Enfranchising Your Own? Experimental Evidence on Polling Officer Identity and Electoral Outcomes in India

Yusuf Neggers*

December 30, 2015

JOB MARKET PAPER

Please find the latest version HERE.

Abstract

The spread of democracy in the developing world has been accompanied by concerns regarding the integrity of election management. This paper estimates the effects of ethnic diversity, or lack thereof, among polling station officials on voting outcomes. I exploit a natural experiment in the 2014 parliamentary elections in India, where the government mandated the random assignment of state employees to the teams that managed polling stations on election day. I find that the presence of officers of minority religious or caste identity within teams led to an average shift in vote share margin of 2.3 percentage points toward the political parties traditionally associated with these groups. Significant spillover effects also occurred across polling stations, and the magnitude of the combined direct and indirect effects is large enough to be relevant to election outcomes. Using survey experiments conducted with more than 5,000 registered voters and election officials, I provide evidence of own-group favoritism in polling personnel and identify the process of voter identity verification as an important channel through which voting outcomes are impacted.

*Harvard University. Email: yneggers@g.harvard.edu. I thank my advisors Alberto Alesina, Rema Hanna, Rohini Pande, and Andrei Shleifer for their encouragement and suggestions over the course of this project. I also thank Leonardo Burzstyn, Michael Callen, Shawn Cole, Melissa Dell, Nilesh Fernando, Leander Heldring, Clément Imbert, Asim Khwaja, Michael Kremer, Horacio Larreguy, John Marshall, Janhavi Nilekani, Benjamin Olken, Vincent Pons, Natalia Rigol, Martin Rotemberg, Monica Singhal, Andrew Weiss, Jeff Wittsoe, David Yanagizawa-Drott, and participants of the Harvard development and political economy lunches for helpful discussions. Shruti Bhimsaria, Azfar Karim, and Akhilesh Kumar provided excellent research assistance. I am grateful to Shashi Thakur, Eric Dodge, Jeffrey McManus, and Ista Zahn for assistance with administrative data acquisition. The field work would not have been possible without the hospitality of the IDinsight Bihar office and the effort of staff at J-PAL South Asia, MORSEL India, and Sunai Consultancy. Financial support for this project was generously provided by the J-PAL Governance Initiative, Harvard Lab for Economic Applications and Policy, Weiss Family Program Fund, Pershing Square Fund for Research on the Foundations of Human Behavior, Harvard South Asia Institute, and Harvard Institute for Quantitative Social Science.

1 Introduction

Electoral malpractice and election day violence are common across the world, as an examination of the most recent round of the World Values Survey suggests.[1] Figure 1 shows that at least twenty-five percent of respondents in more than one half of survey countries indicate that violence at the polling station and the unfair counting of votes are often problems, and in nearly three quarters of countries that election officials are often unfair. The provision of well-functioning elections constitutes a critical public service. The ability of a country's citizens to cast votes in a free and fair setting is desirable in its own right, but is additionally important to the extent that it increases the accountability of elected officials, with subsequent impacts on policy decisions and citizen welfare (Besley and Case 1993, Maskin and Tirole 2004).

Election management and voting technology vary widely across countries, including a fundamental aspect of electoral administration, the staffing of polling stations on election day. Volunteers manage polling stations in the United States, while in Argentina randomly selected citizens work as polling station officials within their own municipalities. In India polling officials are randomly drawn from pools of government employees, and in Kenya polling officers are temporary, paid positions staffed through an open application process. Given the variety in election administration across countries and the frequent dissatisfaction of citizens with elections, there is clear need for causal evidence on what works in election reforms, particularly in relation to personnel management.

In this paper, I ask how diversity, or lack thereof, in the teams of officers that manage polling stations on election day affects voting outcomes. I take advantage of a natural experiment in the Indian electoral setting in which government employees were randomly assigned to teams of polling station officials. The government's method of assigning offi-

[1] Round 6 was administered to representative samples of individuals across sixty countries between 2010 and 2014, but in only forty-two were election-related questions asked. This round was the first to include such questions.

cers generates random variation in team composition in terms of religious and caste identity, both of which have strong connections with political affiliation. Largely in opposition to upper-caste Hindu influence, Muslims and Yadavs (a low-caste Hindu group) formed a political alliance in the state of Bihar in the mid-1990s, and this coalition remained operative in the most recent 2014 national elections.[2] The teams of officers I study contain at least one Muslim or Yadav approximately one third of the time, so the impacts of shifting from a "homogeneous" to "mixed" team of officers at a polling station can be identified.[3]

The random assignment circumvents otherwise confounding issues of selection in election officer placement at polling stations. A government may assign election personnel with greater experience to manage more troubled locations in an effort to maintain neutrality. Alternatively, the ruling party may station supporters as officers in strategically important areas to influence outcomes in their favor. In either case, the assignment of officers would be endogenous to voting behavior. Conditional on the integrity of the randomization, which I test and confirm, the setting considered in this paper is not subject to problems of this type.

An additional benefit of the study context is that the polling officer assignment policy had already been in place statewide for a decade at the time of the election under consideration. This alleviates concerns that the estimated impacts reflect only partial equilibrium effects that may disappear once the policy is brought to full scale or as the government and political parties adjust to the change over time (Acemoglu 2010, Svensson and Yanagizawa-Drott 2012).

I study two districts in Bihar covering more than 5.6 million registered voters across 5,561 polling stations for the 2014 national elections. I use detailed polling station location information and unique officer assignment data to identify the direct effects within stations of changes in team religious and caste composition on voting outcomes, as well

[2]For a detailed account of the state of the alliance over time see Wittsoe (2013).

[3]Due to the low proportions of Muslims and Yadavs among officers, teams that are fully Muslim/Yadav are not observed in my sample.

as the spillover effects across polling stations. To shed light on underlying mechanisms, I embedded list and vignette experiments in surveys of more than five thousand randomly selected election officials and registered voters from the same elections. The survey data allow me to measure election officer bias and determine whether such bias is reflected in differential treatment of voters on election day. A novel contribution of this paper is the generation of primary survey and experimental data within the same populations of election officers and villagers that were exposed to the election policy experiment.

This paper has four main results. First, I find that shifting from homogeneous to mixed team composition affects voting outcomes both within and across polling stations. The average vote share margin between the two major political coalitions is reduced on average by 2.3 percentage points when the officer team at a given polling station is mixed in composition. This shift is driven by a significant 4.6 percent increase in votes for the minority-oriented coalition and a 4.1 percent decrease in votes for the other coalition. In addition, if a neighboring station is mixed team rather than homogeneous, the vote share margin between coalitions decreases at a given station by an average of 2.6 percentage points. I find no evidence of spillovers over longer distances.

Turning to the survey-based experiments, the second finding is that election officers exhibit religious/caste bias on election day and attempt to influence the behavior of potential voters. The results of list experiments suggest that more than 19 percent of officer respondents and 25 percent of registered voter respondents agree that election officers at their polling stations treated potential voters differently based on religion or caste. Additionally they suggest that 5 percent of officers and 13 percent of registered voters agree that officers tried to influence individuals' choices of candidate or make it more difficult for them to cast a vote.

The third main result is the identification of the voter identity verification process at the polling station as a channel through which team composition interacts with officer bias to impact voting outcomes. The results of a vignette experiment conducted

with polling station officers demonstrate that they are 10 percentage points, or 25 percent, more likely to assess individuals favorably in terms of qualification to vote when they are of the same religious/caste-group type, holding all other available information constant. Next, I show that Muslim/Yadav potential voters at polling stations with all non-minority officers are on average less satisfied with their overall polling station experience and less likely to be able to vote than are non-Muslim/Yadav individuals. These effects disappear, however, at mixed team polling stations or when individuals possess the government voter identity card, a method of identification which reduces the need for officer discretion in determining voter eligibility. I additionally determine using polling-station-level administrative data that the effects of team composition on voting outcomes are concentrated in areas with lower voter identity card coverage. Taken together my results suggest that religious/caste diversity within officer teams and the reduction of the scope for discretion in officer duties function as substitutes in improving the impartiality of election proceedings.

Fourth, counterfactual calculations indicate that the magnitude of the combined direct and indirect team composition effects is large enough to influence the outcomes of elections. Under conservative assumptions, I estimate that alternative officer assignment mechanisms would have changed the identity of the winning coalition in approximately five to ten percent of races in the most recent national and state elections in Bihar. These electoral impacts then suggest that officer team composition may also have downstream effects on citizen well-being. For instance, the two major political coalitions strongly differ in their propensities to field Muslim candidates and recent work finds that the religious identity of Indian legislators significantly impacts health and education outcomes (Bhalotra et al. 2014).

This paper complements the literature examining technology-centered approaches to strengthening elections. While technological innovations in the election setting have been shown to significantly impact electoral fraud (Callen et al. 2015), voter turnout (Marx

et al. 2014), and even subsequent public service delivery and health (Fujiwara 2015), less progress has been made in understanding, holding the electoral setting otherwise constant, how the identities of election personnel matter.

My results additionally relate to work which finds that election observers reduce fraud at their posted polling stations when they represent non-politically affiliated international or domestic organizations (Hyde 2007, Ichino and Schündlen 2012), but may introduce additional bias when they themselves have partisan preferences (Casas et al. 2014). While this literature considers individuals external to the government who are explicitly tasked with monitoring polling stations, I focus on the government officials responsible for election proceedings themselves.

This paper also contributes to the body of research studying the negative impacts of ethnic fractionalization on government decision making and the provision of public goods (Easterly and Levine 1997, Alesina et al. 1999, Miguel 2004, Miguel and Gugerty 2005). I provide micro-econometric evidence on an additional area, the administration of elections, in which heterogeneity in the ethnic composition of a population can lead to adverse effects on the quality of public service provision.

Finally, while I consider a setting in the developing world, the implications of my work are also relevant to developed-country democracies. In the United States, a 2014 government study states that "one of the signal weaknesses of the system of election administration in the United States is the absence of a dependable, well-trained corps of poll workers" (PCEA 2014). Recent research also suggests that minorities in the US have different procedural experiences at polling stations on election day (Ansolabehere 2009, Atkeson et al. 2010, Cobb et al. 2012) and poorer perceptions of poll worker job performance (Hall et al. 2009). Additionally, Faller et al. (2014) find that potential voters of different putative ethnicities receive different information from local election administrators across the US in response to otherwise identical requests about voting requirements. My results further underscore the relevance of dimensions of voter identity such as eth-

nicity to the quality of election-related service provision by bureaucrats.

The paper proceeds as follows. In Section 2, I provide background on the context of the study. Section 3 presents a conceptual framework. Section 4 describes the data and performs a randomization check. Section 5 presents the reduced-form impacts of team composition on voting outcomes. Section 6 provides empirical evidence on causal mechanisms. Section 7 considers alternative explanations and Section 8 concludes.

2 Background

This section provides a brief background on the political relevance of religion and caste, administrative structure, election management practices, and government responses to election fraud in India.

2.1 Religion, caste, and politics

Over the last two decades, the dominant political parties in state-level politics in Bihar have been the RJD, BJP, and JDU. The RJD has traditionally enjoyed the support of an alliance between Muslims and Yadavs, a lower-caste Hindu group, which arose in large part in the mid-1990s in an attempt to counter upper-caste Hindu influence in the state (Wittsoe 2013). Muslims and Yadavs are sizeable constituencies in Bihar, making up approximately 17 percent and 14 percent of the population of registered voters, respectively (CSDS 2010). Between 2005 and 2013, the BJP and JDU parties were joined in a political alliance. The BJP was primarily supported by upper-caste Hindus, while the JDU relied more on the support of non-Yadav lower castes. The BJP-JDU alliance dissolved in the run up to the 2014 parliamentary election and, as a result, religion and caste were widely considered of high electoral relevance (Anuja 2013, Bhaskar 2013, Rukmini 2014).

The RJD and BJP subsequently each formed coalitions with other political parties and the JDU contested alone. Members within each coalition agreed prior to the elections

not to field candidates in the same races. As upper-castes are less than 15 percent of the population in Bihar, the BJP increased its efforts to court low-caste Hindu voters. Post-polls for the 2014 elections indicate that only 19 percent of Muslims and 2 percent of Yadavs voted for the BJP, while approximately 78 percent of upper-caste Hindus and more than 50 percent of other low-caste groups did so. Correspondingly, only 5 percent of upper castes and 10 percent of other low-caste groups, but 64 percent of both Muslims and Yadavs, voted for the RJD (Kumar 2014a).

Given the strong connections between religious and caste identity and party affiliation, non-Muslim/Yadav officers are expected on average to be relatively politically inclined toward the BJP coalition over the RJD coalition, and vice versa for Muslim/Yadav officers. Section 3 discusses the channels through which shifting from a homogeneous to mixed polling officer team in terms of religious/caste composition may influence voting outcomes. For ease of exposition, I hereafter refer to the RJD and BJP coalitions as the RJD and the BJP.

2.2 Administrative structure and randomized officer assignment

Bihar is divided into 40 parliamentary constituencies (PCs), single member jurisdictions electing representatives to the national parliament via plurality rule. The PCs are further sub-divided into 243 assembly constituencies (sub-constituencies), each of which contains roughly 250 polling stations on average. Registered voters receive a specific polling station assignment for each election and are only able to cast a vote at that station. Parallel to the electoral structure, the state's bureaucratic structure is divided into 38 districts. PC and district boundaries generally, but not always, fully overlap.[4]

A polling station is managed on election day by a presiding officer and typically three or four polling officers with distinct administrative responsibilities, detailed below.[5] Prior

[4]District administrators are responsible for managing election personnel assignment in only those sub-constituencies falling within their districts.

[5]Four polling officers are assigned to polling stations with greater than 1200 registered voters in rural

to elections, each district uses a proprietary government software program to randomly draw 120 percent of the total number of required officers. Each polling team position has a distinct district-level pool of state government employees from which the officers are selected. After the completion of polling duty training, a subset of the individuals in each position pool are randomly assigned to a polling officer team in a designated sub-constituency. Officers are not assigned to sub-constituencies where they are registered to vote or are employed full time. The randomization is conducted in the presence of official observers assigned by the national office of the Election Commission of India (ECI), no more than seven days prior to election day.

A second randomization is conducted in which polling teams are assigned to specific polling stations. This assignment occurs the day prior to deployment of the teams to polling stations, timed so that they arrive the night before the election and no one has advance knowledge of who the officers at a given polling station will be. The software program also automatically generates team rosters with photographs in .pdf format.

