Net Influences from Polling Officers

Kensuke Maeba

Northwestern University

July 28th, 2021

Election Officers Are Not Neutral in India

"Election officers" = public employees deployed to polling stations for election duty

Beteille (2009)

- Qualitative study about schoolteachers as polling officers
- Self-reported that they could fake votes at the polling stations
- Support for “teacher-friendly” candidates

Neggers (2018)

- Estimate the effects of the identity of polling officers
- Muslim/Yadav officers (MY) on their supporting party (RJD)
- $\uparrow$ 5% Votes for RJD if MY officers at polling stations than Non-MY ones
- $\uparrow$ 2.5pp vote share margin between RJD and the opposing party

Net Influences on Election

From the previous literature

1. Polling officers have influences on voting at polling stations
2. Motives vary by polling officers

RQ 1: What are the net influences from polling officers on elections?

- "Net" = Aggregated within and across polling officers
- Influences may internally and externally conflict

RQ 2: Are there distinct patterns in the distribution of influences?

- Across candidates, polling officers, geographic areas
- A proxy for corruption

What This Project Does

We estimate net influences using India's national parliamentary elections in 2019

- Construct a discrete choice model with influences as unobservables
- Utility = candidate characteristics + influences + preference shock
- Identify influences as the residual vote share variation

Empirical Challenges

- Endogenous sorting of polling officers to polling stations
- Other unobservables as confounders

Empirical Approach

Sorting by polling officers

- A natural experiment that polling officers assigned randomly to polling stations
- Orthogonal to underlying local political environment
- Verified by data we receive from a district in Tamil Nadu

Other confounders

- Similar to unobserved product characteristics: voter's unobserved valuation
- A control function based on the assignment of local party workers
- Include locality FEs

Plan of Today's Talk

1. Research context
2. Voting model
3. Next steps

Disclaimer

- **The project is at a very preliminary stage**
- **No estimation results**

Random Assignment of Polling Officers

In 2008, the ECI announced the random drafting of polling officers

- To prevent collusion among polling officers and candidates

In each district, the drafting is done in three step randomization by the DEO

1. Select officers with reservation from government employees
2. Split officers into teams (4 ~ 5 members per team)
3. Assign teams to polling stations

Remarks on Assignment Process

Two rules to guarantee the proper mix of polling officers

- No two members from the same department
- No officers assigned to polling stations in a subdistrict where they work/live/vote
- The DEO should certify those to the ECI

Each step is done by the software distributed from the state election officers

- Observers are present at the 2nd and 3rd step

Final assignment is announced before polling officers leave for duty

Research Context: Pudukkottai District in TN

- Population (2011 census): 1,600,000 ($\geq 80\%$ living in rural areas)
- 4 parliamentary constituencies in the 2019 election
  - 6 subdistricts, each of which has about 200 polling stations

- We obtain the final list of polling officer assignment
  - Subdistricts to work/live/vote, department, designation
  - Names and gender
  - Observe reserved teams

Supplementary Data

- Vote share data in 2009, 2014, 2019
  - Available at polling station level
- Candidate affidavit data in 2019
- Electoral roll data in 2019 (under construction)
  - Complete list of voters registered at each polling station
  - Name, age, gender, father/husband name

Check Randomization

We test whether the drafting process generated exogenous assignment

1. Balance check between assigned and non-assigned officers

$$Y_i = \beta_0 + \beta_1 \text{Assigned}_i + \varepsilon_i$$

2. Correlation between assigned teams and pre-election characteristics

$$Y_{as} = \gamma_a + X'_{as}\beta + \varepsilon_{as},$$

where γ_a is a subdistrict fixed effect.

