for wether the election was held on a weekend, as this might affect the cost of turnout. Constituencies within a state typically go to the polls in groups over a period of 2-3 days, so this variable varies at the constituency-year level. In columns (1) and (2) of Table A.15 we find that controlling for the Weekend dummy has no impact on our results.

Second, we include rainfall information on each election day. Several studies document that bad weather can raise the cost of turnout. Columns (3) and (4) in Table A.15 show that our estimates of the impact of NOTA are robust to controlling for rainfall. Columns (5) and (6) include both the weekend indicator and rainfall and yield similar results.

Third, we obtained data on the number of voting stations in each constituency. We divide this by the number of eligible voters in order to proxy for the convenience of voting. For example, a low number of voting stations per voters may lead to long wait times at the voting booth and discourage some people from voting. We include this variable as a control in columns (7) and (8) of Table A.15. These estimates should be interpreted with care since

Table A.14: Effect of NOTA on turnout: distribution of coefficients and p-values dropping one state at a time

NOTA coefficients	
Mean	0.030
Median	0.030
10th percentile	0.024
90th percentile	0.032
fraction $p < 0.05$	0.680
fraction $0.05 < p < 0.10$	0.200
fraction $0.10 < p < 0.19$	0.120

Notes: Estimates of the effect of the NOTA policy on turnout using the repeated cross section sample with basic controls, excluding one state at a time (25 regressions).

the number of voting stations could be endogenous for a number of reasons (for example, areas with historically high turnout may receive more stations). Nevertheless, it is reassuring that controlling for differences in voting costs as proxied by the number of stations per voter actually reinforces our findings. Columns (9) and (10) show the corresponding estimates when we instead divide the number of stations with the number of eligible voters.

3.3.6 Ballot placement

NOTA introduced a new option on the voting machines. Moreover, it was assigned the last button on the machines, below all the regular candidates (see Figure A.1). While we are not aware of previous analyses of being placed *last*, studies have shown that ballot placement more generally can affect voter choices (at least in the case of paper ballots: see, e.g., Ho and Imai (2006) and Shue and Luttmer (2009)). Could the introduction of NOTA have such an impact on voters? Clearly, this would not explain the increase in turnout which we find most NOTA votes are coming from. It could, however, account for some of the substitution away from candidates that we find in the structural model.

Table A.15: Effect of NOTA on turnout with controls for voting costs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NOTA	0.028*	0.030*	0.024*	0.025*	0.024*	0.025*	0.051***	0.052**	0.059***	0.061***
	(0.015)	(0.016)	(0.013)	(0.014)	(0.014)	(0.015)	(0.015)	(0.019)	(0.014)	(0.019)
Weekend	-0.003	-0.008			-0.003	-0.007				
	(0.009)	(0.009)			(0.009)	(0.009)				
Rainfall			0.032	0.033	0.032	0.032				
			(0.022)	(0.024)	(0.022)	(0.024)				
Stations / voters							79.791**	86.529***		
							(30.391)	(30.277)		
Voters / stations									-0.020***	-0.021***
									(0.006)	(0.006)
Basic controls	x		x		x		x		x	
Extended controls		x		x		x		x		x
R^2	0.18	0.19	0.18	0.20	0.18	0.20	0.20	0.22	0.23	0.25
N	6,685	6,685	6,684	6,684	6,684	6,684	6,676	6,676	6,676	6,676
States	25	25	25	25	25	25	25	25	25	25

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

While it is difficult to rule this out completely, we can check whether introducing NOTA resulted in a decline in the vote share of the candidate listed last on the machine (i.e., listed last before the NOTA policy and just above the NOTA option once NOTA was introduced). It is possible to do this because the Election Commission regulates the placement of candidates on the machines.[13] Table A.16 columns (1) and (2) run our main specification using as dependent variable the vote share of the candidate listed last (or immediately above NOTA). The results suggest that NOTA did not have a significant impact on this candidate's vote share, although the estimated coefficients are negative. Of course, we would also find a negative effect if voters substitute to NOTA deliberately (not because of ballot placement). To conduct a sharper test, we take the difference between the vote share of the candidate ranked before last and the vote share of the candidate ranked last. If voters disproportionately substituted to NOTA from the last candidate, then we would expect to find significant positive coefficients on NOTA. Columns (3) and (4) of Table A.16 show that this is not the case, suggesting that ballot placement effects from NOTA may not be quantitatively important.

Table A.16: NOTA and ballot order

Dep. Var:	Last candidate		Last cand. minus penultimate cand.	
	(1)	(2)	(3)	(4)
NOTA	-0.024	-0.005	0.006	-0.022
	(0.020)	(0.018)	(0.018)	(0.024)
Basic controls	x		x	
Extended controls		x		x
R^2	0.01	0.02	0.00	0.00
N	6685	6685	6685	6685

Notes: Estimates of the effect of the NOTA policy on the vote share of the candidate listed last using the repeated cross section sample. The dependent variable is the vote share of the last candidate (sample mean = 0.054, s.d. = 0.120) in columns (1) and (2) and the difference between the penultimate and the last candidate's vote share (sample mean = 0.019, s.d. = 0.159) in columns (3) and (4). All regressions control for state and year fixed effects and the following state-level variables: labor force participation, real weekly household earnings, fraction of illiterates, fraction with primary school or less as highest education. The Extended controls specification also controls for reserved constituencies and the following state level variables: unemployment, sex ratio, fraction urban, and the growth rate of net domestic state product. Standard errors clustered by state in parentheses. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

[13] See http://eci.nic.in/eci_main/ElectoralLaws/HandBooks/Handbook_for_Candidates.pdf, p42-43. Candidates are first ordered by party type (national, state, unrecognized, or independent) and then alphabetically, excluding first initials and titles. Using these rules, we replicated the ordering of candidates based on the candidate names available in the Election Commission database.

3.3.7 Voter registration

Changes in voter registration could impact our findings in two ways.[14] First, it could be that some of the new turnout is due to voters deciding to register and vote after the introduction of NOTA. Since voters failing to register is a form of abstention, this would mean that we are underestimating the impact of NOTA on voter participation.[15]

Second, it could be that voter registration lists contain mistakes (e.g., voters who moved or died may incorrectly appear on the list). If such mistakes exist and if the introduction of NOTA was accompanied by increased efforts to fix them, this could yield a reduction in the number of registered voters and show up as increased turnout in our regressions. In Table A.17 we use the number of eligible voters as a dependent variable and find insignificant positive coefficients. There is no evidence that NOTA affected the number of registered voters, and especially that it did so in a negative way.

Table A.17: NOTA and voter registration

Dep. Var:	Log(eligible voters)	
	(1)	(2)
NOTA	0.069	0.077
	(0.050)	(0.047)
Basic controls	x	
Extended controls		x
R^2	0.10	0.10
N	6685	6685

Notes: Estimates of the effect of NOTA on the number of registered voters (in logs) using the repeated cross section sample. Sample mean of dep. var. = 11.917, s.d. = 0.734. All regressions control for state and year fixed effects and the following state-level variables: labor force participation, real weekly household earnings, fraction of illiterates, fraction with primary school or less as highest education. The Extended controls specification also controls for reserved constituencies and the following state level variables: unemployment, sex ratio, fraction urban, and the growth rate of net domestic state product. Standard errors clustered by state in parentheses. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

[14] Voter Registration is a one time procedure. Except in special cases (such as for convicted criminals), once registered as a voter, a person can vote in all subsequent elections without having to go through any further registration process. Once registered the voter's name is on the voters' list and she gets the identification card which needs to be produced at the polling station before being allowed to vote. The voting age is 18.

[15] For example, suppose there are E eligible voters, R of whom registered, and V of whom voted. Suppose that after NOTA, the $(E - R)$ previously unregistered voters register and vote, and total turnout is $V' = V + (E - R)$. Then our estimated effect of NOTA would be $V'/E - V/R$ while the true effect is $V'/E - V/E$, which is larger.

3.4 NOTA and political competition

Here we investigate whether the presence of NOTA is correlated with ex post election closeness in the reduced form. In Table A.18 we present regressions on three different measures of closeness: the difference in the vote shares of the two frontrunners (columns (1) and (2)), the difference in the number of votes received by the two frontrunners (columns (3) and (4)), and the difference in the log number of votes received by the two frontrunners (columns (5) and (6)). The presence of the NOTA option is not associated with significantly closer elections using any of these measures.

Table A.18: Effect of NOTA on election closeness

Dep. Var:	Vote share difference		Vote count difference		Log vote count difference	
	(1)	(2)	(3)	(4)	(5)	(6)
NOTA	-0.001	-0.019	70.1	-3072.2	-0.020	-0.065
	(0.011)	(0.019)	(2052.0)	(2945.9)	(0.030)	(0.048)
Basic controls	x		x		x	
Extended controls		x		x		x
R^2	0.01	0.02	0.11	0.12	0.01	0.01
N	6685	6685	6685	6685	6685	6685
States	25	25	25	25	25	25

Notes: Estimates of the effect of the NOTA policy on election closeness using the repeated cross section sample. Election closeness between the two frontrunners is measured as the difference in vote shares in columns (1) and (2), as the difference in the number of votes received in columns (3) and (4), and as the difference in the log number of votes received in columns (5) and (6). All regressions control for state and year fixed effects, the log number of eligible voters in a constituency and its square, and the following state-level variables: labor force participation, real weekly household earnings, fraction of illiterates, fraction with primary school or less as highest education. The Extended controls specification also controls for reserved constituencies and the following state level variables: unemployment, sex ratio, fraction urban, and the growth rate of net domestic state product. Standard errors clustered by state in parentheses. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

4 Further details for the BLP model

4.1 GMM details

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

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

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \ \tilde{\boldsymbol{\xi}}(\boldsymbol{\theta})' \mathbf{Z} \mathbf{W}^{-1} \mathbf{Z}' \tilde{\boldsymbol{\xi}}(\boldsymbol{\theta}), \tag{2}$$

where $\mathbf{Z}$ is the matrix of instruments, and $\mathbf{W}^{-1}$ is the weighting matrix. For given $\boldsymbol{\theta}_2$, the linear coefficients $\boldsymbol{\beta}$ (which include the party fixed effects) can be obtained analytically from (2). Unlike Nevo (2001), we are able to estimate the party fixed effects and the other coefficients in $\boldsymbol{\beta}$ in the same step because we have variation in candidate characteristics for a given party across constituencies.