2.3 Polling station procedures

Polling station officials are transported together in teams from the district headquarters to their polling stations. Officer absence is therefore more conspicuous and easier to track than it would be if officers reported individually to polling stations on election day. This centralized transport, as well as the automated generation of officer rosters with photographs, also makes it more difficult for officers to report to a polling station different than that to which they were officially assigned or to have someone else impersonate them. If officers are absent from assigned duty without a documented excuse, they are subject to punishment by the ECI. Despite the attempts of the ECI to impose high costs

areas and 1400 registered voters in urban areas (21.1 percent of polling stations), and only two polling officers are assigned to polling station with fewer than 500 registered voters (0.7 percent of polling stations). In the case of four polling officers, the fourth polling officer shares the duties of the second polling officer. In the case of two polling officers, the presiding officer additionally assumes the duties of the third polling officer.

on officers for non-compliance, it may still be that some proportion of officers do not report to their assigned polling stations on election day.[6] To the extent that this occurs, the estimates in this paper can be interpreted as intent-to-treat effects.

On election day, potential voters wait in line at their polling station and sequentially interact with the first through third polling officers. The first polling officer verifies individuals' identities against the official list of registered voters. This list has each individual's name, age, and, when available, a relative's name, voter identity card number, and photograph. Once a voter successfully confirms her identity with the first officer, her name is read out to the rest of the team. The second polling officer then stamps her finger with ink so that she may not vote more than once, obtains her signature or thumb impression in the official register, and gives her a paper slip with a serial number designating the order in which the voting compartment may be entered. The third officer then checks the voter's finger for ink, allows her into the voting compartment, and activates the electronic voting machine so that a single vote may be cast.

Potential voters at the polling station will not necessarily interact with the presiding officer, who is tasked with the overall management and supervision of station activities. Officially, the ECI also requires that one to two unarmed local police officers be stationed in the catchment area surrounding each polling station, but at a distance from the station itself and the queue of potential voters. Political parties also have the option to place a registered agent at polling stations to observe proceedings.

2.4 Election fraud and policy responses

The problem of "booth capturing", as it is commonly known in India, in which a polling station comes under the control of a political party on election day, was a widespread occurrence as recently as the 2004 national elections (Rohde 2004).[7] The ECI implemented a

[6] Official attendance data is not available.

[7] Capturing may take place in a relatively peaceful manner, with local leaders standing near the voting machine to instruct voters on their choice of candidate and making their decisions public to a nearby crowd

number of policies in an effort to stem this type of election fraud. Elections may be staggered over multiple weeks across different regions within a state to maximize the available coverage of central police and paramilitary forces, observers, and camera recording equipment at sensitive locations. Additionally, electronic voting machines (EVMs), which were first used in Bihar during a 2004 nationwide rollout to all state and national assembly elections, were adopted under the general assumption that they are more secure than the traditional paper ballot.[8] For instance, EVMs have a maximum rate allowed of five votes per minute, meant to increase the difficulty of casting large numbers of false votes, and are more difficult to transport and counterfeit than ballot boxes.

The multi-stage randomized assignment of polling station teams was employed state wide in Bihar beginning in 2004, and has since been adopted in all states, covering more than 814 million registered voters across 543 parliamentary constituencies. Among the assumed benefits of the adoption of randomization was a weakened ability of political parties to coordinate ahead of time with polling station officials or identify which locations would be the easiest targets for capture. These policies are generally viewed as having been successful in reducing the frequency of outright booth capturing. However, issues potentially remain with biased election officer behavior on election day or types of electoral fraud that occur in the longer term prior to elections, such as vote buying or intimidation. I focus in this paper on the former.

3 Conceptual framework

In this section, I outline the mechanisms through which changes in the composition of polling officer teams may impact voting outcomes, both within and across polling sta-

of supporters. Votes may also be cast for absent citizens and certain groups may be prevented from voting. Alternatively, more violent methods may be employed, with armed individuals hired by parties taking control of a polling station to cast false votes or steal the ballot box, or using explosives and gunfire to reduce turnout (Wittsoe 2013).

[8]For a criticism of this assumption in the Indian context, see Wolchok et al. 2010.

tions.

3.1 Within-station effects

First, in a setting where officers may choose to engage in biased behavior at the polling station, a change from homogeneous to mixed team composition could influence voting outcomes through a "checks and balances" channel. Polling station officials have two sets of duties on election day: administration of the identity verification and voting process; and maintenance of a neutral environment in the area immediately surrounding the station. The connection of political affiliation to religion and caste is popularly known in this setting and the type of each potential voter is observable to election officers.[9] Homogeneous team officials, whose political preferences are more likely to be aligned, may then treat potential voters at the polling station differently based on religion and caste, or allow political agents to do so, in an effort to affect either ability to vote or choice of candidate conditional on voting.

The presence of an officer of a different type on an otherwise homogeneous team could increase the probability of detection and punishment of team members that act with bias in their administrative duties, reducing the likelihood of such behavior. Officers within a team are stationed in close proximity, typically sitting adjacent to one another (see Appendix Figure 3). Observability of actions across team members is therefore high and officers can lodge complaints to the ECI directly, with potentially severe consequences for individuals found to have behaved improperly in the conduct of their duties. In addition to strengthening the deterrence effect stemming from the potential for future punishment (i.e. higher expected costs), the presence of a different-type officer on an otherwise homogeneous team may also lower the probability that attempts at influencing voting on election day are successful (i.e. lower expected gains), further reducing the incentives of officers to engage in biased behavior.

[9] Each potential voter's name is read aloud during the identity verification process.

The most important administrative task at the polling station is the verification of voter identity prior to the casting of votes. This process necessarily involves discretionary decision making by election officials and may give them the ability to successfully influence voting outomes with a lower probability of punishment as compared to actions that are be less ambiguously identified as improper if observed.[10] As such, officers may be particularly likely to behave with bias during this step, possibly disenfranchising qualified potential voters or enfranchising unqualified individuals.

The scope for officer discretion in the identity verification process, however, depends critically on the identification documents that potential voters possess. The government-issued voter identity card is the officially preferred and least controvertible form of identification (Appendix Figure 1 provides an example of the card). While eleven other sets of documents are allowed on election day, their use may provide greater discretionary cover to biased officer behavior during voter identity assessment. Potential voters may be less certain about what constitutes a valid alternative means of verifying identity, making them less likely to dispute officer judgement regarding their qualification to vote or increasing their susceptibility to influence in choice of candidate (e.g. if officers make them feel as if they are receiving a favor in being allowed to vote). The potential monitoring benefit provided by a shift from homogeneous to mixed officer team composition may then be particularly important in situations where voter identity cards are less common.

The officer team is also responsible for maintaining a neutral environment in the area immediately surrounding the polling station. More specifically, any activities which may influence potential voters, such as canvassing of votes or disorderly behavior, are officially prohibited within one hundred meters of the polling station. If all officers on a team are of the same type, they may selectively allow agents of the political coalition with which they are aligned to engage in such behavior within that range of the station. As

[10]Guidelines from the ECI on election day management of polling stations even state that "minor errors in the EPIC [voter identity card] and electoral roll may be ignored and overlooked."

mixed team composition may reduce the incentives of officers to behave with bias, the likelihood that agents from both coalitions are prevented from violating neutrality could increase. In sum, if a homogeneous officer team relatively favors one coalition, shifting to a mixed team would be expected to decrease votes for the previously favored coalition and/or increase votes for the other coalition, with ambiguous predictions on total turnout.

Second, in the absence of biased behavior on the part of officers, introducing heterogeneity into polling station teams may influence voting through a "team performance" channel. The literature on teams and heterogeneity has highlighted the potential tradeoff of benefits associated with a greater diversity of skills and information against increased communication and coordination costs and reduced motivation (Prat 2002, Hamilton et al 2003, Marx et al. 2015). Changes in the overall productivity of the officer team may affect the length of waiting time and consequently the proportion of potential voters willing to incur this cost of voting. In this case, impacts would be expected on total turnout, with effects on each coalition in the same direction.

Finally, the identities of the election officials with whom potential voters interact at the polling station may impact voting behavior through an "identity salience" channel. The behavior of voters has been shown to be sensitive to small changes of different types (Gerber and Rogers 2009, Shue and Luttmer 2009, Bryan et al. 2011). Even if officer actions are unaffected by changes in team composition, the religion and caste of election officials at the polling station on election day may be discerned by potential voters and influence their voting behavior. Effects of this type would be expected to influence primarily the choice of candidate, rather than the extensive voting margin.

3.2 Cross-station spillovers

The discussion above has considered the channels through which officer team composition may influence voting outcomes within a given polling station. The composition of the team may influence other stations as well, especially in settings where multiple stations are located within a short distance of one another (Appendix Figure 2 provides an example).

It is important to account for the possibility of these cross-station effects when calculating the total impact of changes in team composition, as their exclusion could bias the overall estimates downward or upward. If a polling station is more strictly managed in terms of maintaining a neutral environment under mixed officer composition, the ability of local political agents to influence proceedings there may be reduced. These individuals could then intensify their focus on other stations which are more amenable to their activity, leading to *displacement effects* (Ichino and Schündlen 2012) that reduce the magnitude of the total impact on voting outcomes.

The effects of more impartial management may alternatively spill over positively to nearby stations. Informational spillovers about what constitutes sufficient documentation for identity verification may take place across potential voters in neighboring polling stations, or the presence of officers of different types on teams in close proximity may serve a monitoring role as within teams. It may also be that the surrounding areas of nearby polling stations overlap such that improved neutrality in the management of one area impacts those of other stations as well. In these cases, mixed team composition could yield additional *chilling effects* (Callen and Long 2015) that increase the magnitude of the total effect. It is also possible that both displacement and chilling effects occur, but over different distances from a given polling station. Chilling effects would be expected to occur across polling stations within closer proximity, while displacement effects could take place over longer distances.

4 Data

4.1 Administrative data

Administrative data on polling officers was acquired for two districts in Bihar for the 2014 elections, covering 23,384 officials posted across 5,561 polling stations. The data include officer name as well as team and position assignment. This information allows me to infer the religious and caste composition of each polling station team, described in greater detail in Section 4.3. Polling stations with at least one Muslim or Yadav officer are defined as "mixed" team polling stations, as opposed to "homogeneous". Polling station level electoral returns were obtained from the website of the Office of the Chief Electoral Officer (CEO), Bihar. The main outcomes of interest generated from this data are the log number of votes received by each of the two main coalitions, the log of total votes cast, and the vote share margin between the coalitions. Sub-constituency-level measures of voter identity card possession were also available on the CEO website. In order to generate measures of electorate religious and caste composition at the polling station level, I additionally scraped publicly available lists of approximately 5.6 million registered voters across the polling stations in the two districts for which officer assignment data was available.

For the analysis of cross-station spillover effects, I obtained polling station GPS coordinates from the dataset of Susewind (2014). As polling station identifier numbers are not constant over time and those in the dataset reflect the 2010 election cycle in Bihar, I then hand matched stations by name, achieving a 94.5 percent match rate. The non-matches occur almost entirely due to the creation of new polling stations in the interim period due to increases in the number of registered voters. I also acquired 2011 census village shapefiles from ML InfoMap to match polling stations to villages.

4.2 Survey sampling and data

Between May and September 2015, I conducted surveys of registered voters and election officers from the 2014 elections. The surveys collected information on socio-demographic characteristics and election-related experiences. I additionally included experimental modules, which are discussed in more detail in Section 6. The surveys took place in one of the districts for which officer assignment data was available.

Registered voters survey

A total of 4,320 individuals across 360 polling stations were chosen for the registered voters survey. In each of the 5 sub-constituencies in the district, 36 mixed team and 36 homogeneous polling stations were randomly selected, stratifying by whether the Muslim-Yadav proportion of the population was above or below the district-level median (Appendix Table 2 provides additional details). For each of these polling stations, three Muslim and two Yadav registered voters were randomly chosen from the list of registered voters, if possible, along with seven randomly selected registered voters inferred as neither Muslim nor Yadav.

Election officers survey

A total of 915 officers across 610 polling stations were sampled for the survey of election officers. 61 mixed team and 61 homogeneous team polling stations in each of the 5 ACs were chosen randomly. One Muslim or Yadav officer and one non-Muslim, non-Yadav officer were then randomly selected from each mixed team, while a single non-Muslim, non-Yadav officer was randomly chosen from each homogeneous team (Appendix Table 2 provides additional details).

4.3 Inference of religious and caste identity

The categorization of election officers and registered voters as Muslim, Yadav, or neither is inferred from name. The Anthropological Survey of India's *People of India* (POI) series lists common surnames as well as religion and caste for 261 distinct communities identified as inhabiting Bihar. A surname may be associated with multiple communities, potentially of different religious or caste affiliations. I categorized individuals as Muslim if their surnames matched one listed in the POI that is associated only with Muslim communities. I subsequently categorized individuals as Muslim if their name had components of clear Islamic origin, e.g., "Haiderali", "Raiyaz", or "Mohammed". I categorized as Yadav those individuals with the surname "Yadav", as the majority of the members of the caste are so named and the surname is not associated with other communities.[11] The lists of registered voters also provide the name of a relative for each individual (typically a father in the case of males or unmarried females, and husband in the case of married females). Given strong norms of marrying within religion and caste group in the region, I also categorized registered voters as Muslim or Yadav if their listed relative was inferred as falling into one of these categories.[12]

4.4 Randomization check

In the two districts in my sample, between 8.3 and 9.3 percent of officers in each position are inferred to be Muslim/Yadav, giving a total of 32.3 percent of polling stations with at least one Muslim/Yadav officer (i.e. mixed team). Given that, within a district, officers are not assigned to sub-constituencies in which they are registered to work or vote full time, a sub-constituency with a larger population proportion of Muslim/Yadav officers rela-

[11] Other Yadav-associated surnames include Rai, Ram, and Singh, which are common among other communities as well. To the extent that individuals are misclassified, my estimates of the impact of Muslim/Yadav officer presence will be biased toward zero.

[12] In Bihar, nearly all Yadav females adopt the surname Devi upon marriage, so use of the relative name information is important in identifying this population. The issue of marriage-related name change is not relevant for the population of polling station officers, as they are nearly always male.

tive to other constituencies within the same district could then receive a lower proportion of Muslim/Yadav officers assigned to its polling stations, potentially mechanically leading to correlations between team composition and voting outcomes. However, it is still the case that polling stations *within* a sub-constituency are equally likely to have Muslim/Yadav officials posted to the officer team. I therefore include sub-constituency-level fixed effects in all regression specifications in my subsequent analysis.