Assigned vs Non-assigned Officers

Balance check: Assigned vs Non-assigned officers

	Assigned	Non-assigned	Difference	p-val
= 1 if female officer	0.569	0.561	0.008	0.604
= 1 if work in Pudukkottai	0.993	0.989	0.003	0.260
= 1 if live in Pudukkottai	0.745	0.740	0.005	0.704
= 1 if vote in Pudukkottai	0.690	0.686	0.004	0.782
= 1 if Hindus	0.700	0.708	-0.008	0.595
= 1 if Muslim	0.051	0.054	-0.003	0.714
Obs.	6191	1232		

Correlation with Pre-election Variables

Correlation with previous election results

	ln(total votes)	ln(votes for UPA)	ln(votes for NDA)
Number of female officers	0.00531	0.00308	0.00972
	(0.00767)	(0.0100)	(0.0105)
Number of local working officers	0.00796	-0.0255	0.0314
	(0.0358)	(0.0547)	(0.0443)
Number of local living officers	-0.0252	-0.0341	-0.0304
	(0.0169)	(0.0219)	(0.0218)
Number of local voting officers	0.00697	0.00656	0.0115
	(0.0152)	(0.0201)	(0.0198)
Number of Hindu officers	-0.0110	-0.0132	-0.00871
	(0.00839)	(0.0115)	(0.0109)
Number of Muslim officers	0.0144	0.00358	0.0282
	(0.0184)	(0.0235)	(0.0226)
Obs.	1536	1536	1536
Pr(Jointly no diff.)	0.148	0.157	0.267

Note: Robust standard errors in parentheses.
Subdistrict-team-size FEs are included.
*** $p<0.01$ ** $p<0.05$ * $p<0.1$

Voting Model: Overview

Goal: Estimate net influences on voters using vote share data

- Potentially various motives for influences
- Agnostic view of underlying mechanisms

Identification: Residual vote share variation

- Construct a discrete choice model (a la Berry, Levinsohn and Pakes (1995))
- Add a control function for unobserved preferences (a la Olley and Pakes (1996))
- Influences $\equiv$ Model prediction $-$ data

Voting Model: Utility Function

Index: voter i voting at polling station s for candidate k

Consider a discrete choice model where each voter casts a ballot to one candidate

$$U_{isk} \quad = \quad \underbrace{u(x_i, w_k) + \mathrm{BLA}_{sk}}_{\text{Observed}} + \underbrace{\xi_{sk} + \eta_{sk} + \varepsilon_{isk}}_{\text{Unobserved}}.$$

- x_i : a vector of demographics of the voter
- w_k : a vector of the candidate characteristics
- BLA_{sk}: local party workers
- ξ_{sk} : unobserved local political environment
- η_{sk} : polling officer influences
- ε_{isk} : idiosyncratic preference shocks

Vote Share at Polling Station

Given the assumption on ε_{isk}, the probability of choosing candidate k is

$$Pr\left(k = \text{argmax}\ U_{isk}\right) \quad = \quad \frac{\exp\left(u(x_i, w_k) + \text{BLA}_{sk} + \xi_{sk} + \eta_{sk}\right)}{\sum_l \exp\left(u(x_i, w_l) + \text{BLA}_{sl} + \xi_{sl} + \eta_{sl}\right)},$$

Vote share of candidate k at polling station s is

$$\begin{aligned} \mu_{sk} &= \int Pr\left(k = \text{argmax}\ U_{isk}\right) dF_{x,s}\left(x_{is}\right), \\ &= \int \frac{\exp\left(u(x_i, w_k) + \text{BLA}_{sk} + \xi_{sk} + \eta_{sk}\right)}{\sum_l \exp\left(u(x_i, w_l) + \text{BLA}_{sl} + \xi_{sl} + \eta_{sl}\right)} dF_{x,s}\left(x_{is}\right), \end{aligned}$$

We can take this to the data to obtain $\delta_{sk} \equiv u(x_s, w_k) + \text{BLA}_{sk} + \xi_{sk} + \eta_{sk}$

Separating η_{sk} from ξ_{sk}

η_{sk} and ξ_{sk} are both unobserved

- Regress δ_{sk} on the observables does not identify η_{sk}
- Unlike consumer demand estimation, IV does not help separation