To compute the estimate in (2), we use the standard two-step GMM procedure. We first set $\mathbf{W} = \mathbf{Z}'\mathbf{Z}$ and compute an initial estimate of the parameters, $\boldsymbol{\theta}^{Step1}$. We then use this initial estimate to compute a robust weight matrix and use this updated weight matrix to compute the final parameter estimates. For the robust weight matrix we use the cluster-robust formula $\mathbf{W} = \sum_{c}^{C} \mathbf{Z}_c' \tilde{\boldsymbol{\xi}}_c(\boldsymbol{\theta}^{Step1}) \tilde{\boldsymbol{\xi}}_c(\boldsymbol{\theta}^{Step1})' \mathbf{Z}_c$, i.e. we allow for both heteroskedasticity and correlation of the errors ξ_{jc} across candidates within a constituency. As discussed in Section 6.1 in the paper, this is important if the expected closeness of the race results in correlation between unobserved voter preferences for some of the candidates.

The covariance matrix of the estimated parameters is computed using the standard formulas (e.g., Cameron and Trivedi, 2005, p194-195). Letting $\boldsymbol{\theta}^{Step2}$ denote the final vector of parameter estimates, we compute the derivatives of the GMM error term, $\mathbf{D} = \partial \tilde{\boldsymbol{\xi}}(\boldsymbol{\theta}^{Step2}) / \partial \boldsymbol{\theta}$, and the (scaled) covariance matrix of the moment conditions, $\mathbf{S} = \sum_{c}^{C} \mathbf{Z}_c' \tilde{\boldsymbol{\xi}}_c(\boldsymbol{\theta}^{Step2}) \tilde{\boldsymbol{\xi}}_c(\boldsymbol{\theta}^{Step2})' \mathbf{Z}_c$. The estimated covariance matrix of the parameters is then

$$[\mathbf{D}'\mathbf{Z}\mathbf{W}^{-1}\mathbf{Z}'\mathbf{D}]^{-1}[\mathbf{D}'\mathbf{Z}\mathbf{W}^{-1}\mathbf{S}\mathbf{W}^{-1}\mathbf{Z}'\mathbf{D}][\mathbf{D}'\mathbf{Z}\mathbf{W}^{-1}\mathbf{Z}'\mathbf{D}]^{-1},$$

which yields standard errors robust to heteroskedasticity and constituency-level clustering.

4.2 Aggregation

Parties play an important role in our specification. A difficulty arises because of the presence of many small parties. There are a total of 200 parties in the data, but half of them field candidates in only 1 of every 40 constituency within a state. A second, related difficulty is the presence of independent candidates (candidates not affiliated with any party). There are 6751 of these candidates in the data, but 70% of them receive less than 1% of the votes in a constituency and only 3% receive more than 10%. Each of these parties and candidates adds a new fixed effect that is difficult to identify due to the small number of constituencies where the party is represented (in the extreme case of an independent candidate running in only one year, identifying the fixed effect is not possible).[16]

To deal with this difficulty, we create a "Small Party" category comprising parties fielding candidates in less than one third of the constituencies in any given state and we average all small party candidates' characteristics within a constituency (we do this after constructing the instruments so that the individual IVs are aggregated also). Below, we explore alternative aggregation thresholds requiring fielding a candidate in 1/4 or 1/2 of the constituencies in a given state. We also create an "Independent Party" containing all independent candidates, and aggregate them within constituencies in the same way. After this aggregation, we are left with a total of 22 parties.

Table A.19 shows summary statistics before and after the aggregation of Independent and Small party candidates. Here, "small parties" are parties fielding candidates in less then 1/3 of the constituencies in a given state. Tables A.20 and A.21 present corresponding figures using thresholds of 1/2 and 1/4, respectively.

4.3 Instruments and identification

4.3.1 Instrument sets

We compare the results using 3 different sets of instruments, summarized in Table A.22. The first set includes the average of the gender and age characteristics. As described in the text, we compute the average characteristic of a party's candidates in all other constituencies in the state. We also take the interaction of these averages with the state dummies to allow for more flexibility in the instruments' impact by state. Similarly, we compute the average

[16] In IO applications, the solution to the analogous problem of many small products is typically to assume that these are part of the broadly defined "outside option." We cannot follow this route here since the outside option is a well-defined choice (abstention) and the focus of our study.

Table A.19: Aggregating Independent and Small party candidates

Variable	N	Mean	Std. Dev.	10%	90%
Before aggregation					
Number of candidates	1446	11.370	4.909	6	17
Number of independents	1446	4.667	3.729	1	9
Number of small-party candidates	1446	2.014	1.763	0	4
Vote share of independents	6748	0.018	0.048	0.002	0.027
Vote share of small-party candidates	2912	0.018	0.045	0.002	0.029
After aggregation					
Number of candidates	1446	6.439	1.403	5	8
Vote share of independents	1354	0.090	0.115	0.015	0.252
Vote share of small-party candidates	1176	0.021	0.053	0.002	0.034

Notes: Candidate counts are for the 1446 constituencies in the data and exclude the NOTA option. Vote shares are for all independent/small party candidates. Independent candidates are not affiliated with any party. Small parties are parties fielding candidates in less than 1/3 of the constituencies in a state. For each of these categories we aggregate candidates in a constituency as described in the paper.

Table A.20: Aggregating Independent and Small party candidates (small party threshold: 1/2)

Variable	N	Mean	Std. Dev.	10%	90%
Before aggregation					
Number of candidates	1446	11.370	4.909	6	17
Number of independents	1446	4.667	3.729	1	9
Number of small-party candidates	1446	2.385	1.970	0	5
Vote share of independents	6748	0.018	0.048	0.002	0.027
Vote share of small party candidates	3448	0.018	0.044	0.002	0.030
After aggregation					
Number of candidates	1446	6.101	1.233	5	8
Vote share of independents	1354	0.090	0.115	0.015	0.252
Vote share of small party candidates	1223	0.020	0.050	0.002	0.032

Notes: Candidate counts are for the 1446 constituencies in the data and exclude the NOTA option. Vote shares are for all independent/small party candidates. Independent candidates are not affiliated with any party. Small parties are parties fielding candidates in less than 1/2 of the constituencies in a state. For each of these categories we aggregate candidates in a constituency as described in the paper.

Table A.21: Aggregating Independent and Small party candidates (small party threshold: 1/4)

Variable	N	Mean	Std. Dev.	10%	90%
Before aggregation					
Number of candidates	1446	11.370	4.909	6	17
Number of independents	1446	4.667	3.729	1	9
Number of small-party candidates	1446	1.569	1.554	0	4
Vote share of independents	6748	0.018	0.048	0.002	0.027
Vote share of small party candidates	2269	0.017	0.046	0.002	0.027
After aggregation					
Number of candidates	1446	6.804	1.482	5	9
Vote share of independents	1354	0.090	0.115	0.015	0.252
Vote share of small party candidates	1061	0.020	0.049	0.002	0.032

Notes: Candidate counts are for the 1446 constituencies in the data and exclude the NOTA option. Vote shares are for all independent/small party candidates. Independent candidates are not affiliated with any party. Small parties are parties fielding candidates in less than 1/4 of the constituencies in a state. For each of these categories we aggregate candidates in a constituency as described in the paper.

characteristics of all candidates in other constituencies, and we do this for both elections in the data (2008 and 2013). We increase this set with further instruments in order to help identify the nonlinear parameters of the model. The second set of instruments augments the first set by adding interactions of the NOTA indicator with the constituency-level average of the demographics used in the estimation (minority population, literacy rate, share of rural workers). For the third set, we augment the first set with instruments for the minority candidate characteristic, constructed similarly to those for gender and age. Table A.23 presents detailed summary statistics of these instruments.

To get a sense of the strength of the different instrument sets, Table A.24 reports "first stage" F statistics from linear (Logit) specifications. Because the full model we will estimate is nonlinear, we emphasize that these tests are merely suggestive.[17] The first instrument set appears to be relatively stronger than the second and third sets.[18] The first two sets pass the conventional weak IV test ($F > 10$) while the third one does not. Because in the specifications below the third instrument set does not pass the overidentification J test, we will not use it in our counterfactual analysis.

[17]Weak identification in nonlinear GMM is an active area of current research and we know of no test directly applicable to our setting. Some recent studies tackle this issue by using Chamberlain-type optimal instruments but this would require specifying the exact form of endogeneity, i.e., parametrizing the supply model of candidate characteristics.

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

Table A.22: Description of the Instrumental Variables

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

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

4.3.2 Identification: examples

The identifying assumption for estimation, expressed in the moment conditions, is that voters' valuation for a candidate's unobserved characteristics in a constituency (ξ_{jc}) is conditionally independent of the average (observed) characteristics of candidates in *other* constituencies (i.e., the instruments). We now discuss several examples where this assumption is likely or unlikely to hold.

First, suppose that parties do not condition their choice of candidate characteristics on the popularity shocks ξ_{jc} in equation (6) in the paper. For example a party with an SC base may find it impossible to respond to a popularity shock by finding a candidate from a different caste in time for the election. In this case, the mix of candidate characteristics offered by a party would reflect only the supply of characteristics in the relevant population, and would not be affected by the popularity shocks among voters, and the identification assumption would be satisfied.