Potential concerns with my identification strategy then are that either the random assignment of individuals to polling team positions or of polling teams to polling stations is compromised. In Panels A and B of Table 1, I test whether mixed composition polling teams are assigned to polling stations which vary systematically along pre-election dimensions potentially correlated with voting outcomes. First, I examine whether the size or composition of the electorate differs across homogeneous and mixed team polling stations. The average polling station has roughly 1,000 registered voters of which 46 percent are female and 13 percent are Muslim or Yadav, with no significant differences by team composition. I next consider station-level electoral results from the previous 2010 elections to the state assembly. As the number of polling stations increases over time due to growing numbers of registered voters, it is not always possible to match polling stations across elections. For instance, a location which previously had two polling stations in 2010 may have three polling stations in 2014.[13] For each 2010 election-related variable, I therefore take the average value across all polling stations in a given location in 2010 and assign that to each polling station in that location in 2014. In the previous example, the average value across the two polling stations in 2010 would be applied to each of the three polling stations in 2014. Additionally, a small proportion of polling stations were established in new locations for the 2014 election and so cannot be matched to previous elections.[14] I observe no significant differences in the log votes previously received by

[13]Section 5.2 defines "location" in more detail.

[14]The total number of polling stations across Bihar increased by 5.9 percent between the 2010 and 2014 elections. As a result, the electorate assigned to each previously existing polling station potentially also

each coalition or total, or in the vote share margin between the two coalitions.[15]

Appendix Table 3 presents the results of tests for differences in the spatial distribution of polling stations. In no instance do I observe significant differences by team composition of a given polling station when considering the average numbers of total or mixed team neighboring polling stations in the same location, within 0.25 or between 0.25 and 0.75 kilometers, or within the same or neighboring villages. I also test whether the assignment of a Muslim/Yadav officer to a given position within a polling party is significantly correlated with the assignment of Muslim/Yadav officers to other positions within that party.

I next turn to the samples of surveyed election officers and registered voters in Panels C and D of Table 1. None of the officer characteristics differ significantly by polling station team composition.[16] While the sample of registered voters from mixed team polling stations is significantly less likely to be female, the difference in terms of magnitude is not large (58 percent versus 53.7 percent) and a control for gender is included in all registered-voter-level regressions.

5 Reduced-form impacts on voting outcomes

5.1 Within-station effects of team composition

Figure 4 plots kernel density estimates of the distributions of the vote share margin between the RJD and BJP separately for homogeneous and mixed team polling stations. The plots demonstrate that the average vote share of the RJD relative to that of the BJP is higher for mixed teams, where the equality of the distribution functions is rejected at the

shifted.

[15]Observation numbers change across the 2010 election outcomes because coalitions fielded candidates in different numbers of constituencies.

[16]By definition, homogeneous officer teams do not contain Muslim/Yadav officers. Therefore balance tests across team types of officer characteristics are necessarily restricted to the sample of non-Muslim/Yadav officers.

5 percent level. I then turn to the corresponding regression specification:

$$Y_{pc} = \mu_c + \theta_o + \beta Mixed_{pc} + \mathbf{X'_{pc}}\lambda + \epsilon_{pc}, \quad (1)$$

where p is a polling station in sub-constituency c, μ_c are sub-constituency-level fixed effects, and θ_o are fixed effects for the number of polling team members.[17] $Mixed_{pc}$ is an indicator variable taking value 1 if at least one polling team member is Muslim/Yadav and 0 otherwise. $\mathbf{X_{pc}}$ is a vector of polling station characteristics – the log number of registered voters and the share categorized as Muslim/Yadav. Given randomized officer assignment, the polling station-level covariates are included to improve statistical precision. Y_{pc} is a voting outcome, either the log votes received by a coalition or in total, or the vote share margin between the RJD and BJP.

Panel A of Table 2 reports the Equation 1 estimates. Column (1) shows a significant 4.6 percent increase in the votes received by the RJD. I also observe a significant 4.1 percent decrease in BJP votes in column (2). In column (3), I consider the combined impact on the vote share margin between the RJD and BJP and find that mixed team composition significantly narrows the gap between the coalitions by 2.3 percentage points, or 12.7 percent.[18] Additionally, a 1 percentage point increase in the Muslim/Yadav share of registered voters at a polling station leads to a 3 percent increase in RJD votes and 3 percent decrease in BJP votes, consistent with the previously asserted connections of political affiliation with religion and caste in this setting.[19] While column (4) indicates that mixed team composition has no average effect on the log total votes cast, I am unable to rule out effects of approximately 1.6 percentage points in magnitude in either direction.[20]

[17] The likelihood of the presence of a Muslim/Yadav officer on a polling station team mechanically increases with team size. I therefore include fixed effects for the number of officers.

[18] No significant effects are observed on log JDU or total non-RJD/BJP votes.

[19] Average vote shares for the RJD and BJP coalitions are 0.285 and 0.457, implying roughly a 1-to-1 effect.

[20] Appendix Table 5 considers whether impacts vary significantly by: the position within a team in which Muslim/Yadav officer presence occurs, or the presence of single versus multiple Muslim/Yadav officers. Significant differences are not found across positions or by number. Appendix Table 6 additionally shows the absence of signficant heterogeneity in impacts by share Muslim/Yadav registered voters.

5.2 Spillovers across polling stations

I next consider the existence of spillover effects of officer team composition across polling stations. First, I define "neighboring" polling stations as those situated within the same building/compound location. For example, a group of polling stations may be listed in the administrative data as situated in "K L Primary School (South Part)", "K L Primary School (North Part)", and "K L Primary School (Middle Part)" and would be categorized as neighbors. I observe in the data that the number of neighbors for a given station ranges from 0 to 8 with a mean of 1.169, and the number of neighbors that are mixed team ranges from 0 to 4 with a mean of 0.385.

In addition to providing exogenous variation in team composition within polling stations, the officer assignment mechanism generates random variation in the local density of Muslim/Yadav officer presence in neighboring polling stations.[21] This fact allows me to adapt the approach of Miguel and Kremer (2004) to estimate cross-station spillovers. I use the regression specification:

$$Y_{pc} = \mu_c + \theta_o + \beta Mixed_{pc} + \gamma T_{pc} + \phi N_{pc} + \mathbf{X'_{pc}}\lambda + \epsilon_{pc}, \tag{2}$$

where N_{pc} is the number of neighbors of polling station p in constituency c, and T_{pc} is the number of these neighbors with a mixed composition polling team. Standard errors are clustered at the location level. Impacts associated with polling station density are captured by N_{pc}. Controlling for this density, the number of neighboring polling stations with mixed composition teams is randomly determined.

The within-station direct effects of mixed team composition on voting outcomes are captured by β, while γ gives the average cross-station spillover effect of a mixed team neighbor. As the treatment density in surrounding polling stations is orthogonal to the

[21] Consider, for example, a polling station with four neighboring polling stations. The proportion of these four stations that is mixed team versus homogeneous would be randomly determined by the assignment procedure.

treatment status of a given polling station, the point estimates of the within-station impacts of mixed team composition should be unchanged from Equation 1.

As chilling and displacement effects may be observed at different distances, I extend the range over which spillovers are considered using two approaches. First, I supplement Equation 2 with the variables $N_{pc}^{0.25km}$ and $N_{pc}^{0.25-0.75km}$, the number of polling stations in different locations within 0.25km or between 0.25-0.75km of polling station p, and $T_{pc}^{0.25km}$and $T_{pc}^{0.25-0.75km}$, the numbers of such polling stations with mixed composition teams.[22] Second, while this specification allows the impact of team composition on other stations to vary by linear distance from a polling station, a more meaningful distinction may be captured by administrative boundaries (e.g. if canvassing is organized by agents at the village level). I then employ a specification which augments Equation 2 with variables for the total and mixed team composition numbers of polling stations in different locations within the same village, N_{pc}^{vill} and T_{pc}^{vill}, and neighboring villages, N_{pc}^{nei} and T_{pc}^{nei}. As the top 1 percent of the distribution of villages in terms of polling stations has a mean of 98.8 as compared to the overall mean of 2.4, I trim the sample for this specification to exclude polling stations located in or neighboring these villages.[23]

Panel B of Table 2 presents the Equation 2 estimates. Columns (1) and (2) show an imprecisely estimated 3.1 percent increase in RJD votes and a significant 4.2 percent decrease in BJP votes associated with a change in a neighboring polling station from homogeneous to mixed team composition. In column (3), the combination of these two effects yields a highly significant 2.6 percentage point cross-polling-station change in vote share toward the RJD away from the BJP. As expected, the point estimates on the within-polling station mixed team indicator are unchanged as compared to those of Equation 1. These results demonstrate the occurrence of chilling effects across polling stations within the same location. Panels A and B of Appendix Table 7 present the results of tests for spillover effects

[22]The sample for this specification is slightly reduced, as it excludes polling stations which could not be matched to the 2010 polling station GPS coordinates.

[23]The outlier villages are also urban and very large in area relative to typical villages.

over greater distances, defined either in linear distance or by village boundaries. The estimates show no evidence of chilling or displacement effects over longer ranges.

6 Officer bias in voter identity verification

In this section, I examine the extent to which the process of voter identity verification functions as a channel through which team composition and officer bias interact to impact voting outcomes at polling stations.

6.1 Vignette experiment: own-group bias in election officers

First, I test for election officer own-type bias in the evaluation of voting eligibility, using a vignette experiment embedded within the survey of officers. I examine whether, holding all other information constant, potential voters are more likely to be assessed by an election officer as qualified to vote if they are of the same type as that official. Vignette experiments have been used previously to address research questions in the electoral setting (Carlson 2010, Banerjee et al. 2014) and are methodologically similar to the randomized CV experiment approach that has been employed in the labor market discrimination literature (Bertrand and Mullainathan 2003, Banerjee et al. 2009).

Each respondent was read a vignette describing a hypothetical individual attempting to vote, with the wording identical across respondents with the exception of the individual's name, which was randomly assigned.[24] At the conclusion of the vignette, respondents were asked to indicate the likelihood on a 4-point scale that the individual in the vignette would be able to cast a vote. Each officer respondent was randomly assigned one of nine possible voter names. Three names each were chosen to signal Muslim (ex-

[24]The vignette question was worded as: "Please consider the following situation: A voter named [RANDOMLY ASSIGNED] arrives at the polling station without an EPIC card but has a government voter's slip without a photograph. He can recite his name and other particulars. On a scale of 1 to 4, how likely do you think it is that he would be allowed to cast a vote based on this information?", where the potential responses are "Very unlikely (1)", "Unlikely (2)", "Likely (3)", "Very likely (4)".

ample: "Mustak Ansari"), Yadav (example: "Ajay Yadav"), or Brahmin (example: "Alok Chaturvedi") identity in the hypothetical voter.[25] To examine whether an officer's evaluation of the likelihood of a potential voter's ability to cast a vote is influenced by whether that individual is of the same type as the officer, I use regression specifications of the form:

$$Y_{qpc} = \mu_c + \varphi_n + \pi_v + \theta Match_{qpc} + \mathbf{X'_{qpc}}\lambda + \epsilon_{qpc}, \tag{3}$$

where Y_{qpc} is an outcome of officer q in polling station p in sub-constituency c, and μ_c signifies sub-constituency fixed effects. Additionally included are fixed effects for randomly assigned potential voter name, φ_n, and election officer type (Muslim, Yadav, Brahmin, or other), π_v. $Match_{qpc}$ is an indicator variable taking value 1 if the election officer's group type and that of the potential voter are the same (e.g. Yadav and Yadav) and 0 otherwise. As the potential-voter-name and officer-type fixed effects control for the average differences in assessed likelihood of the potential voter's ability to vote across the different hypothetical names and by officers of different types, the coefficient of interest, θ, captures the average change in officer assessment caused by the officer-voter type match. Further controls are fixed effects for polling team composition and a set of officer-level covariates: age, log monthly salary, an indicator for first term of service at a polling station, and fixed effects for occupation type, education level, and polling team position. A second specification additionally includes polling-station-level controls for log total registered voters, share Muslim/Yadav registered voters, and fixed effects for station location type and number of officer team members.

I consider as outcomes both an indicator variable taking value 1 if the officer chooses "Likely" or "Very Likely" and a continuous variable taking the 1-to-4 scale value. Figure 5 shows that for both variables the average assessed likelihood of voting increases significantly when the hypothetical individual is of the same type as the election officer.

[25] Appendix Table 1 provides the full list of names in each category.

Table 3 presents the underlying estimates from Equation 3. Considering the binary outcome variable, columns (1) and (2) show a significant 10 percentage point, or more than 25 percent, increase in the probability that an individual is assessed as likely able to cast a vote. Similarly, using the 4-point-scale measure as the outcome in columns (3) and (4), a significant increase of approximately 0.24 points is observed.

6.2 List experiment: biased officer behavior on election day

I next consider whether biased officer behavior was perceived by voters and election officers as a relevant election day phenomenon in this setting. As direct elicitation of survey respondents may yield unreliable estimates of the occurrence of potentially sensitive topics such as biased officer behavior during elections, I employed list experiments in both the surveys of registered voters and election officials. This method of indirect elicitation has been used to generate measures of sensitive topics related to political and electoral behavior in a number of recent papers (Gonzalo-Ocantos 2010, Corstange 2012, Kramon and Weghorst 2012, Ahlquist et al. 2013, Burzstyn et al. 2014).

For each list experiment question, respondents were randomly assigned to either a control or treatment group. Members of each group were asked to indicate only the total number of statements that occurred at their polling station during the 2014 elections from a list of statements read to them. Control respondents were given a list of four statements on non-sensitive election day topics, while treatment respondents were given the same list but with an additional sensitive statement included in the third position. This approach prevents individual-level determination of which statements were chosen, but allows for the population-level prevalence of the sensitive occurrence to be estimated as follows:

$$N_{ipc} = \alpha_c + \phi Treat_{ipc} + \mathbf{X'_{ipc}}\lambda + \epsilon_{ipc}, \tag{4}$$

where N_{ipc} is the number of statements indicated as occurring at polling station p by

respondent i, $Treat_{ipc}$ is an indicator variable for assignment to the group additionally receiving the sensitive statement, and $\mathbf{X_{pc}}$ is a vector of polling station and individual characteristics. Assuming that respondents assess the sensitive item truthfully and the inclusion of the sensitive topic does not influence their evaluation of the non-sensitive items, ϕ gives an unbiased estimate of the population proportion for whom the sensitive item occurred. Standard errors are clustered at the polling station level. Additionally included are polling-station-level controls for log registered voters, share Muslim/Yadav registered voters and fixed effects for polling station number of officers and location type and for the remaining respective survey sampling strata. For officer-respondent regressions, individual controls for age, log monthly salary and fixed effects for occupation type, education level, first term of service, and team position are included. For voter-respondent regressions, controls are included for age, gender, log monthly household income, and household head status and fixed effects for occupation category and education level are included.