Our approach: **Control function a la OP**

- Assignment of local party workers likely endogenous
- If so, informative about local political environment
- Can directly control ξ_{sk}

Assignment of BLAs

BLAs as party workers at the polling station level

- Help election officers scrutinize electoral rolls by their local familiarity
- Help voters on the election day to figure out to whom vote for
- Check who don't vote and encourage their turnout

Between-polling-station variation in effective number of BLAs

- Officially, BLAs should be a voter of the assigned station
- In data, same names appear multiple times
- Variation in effective number of BLAs across party and polling stations

Suggestive Evidence on Strategic Assignment

Correlation with votes at polling station level

	BLA:UP	BLA:UP	BLA:ND	BLA:ND
ln(Vote:own)	0.0663***	0.0845***	0.108***	0.102***
	(0.0175)	(0.0279)	(0.0254)	(0.0287)
ln(Vote cast)		-0.0298		0.0267
		(0.0390)		(0.0638)
Obs.	1537	1537	1537	1537

Note: Robust standard errors are in parentheses.

Parliamentary constituency FEs are included.

*** $p<0.01$ ** $p<0.05$ * $p<0.1$

Control Function for ξ_{sk}

Assumptions

- Assignment is based on $\left(x_s, \xi_{sk}, \mathrm{VS}_{sk}^{t-1}\right)$
- Complementarity between BLA_{sk} and ξ_{sk}

Then we have

$$\begin{aligned} \mathrm{BLA}_{sk} &= h\left(x_s, \xi_{sk}, \mathrm{VS}_{sk}^{t-1}\right) \\ \Rightarrow \xi_{sk} &= h^{-1}\left(x_s, \mathrm{VS}_{sk}^{t-1}, \mathrm{BLA}_{sk}\right) \end{aligned}$$

Identification of η_{sk}

With the control function, our model is now

$$\begin{aligned}
\delta_{sk} &= u(x_s, w_k) + \text{BLA}_{sk} + \xi_{sk} + \eta_{sk} \\
&= u(x_s, w_k) + \text{BLA}_{sk} + h^{-1}\left(x_s, \text{VS}_{sk}^{t-1}, \text{BLA}_{sk}\right) + \eta_{sk} \\
&= \phi\left(x_s, w_k, \text{BLA}_{sk}, \text{VS}_{sk}^{t-1}\right) + \eta_{sk},
\end{aligned}$$

where $\phi(\cdot)$ is an unknown function. Thus, influences are identified as

$$\eta_{sk} = \delta_{sk} - \phi\left(x_s, w_k, \text{BLA}_{sk}, \text{VS}_{sk}^{t-1}\right).$$

We (implicitly) assume $E\left[\eta_{sk} | x_s, w_k, \text{BLA}_{sk}, \text{VS}_{sk}^{t-1}\right] = 0$, which should hold because of the random assignment.

Next Steps

Short-run

- Demonstrate our empirical approach using our data
- More data on polling officer list (previous election, different districts)

Long-run

- Expand analysis across India **(as long as polling officers are randomly assigned)**
- Descriptive exercise using influences as a proxy for corruption

Limitations

1. Only applicable to parties with positive number of BLAs
   - In our data, this restricts to the two major political alliances
2. Hard to establish robustness
   - The residual can be unreasonably large
   - Model misspecification, measurement errors, etc.
3. Level of influences are not identified
   - Not separable if all voters prefer one candidate
   - Hard to think about counterfactuals removing “influences”

Conclusion

- We study the influences of election officers on voters in India
- We construct a discrete choice model with polling officer influences
- We show identification of influences by leveraging the random assignment of polling officers and the endogenous assignment of local party workers

References I

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841.

Beteille, Tara. 2009. *Absenteeism, Transfers and Patronage: The Political Economy of Teacher Labor Markets in India.* Citeseer.

Neggers, Yusuf. 2018. "Enfranchising Your Own? Experimental Evidence on Bureaucrat Diversity and Election Bias in India." *American Economic Review*, 108(6): 1288–1321.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263.