Second, suppose that parties' choices of candidate characteristics do respond to the popularity shocks ξ_{jc} but, controlling for party-specific means and demographics, these shocks are independent across constituencies (but may be correlated for a given constituency over time). In this case, candidate characteristics in two different constituencies, $\mathbf{x}_{jc}$ and $\mathbf{x}_{jc'}$ have both a common source (the supply effect) and separate sources (the shocks ξ_{jc}). Because of this, the variation in $\mathbf{x}_{jc'}$ can be used to separate out variation in $\mathbf{x}_{jc}$ that is independent of ξ_{jc}, and the identification assumption is again satisfied. This is the leading example discussed in the consumer demand literature, where valuation shocks for a product are assumed to

Table A.23: Summary statistics of the instruments

Instrument	N	Mean	Std. Dev.	10%	90%
Regular candidates					
Female in other constituencies 2008	9311	0.076	0.021	0.048	0.104
Female in other constituencies 2013	9311	0.079	0.022	0.049	0.096
Age in other constituencies 2008	9311	0.453	0.021	0.422	0.474
Age in other constituencies 2013	9311	0.462	0.012	0.447	0.473
Minority in other constituencies 2008	9311	0.395	0.103	0.289	0.524
Minority in other constituencies 2013	9311	0.408	0.087	0.340	0.518
NOTA					
Minority population	520	0.385	0.194	0.207	0.720
Literacy	520	0.567	0.090	0.468	0.670
Rural workers	520	0.683	0.164	0.463	0.857
State 1 (Karnataka)					
Own party's female	2814	0.048	0.012	0.035	0.067
Own party's age	2814	0.473	0.040	0.426	0.526
Own party's minority	2814	0.317	0.100	0.250	0.435
State 2 (Madhya Pradesh)					
Own party's female	2889	0.089	0.021	0.062	0.116
Own party's age	2889	0.440	0.036	0.397	0.492
Own party's minority	2889	0.441	0.083	0.358	0.558
State 3 (Mizoram)					
Own party's female	98	0.020	0.039	0.000	0.100
Own party's age	98	0.493	0.033	0.442	0.537
State 4 (Rajasthan)					
Own party's female	2301	0.089	0.037	0.047	0.131
Own party's age	2301	0.469	0.043	0.412	0.530
Own party's minority	2301	0.368	0.063	0.313	0.440
State 5 (Chhattisgarh)					
Own party's female	1209	0.101	0.048	0.051	0.170
Own party's age	1209	0.433	0.039	0.402	0.497
Own party's minority	1209	0.526	0.109	0.415	0.639

Notes: Female in other constituencies 2008 is the average of Female for all candidates in other constituencies in the state in the 2008 election. Variables for other elections and candidate characteristics are constructed similarly for the 9311 non-NOTA candidates. The NOTA indicator is interacted with 3 average demographics of the constituency (minority pop., literacy, rural workers). For each state and each candidate characteristic (female, age, and minority) an instrument is created by interacting the state indicator with the average of a party's candidates in other constituencies within the state in the given election. Summary statistics for these instruments are listed by state with the number of candidates for the state given under N. The variable State3*Minority is not created because all constituencies in state 3 (Mizoram) are reserved for minority candidates. In the estimation we use 3 different subsets of the listed instruments, see the paper for details.

Table A.24: Logit IV estimates

Instrument set:	First	Second	Third
	(1)	(2)	(3)
Female	0.805***	0.512**	0.289*
	(0.269)	(0.200)	(0.175)
Age	2.936**	3.313***	1.828***
	(1.174)	(0.857)	(0.664)
Minority	-4.177***	-1.382***	-0.930***
	(0.639)	(0.145)	(0.174)
NOTA	-4.683***	-3.798***	-3.653***
	(0.229)	(0.072)	(0.073)
Ran	0.287***	0.288***	0.297***
	(0.085)	(0.063)	(0.059)
Won	0.573***	0.630***	0.634***
	(0.076)	(0.067)	(0.064)
N	9831	9831	9831
Weak IV F stat	39.09	10.43	8.95
J	84.89	200.42	292.60
df	11	14	20
p-value	0.000	0.000	0.000

Notes: Two-step GMM estimates of the linear (Logit) model. All regressions include state, year, and party fixed effects, reservation status, broadcast allowance and rainfall. The Weak IV F statistic is the Kleibergen-Paap statistic computed by ivreg2 in Stata. Observations weighted by the number of eligible voters. Standard errors clustered by constituency in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively. N = 9831.

be independent across markets (e.g., cities) (see Hausman (1996) and Nevo (2000, 2001) for discussions). This assumption rules out a popularity shock to *some* of a party's candidates as would be caused, e.g., by a regionally coordinated advertising campaign (a campaign raising the popularity of *all* candidates of a party would be captured by the party dummies).

In our setting, a natural possibility is the correlation of preference shocks across constituencies within a particular state or a particular year, as would be the case, e.g., if a party conducted a particularly effective campaign in that state or that year only. Two features of the exercise mitigate this possibility. First, in our sample 65% of the parties field candidates in only one of the states and 60% field candidates in only one year. For these parties, controlling for a party fixed effect captures any correlation between the voter valuations ξ_{jc} across constituencies in a given state or a given election. Second, the included Broadcast allowance control (which varies at the party/state/year level) should capture some of the local advertising effects described above.[19]

As another example, suppose there exists some unobserved constituency characteristic (e.g., culture or history) that affects the valuation of a candidate's unobserved characteristic (e.g., the political past of his family). If multiple constituencies share the same history, then the valuation of a candidate's family background will be correlated, which would suggest that the ξ_{jc} will be correlated across constituencies. However, this is not necessarily the case. For example, the ξ_{jc} may be uncorrelated if parties have no control over the unobserved characteristic (the family background) of their candidates.

To see this formally, suppose that there exists some unobserved constituency characteristic, H, (such as history) that affects the valuation of a candidate's unobserved characteristic, f (such as family background). Consider two constituencies, 1 and 2, with related histories, so that $Cov(H_1, H_2) \neq 0$, and take a candidate of the same party in the two constituencies (so we can drop the j index for simplicity). Suppose the candidate's valuation shock is given by $\xi_c = f_c H_c$ in constituency $c = 1, 2$ (this is consistent with the interactive specification for *observed* candidate characteristics used in the paper). Is it necessarily the case that the valuation shocks will be correlated, i.e., that $Cov(\xi_1, \xi_2) \neq 0$?

The answer is no. In general, the formula for $Cov(\xi_1, \xi_2) = Cov(f_1 H_1, f_2 H_2)$ turns out to be quite involved: see Bohrnstedt and Goldberger (1969, p11). Consider a special case, where the f's and H's are jointly normal with mean 0. Then

$$Cov(f_1 H_1, f_2 H_2) = Cov(f_1, f_2) Cov(H_1, H_2) + Cov(f_1, H_2) Cov(H_1, f_2) \tag{3}$$

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

from expression (13) in Bohrnstedt and Goldberger (1969). Thus, $Cov(H_1, H_2) \neq 0$ is not sufficient for $Cov(\xi_1, \xi_2) \neq 0$. For example, suppose that in the first constituency, the party has no control over the family background of its candidate so that f_1 is independent of both f_2 and H_2. Then $Cov(f_1, f_2) = Cov(f_1, H_2) = 0$ so that under (3), $Cov(\xi_1, \xi_2) = 0$ regardless of $Cov(H_1, H_2)$.[20]

Finally, suppose that not only do choices of candidate characteristics $\mathbf{x}$ respond to the popularity shocks ξ_{jc}, but these shocks are also correlated across constituencies. This does not necessarily invalidate our identifying assumptions, because those rely on conditional independence of ξ_{jc} from the (averages of) $\mathbf{x}_{c'}$ in other constituencies. The latter are functions of the shocks $\xi_{jc'}$ and other variables (equation (6) in the paper), so that $Cov(\xi_{jc}, \xi_{jc'})$ has no direct implication for $Cov(\xi_{jc}, \mathbf{x}_{c'})$. One way that the moment conditions could be violated is if a shock ξ_{jc} in a constituency changes the prices $\mathbf{q}_{jc'}$ in other constituencies. For example, suppose that, in response to a shock, a party is forced to move its only female candidate from constituency c' to c.[21] If this raises the price of finding another female candidate to run in c', the party may decide to run a male candidate instead, so that the shock in c affects the vector of characteristics $\mathbf{x}_{c'}$. The impact on the validity of the moment conditions is mitigated by the fact that we use as instruments the *average* $\mathbf{x}_{c'}$ across constituencies other than c. But if the shock in c causes price changes that affect this average, then the moment condition will not hold.

4.4 Simulating the voters

The BLP algorithm requires numerically solving the integral in equation (5) in the paper in order to obtain the predicted market shares. We do this in the standard way by drawing individual voters from the distribution of demographics in each constituency, computing the predicted individual probabilities of voting for each candidate, and averaging across simulations to obtain the simulator for the integral. To simulate the individual voters, we proceed as follows.

First, we match to each constituency the tehsils (or sub-districts) that it overlaps using the GIS boundary files for the electoral and the administrative divisions. We compute the fraction of the constituency's area that falls in each tehsil. The simplest approach would be to use tehsil-level demographics from the Census and take the area-weighted average of these

[20] Note that $Cov(\xi_1, \xi_2) = 0$ is only sufficient, and *not* necessary for identification. The moment conditions are of the form $E(\xi_1 z_2) = 0$ where z_2 is an *observed* characteristic in constituency 2 (and therefore different from the unobserved f_2). In the above specification of ξ_c, the moment condition we rely on would be $E(f_1 H_1 z_2) = 0$.

[21] The lack of residency requirements in India makes such "reshuffling" possible until the deadline for candidate nominations (around 3 weeks before the election).

for each constituency. The disadvantage of this approach is that it ignores the within-tehsil correlation of demographic variables (e.g., if rural villages also tend to be less literate). To preserve the correlation of demographics across villages, we instead proceed as follows.

In the census data, we compute the fraction of each tehsil's population in the various villages. For each simulated voter in a constituency, we first randomly pick a tehsil using the distribution of the constituency's area across tehsils. Next, from the chosen tehsil we randomly pick a village using the distribution of the tehsil's population across villages. Finally, from the chosen village we pick the voter's demographics using the village characteristics given in the Census. We repeat this procedure 1000 times for each constituency.

4.5 Algorithms and codes

We implement the BLP procedure in MATLAB using the Nested Fixed Point (NFP) algorithm proposed by Berry et al. (1995). As emphasized by Dube et al. (2012), implementing the procedure requires care in order to avoid numerical instability, local optima, and biased standard errors. In particular, Dube et al. (2012) show that inaccuracies in the computation of the mean utilities δ_{jc} in BLP's contraction mapping (see section 6.1 in the main text) can make the parameter estimates unreliable. This is especially the case for optimizers that use user-supplied derivatives because here the computed δ_{jc} enter *both* in the evaluation of the GMM objective function and its gradient.