The survey of election officers included two list experiments.[26] The sensitive statements for the treatment group in the first and second experiments were: "One or more of the election officers at the polling station treated some voters differently based on the voters' religion or caste" and "One or more of the election officers tried to influence some voters' choice of candidate or make it harder for them to vote". Similarly, the two list experiments included in the survey of registered voters had the sensitive statements: "One or more of the election officers at your polling station treated you or others differently based on your religion or caste" and "One or more of the election officers at your polling station tried to influence how you or others voted or to make it more difficult for you or them to cast votes".

Table 4 presents the results of the list experiments. The estimates in column (3) of Panel A suggest that 19 percent of officers agree that at least one of the officers at their

[26] Appendix Table 1 provides the introductory prompt used in these experiments.

polling station treated voters differently based on religion or caste and that 5 percent indicated that at one least member of their polling station team tried to influence voter behavior, whether choice of candidate or making it more difficult to vote. Turning to the registered voter results in Panel B, estimates imply that 25 percent of respondents indicate that election officials at their polling station treated voters differently based on religion or caste and that 13 percent agree that election officers tried to influence voting behavior at their polling station. These results suggest that a subset of officers do attempt to influence voting behavior on election day, and likely do so at least in part along religious and caste lines.

6.3 Election day experiences of potential voters

In this section I examine how the election day experiences of potential voters vary by officer team composition, both in terms of overall station area management and the individual-specific identity verification process. I use the following specification to test whether the frequency of canvassing or disorderly behavior at the polling station, as reported by registered voter survey respondents, varies by team composition:

$$Y_{wpc} = \mu_c + \theta_o + \beta Mixed_{pc} + \mathbf{X'_{wpc}}\lambda + \epsilon_{wpc} \tag{5}$$

where Y_{wpc} is an outcome for respondent w in polling station p in sub-constituency c. The included individual and polling station controls are the same as in Equation 4. I additionally examine whether the reported impact of team composition differs with respondent type by including an interaction with an indicator for whether the respondent is Muslim or Yadav. Standard errors are clustered at the polling station level.

Columns (1) and (3) of Table 5 show that canvassing and disorderly behavior are uncommon on election day and that the likelihoods of their occurrence do not vary with team composition. Columns (2) and (4) further demonstrate that the absence of signifi-

cant differences by team composition holds regardless of respondent type. These results suggest that stricter management of the area surrounding the polling station is not the primary channel through which a shift from homogeneous to mixed team composition impacts voting outcomes.

Turning to individual-specific experiences at the polling station, I employ the following regression:

$$Y_{wpc} = \alpha_{pc} + \phi MY_{wpc} + \lambda ID_{wpc} + \psi \left(MY_{wpc} * ID_{wpc} \right) + \mathbf{X'_{wpc}}\lambda + \epsilon_{wpc} \tag{6}$$

where α_{pc} are polling station fixed effects and ID_{wpc} is an indicator for voter identity card possession by individual w at polling station p. I estimate this regression separately for individuals in the randomly determined samples of mixed team and homogeneous team polling stations. The same set of registered-voter-level controls from Equation 4 are included and standard errors are clustered at the polling station level.

I first consider a potential voter's ability to cast a vote as an outcome. I find in column (1) of Table 6 that, at polling stations with homogeneous officer teams, individuals are significantly less likely to be able to vote if they are Muslim/Yadav. This difference disappears among individuals with voter identity cards. In addition, for non-Muslim-Yadav individuals, possession of a voter identity card does not significantly change the likelihood of being able to cast a vote. Column (2) shows that, at polling stations with mixed officer teams, voter identity card possession significantly increases the likeilhood of being allowed to cast a vote, but that this no longer varies with Muslim/Yadav identity. These regressions include a variety of individual-level controls, reducing concerns that the effects are driven by correlations between Muslim/Yadav identity or voter identity card posession with unobservables that influence voting ability. To summarize, at homogeneous team stations, voter identity cards matter in terms of voting ability only for Muslim/Yadav potential voters, while at mixed team stations they are important for

potential voters of all types.

I subsequently examine as an outcome the likelihood of a potential voter having a satisfactory overall experience at the polling station on election day.[27] Column (3) shows that Muslim/Yadav potential voters facing homogeneous teams of officers have worse polling station experiences on average, but only in the absence of voter identity cards. Non-Muslim/Yadav voters, however, express lower satisfaction if they possess a voter identity card. The latter effect could reflect that under-qualified potential voters are relatively more appreciative of being allowed to vote than those with voter identity cards, who are more certain that they should be allowed to do so. I find in column (4) that these impacts are absent at mixed team polling stations. Overall, the results are consistent with mixed team composition and voter identity card provision each reducing the differential treatment of potential voters at polling stations, where homogeneous teams are relatively more stringent toward Muslim/Yadavs.

6.4 Heterogeneity in effects by voter identity card coverage

Using survey and experimental data, the previous sections established that election officers are relatively biased in favor of potential voters of their own type and that individuals' religious and caste identities can influence their ability to vote if they do not have a voter identity card. If in general identity card possession reduces the scope of potentially discriminatory discretion available to officers and mixed team composition shifts station administration to be more neutral, a substitute relationship in the impacts of the two on polling-station-level voting outcomes would also be expected. Returning to the polling station administrative data on voting outcomes, I test for this relationship using

[27] This variable takes value 1 if a respondent indicates that her overall voting experience at the polling station on election day was "Excellent","Good", or "Fair", as opposed to "Poor".

specifications of the form:

$$Y_{pc} = \mu_c + \theta_o + \beta Mixed_{pc} + \gamma(Mixed_{pc} * ID_c) + \mathbf{X'_{pc}}\lambda + \epsilon_{pc}, \tag{7}$$

where ID_c is a continuous variable for the proportion of registered voters in sub-constituency c without a voter identity card.[28] Polling-station-level controls included are the log number of registered voters and the Muslim/Yadav share of registered voters. The top one percent of observations in terms of the absolute value of the vote share margin between the RJD and BJP are trimmed.[29] The main effect for ID_c is absorbed by the sub-constituency-level fixed effects, and the coefficient of interest is γ, where an estimated sign opposite that of β indicates that polling station composition and voter identity card coverage exhibit substitutability in their impacts on voting outcomes. Given that sub-constituency-level voter identity card coverage is not randomly determined, as a robustness check I also consider a specification where I further interact officer team composition with sub-constituency-level measures of the population proportions that are literate, Scheduled Caste/Scheduled Tribe, and Muslim/Yadav.

In columns (1) and (2) of Table 7, I find that the positive impact of mixed team composition on RJD votes decreases by a significant 0.9 percentage points per 1 percentage point increase in voter identity card possession. The results for BJP votes in columns (3) and (4) also indicate that the team composition effects are strongest in areas with low voter identity coverage. Columns (5) and (6) show that the vote share margin shift toward the RJD caused by changing from a homogeneous to a mixed team is approximately 0.5 percentage points smaller per 1 percentage point increase in voter identity card coverage. Voter identity card coverage in my sample of sub-constituencies ranges from 76.3 to 93.9 percent. Figure 6 plots the implied effect of mixed team composition over a similar range of voter identity coverage and demonstrates that the signficant impact observed at low cov-

[28] Data on voter identity card coverage at the polling station level is not available.

[29] These are polling stations where one coalition won by a margin of at least 88 percent.

erage levels becomes insignificant as full coverage is approached. These estimates, taken together with the earlier survey and experimental results, suggest that mixed team composition and identity card coverage serve substitute roles in strengthening the neutrality of the identity verification process for potential voters.

7 Alternative explanations

A possible concern in attributing the previously identified impacts to changes in officers' religious and caste identities is that there may exist other characteristics that correlate with these identities and also influence voting outcomes. This is unlikely to be the case for two reasons: the previous analysis captures the effects of the presence on otherwise homogeneous teams of officers that are either Muslim or Yadav, two groups which are not particularly similar outside of their political alliance; and indviduals of different religions and castes within the population of polling station officers are more likely to be similar along other dimensions than would be their populations in general.

First, Yadavs are a lower-caste Hindu group in Bihar and, other than in political orientation, it is unclear along what dimensions they would be systematically more similar to Muslims than to other Hindu groups, especially given the dispersed support for the BJP across upper- and lower-castes in these elections.[30] In Appendix Table 8, I examine the influence of Muslim and Yadav officer presence separately using a regression specification analogous to that of Equation 1. The estimates across columns (1) through (4) reveal similar impacts for Muslim and Yadav officers. The coefficients for the two groups are statistically indistinguishable in each case, and the shift in vote share margin toward the RJD is significant at the 5 percent level for both Muslim and Yadavs.

Second, while Muslims in Bihar are on average poorer and less educated than the

[30]Highlighting the differences between the two groups, Lalu Prasad Yadav, the politician responsible for the creation of the Muslim/Yadav coalition, has even said "I have made an alliance between those who worship the cow [Yadavs] and those who eat the cow [Muslims]." (Wittsoe 2013, p.60)

general population (Ghosh 2004), polling station officers are selected from pools of government employees who are likely more similar than would be average individuals from different religious and caste groups. I explicitly test for differences by Muslim/Yadav status in the sample of surveyed polling station officers across a number of characteristics proxying for experience, knowledge, and capability: age, log monthly salary, college graduation, and prior election officer experience. I regress each of these outcomes on an indicator variable for Muslim/Yadav identity and fixed effects for sub-constituency and team position. As a further check, I also construct measures of age and log monthly salary based on separate administrative data available for the full population of election officers in the district in which the officer survey was conducted. The results in columns (1) through (6) of Table 8 show that in no case are there significant differences by Muslim/Yadav status.

8 Conclusion

Having identified within- and cross-polling station impacts on voting outcomes, a natural question is the extent to which changes in the composition of polling officer teams can influence who wins elections. I therefore conduct counterfactual calculations of the effects of alternative officer assignment mechanisms on the identities of winners in the 2014 parliamentary elections. First, I use administrative data available across the state of Bihar to calculate the sub-constituency-level average numbers of polling station neighbors. Second, the observed margins of victory from these elections already reflect the effects of the underlying (but unobserved outside of the two study districts) proportions of mixed team polling stations in each parliamentary constituency. Taken together, this information allows me to estimate the magnitudes of the shifts in the proportions of homogeneous and mixed team polling stations necessary to change the outcome of each election in which the RJD and BJP coalitions were both either winner or runner up (Appendix Table 9 provides

additional details).[31]

I consider the effects of two alternatives to the current method of randomized officer assignment: (1) the imposition of mixed team composition on all polling officer teams, and (2) the exclusion of Muslim/Yadav officers from teams. During the 2014 elections, the RJD and BJP fielded the top two candidates in 29 of the 40 parliamentary constituencies in Bihar (Appendix Figure 4 provides the distribution of vote share margins). I assume that the proportion of mixed team polling stations in each sub-constituency is the same as the average value (0.324) across the two districts for which it can be directly observed in my data. As shown in Table 9, a shift to Alternative 1 is then estimated to switch one election outcome in favor of the RJD and a shift to Alternative 2 to change one outcome to a BJP victory. I repeat this exercise for the most recent 2010 state assembly elections, where the RJD and BJP fielded the top two candidates in 185 of 243 races. Twelve races are estimated to change to an RJD victory under Alternative 1 and two elections to switch in favor of the BJP under Alternative 2, assuming 2014 levels of voter identity card coverage. Under the actual 2010 levels of coverage, the number of influenced races in the 2010 elections would more than double, with Alternative 1 switching thirty-three races in favor of the RJD and Alternative 2 shifting six outcomes in favor of the BJP, or a combined 16 percent of the total.

These results demonstrate that the way in which officer teams are constituted matters for election outcomes. In addition, the religious and caste composition of candidates put forward in elections differs considerably across the coalitions,[32] and recent work has shown that increasing Muslim representation in state legislatures in India results in significant reductions in child mortality rates and gains in educational attainment across both

[31]The vote share margin between the runner-up candidate and the remainder of the field is generally large enough that having a third place or lower candidate shift to become the winner could not feasibly occur as a result of changes in team composition.

[32]For example, 17.5 percent of RJD coalition candidates in the 2014 Bihar elections were Muslim, as compared to 2.5 percent for the BJP coalition. The previous counterfactual calculations suggest that a shift to all mixed team polling stations would have led to an increase in Muslim state legislators from 19 to 24 in the 2010 elections and in Muslim members of parliament from 4 to 5 in the 2014 elections.

Muslim and non-Muslim households (Bhalotra et al. 2014). Officer team composition may therefore also have non-trivial downstream impacts on outcomes directly relevant to citizen well-being.

An additional possibility is that the effects of officer team composition on citizens' election day experiences influence their expectations and behavior in subsequent elections, for instance in whether to turn out at the polling station. I use polling-station-level data for the 2015 state assembly elections in Bihar and find in Appendix Table 10 that voting outcomes in these elections are unaffected by officer team composition from 2014, providing no evidence of cross-election-cycle impacts on voter behavior.

Fair and well-functioning elections are critical to maintaining the responsiveness of elected officials to citizens in democracies. While the related literature on election reforms has focused more heavily on the effects of technological advances in monitoring and voting technology, this paper provides, to my knowledge, the first well-identified evidence of the importance of the identities of election personnel. Additionally, Indian elections are technologically advanced and their administration is highly regulated, indicating that these effects remain important even at the present frontier of election practice.

Though my findings suggest that diversity within teams of election officers can improve the impartiality of polling station management, it may not always be politically or administratively feasible to mandate that such mixed composition occur. It could also be difficult in other contexts to determine the dimensions of identity along which diversity should be defined. My results, however, also indicate that policies which reduce the scope for officer discretion in the election process, such as the widespread provision of voter identity cards, may be promising alternatives in reducing the ability of local-level election officials to influence voting outcomes. Finally, and more generally, my findings demonstrate that institutions which require greater discretionary decision making by bureaucrats or other government employees may be more susceptible to adverse impacts of these individuals' underlying preferences and biases on the quality of public services.

References

[1] Acemoglu, Daron (2010). "Theory, General Equilibrium, and Political Economy in Development Economics." *Journal of Economic Perspectives* 24(3): 17-32.

[2] Ahlquist, John S., Kenneth R. Mayer, and Simon Jackman (2013). "Fraudulent Voters, Voter Identification, and the 2012 US General Election: Evidence from a Survey List Experiment." Working paper.

[3] Alesina, Alberto, Reza Baqir, and William Easterly (1999). "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics* 114(4): 1234-84.

[4] Ansolabehere, Stephen (2009). "Effects of Identification Requirements on Voting: Evidence from the Experiences of Voters on Election Day." *PS: Political Science & Politics* 42(1): 127-130.

[5] Anuja (2013). "Bihar poised to return to politics of caste, religion." *Mint*, August 5.

[6] Atkeson, Lonna Rae, Lisa A. Bryant, Thad E. Hall, Kyle L. Saunders and R. Michael Alvarez (2010). "A new barrier to participation: Heterogenous application of voter identification policies." *Electoral Studies* 29(1): 66-73.

[7] Banerjee, Abhijit, M. Bertrand, S. Datta, and S. Mullainathan (2009). "Labor Market Discrimination in Delhi: Evidence from a field experiment." *Journal of Comparative Economics* 37(1): 14-27.

[8] Banerjee, Abhijit, Donald Green, Jeffrey McManus, and Rohini Pande (2014). "Are Poor Voters Indifferent to Whether Elected Leaders are Criminal or Corrupt? A Vignette Experiment in Rural India." *Political Communications* 31(3): 391-407.

[9] Bertrand, Marianne and Sendhil Mullainathan (2004). "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94(4): 991-1013.

[10] Besley, Timothy and Anne Case (1995). "Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits." *Quarterly Journal of Economics* 110(3): 769-798.

[11] Bhalotra, Sonia, Guilhem Cassan, Irma Clots-Figueras, and Lakshmi Iyer (2014). "Religion, Politician Identity, and Development Outcomes: Evidence from India." *Journal of Economic Behavior and Organization* 104: 4-17.

[12] Bhaskar, Utpal (2013). "Bihar riots spark blame game, fears of return to past." *Mint*, August 19.

[13] Bjornlund, Eric (2013). "Final Panel Address." The Electoral Integrity Project Annual Workshop. Weatherhead Center, Harvard University, June 4th.

[14] Bryan, Christopher J., Gregory M. Walton, Todd Rogers, and Carol S. Dweck (2011). "Motivating voter turnout by invoking the self." *Proceedings of the National Academy of Sciences* 108(31): 12653-12656.

[15] Burzstyn, Leonardo, Michael Callen, Bruno Ferman, Ali Hasanain, and Noam Yuchtman (2014). "Identifying Ideology: Experimental Evidence on Anti-Americanism in Pakistan." Working paper.

[16] Callen, Michael J. and James D. Long (2015) "Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan." *American Economic Review* 105(1): 354-381.

[17] Casas, Agustín, Guillemo Díaz, and Andre Trindade (2014) "Who monitors the monitor? Effect of Party Observers on Electoral Outcomes." UC3M working papers.

[18] Carlson, Elizabeth (2010). "Great Expectations: Ethnicity, Performance, and Ugandan Voters." Working paper.

[19] Centre for Study of Developing Societies (CSDS) (2010) "Bihar Assembly Election Study 2010."

[20] Cobb, Rachael V., D. James Greiner and Kevin M. Quinn (2012). "Can Voter ID Laws Be Administered in a Race-Neutral Manner? Evidence from the City of Boston in 2008." *Quarterly Journal of Political Science* 7:1–33.

[21] Corstange, Daniel (2012). "Vote Trafficking in Lebanon." *International Journal of Middle East Studies* 44(3): 483-505.

[22] Easterly, William and Ross Levine (1997). "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4): 1203-1250.

[23] The Electoral Integrity Project (2013). Annual Workshop on "Concepts and Indicators of Electoral Integrity". Weatherhead Center, Harvard University, 3-4th June.

[24] Faller, J.K., Nathan, N.L. & White, A.R. (2014). "What Do I Need to Vote? Bureaucratic Discretion and Discrimination by Local Election Officials." *American Political Science Review*, available on CJO2014.

[25] Finan, Frederico and Laura Schechter (2012). "Vote-Buying and Reciprocity." *Econometrica* 80(2): 863-881.

[26] Fisman, Raymond, Daniel Paravasini, and Vikrant Vig (2012). "Cultural Proximity and Loan Outcomes." NBER Working Paper No.18096.

[27] Fujiwara, Thomas (2015). "Voting Technology, Political Responsivenss, and Infant Health: Evidence from Brazil." *Econometrica* 83(2): 423-464.

[28] Gerber, Alan S. and Todd Rogers (2009). "Descriptive Social Norms and Motivation to Vote: Everybody's Voting and So Should You." *Journal of Politics* 71(1): 178-191.

[29] Gonzalo-Ocantos, Ezequiel, Chad Kiewiet de Jonge, Carlos Melendez, Javier Osorio, and David W. Nickerson (2012). "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua." *American Journal of Political Science* 56(1): 202-217.

[30] Ghosh, Prabhat P. (2004). *Socio-Economic and Educational Status of Muslims in Bihar*. Monograph 01/2004. Patna: Asian Development Research Institute.

[31] Gupta, Saurabh (2008). Election Confidential (ELECON). In Piyush Gupta and R.K. Bagga (Eds.), *Compendium of eGovernance Initiatives* (288-298). Hyderabad, India: Universities Press.

[32] Hall, Thad E., Quin Monson, and Kelly D. Patterson (2009). "The Human Dimension of Elections: How Poll Workers Shape Public Confidence in Elections." *Political Research Quarterly* 62(3): 507-522.

[33] Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan (2003). "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation." *Journal of Political Economy* 111(3): 465-497.

[34] Hjort, Jonas (2014). "Ethnic Divisions and Production in Firms." *Quarterly Journal of Economics* 129(4): 1899-1946.

[35] Hyde, Susan (2007). "The Observer Effect in International Politics: Evidence from a Natural Experiment." *World Politics* 60(1): 37-63.

[36] Ichino, Nahomi and Matthias Schündlen (2012). "Deterring or displacing electoral irregularities? Spillover effects of observers in a randomized field experiment in Ghana." *Journal of Politics* 74(1): 292-307.

[37] Kramon, Eric and Keith R. Weghorst (2012). "Measuring Sensitive Attitudes in Developing Countries: Lessons from Implementing the List Experiment." *The Experimental Political Scientist* 3(2): 14-24.

[38] Kumar, Hari (2014). "The Device that Runs the World's Biggest Election." *The New York Times*, May 6.

[39] Kumar, Sanjay (2014a). "Bihar: interpreting the massive mandate." *The Hindu*, May 23.

[40] Kumar, Sanjay (2014b). "BJP Crafts a New Social Coalition in Bihar." *Economic & Political Weekly* 49(39): 95-98.

[41] Maskin, Eric and Jean Tirole (2004). "The Politician and the Judge: Accountability in Government." *The American Economic Review* 94(4): 1034-1054.

[42] Marx, Benjamin, Vincent Pons, and Tavneet Suri (2014). "Six Tests, Six Votes: Mobilizing voters in Kenya." Paper in progress.

[43] Marx, Benjamin, Vincent Pons, and Tavneet Suri (2015). "Diversity and Team Performance in a Kenyan Organization." Working paper.

[44] Martin, Elizabeth (2006). "Vignettes and Respondent Debriefings for Questionnaire Design and Evalution." U.S. Bureau of Census, Research Report Series: Survey Methodology 2006-2008.

[45] Miguel, Edward (2004). "Tribe or Nation? Nation Building and Public Goods in Kenya Versus Tanzania." *World Politics* 56 (3): 327-362.

[46] Miguel, Edward and Mary Kay Gugerty (2005). "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of Public Economics* 89(11-12): 2325-2368.

[47] Prat, Andrea (2002). "Should a Team Be Homogeneous?" *European Economic Review* 46(7): 1187-1207.

[48] Presidential Commisson on Election Administration (2014). "The American Voting Experience: Report and Recommendations of the Presidential Commission on Election Administration." Washington, DC.

[49] Rohde, David (2004). "On New Voting Machine, the Same Old Fraud." *The New York Times*, April 27.

[50] Rukmini S. (2014). "The continuing grip of caste." *The Hindu*, March 16.

[51] Shayo, Moses and Asaf Zussman (2011). "Judicial Ingroup Bias in the Shadow of Terrorism." *Quarterly Journal of Economics* 126(3): 1447-1484.

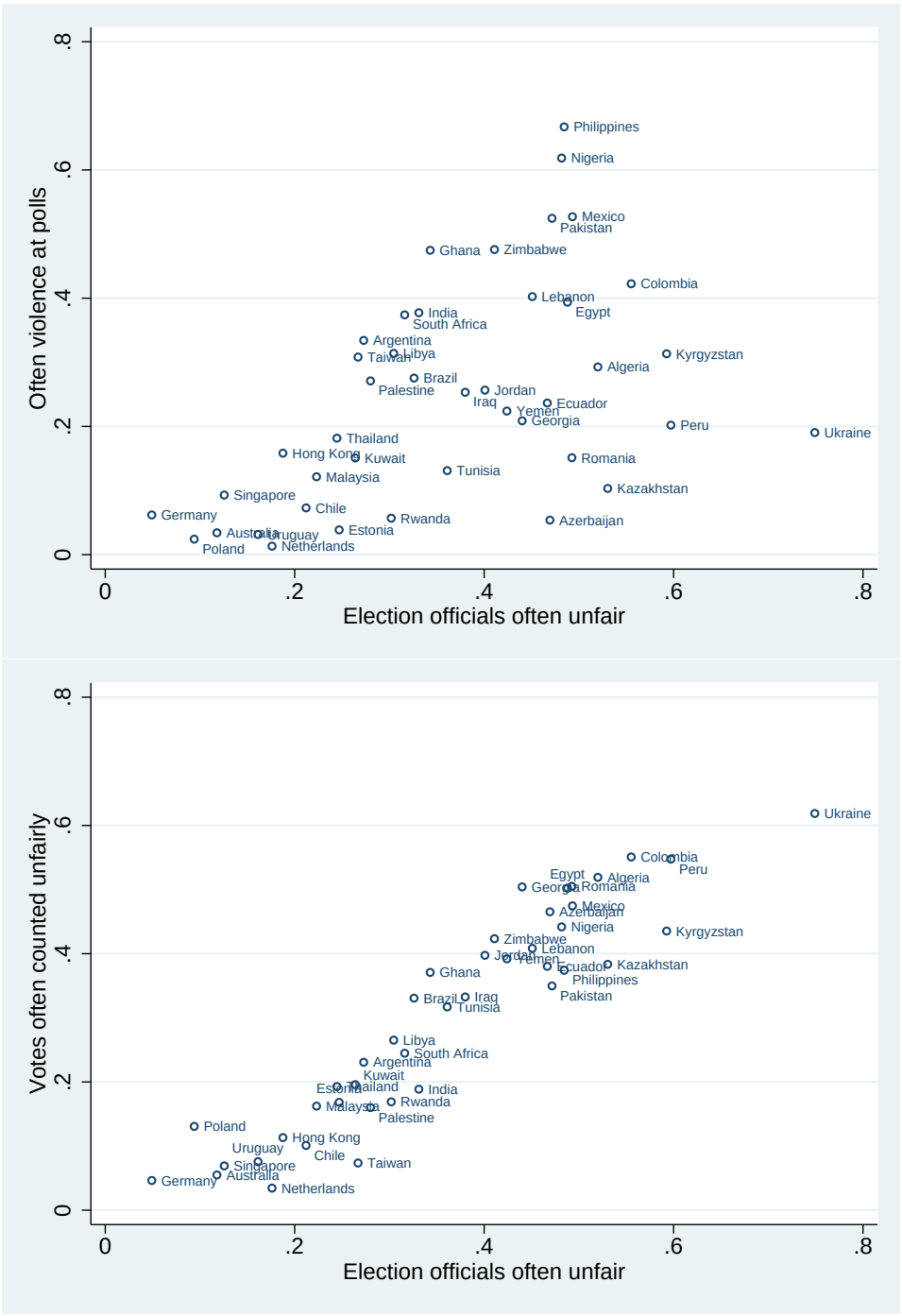
[52] Shue, Kelly and Erzo F.P. Luttmer (2009). "Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes." *American Economic Journal: Economic Policy* 1(1): 229-257.

[53] Susewind, Raphael (2014). GIS shapefiles for India's parliamentary and assembly constituencies including polling booth localities. Bielefeld University. doi:10.4119/unibi/2674065.

[54] Svensson Jakob and David Yanagizawa Drott (2012). "Estimating Impact in Partial vs. General Equilibrium: A Cautionary Tale from a Natural Experiment in Uganda." Working paper.