Apart from following Dube et al.'s (2012) recommendation of using a tight convergence criterion for the contraction mapping (we use 10^{-12}), we took two additional steps in order to avoid these potential pitfalls.

First, we eliminated a source of numerical instability for applications with many markets in the typical codes used to compute market shares. Specifically, computing the market shares requires aggregating the utilities corresponding to the various options within a market (see the denominator of equation (5) in the paper). It is common to code this by first aggregating across all constituencies using the "cumsum" function, then taking differences for each constituency using the "diff" function. For example, with 3 markets and 5 possible options in each, the code would compute the sum for the 3rd market by summing over the 10 options in markets 1-2, then summing over all 15 options, and finally subtracting the former from the latter.[22] While this procedure is perfectly fine in many applications, with 1446 markets and 9831 options, aggregating across options quickly results in very large numbers, and MATLAB runs out of precision to accurately compute the small differences between these large numbers. To circumvent this, we use the more recent "accumarray" function,

[22] All publicly available codes that we are aware of use this procedure when computing the market shares.

which allows aggregating each market separately and thus yields numerically precise market shares. Precision in the computed market shares is crucial for the precise computation of δ_{jc}.

Second, we use a derivative-free procedure for optimizing the GMM objective. While methods that allow for a user-specified gradient can be much faster, they are susceptible to error if the gradient is not computed precisely. As highlighted by Dube et al. (2012), any error in δ_{jc} is likely to be magnified when it shows up *both* in the objective function and its user-supplied gradient. To avoid this loss of precision at the cost of giving up speed, we use a derivative-free optimizer. We used the "patternsearch" algorithm, which performs a grid search without evaluating the GMM gradient.

For our preferred specification, upon which our counterfactual analysis is based, we verified that neither of the alternative optimizers "fminsearch" or "fminunc" could improve on the estimates, either holding the GMM weighting matrix constant (i.e., running the second step only) or re-running the entire estimation routine from the beginning. We also verified the "patternsearch" results using various starting values, including a set of randomly chosen starting values.[23]

[23] An alternative to NFP used in the literature is mathematical programming with equilibrium constraints (MPEC) (see Dube et al. (2012)). This procedure uses the market share equations as constraints in the GMM program, and uses constrained optimization. With 9831 candidates, in our case the optimizer would need to handle 9831 constraints. Assuming an optimizer would be able to handle this many constraints, implementing it would require more extensive computing resources than we have access to. Given our careful implementation of NFP described above, it is unclear whether the gains from MPEC would exceed its costs in this particular case.

5 Specifications and counterfactual results

5.1 Random coefficients specifications

Table A.25 presents estimation results for Normal random coefficients specifications using the second instrument set. In column (1) we include random coefficients on the candidate characteristics female, age, minority, NOTA, as well as the utility of abstention (the constant). As can be seen, many of the linear parameters on the candidate characteristics are statistically significant, indicating that these variables are relevant determinants of average voter utility in a constituency over and above the party labels (since the mean utility always includes party fixed effects $\bar{\xi}_j$). Turning to the nonlinear parameters of the random coefficients, we see that the coefficients on female and age are larger and statistically significant while for the other characteristics they are small and insignificant. This suggests the presence of significant heterogeneity in voter preferences for female and age but not for other characteristics. In columns (2) - (4) we experiment with random coefficients on other candidate characteristics, including the party dummies, and always find similar results. For example, in column (4) we allow for random coefficients on the two largest parties, INC and BJP, as well as on independent candidates and the "Small party" category. The estimates for these coefficients are all close to 0. 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.

In columns (5) and (6) we include the additional candidate characteristics education, criminal history, and assets (along with their missing-indicators as discussed in the text). These do not affect the above conclusions, in particular the random coefficients on female and age remain statistically significant.

Table A.25 also shows that these Normal random coefficients specifications of the model are inadequate: the J-test always rejects the validity of these specifications.

Tables A.26 and A.27 present estimation results for random coefficients specifications using the first and third instrument set, respectively. The findings are broadly similar.

Table A.25: Parameter estimates using Normally distributed random coefficients

	(1)	(2)	(3)	(4)	(5)	(6)
Linear parameters						
Female	-2.264	-2.217	-2.263	0.517**	-1.787	-1.976
	(1.441)	(1.444)	(1.739)	(0.235)	(1.573)	(1.426)
Age	7.825***	7.814***	7.801**	3.759***	7.964***	5.119**
	(2.196)	(2.396)	(3.371)	(1.241)	(1.880)	(2.666)
Minority	-1.294	-1.287***	-1.287***	-1.332***	-1.185***	-1.211***
	(0.311)	(0.223)	(0.187)	(0.229)	(0.356)	(0.285)
Ran	0.146	0.156	0.147	0.285***	0.183	0.221**
	(0.107)	(0.413)	(0.117)	(0.071)	(0.132)	(0.120)
Won	0.555***	0.554***	0.556***	0.620***	0.501***	0.572***
	(0.119)	(0.127)	(0.154)	(0.088)	(0.109)	(0.118)
NOTA	-4.254***	-4.245***	-4.231***	-3.796	-4.108***	-3.964***
	(0.840)	(0.698)	(0.153)	(4.869)	(0.869)	(0.360)
Reserved SC	0.842***	0.836***	0.838***	1.113***	0.847***	0.996***
	(0.258)	(0.208)	(0.233)	(0.210)	(0.268)	(0.235)
Reserved ST	1.178***	1.172***	1.172***	1.382***	1.169***	1.233***
	(0.266)	(0.186)	(0.108)	(0.159)	(0.269)	(0.229)
Rainfall	-0.018	-0.019	-0.018	-0.037	-0.037	-0.033
	(0.066)	(0.065)	(0.071)	(0.040)	(0.073)	(0.082)
Broadcast	0.069	0.067	0.068	-0.059	0.058	0.032
	(0.162)	(0.200)	(0.174)	(0.105)	(0.152)	(0.164)
Education					0.405***	0.376***
					(0.095)	(0.098)
Crime					0.353***	0.366***
					(0.079)	(0.065)
Missing educ/crime					-0.011	0.028
					(1.565)	(0.123)
Assets						0.258
						(0.427)
Missing assets						0.261
						(1.829)

Cont'd on next page

Table A.25 cont'd

	(1)	(2)	(3)	(4)	(5)	(6)
Nonlinear parameters (Σ)						
Female	4.035**	3.964**	4.034*		3.844*	4.698**
	(1.823)	(1.795)	(2.169)		(1.965)	(1.826)
Age	7.823***	7.827***	7.791***		6.781***	4.263**
	(1.484)	(1.681)	(2.012)		(1.774)	(1.888)
Minority	0.018	0.016	0.017		-0.219	0.019
	(11.122)	(11.186)	(11.297)		(3.651)	(15.156)
Constant	0.215	-0.185		-0.058	0.200	0.122
	(5.258)	(5.142)		(10.975)	(5.581)	(7.756)
NOTA	0.150	-0.096		0.178	0.163	-0.034
	(4.552)	(8.222)		(27.380)	(4.561)	(6.874)
Ran		-0.229				
		(4.499)				
Won		-0.008				
		(30.999)				
INC			-0.046	-0.004		
			(20.750)	(32.247)		
BJP			0.040	0.043		
			(12.852)	(38.006)		
Indep.				-0.010		
				(39.787)		
Small				-0.103		
				(25.577)		
Education					0.015	
					(13.509)	
Crime					0.008	
					(35.389)	
Missing educ/crime					-0.100	
					(26.58)	
Assets						-0.091
						(11.361)
Missing assets						0.072
						(40.091)
J	89.578	90.129	89.961	188.566	81.178	99.934
df	9	7	9	8	6	7
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Newey-West D	18.351	17.892	18.587	0.043	15.708	8.272
p-value	0.003	0.013	0.002	1.000	0.047	0.309

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

Table A.26: Parameter estimates using Normally distributed random coefficients, first instrument set

	(1)	(2)	(3)	(4)
Linear parameters				
Female	-0.355	-0.357	-0.354	0.822**
	(1.431)	(1.765)	(1.854)	(0.386)
Age	3.631	3.651	3.642	3.095
	(3.511)	(3.806)	(3.990)	(2.093)
Minority	-4.018***	-4.015**	-4.015***	-4.205***
	(1.480)	(1.577)	(1.295)	(0.891)
Ran	0.260**	0.260*	0.260**	0.284***
	(0.132)	(0.149)	(0.105)	(0.101)
Won	0.593***	0.592***	0.592**	0.565***
	(0.111)	(0.124)	(0.257)	(0.111)
NOTA	-4.730	-4.704*	-4.700***	-4.700
	(8.543)	(2.562)	(0.375)	(11.195)
Reserved SC	3.321*	3.318*	3.317**	3.468***
	(1.744)	(1.905)	(1.301)	(0.750)
Reserved ST	3.492**	3.489**	3.489***	3.671***
	(1.670)	(1.752)	(1.238)	(0.698)
Rainfall	-0.064	-0.064	-0.064	-0.063
	(0.083)	(0.067)	(0.050)	(0.053)
Broadcast	-0.001	0.000	-0.001	0.018
	(0.127)	(0.240)	(0.212)	(0.138)
Nonlinear parameters (Σ)				
Female	2.964	2.961	2.961	
	(1.854)	(1.930)	(2.498)	
Age	1.625	1.649	1.639	
	(7.882)	(6.224)	(4.600)	
Minority	0.044	0.043	0.044	
	(21.648)	(30.731)	(16.593)	
Constant	0.014	-0.007		-0.091
	(47.325)	(50.461)		(15.336)
NOTA	-0.246	0.004		0.178
	(36.242)	(93.134)		(64.525)
Ran		-0.066		
		(21.434)		
Won		0.083		
		(35.835)		
INC			-0.028	-0.001
			(39.315)	(50.013)
BJP			-0.012	0.024
			(30.408)	(66.445)
Indep.				-0.065
				(50.922)
Small				-0.101
				(31.923)
J	64.127	63.967	64.135	77.391
df	6	4	6	5
p-value	0.000	0.000	0.000	0.000
Newey-West D	1.916	1.917	1.916	0.028
p-value	0.861	0.964	0.861	1.000