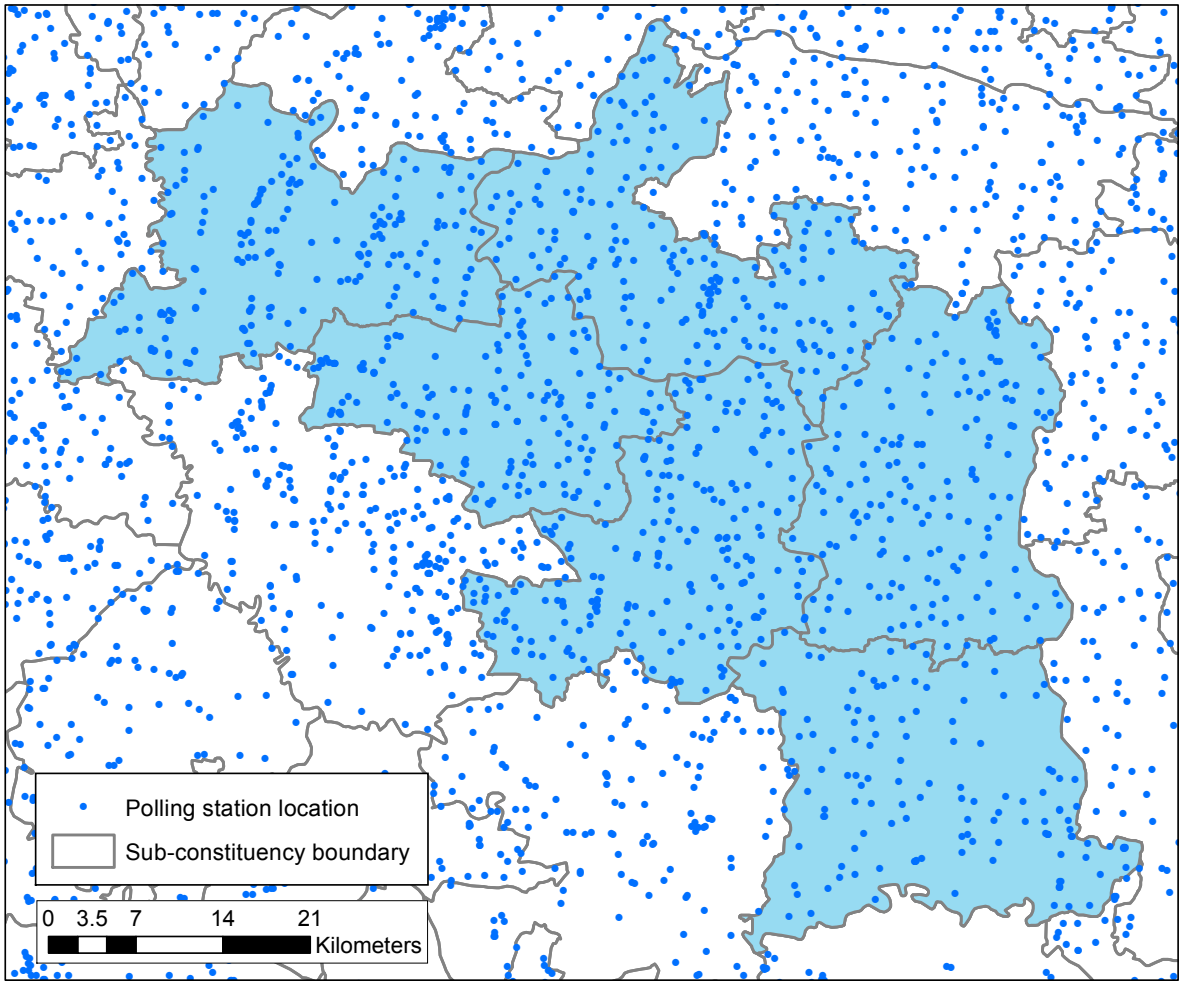
[55] Wittsoe, Jeffrey (2013). *Democracy Against Development: Lower-Caste Politics and Political Modernity in Postcolonial India.* Chicago: University of Chicago Press.

[56] Wolchok, Scott, Eric Wustrow, J. Alex Halderman, Hari K. Prasad, Arun Kankipati, Sai Krishna Sakhamuri, Vasavya Yagati, and Rop Gonggrijp (2010). Security Analysis of India's Electronic Voting Machines. In *Proceedings of the 17th ACM Conference on Computer and Communications Security (CCS '10).* Chicago: October 4-8.



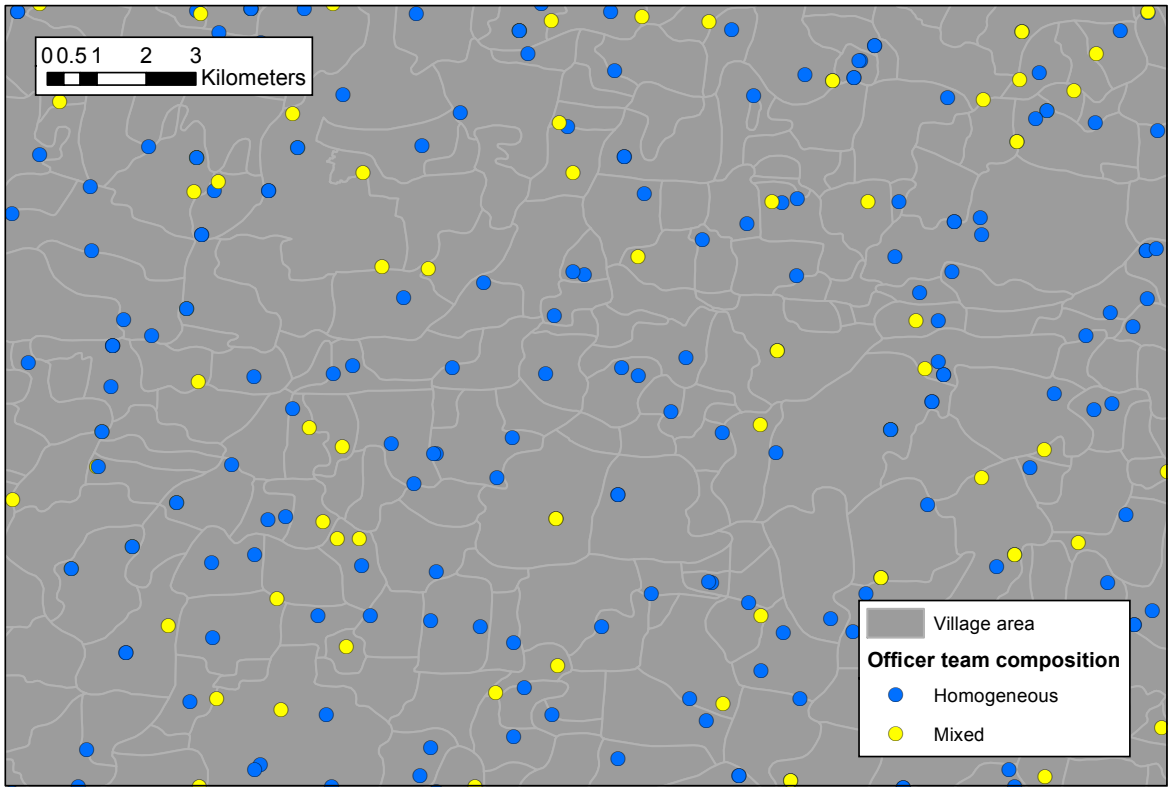
Notes: Measures computed using World Values Survey Wave 6 (2010-2014). "Election officials often unfair" is the weighted percentage of respondents in each country, when asked "In your view, how often do the following things occur in this country's elections?", answering "Not at all often" or "Not often" to "Election officials are fair", against the alternatives of "Very often", "Fairly often", or "Don't know/Not answer". "Often violence at polls" is the percentage answering "Very often" or "Fairly often" to "Voters are threatened with violence at the polls." "Votes often counted unfairly" is the percentage answering "Not at all often" or "Not often" to "Votes are counted fairly."

Figure 1: Election administration difficulties by country



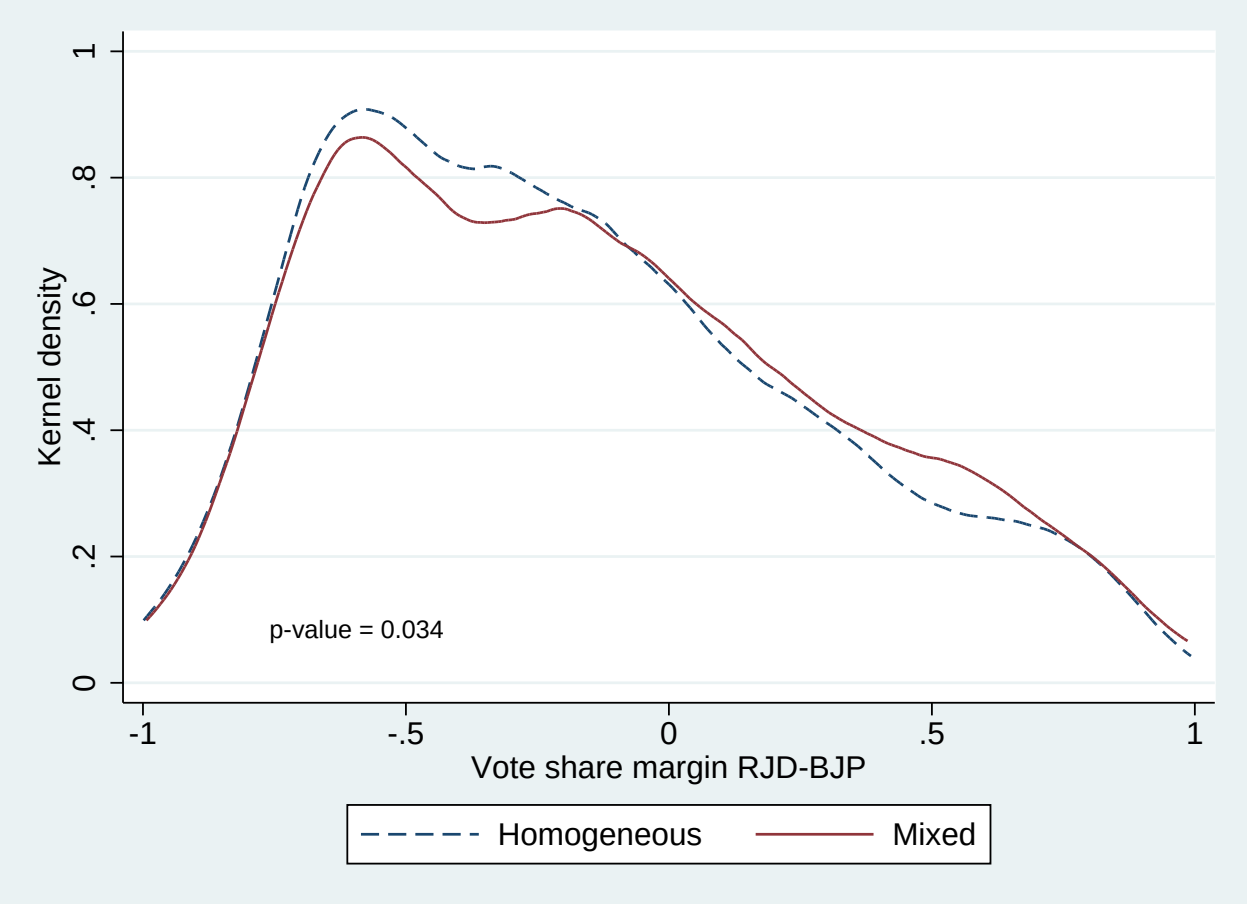
Notes: Shaded area in the figure indicates the extent of an example parliamentary constituency.

Figure 2: Polling station distribution across example parliamentary constituency



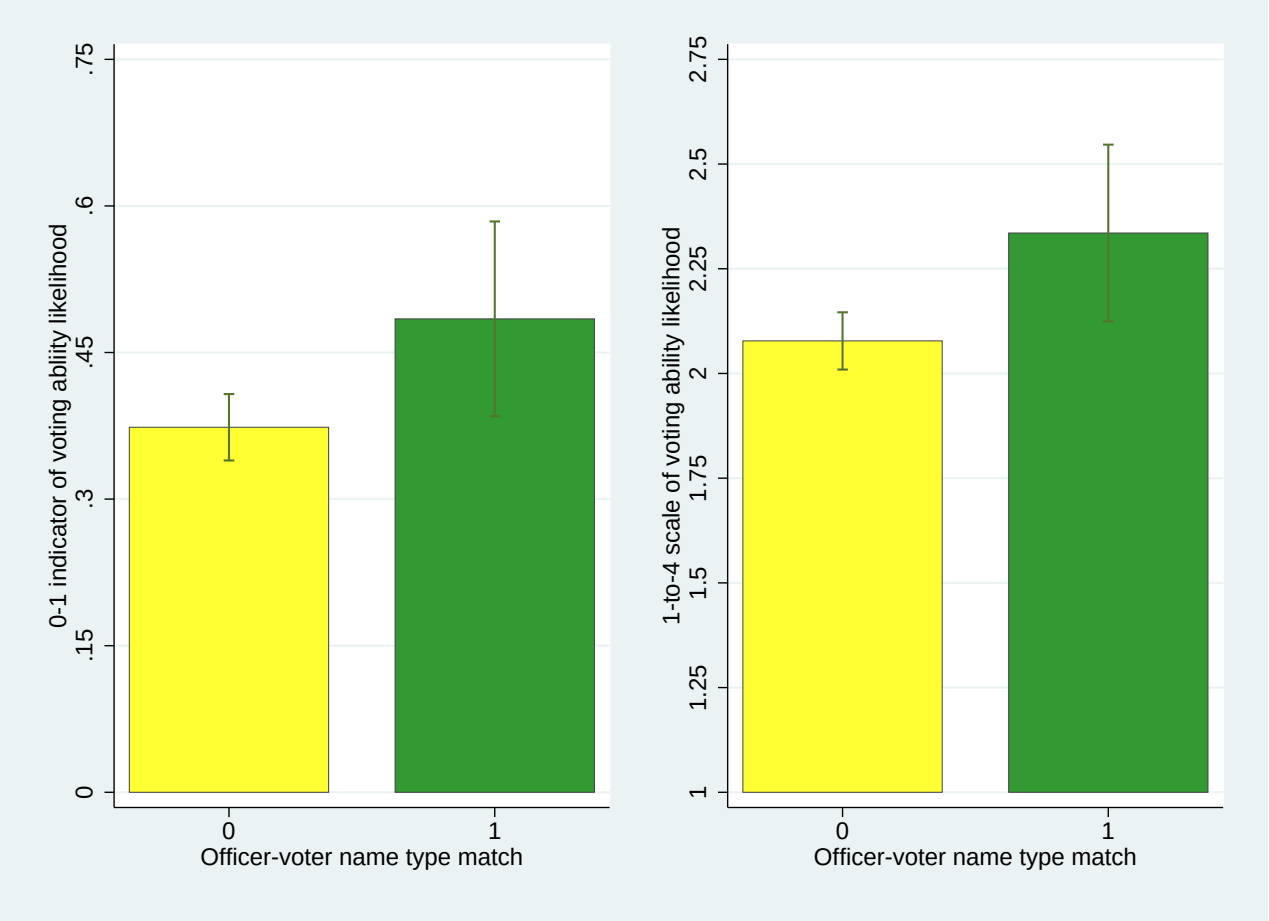
Notes: Each circle represents a polling station, with the color signifying whether the officer team was homogeneous or mixed in composition.