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

Table A.27: Parameter estimates using Normally distributed random coefficients, third instrument set

	(1)	(2)	(3)	(4)
Linear parameters				
Female	-2.268**	-2.260*	-2.262*	0.411**
	(1.112)	(1.233)	(1.223)	(0.203)
Age	5.813***	5.783**	5.782**	1.814**
	(2.093)	(2.435)	(2.261)	(0.873)
Minority	-0.466	-0.473	-0.473	-1.064***
	(0.410)	(0.511)	(0.445)	(0.211)
Ran	0.203**	0.203	0.203**	0.314***
	(0.097)	(0.150)	(0.100)	(0.065)
Won	0.589***	0.589***	0.589***	0.642***
	(0.099)	(0.171)	(0.123)	(0.072)
NOTA	-3.906*	-3.904**	-3.901***	-3.702
	(2.283)	(1.573)	(0.148)	(3.927)
Reserved SC	0.233	0.240	0.241	0.887***
	(0.412)	(0.480)	(0.437)	(0.179)
Reserved ST	0.496	0.501	0.501	1.117***
	(0.359)	(0.462)	(0.384)	(0.187)
Rainfall	-0.001	-0.002	-0.002	-0.026
	(0.055)	(0.058)	(0.058)	(0.037)
Broadcast	-0.046	-0.047	-0.048	-0.083
	(0.142)	(0.201)	(0.146)	(0.090)
Nonlinear parameters (Σ)				
Female	3.964***	3.964***	3.968**	
	(1.310)	(1.390)	(1.516)	
Age	6.054***	6.021***	6.019***	
	(1.855)	(1.785)	(1.883)	
Minority	0.003	0.016	0.001	
	(9.377)	(10.664)	(9.429)	
Constant	0.165	0.038		-0.074
	(6.818)	(9.043)		(8.432)
NOTA	-0.050	-0.018		0.127
	(50.438)	(33.687)		(31.699)
Ran		-0.037		
		(19.638)		
Won		0.082		
		(16.773)		
INC			-0.040	-0.001
			(14.164)	(21.240)
BJP			0.015	0.026
			(11.672)	(33.256)
Indep.				0.012
				(22.518)
Small				-0.097
				(21.675)
J	192.940	193.666	193.748	249.125
df	12	10	12	11
p-value	0.000	0.000	0.000	0.000
Newey-West D	22.161	22.242	22.199	0.054
p-value	0.000	0.002	0.000	1.000

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

5.2 Specifications using voter demographics

Table A.28 introduces the interactions with voter demographics ($\mathbf{\Sigma} = \mathbf{0}$, $\mathbf{\Pi} \neq \mathbf{0}$). Columns (1) - (3) differ in the instrument sets used for identification. We find that specifications (1) and (2) perform much better than the Normal random coefficients specifications above. They pass the J-test for the validity of the moment conditions, and the Newey and West D-test always rejects the null that the nonlinear parameters included in the specifications are jointly 0. The third instrument set yields somewhat different patterns for mean voter valuations and heterogeneity, but the J-test rejects this specification. In the counterfactual exercise we use column (2) as our preferred specification.

Columns (4) and (5) include as additional candidate characteristics education, criminal history and assets. For the counterfactual exercises below we also present results using the specification in column (5) as a robustness check.

5.3 Criminal histories involving serious crimes

In the paper, we find evidence of voter preference for candidates with a criminal history. As described in section 7 in the paper, these findings are consistent with Vaishnav's (2017) argument that criminality, and serious crimes in particular, signal to voters that a candidate is willingness to do "whatever it takes" to protect the interests of political supporters. To further probe the correspondence with Vaishnav's findings, we use information on crimes coded by ADR as being "serious." For the time period we consider, these involve crimes that satisfy any of the following criteria: the maximum punishment is at least five years in jail; no possibility of bail; pertains to an electoral violation, a loss to the exchequer, assault, murder, kidnapping, rape, or a crime against women; the offence is mentioned in the Representation of the People Act (Section 8) or the Prevention of Corruption Act (Vaishnav, 2017, p320).[24] In our data, 7.8 percent of candidates have a criminal history involving a serious crime (or just over half of all candidates with a criminal history).

Table A.29 presents estimates corresponding to column (4) and (5) of Table A.28, but use as an additional candidate characteristic a variable equal to 1 if the candidate's criminal history includes at least one serious crime. In both specifications, the Serious crime variable is positive and statistically significant, while the original Crime variable is positive but much smaller and statistically insignificant. Consistent with Vaishnav's argument, voters appear to value criminal histories involving serious crimes rather than all criminal histories.

[24] In his work, Vaishnav also uses an alternative definition of "serious crimes" but he only created this variable for the period 2003-2009.

Table A.28: Parameter estimates of the full model using voter demographics

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

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

Table A.29: Parameter estimates using voter demographics, with serious crimes

	(1)	(2)
Linear parameters		
Female	7.251***	7.387***
	(2.387)	(2.523)
Age	-0.078	-1.363
	(3.282)	(3.267)
Minority	-5.082***	-5.123***
	(0.767)	(0.812)
Ran	0.201	0.219
	(0.175)	(0.173)
Won	0.534***	0.629***
	(0.170)	(0.159)
NOTA	-3.951***	-4.042***
	(0.200)	(0.206)
Reserved SC	3.761***	3.769***
	(0.544)	(0.562)
Reserved ST	1.652***	1.679***
	(0.274)	(0.272)
Rainfall	-0.128	-0.109
	(0.085)	(0.080)
Broadcast	-0.256	-0.265
	(0.263)	(0.258)
Education	0.458***	0.479***
	(0.113)	(0.118)
Crime	0.079	0.115
	(0.142)	(0.133)
Serious crime	0.361**	0.310*
	(0.176)	(0.170)
Missing educ/crime	-0.002	0.041
	(0.128)	(0.165)
Assets		-0.855*
		(0.459)
Missing assets		-1.318**
		(0.657)
Nonlinear parameters (Π)		
Female x Minority pop.	-5.813	-7.375
	(4.564)	(5.069)
Female x Literacy	-14.625**	-14.750**
	(6.838)	(6.455)
Age x Minority pop.	17.250***	16.750***
	(2.668)	(2.710)
Age x Rural workers	8.750***	8.125***
	(2.610)	(2.696)
NOTA x Rural workers	-0.029	-0.045
	(0.439)	(0.440)
J	9.532	14.054
df	9	9
p-value	0.390	0.120
Newey-West D	35.449	47.680
p-value	0.000	0.000

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

5.4 Further specifications and corresponding counterfactuals

Table A.30 contains further estimation results for specifications with voter demographics. Columns (1) - (4) include different nonlinear parameters, and column (5) is for the preferred specification discussed in the paper but excluding the state of Mizoram.

Figure A.6 and Table A.31 describe the results from the counterfactual (no-NOTA) exercise using the various specifications. Results from the preferred specification discussed in the text are in column (1).

Table A.32 shows estimates of the preferred specification when the aggregation of "small parties" is based on the alternative definitions discussed in section 4.2 above (using cutoffs of 1/2 or 1/4). Tables A.33-A.35 present the counterfactual results corresponding to these estimates. These alternative definitions cause little change in the results. The share of protest voters who normally abstain remains around 2/3 (Table A.33), and protest voters make up small shares of any given party's supporters (Tables A.34 and A.35). When the smaller threshold for aggregation is used in Table A.35, we find that the party from which voters are relatively most likely to switch to NOTA is the SHS or "Army of Shivaji" (with the 1/3 threshold this party was part of the Small party category). This is a far-right nationalist party advocating preferential treatment for the Marathi ethnic group and intolerance towards others, especially the non-Hindi. It has been associated with a number of violent ethnic riots.[25]

Figure A.7 shows the geographic distribution of the counterfactual results (the share of the NOTA voters who would abstain without NOTA). There do not seem to be any obvious geographic patterns in the behavior of NOTA voters.

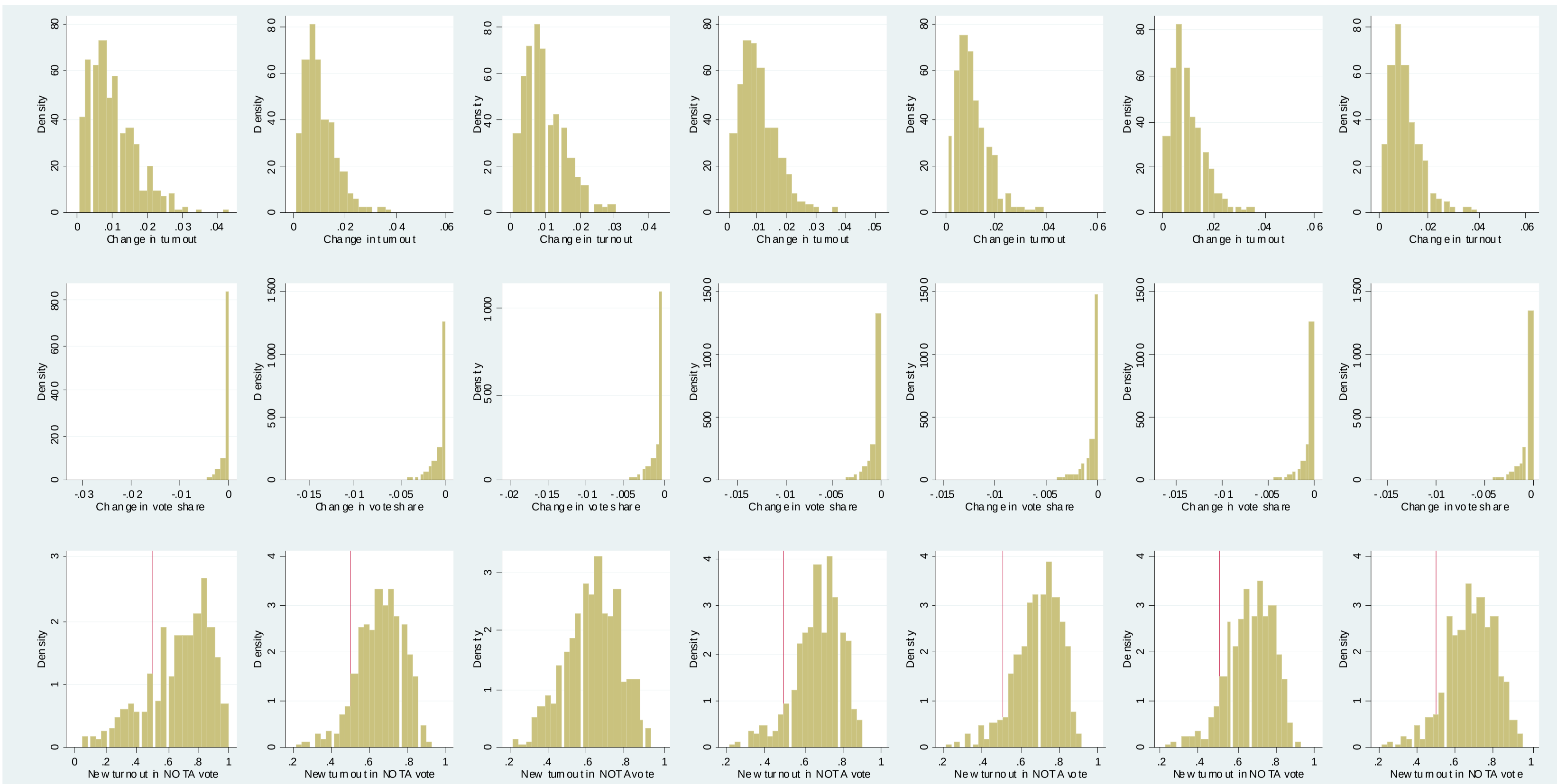
[25] http://www.elections.in/political-parties-in-india/shiv-sena.html

Table A.30: Parameter estimates using voter demographics, additional specifications

	(1)	(2)	(3)	(4)	(5)
Linear parameters					
Female	6.980***	8.852***	6.926***	7.253***	6.677***
	(2.289)	(3.135)	(2.339)	(2.408)	(2.340)
Age	-1.201	-10.822	-0.825	-0.603	-0.828
	(3.086)	(8.554)	(3.287)	(3.251)	(3.608)
Minority	-5.083***	-3.750***	-5.405***	-5.435***	-6.071***
	(0.784)	(0.890)	(0.842)	(0.825)	(2.111)
Ran	0.216	0.149	0.243	0.221	0.291
	(0.172)	(0.182)	(0.190)	(0.181)	(0.295)
Won	0.562***	0.587***	0.544***	0.549***	0.542***
	(0.167)	(0.188)	(0.182)	(0.178)	(0.186)
NOTA	-3.960***	-3.841***	-3.910***	-3.424***	-4.024***
	(0.199)	(0.197)	(0.234)	(0.509)	(0.258)
Reserved SC	3.741***	2.666***	3.929***	4.006***	4.503***
	(0.563)	(0.691)	(0.611)	(0.592)	(1.688)
Reserved ST	1.641***	1.397***	1.444***	1.753***	1.791***
	(0.275)	(0.217)	(0.509)	(0.309)	(0.413)
Rainfall	-0.125	-0.151	-0.141	-0.137	-0.141
	(0.080)	(0.100)	(0.090)	(0.087)	(0.088)
Broadcast	-0.208	-0.264	-0.280	-0.266	-0.315
	(0.258)	(0.279)	(0.284)	(0.275)	(0.348)
Nonlinear parameters (pi)					
Female x Minority pop.	-5.406	-10.658	-4.438	-4.750	-3.609
	(4.191)	(7.566)	(4.120)	(4.191)	(4.780)
Female x Literacy	-13.938**	-14.516**	-13.594*	-14.625**	-12.875*
	(6.639)	(5.791)	(7.024)	(7.204)	(7.224)
Age x Minority pop.	17.328***	11.258***	20.063***	18.125***	21.922*
	(2.814)	(4.198)	(4.016)	(2.828)	(11.326)
Age x Rural workers	8.625***	27.484***	10.063***	10.063***	8.922***
	(2.557)	(17.000)	(2.698)	(2.651)	(3.181)
NOTA x Minority pop.			0.998		
			(1.299)		
NOTA x Rural workers		8.076			-0.102
		(6.917)			(0.538)
NOTA x Literacy				-0.820	
				(0.784)	
Constant x Rural worker	-0.047	-7.980	-0.156	-0.406	
	(0.448)	(6.940)	(0.495)	(0.637)	
J	8.204	11.266	9.613	8.892	11.387
df	9	8	8	8	8
p-value	0.514	0.187	0.294	0.352	0.1807
Newey-West D	17.900	17.854	38.156	34.162	21.559
p-value	0.003	0.007	0.000	0.000	0.000

Notes: Parameter estimates from the BLP model using voter demographics ($\Pi \neq 0$). Second instrument set. Column (5) excludes the state of Mizoram. The linear parameters also include indicators for parties, states, and years. Standard errors robust to heteroskedasticity and intra-constituency correlation in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively. J is the overidentification test statistic with corresponding degrees of freedom and p-value. Newey-West D is a likelihood ratio test for the null hypothesis that the nonlinear parameters are jointly 0 with the corresponding p-value. N = 9831 in columns (1) - (4), N = 9720 in column (5).

Figure A.6: Counterfactual results from different specifications



Notes: Each column presents the simulated impact of NOTA from a different specification. Graphs in the first row show changes in turnout across constituencies. Graphs in the second row show the change in vote shares (as a fraction of eligible voters) across individual candidates. Graphs in the third row show the share of the NOTA voters in a constituency who would abstain without NOTA (vertical lines are at 0.5). From left to right, the specifications are for column 1, Table A.28 (first instrument set), columns 1-4, Table A.30 (alternative specifications), column 5, Table A.28 (education, criminal history and assets), and column 5, Table A.30 (no Mizoram). For each column except the last, N = 520, N = 3073, and N = 520 for rows 1-3, respectively. For the last column the corresponding numbers are N = 507, N = 3031, and N = 507.

Table A.31: Counterfactual results from different specifications

Specification (Table, column)	A.28(2)	A.28(1)	A.30(1)	A.30(2)	A.30(3)	A.30(4)	A.28(5)	A.30(5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Change in turnout (ppoint)	1.075	1.057	1.047	0.982	1.067	1.086	1.053	1.097
Standard deviation	0.700	0.734	0.688	0.610	0.667	0.694	0.687	0.722
Change in candidate vote shares (ppoint)	-0.084	-0.087	-0.088	-0.099	-0.085	-0.082	-0.087	-0.084
Standard deviation	0.140	0.204	0.146	0.167	0.145	0.136	0.145	0.144
Largest change in candidate vote share (ppoint)	-0.280	-0.343	-0.295	-0.326	-0.287	-0.271	-0.292	-0.291
Standard deviation	0.217	0.379	0.223	0.262	0.225	0.209	0.221	0.225
Share of NOTA vote due to new turnout	0.679	0.668	0.660	0.628	0.678	0.688	0.664	0.679
Standard deviation	0.121	0.207	0.119	0.139	0.123	0.120	0.119	0.125
Elections where winner changes	2	2	2	2	2	2	2	2

Notes: Means and standard deviations of the simulated impact of NOTA from different specifications. Column (1) is the preferred specification discussed in the paper. Columns (2) - (8) correspond to the graphs in Figure A.6 above (2: different instruments, 3-6: different demographic interactions, 7: education, criminal history and assets, 8: excluding the state of Mizoram). See notes to Figure A.6 for details.

Table A.32: Parameter estimates using alternative aggregation thresholds for small parties

	(1)	(2)
Linear parameters		
Female	5.913***	6.124***
	(2.033)	(2.215)
Age	-1.798	1.078
	(2.739)	(3.353)
Minority	-5.159***	-5.055***
	(0.693)	(0.833)
Ran	0.500***	0.198
	(0.158)	(0.168)
Won	0.595***	0.483***
	(0.147)	(0.178)
NOTA	-3.940***	-3.873***
	(0.189)	(0.200)
Reserved SC	3.866***	3.710***
	(0.507)	(0.598)
Reserved ST	1.777***	1.594***
	(0.275)	(0.284)
Rainfall	-0.119	-0.130
	(0.073)	(0.088)
Broadcast	-0.526**	-0.068
	(0.248)	(0.265)
Nonlinear parameters (Π)		
Female x Minority pop.	-2.789	-0.758
	(3.000)	(5.127)
Female x Literacy	-11.188*	-12.625*
	(5.890)	(7.152)
Age x Minority pop.	16.188***	17.469***
	(3.200)	(5.127)
Age x Rural workers	7.500***	9.531***
	(2.218)	(2.677)
NOTA x Rural workers	-0.236	0.063
	(0.464)	(0.425)
J	11.539	13.0895
df	9	9
p-value	0.2405	0.1586
Newey-West D	48.590	39.308
p-value	0.000	0.000

Notes: Parameter estimates from the BLP model using voter demographics ($\Pi \neq 0$). Second instrument set. The aggregation threshold for small parties is 1/2 in column (1) and 1/4 in column (2). The linear parameters also include indicators for parties, states, and years. Standard errors robust to heteroskedasticity and intra-constituency correlation in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively. J is the overidentification test statistic with corresponding degrees of freedom and p-value. Newey-West D is a likelihood ratio test for the null hypothesis that the nonlinear parameters are jointly 0 with the corresponding p-value. N = 9342 in column (1) and 10359 in column (2).

Table A.33: Counterfactual results for different aggregation thresholds for small parties

Specification	Table A.32(1) (1)	Table A.32(2) (2)
Change in turnout (ppoint)	1.034	1.054
Standard deviation	0.686	0.691
Change in candidate vote shares (ppoint)	-0.095	-0.080
Standard deviation	0.151	0.142
Largest change in candidate vote share (ppoint)	-0.300	-0.293
Standard deviation	0.223	0.226
Share of NOTA vote due to new turnout	0.650	0.666
Standard deviation	0.117	0.121
Elections where winner changes	2	2

Notes: Means and standard deviations of the simulated impact of NOTA obtained from the specifications in Table A.32. The aggregation threshold for small parties is 1/2 in column (1) and 1/4 in column (2).