Figure 3: Variation in officer team composition across polling stations



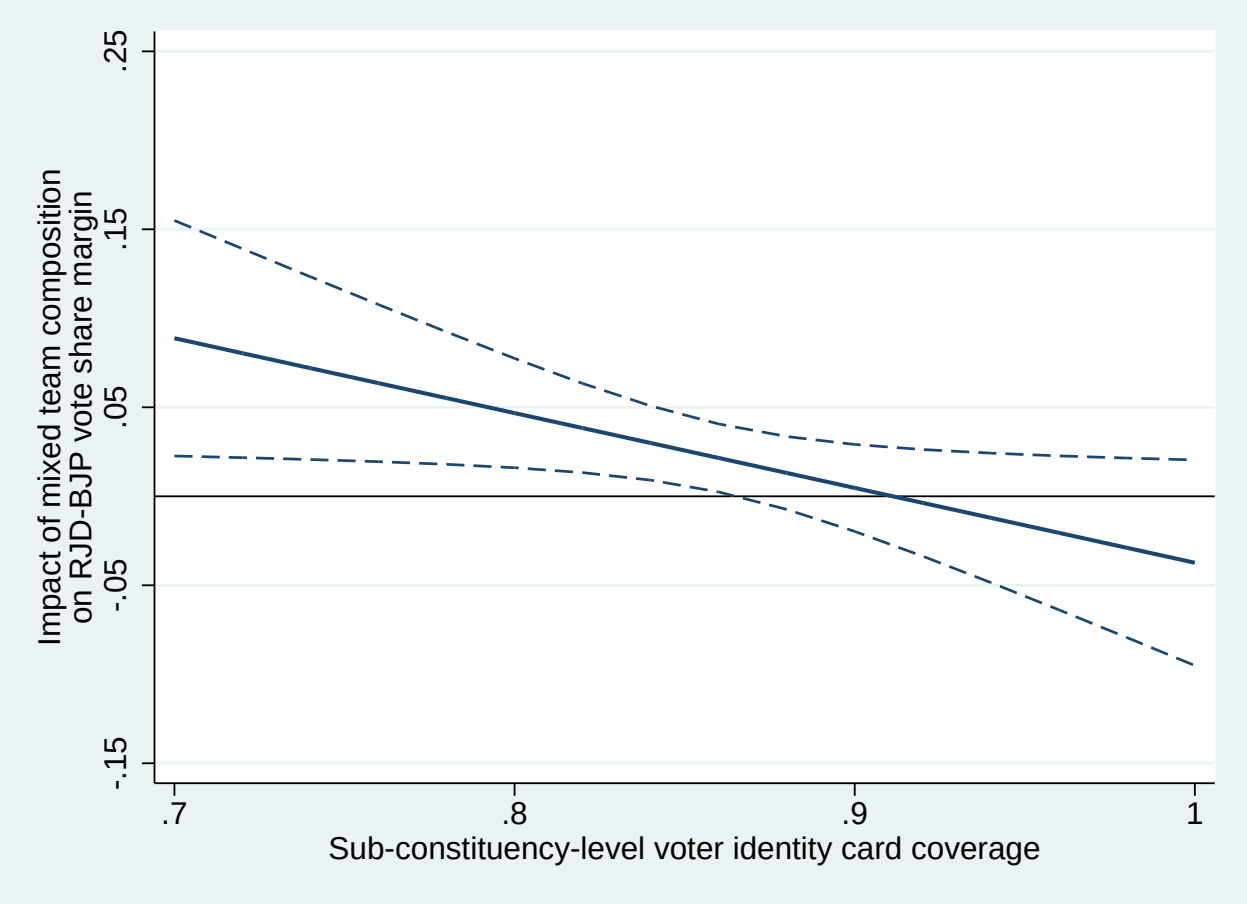
Notes: The figure plots kernel density estimates of the polling station-level vote share margin between the RJD and BJP coalitions, separately for polling stations with homogeneous (dashed line) and mixed (solid line) teams of polling stations officers. Estimated using an Epanechnikov kernel. The p-value is computed using the Kolmogorov-Smirnov equality-of-distributions test for the two groups of polling stations.

Figure 4: Empirical distribution of coalition vote share margins by team composition



Notes: The figure in the left panel depicts the estimated probabilities of an officer respondent indicating that a hypothetical individual described in the officer's survey vignette would be "(3) Likely" or "(4) Very Likely" able to cast a vote, as opposed to "(2) Unlikely" or "(1) Very Unlikely". The estimates are based on the regression in column (2) of Table 3, assuming mean values of all control variables. The left bar represents the randomly assigned subset of officer respondents for whom the hypothetical individual's type (Muslim, Yadav, Brahmin) did not match the officer's own type, while the right bar represents the subset for whom the types match. The figure in the right panel depicts the estimated 4-point scale values of a hypothetical individual's voting ability likelihood, based on the same question as the left panel. The estimates are based on the regression in column (4) of Table 3, assuming mean values of all control variables. The notes to Table 3 provide the full vignette question text. Error bars signify 95 percent confidence intervals.

Figure 5: Own-type bias in officer assessment of voting qualification



Notes: Figure plots the estimated polling-station-level impact of mixed team composition on the vote share margin between the RJD and BJP coalitions at different levels of sub-constituency-level voter identity card coverage. Dashed lines signify 95 percent confidence intervals. Calculated using the estimates from Column (5) of Table 7.

Figure 6: Heterogeneity by voter identity card coverage in impact of team composition

Table 1. Randomization check

	Homog. team (1)	Mixed team (2)	Difference (3)	*p*-value (4)	Obs. (5)
Panel A. Electorate characteristics					
Ln total registered voters	6.873 [0.314]	6.905 [0.305]	0.009 (0.007)	0.160	5,561
Share female registered voters	0.463 [0.023]	0.463 [0.022]	0.000 (0.001)	0.864	5,561
Share Muslim/Yadav registered voters	0.128 [0.172]	0.135 [0.175]	0.005 (0.005)	0.312	5,561
Panel B. Prior election (2010) characteristics					
Ln total votes	6.061 [0.332]	6.057 [0.319]	-0.007 (0.009)	0.412	5,275
Vote share margin RJD-BJP coalition	-0.287 [0.378]	-0.272 [0.376]	0.000 (0.009)	0.992	3,947
Ln votes RJD coalition	3.941 [1.424]	3.945 [1.403]	-0.009 (0.023)	0.694	5,246
Ln votes BJP coalition	4.940 [0.995]	4.901 [1.019]	-0.003 (0.025)	0.899	3,946
Panel C. Officer characteristics					
Age	42.313 [9.781]	43.264 [9.677]	0.910 (0.866)	0.294	517
College graduate	0.695 [0.462]	0.675 [0.469]	-0.014 (0.041)	0.728	516
Ln monthly salary	9.539 [0.609]	9.584 [0.562]	0.047 (0.053)	0.371	503
First time officer	0.342 [0.475]	0.325 [0.469]	-0.017 (0.042)	0.686	511
Panel D. Registered voter characteristics					
Muslim/Yadav	0.430 [0.495]	0.446 [0.497]	0.016 (0.016)	0.306	3,903
Age	45.402 [16.844]	45.453 [16.429]	0.073 (0.577)	0.900	3,877
Female	0.580 [0.494]	0.537 [0.499]	-0.043 (0.016)	0.008	3,903
Literate	0.385 [0.487]	0.413 [0.493]	0.028 (0.018)	0.107	3,901
Household head	0.458 [0.498]	0.463 [0.499]	0.006 (0.018)	0.730	3,903
Ln monthly household income	8.212 [0.828]	8.254 [0.827]	0.038 (0.035)	0.286	3,326
Voter identity card possession	0.945 [0.228]	0.940 [0.237]	-0.004 (0.009)	0.647	3,903

Notes: Columns (1) and (2) report variable means with standard deviations in brackets for homogeneous and mixed officer teams. Column (3) reports the coefficient from an OLS regression where the listed outcome is regressed on an indicator for polling station mixed team composition and column (4) reports the associated p-value. Panels A and B also include sub-constituency and number of officer fixed effects. Prior election characteristic outcome values are based on the average value across all polling stations from 2010 in the same location as the 2014 polling station, as the total numbers and locations of polling stations change across election cycles. 2014 coalition definitions are used. Panel C is restricted to non-Muslim/Yadav officer respondents, due to the definition of mixed teams. Additionally included are sub-constituency fixed effects. Panel D considers registered voter respondents and additionally includes strata fixed effects (sub-constituency and above-below district-level Muslim/Yadav registered voter percentage median). *Significant at 10% level **5% level ***1% level.

Table 2. Impacts of randomized officer team composition on voting outcomes

	Ln votes RJD (1)	Ln votes BJP (2)	Vote share margin RJD-BJP (3)	Ln total votes (4)
Panel A. Within-station effects				
Mixed team	0.046*	-0.041*	0.023**	0.001
	(0.027)	(0.021)	(0.010)	(0.008)
Muslim/Yadav registered voter %	0.031***	-0.030***	0.015***	-0.000**
	(0.001)	(0.001)	(0.003)	(0.000)
Ln total registered voters	1.008***	1.177***	-0.060***	0.935***
	(0.060)	(0.048)	(0.023)	(0.018)
Observations	5,535	5,549	5,552	5,552
Homogeneous team mean [SD]	4.451	5.143	-0.181	6.180
	[1.198]	[0.969]	[0.452]	[0.402]
Panel B. Cross-station spillovers				
Mixed team	0.045*	-0.040*	0.023**	0.000
	(0.027)	(0.021)	(0.010)	(0.008)
Number mixed team neighbor stations	0.031	-0.042**	0.026***	0.003
	(0.025)	(0.019)	(0.010)	(0.008)
Total neighbor stations	-0.044***	0.046***	-0.032***	-0.017***
	(0.014)	(0.010)	(0.005)	(0.005)
Observations	5,535	5,549	5,552	5,552
Number locations	3,619	3,619	3,619	3,619

Notes: All columns report OLS estimates from regressions at the polling station level of the listed variable on an indicator for mixed team composition. Additionally included are sub-constituency and number of officer fixed effects and controls for Muslim/Yadav share of registered voters and log total registered voters. Standard errors clustered at the station level. In Panel B, variables for the numbers of total and mixed composition team neighboring polling stations are also included. Neighbor stations are polling stations within the same location (building/compound) as a given polling station. Standard errors in Panel B clustered at the location level. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table 3. Vignette experiment: own-type bias in officer assessment of voting qualification

	Ability to cast vote			
	0-1 indicator		4-point scale	
	(1)	(2)	(3)	(4)
Officer-potential voter name type match	0.103*	0.111**	0.237**	0.258**
	(0.056)	(0.055)	(0.117)	(0.116)
Observations	871	869	871	869
Name fixed effects	X	X	X	X
Officer type fixed effects	X	X	X	X
Individual controls	X	X	X	X
Polling station controls		X		X
Non-match group outcome mean [SD]	0.382	0.380	2.096	2.092
	[0.486]	[0.486]	[0.974]	[0.974]

Notes: Columns (1) and (2) report OLS estimates from regressions at the officer level of an indicator variable taking value 1 if the respondent answers "Very likely (4)" or "Likely (3)" as opposed to "Unlikely (2)" or "Very unlikely (1)" to the question: "A voter named [RANDOMLY ASSIGNED] arrives at the polling station without an EPIC card but has a government voter's slip without a photograph. He can recite his name and other particulars. On a scale of 1 to 4, how likely do you think it is that he would be allowed to cast a vote based on this information?" and 0 otherwise, on an indicator variable for whether the officer's own type matches that (Muslim, Yadav, Brahmin) of the randomly assigned voter name. Columns (3) and (4) report OLS estimates from regressions with the 1-4 scale value as the outcome. Columns (1) and (3) include fixed effects for respondent name and officer type, the stratification variables (sub-constituency in which officer was assigned to a polling station and officer category [Muslim/Yadav at mixed polling station, non-Muslim/Yadav at mixed polling station, non-Muslim/Yadav at homogeneous polling station] plus the following individual level controls: age, log monthly salary, an indicator for first term of service at a polling station, and fixed effects for occupation type, education level, and polling station position. Columns (2) and (4) further include polling station level controls for log total registered voters and proportion Muslim/Yadav registered voters, and fixed effects for station location type and number of team officers. Standard errors are clustered at the polling station level. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table 4. List experiments: biased officer behavior on election day

	Control (1)	Treatment (2)	Difference (3)	Obs. (4)
Panel A. Election officers				
"One or more of the election officers at the polling station treated some voters differently based on the voters' religion or caste."	1.915 [0.777]	2.086 [0.873]	0.192*** (0.048)	878
"One or more of the election officers tried to influence some voters' choice of candidate or make it harder for them to vote."	2.883 [0.376]	2.960 [0.398]	0.047** (0.021)	877
Panel B. Registered voters				
"One or more of the election officers at your polling station treated you or others differently based on your religion or caste."	2.036 [0.758]	2.280 [0.913]	0.254*** (0.026)	3,532
"One or more of the election officers at your polling station tried to influence how you or others voted or to make it more difficult for you or them to cast votes."	2.396 [0.682]	2.539 [0.809]	0.128*** (0.023)	3,547

Notes: Columns (1) and (2) report unconditional means and standard deviations of the control (individuals receiving a list of four questions with the listed statement omitted) and treatment (individuals receiving a list of the same four questions plus the listed statement included). Column (3) reports the coefficient of an OLS regression at the individual level of the total number of statements the respondent indicated occurred at the polling station during the 2014 elections and sub-constituency fixed effects. Additionally included are polling-station-level controls for log total registered voters, share Muslim/Yadav registered voters, and fixed effects for location type and number of officers. In Panel A, additional officer-level controls are age and log monthly salary and fixed effects for occupation and education and controls for log monthly salary and prior election experience. In Panel B, additional registered-voter-level controls are fixed effects for education level and occupation type and controls for age, sex, household head status, and log monthly household income. Standard errors are clustered at the polling station level. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table 5. Overall polling station management

	Canvassing at station		Disorderly behavior at station	
	(1)	(2)	(3)	(4)
Mixed team	0.005	0.002	-0.001	0.006
	(0.005)	(0.007)	(0.011)	(0.014)
Muslim/Yadav		-0.003		0.010
		(0.006)		(0.014)
Muslim/Yadav * Mixed team		0.006		-0.018
		(0.010)		(0.019)
Observations	3,733	3,733	3,775	3,775
Polling stations	351	351	351	351
Outcome mean [SD]	0.020	0.020	0.068	0.068
	[0.141]	[0.141]	[0.256]	[0.256]

Notes: All columns report OLS estimates from regressions at the individual level of the listed variable on an indicator for mixed officer team composition at the polling station. Even-numbered columns include an interaction with an indicator for whether the respondent is Muslim/Yadav. Additionally included in all columns are fixed effects for the stratification variables (sub-constituency and above-below district level median in terms of MY elector percentage) and individual-level controls for age, sex, education level, household head status, household structure type, occupation type, and log monthly household income. Polling-station-level controls are included for log total registered voters, share Muslim/Yadav registered voters, and fixed effects for the number of officers stationed at the polling station. Standard errors clustered at the polling station level. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table 6. Identity verification experience of potential voters

	Able to cast vote		Satisfactory overall station experience	
	Homog. team (1)	Mixed team (2)	Homog. team (3)	Mixed team (4)
Muslim/Yadav	-0.104*	-0.047	-0.066+	0.001
	(0.062)	(0.066)	(0.041)	(0.031)
Possess voter identity card	0.011	0.105**	-0.019*	0.013
	(0.025)	(0.046)	(0.010)	(0.024)
Muslim/Yadav * Possess voter identity card	0.109*	0.033	0.072*	-0.004
	(0.062)	(0.066)	(0.043)	(0.033)
Observations	1,929	1,946	1,907	1,900
Polling stations	175	176	175	176
Outcome mean [SD]	0.981	0.980	0.981	0.982
	[0.137]	[0.138]	[0.136]	[0.133]

Notes: All columns report OLS estimates from regressions at the individual level of the listed variable on an interaction of the Muslim-Yadav respondent indicator with an indicator for voter identity card possession, for the sample of polling stations indicated in each column. Additionally included are polling station-level fixed effects and individual-level controls for age, gender, education level, household head status, household structure type, occupation type, and log monthly household income. "Satisfactory overall station experience" is an indicator for whether the respondent indicated that their overall voting experience at the polling station on election day was "Excellent"/"Good"/"Fair", as opposed to "Poor". Standard errors clustered at the polling station level. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent. + p-value = 0.112.