Table A.34: Impact of NOTA on vote shares by party with 1/2 aggregation threshold for small parties

Choice	N. of candidates	Elections won	Percent of all voters	Change due to NOTA Full model (percentage points)	Full model (percent)
	(1)	(2)	(3)	(4)	(5)
BJP	507	361	32.910	-0.221	-0.671
BSP	499	8	3.623	-0.030	-0.842
BYS	102	0	0.102	-0.002	-1.532
CSM	54	0	0.223	-0.004	-1.830
INC	519	127	26.732	-0.185	-0.692
Independents	469	15	4.963	-0.048	-0.966
MNF	10	1	0.057	0.000	-0.246
NPEP	133	4	1.294	-0.010	-0.756
NOTA			1.578		
SP	154	0	0.395	-0.002	-0.618
Small parties	458	4	2.993	-0.032	-1.069
ZNP	11	0	0.016	0.000	-0.259
Abstention			25.114	-1.044	-4.157

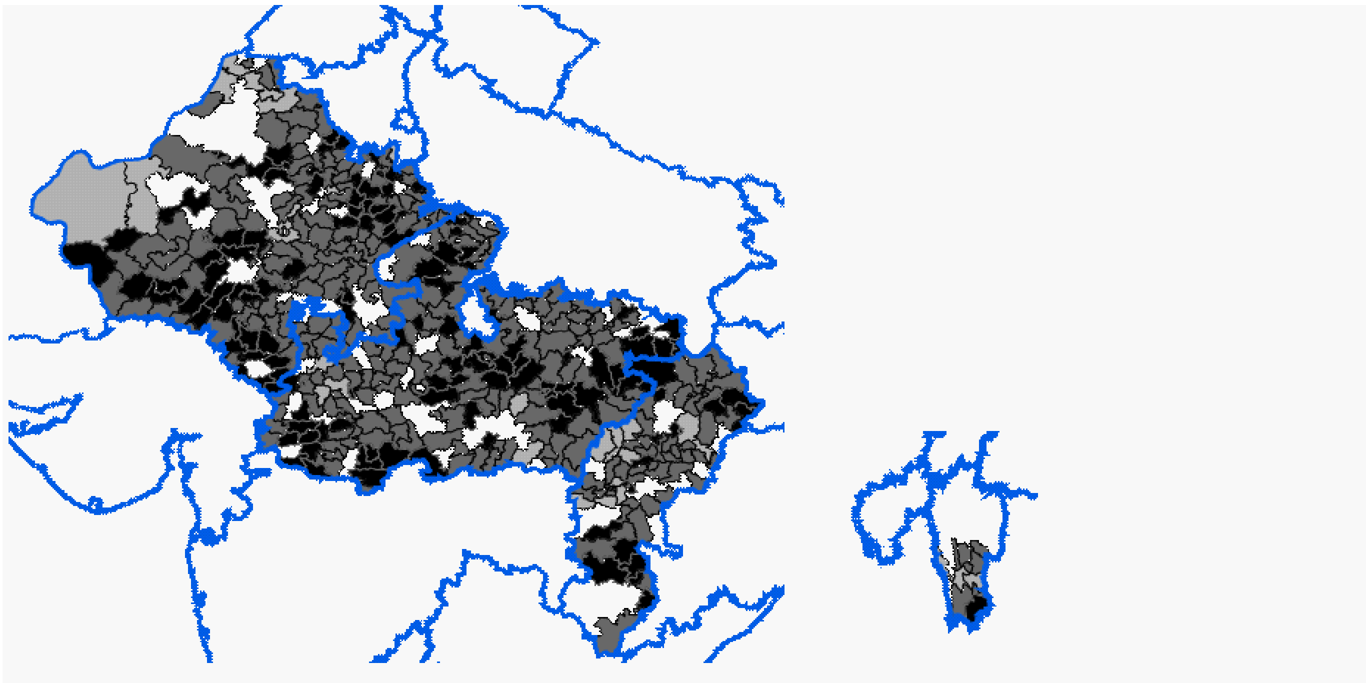
Notes: Tabulation of all the choices available in the data used for the counterfactual exercise. For each party column (1) shows the total number of candidates and (2) the number of constituencies won. Column (3) is the share of all voters (out of 101.384 million eligible voters) choosing each option in the data. Column (4) gives the simulated effect of introducing NOTA in the full model, (5) is (4) divided by (3), times 100.

Table A.35: Impact of NOTA on vote shares by party with 1/4 aggregation threshold for small parties

Choice	N. of candidates	Elections won	Percent of all voters	Change due to NOTA Full model (percentage points)	Full model (percent)
	(1)	(2)	(3)	(4)	(5)
BASD	62	0	0.163	-0.001	-0.570
BJP	507	362	32.910	-0.212	-0.645
BSP	499	8	3.623	-0.030	-0.833
BYS	102	0	0.102	-0.002	-1.496
CSM	54	0	0.223	-0.004	-1.875
GGP	106	0	0.536	-0.007	-1.346
INC	519	127	26.732	-0.177	-0.661
Independents	469	15	4.963	-0.045	-0.915
JGP	85	0	0.060	-0.001	-1.237
MNF	10	1	0.057	0.000	-0.225
NCP	68	0	0.096	-0.001	-1.092
NPEP	159	4	1.331	-0.010	-0.776
NOTA			1.578		
SHS	27	0	0.037	-0.001	-2.279
SP	250	0	0.550	-0.004	-0.661
Small parties	406	3	1.907	-0.019	-0.972
ZNP	11	0	0.016	0.000	-0.256
Abstention			25.277	-1.065	-4.214

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.

Figure A.7: Geographic distribution of the share of NOTA voters who abstain without NOTA



Notes: Share of NOTA voters who abstain when NOTA is removed based on the main counterfactual analysis in the paper. Light grey: <50 %, dark grey: 50-75 %, black: >75 %.

5.5 Heterogeneity among NOTA voters

Table A.36 further explores the heterogeneity in the simulated behavior of protest voters. What type of NOTA voters substitute to abstention, major parties, or non-major parties when NOTA is not available? To explore this question, we use the 1000 simulated voters in the 520 constituencies in our counterfactual exercise. We restrict attention to those voters who have a non-zero probability of choosing NOTA (with an 8-decimal precision), and we drop 1 constituency which only had major party candidates running, yielding 515,755 observations. In Table A.36, we regress the probability of substituting to abstention, major parties (INC or BJP), or non-major parties in the counterfactual, conditional on choosing NOTA. (Since we account for all possible substitution opportunities, the coefficients in the 3 regressions sum to 0.)

According to these correlations, using abstention as a substitute for NOTA is more likely among literate voters, while rural worker protest voters are more likely to choose one of the candidates when NOTA is not available. Abstention among NOTA voters is also more likely in reserved constituencies.

Comparing reserved and general constituencies, we see that general caste NOTA voters are more likely to abstain and less likely to vote for one of the candidates in the former. For example, general caste voters are 24.3 percentage points more likely to abstain and 18 (6.3) percentage points less likely to choose a major (non-major) party candidate in ST-reserved constituencies. Minority voters behave differently: their propensity to abstain is relatively weaker in reserved constituencies, and their likelihood of choosing a candidate is relatively higher. For example, minority voters are only 12 percentage points ($= 0.243 - 0.123$) more likely to abstain in ST-reserved constituencies than in general constituencies, and are no less likely to choose minor party candidates. In SC-reserved constituencies, minority NOTA voters are actually weakly more likely to choose minor party candidates (by $-0.029 + 0.039 =$ 1 percentage point). These patterns are consistent with the idea that reserved constituencies offer fewer opportunities for general caste voters to obtain utility by voting for a candidate.

Table A.36: Heterogeneity in the behavior of simulated NOTA voters

Dep. Var.:	Probability of NOTA voters switching to		
	Abstention	Major parties	Other parties
Literate	0.054***	-0.047***	-0.007
	(0.007)	(0.009)	(0.006)
Rural worker	-0.388***	0.307***	0.081***
	(0.010)	(0.010)	(0.006)
Minority	-0.341***	0.349***	-0.008
	(0.018)	(0.017)	(0.014)
Reserved SC	0.044***	-0.015	-0.029**
	(0.016)	(0.021)	(0.015)
Reserved ST	0.243***	-0.180***	-0.063***
	(0.025)	(0.025)	(0.013)
Minority x Reserved SC	-0.001	-0.038	0.039
	(0.028)	(0.034)	(0.025)
Minority x Reserved ST	-0.123***	0.060*	0.063***
	(0.031)	(0.033)	(0.020)
Size	0.120**	0.017	-0.137***
	(0.049)	(0.058)	(0.039)

Notes: Regressions of the probability of NOTA voters' switching to abstention, major parties (INC or BJP) or other parties in each constituency in the no-NOTA counterfactual. Size is log(number eligible voters). Each regression includes state fixed effects. Standard errors clustered by constituency in parentheses. ***, **, and * indicates significance at 1, 5, and 10 percent, respectively. N=520.

5.6 Additional results on compulsory voting with and without NOTA

This section presents further details on the counterfactual results with compulsory voting discussed in the paper. Table A.37 summarizes the distribution of counterfactual results across constituencies. When NOTA is available, compulsory voting increases its votes share by 6.8 percent of eligible voters in the average constituency. Compulsory voting increases the share of NOTA votes among votes cast by a factor of four (6.3 percentage points). Some candidates also experience large losses in vote shares: as a fraction of votes cast, the largest loss in the average constituency is 6.7 percentage points.

The lower half of Table A.37 presents corresponding results when compulsory voting is introduced in an environment without NOTA. We find that compulsory voting leads to similar losses in vote shares as it did with NOTA, but the gains are now more concentrated, and some candidates see a large increase in their vote share (the largest gain is 5.3 percentage points on average).

Table A.37: The impact of compulsory voting with and without NOTA

	Mean	Std. Dev.	Median	10%	90%	N
With NOTA						
Increase in NOTA votes (fraction of eligible voters)	0.068	0.045	0.061	0.016	0.126	520
Increase in NOTA votes (fraction of votes cast)	0.063	0.043	0.057	0.012	0.120	520
Largest drop in candidate's vote share (fraction of votes cast)	-0.067	0.029	-0.063	-0.106	-0.034	520
Largest increase in candidate's vote share (fraction of votes cast)	0.023	0.024	0.016	0.002	0.053	520
Election overturned (0/1)	0.150					520
Without NOTA						
Largest drop in candidate's vote share (fraction of votes cast)	-0.062	0.032	-0.058	-0.106	-0.022	520
Largest increase in candidate's vote share (fraction of votes cast)	0.053	0.033	0.046	0.017	0.101	520
Election overturned (0/1)	0.225					520

Notes: Results from a counterfactual simulation removing the possibility of abstention when NOTA is available (upper panel), and when it is not (lower panel). Election overturned is equal to 1 if removing abstention changes the winner of the election.

Table A.38 shows the impact of compulsory voting on elections ignoring counterfactual wins by aggregated small party or independent candidates. With NOTA, compulsory voting changes the winner in 13.8% of elections. With NOTA, this figure is 21%. The BJP is a net loser and the INC a net winner in both cases, with larger differences when NOTA is not available.

Table A.38: The impact of compulsory voting on parties ignoring aggregated candidates

Choice	Change due to compulsory voting			
	with NOTA		without NOTA	
	Extra wins	Extra losses	Extra wins	Extra losses
BJP	18	46	28	71
BSP	7	3	13	3
BYS	0	0	0	0
CSM	0	0	0	0
GGP	0	0	0	0
INC	45	19	64	30
Independents	0	2	0	2
JGP	0	0	0	0
MNF	0	0	1	0
NPEP	0	1	1	1
NOTA	1	0	0	0
SP	0	0	0	0
Small parties	0	0	0	0
ZNP	0	0	0	0
Total	71	71	107	107
Share of all	13.8	13.8	21.0	21.0

Notes: Number of additional constituencies won and lost by each party as a result of compulsory voting, with or without NOTA, Counterfactual wins by aggregated small party or independent candidates are ignored, leaving 513 elections in the with-NOTA and 510 in the without-NOTA case.

6 Costs and benefits of NOTA

Our main counterfactual exercise implies that NOTA reduced abstention by 4.7 *percent* in the average constituency. According to estimates available in the literature, even highly personalized interventions like phone calls to voters typically reduce abstention by under 10 percent (Nickerson, 2006). In Gerber et al. (2008), informing voters that their voting behavior will be scrutinized by researchers reduced abstention by 3.6 percent; threatening to publicly reveal whether or not they voted reduced abstention by an additional 8.2 percent. In the Indian context, Banerjee et al. (2011) find a turnout increase of 3.5 percent from providing households with information on legislators' responsibilities and their performance, and George et al. (2018) a 1.6 percent increase from voice and text messages informing voters about candidates' criminal history. Compared to these more costly interventions, the impact of NOTA seems remarkably large in a setting where abstention rates were low to begin with.

What were the costs of NOTA? Including a NOTA option on an existing voting machine simply involves labeling one of the buttons, exactly as would be done if a new candidate was added to the ballot. The cost of this is negligible. The direct costs of NOTA are higher if the voting machine has to be modified. The voting machines used in India have two parts, a Control Unit, which is operated by the election official to authorize a vote to be cast, and one or more Balloting Units, on which the actual votes are cast.[26] Each balloting unit has buttons for 16 different candidates, and each control unit can operate up to 4 balloting units. This means that if the number of candidates before NOTA is either 16 or 32, a new balloting unit has to be linked to an existing control unit in order to accommodate the NOTA option. If the number of candidates before NOTA is 64, adding NOTA requires both a new control unit and a new balloting unit.

In the 520 constituencies in the counterfactual exercise, the highest number of candidates is 38 so introducing NOTA never requires a new control unit. The number of candidates is 16 in 13 constituencies and 32 in 2 constituencies. Thus, a possible estimate of the direct cost of NOTA in this sample is the cost of 13 + 2 =15 new balloting units.

While the cost of a separate balloting unit is unknown, as of 2014 the cost of one control unit plus one balloting unit was estimated at \$175.[27] Assuming that the cost of a separate balloting unit is half of that, the direct cost of introducing NOTA in this sample would be $175/2 \times 15 = 1312.50$ dollars, or about \$13 for every 1 million eligible voters.

Based on this calculation and the paper's results, NOTA appears to have fulfilled the Supreme Court's stated goal of increased voter participation at very low cost. In most cases, NOTA generated consumption utility for voters without affecting the winner of the election.[28]

Of course, NOTA may also have externalities, both negative and positive. On the negative side, the several millions of extra voters showing up at the polls may have created extra costs in terms of election administration, or increased the waiting time for other voters. On the positive side, increased civic participation in elections may spill over to improved civic participation in other areas. In the long run, it may help hold politicians accountable,

[26] http://eci.nic.in/eci_main1/evm.aspx

[27] https://www.theatlantic.com/international/archive/2014/04/indian-democracy-runs-on-briefcase-sized-voting-machines/360554/

[28] The welfare calculation would be very different if NOTA had overturned more election results.

or crowd out less desirable forms of political expression (e.g., violence). Quantifying these effects is beyond the scope of this paper, but some of these possibilities would be interesting to explore in future research.

7 Using NOTA to identify an option-specific consumption utility from voting

In the "calculus of voting" model (Downs, 1957; Riker and Ordeshook, 1968), voters consider both instrumental and consumption benefits. They vote for candidate j if

$$P_j B_j + (U_j + U_0) > c \tag{4}$$

and abstain otherwise, where $j \in \arg\max_{j'}(P_{j'} B_{j'} + U_{j'})$ is (one of) the voter's preferred candidate(s). The first term on the left-hand side of (4) is the expected instrumental benefit, where P_j is an individual's probability of being pivotal in the election of candidate j and B_j is the benefit of the candidate winning. The second term is the consumption utility of voting, which captures a wide range of factors sometimes referred to as "expressive utility" or "civic duty": "1. the satisfaction from compliance with the ethic of voting [...] 2. the satisfaction from affirming allegiance to the political system [...] 3. the satisfaction from affirming a partisan preference [...] 4. the satisfaction of deciding, going to the polls, etc. [...] 5. the satisfaction of affirming one's efficacy in the political system" (Riker and Ordeshook, 1968, p28). Voters who show up at the polls are assumed to derive utility from some or all of these factors, which we separate into two groups. Part of the utility (U_j) depends on voting for the specific candidate j (e.g., the satisfaction from expressing partisan support), while part of it (U_0) only depends on showing up at the polls regardless of who one votes for (e.g., satisfaction from compliance with an ethical norm to vote). Finally, on the right-hand side of (4) c represents any direct or opportunity costs from voting.

The NOTA policy creates an option that voters can vote for but that, by design, cannot affect the electoral outcome. Because $P_{NOTA} B_{NOTA} = 0$, from equation (4) a voter who chooses NOTA must have

$$U_{NOTA} + U_0 > c, \tag{5}$$

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

In the paper, we study the counterfactual behavior of NOTA voters, and find that approximately 2/3 of them would abstain if NOTA was not available. We have the following result.

Proposition 1 *Voters who choose NOTA when it is available and abstain when it is not must derive an option-specific consumption utility from voting: it cannot be that* $U_{NOTA} = U_{j'} = 0$ *for all candidates* j'.

Proof. As above, let $j \in \arg\max_{j'}(P_{j'}B_{j'} + U_{j'})$ denote the voter's favorite candidates(s). Suppose that $U_{j'} = 0 \ \forall j'$, so that $j \in \arg\max_{j'} P_{j'}B_{j'}$. Since B is a difference, there have to be some candidates with $P_{j'}B_{j'} \geq 0$, and therefore $P_jB_j \geq 0$.

Consider now the voter described in the proposition. Because the voter chooses NOTA when it is available, inequality (5) must hold. Because without NOTA the voter would have abstained, it must be that $c \geq P_jB_j + (U_j + U_0)$. Combining this with (5), we have

$$U_{NOTA} > P_jB_j + U_j. \tag{6}$$

But if $U_{NOTA} = U_{j'} = 0$ for all j', then (6) would mean $P_jB_j < 0$, a contradiction. ■

Proposition 1 shows that a NOTA voter who would abstain without NOTA (all else equal) must derive a non-zero option-specific consumption utility from voting. For example, the voter may strongly dislike all the candidates ($U_j < 0$) and/or may gain utility from expressing this by voting for NOTA ($U_{NOTA} > 0$). The proposition also implies that the fraction of NOTA voters who would otherwise abstain provides a lower bound on the fraction of voters who are motivated at least in part by an option-specific consumption utility.

References

[1] Banerjee, A.V., S. Kumar, R. Pande, and F. Su (2011): "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India," working paper.

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

[3] Bohrnstedt, G.W., and A.S. Goldberger (1969): "On the Exact Covariance of Products of Random Variables," *Journal of the American Statistical Association* 64(328), 1439-1442.

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

[5] Cameron, C., and P. Trivedi (2005): *Microeconometrics: Methods and Applications.* Cambridge University Press, New York, NY.

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

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

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

[9] Dubé, J.-P., J. Fox and C.-L. Su (2012): "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation," *Econometrica*, 80(5), 2231-2267.

[10] Election Commission of India (2014): *Systematic Voters' Education & Electoral Participation. India National Document (2009-2014)*, United Nations Development Programme.

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

[12] 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.

[13] George, S., S. Gupta, and Y. Neggers (2018): "Coordinating Voters against Criminal Politicians: Evidence from a Mobile Experiment in India," working paper.

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

[15] 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.

[16] Ho, D.E., and K. Imai (2006): "Randomization Inference With Natural Experiments: An Analysis of Ballot Effects in the 2003 California Recall Election," *Journal of the American Statistical Association* 101(475), 888-900.

[17] MacKinnon, J.G., and M.D. Webb (2018): "Randomization Inference for Difference-in-Differences with Few Treated Clusters," working paper, Queen's University.

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

[19] 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.

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

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

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

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

[24] Shue, K., and E.F.P. Luttmer (2009): "Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes," *American Economic Journal: Economic Policy* 1(1), 229-257.

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

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

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