Table 7. Heterogeneity in effects of team composition by voter identity card coverage

	Ln votes RJD		Ln votes BJP		Vote share margin RJD-BJP	
	(1)	(2)	(3)	(4)	(5)	(6)
Mixed team *	-0.009*	-0.014*	0.005	0.010*	-0.004**	-0.006**
Voter identity card coverage %	(0.005)	(0.008)	(0.004)	(0.006)	(0.002)	(0.003)
Mixed team	0.809	1.072	-0.453	-0.960*	0.383**	0.558*
	(0.457)	(0.768)	(0.332)	(0.579)	(0.173)	(0.314)
Observations	5,429	5,429	5,439	5,439	5,442	5,442
Polling station controls	X	X	X	X	X	X
Sub-constituency fixed effects	X	X	X	X	X	X
Interacted sub-constituency controls		X		X		X
Implied effect: minimum sample	0.129**	0.152***	-0.079**	-0.086**	0.063***	0.072***
card coverage sub-constituency	(0.054)	(0.057)	(0.040)	(0.043)	(0.022)	(0.024)

Notes: All columns report OLS estimates from regressions at the polling station level of the listed variable on an indicator for mixed team composition interacted with the sub-constituency-level percentage of registered voters with a voter ID card. Also included are sub-constituency and number of officer fixed effects and controls for Muslim/Yadav share of registered voters and log total registered voters. Even-numbered columns additionally include interactions (not shown) with sub-constituency-level measures of the population proportions that are literate and Schedule Caste/Schedule Tribe, and the share of registered voters that are Muslim/Yadav (none of these interaction effects are statistically significant). The implied effect given in each column reflects the estimated impact of mixed team composition for the sub-constituency with the lowest level of voter identity card coverage observed in the sample. Coverage ranges between 76.3 and 93.9 percent in sample sub-constituencies. The sample trims the top one percent of observations in terms of absolute value of coalition vote share margin (polling stations with a margin greater than 88 percentage points). *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table 8. Variation in other officer characteristics by Muslim/Yadav identity

	Survey data				Administrative data	
	Age	Ln monthly salary	College graduate	First time officer	Age	Ln monthly salary
	(1)	(2)	(3)	(4)	(5)	(6)
Muslim/Yadav officer	-0.340	0.018	-0.025	-0.019	0.439	0.001
	(0.554)	(0.029)	(0.030)	(0.031)	(0.373)	(0.012)
Observations	912	888	911	903	5,983	6,198
Non-Muslim/Yadav mean [SD]	42.822	9.563	0.684	0.333	44.975	9.291
	[9.727]	[0.584]	[0.465]	[0.472]	[9.802]	[0.363]

Notes: All columns report OLS estimates from regressions at the officer level of the listed variable on an indicator for Muslim/Yadav identity. Additionally included are sub-constituency and officer-position fixed effects. Columns (1) through (4) are based on reported data from the survey of officers. Columns (5) and (6) are based on full sample of administrative data available for the same district in which the surveys were conducted. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table 9. Changes in election outcomes under alternative officer assignment mechanisms

	Alternative 1: All mixed teams		Alternative 2: No mixed teams			
	BJP to RJD victory (1)	Vote share margin range (2)	RJD to BJP victory (3)	Vote share margin range (4)	RJD/ BJP top two parties (5)	Total races (6)
National parliament, 2014	1	-0.024	1	0.010	29	40
State assembly elections, 2010 (2014 voter identity card coverage)	14	[-0.037, -0.0003]	2	[0.004, 0.006]	185	243
State assembly elections, 2010 (2010 voter identity card coverage)	33	[-0.066, -0.0003]	6	[0.004, 0.023]	185	243

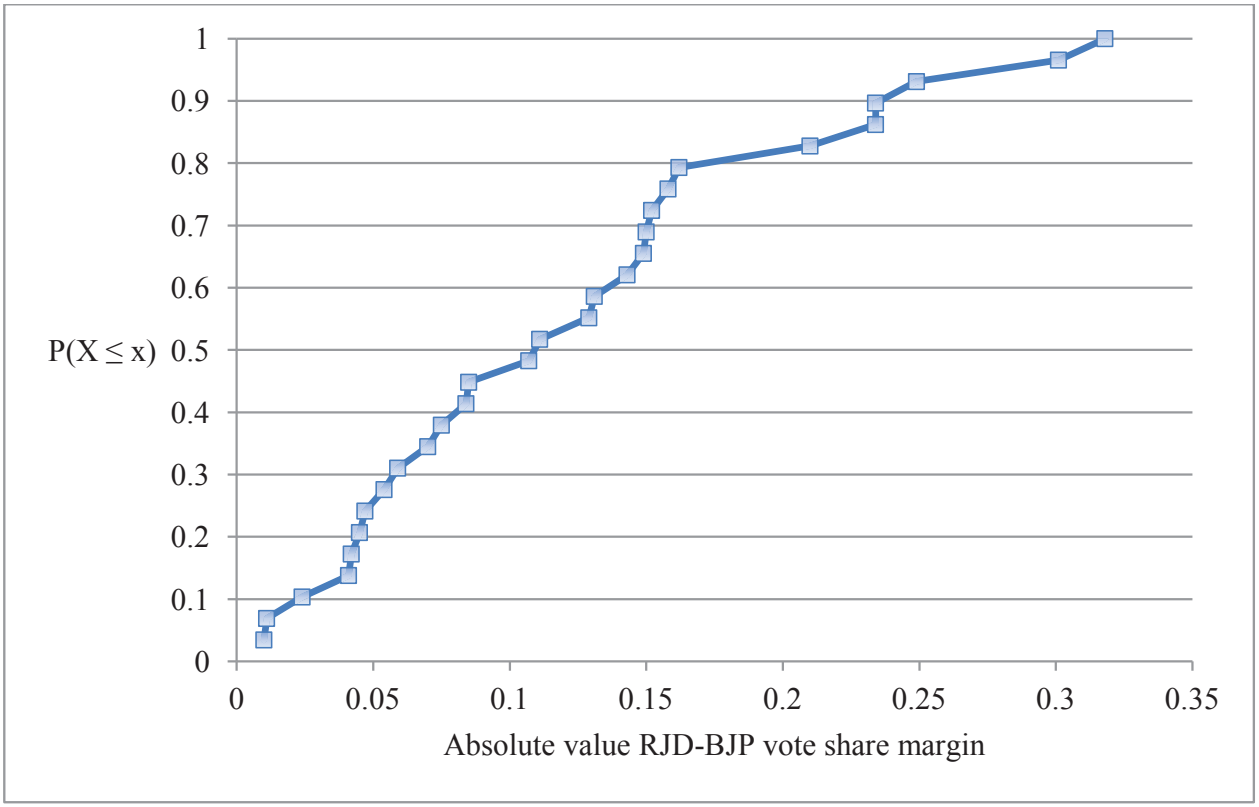
Notes: This table reports estimates of the potential number of races for which the winning candidate would have switched between the RJD coalition and the BJP coalition, under two alternative officer assignment scenarios. Alternative 1 is the absence of any mixed composition officer teams and Alternative 2 is the presence of all mixed composition teams, assuming an initial 0.324 proportion of mixed teams (that observed in the available 2014 data). Columns (1) and (3) give the number of races for which the winning party would change as indicated. Columns (2) and (4) give the range of the RJD-BJP coalition vote share margins observed in the impacted constituencies. Column (5) gives the number of races in which the RJD and BJP coalitions fielded the top two candidates, and column (6) the total number of races in Bihar for that election cycle. The calculation accounts for spillover effects from neighboring mixed team polling stations and heterogeneity in impact by voter identity card coverage (at the sub-constituency level). For the 2010 state assembly elections, two sets of results are presented: using the voter identity coverage from 2014; and using the (lower) voter identity coverage from 2010.

Appendix

Figure A1: Government-issued voter identity card

Figure A2: Neighboring polling stations in close proximity

Figure A3: Polling officer team during election day proceedings



Notes: Figure plots the empirical cumulative distribution function of the absolute value of the parliamentary-constituency-level vote share margin between the RJD and BJP coalitions, for the 29 of 40 races where these two coalitions fielded the top two candidates.

Figure A4: Cumulative distribution function of coalition vote share margins

Table A1. Vignette experiment names and list experiments prompt

Vignette experiment names

Muslim: Najam Uddin, Mustak Ansari, Mohammed Alam

Yadav: Ajay Yadav, Kailesh Yadav, Surendra Yadav

Brahmin: Arjun Tripathi, Rohit Mishra, Alok Chaturvedi

List experiments prompt

"I'm going to read you a list of various statements, and I would like for you to tell me how many of them occurred during the previous 2014 Lok Sabha election. Please, count to yourself. Do not tell me which ones, only HOW MANY IN TOTAL. For example, it might be that none of them occurred, all of them occurred, or any number in between."

Table A2. Survey sampling

Registered voters survey

Polling stations in urban areas, where locating specific individuals based on the information available in the electoral roll would not have been feasible, were excluded from the sample (8.3 percent). Additionally excluded were polling stations with only three election officers (0.7 percent), as were polling stations that were split across a main polling station and an extension station (9.8 percent). The list of registered voters was at the (main+extension) level, so it was not possible to determine to which of the main station or extension individuals were assigned. The only difference between having a main and extension station versus two polling stations in the same location is whether the threshold for maximum registered voters at a single station was reached after the formal yearly deadline to split polling stations. Administration is otherwise identical.

In some locations, fewer than three Muslims or two Yadavs were identified in the list. If too few Muslims were available, Yadavs were randomly drawn to fill the positions when possible, and vice versa. If fewer than five Muslims and Yadavs in total were identified, individuals that were neither Muslim nor Yadav were randomly drawn to fill the position.

Seasonal migration is common in the survey area and the electoral rolls contain errors (e.g. listed individuals may be duplicates or have moved and registered at another polling station without being deleted from the list at the previous station). Therefore, randomly drawn backup respondents were also identified for each primary respondent. In the final sample, 36.6 percent of respondents were from the primary sample, 22.6 percent

were the first backup, 14.6 percent were the second backup, 11.2 percent were the third backup, and 15 percent were fourth backup or higher. These rates of replacement are similar to those of other surveys in the region which identified respondents based on the electoral roll (Banerjee 2014). The rate of primary versus backup respondents does not differ significantly by whether the polling station is mixed versus homogeneous team. The consent rate among located respondents was very high, with more than 98.5 percent of individuals agreeing to participate. If an individual indicated that they did not go to the polling station to attempt to vote on election day, the next backup individual was then substituted.

Election officers survey

A total of 6,251 officers served at polling stations during the 2014 election in the district in which the survey was conducted. Out of these officers, 6,045 had phone numbers listed in the administrative data which were not obviously incorrect (i.e. having the wrong number of digits or all zero numerals). Of these 6,045 individuals, 614 officers were inferred as Muslim or Yadav. Each of these individuals was attempted to be reached by phone. One non-Muslim/Yadav officer was randomly selected for calling from each of the mixed composition teams of which the previous 614 Muslim/Yadav officers were a member. If the officer could not be reached or did not consent, another non-Muslim, non-Yadav officer was selected as a replacement, if possible. An additional 600 homogeneous polling teams were randomly chosen and an officer from within the team was randomly selected. Again, if the officer could not be reached or did not consent, another officer was selected as a replacement, if possible. A total of 2,350 officers were called in total. In 30 percent of instances the individual was not reachable (in the vast majority of cases due to the listed phone number not being functional). Willingness to participate was very high among the officers who were reachable, with only 2 percent (33) of officers not consenting to be surveyed in the future. Calling yielded 380 mixed team polling stations with at least one M-Y officer and non-MY officer each confirmed as consenting and 436 homogeneous polling stations with at least one officer confirmed as consenting, from which 305 mixed team and homogeneous pollling stations each were randomly selected as described in the main text.

Table A3. Randomization check – spatial characteristics

	Homog. team (1)	Mixed team (2)	Difference (3)	*p*-value (4)	Obs. (5)
Number mixed team neighbor stations	0.385 [0.746]	0.386 [0.719]	-0.012 (0.018)	0.493	5,561
Total neighbor stations	1.200 [1.614]	1.191 [1.647]	-0.027 (0.034)	0.420	5,561
Number mixed team stations within 0.25km	0.420 [1.078]	0.452 [1.159]	0.026 (0.030)	0.392	5,097
Number mixed team stations within 0.25-0.75km	2.536 [4.263]	2.622 [4.470]	0.066 (0.084)	0.430	5,097
Total stations within 0.25km	1.357 [2.930]	1.336 [2.904]	-0.025 (0.073)	0.735	5,097
Total stations within 0.25-0.75km	7.893 [12.830]	7.958 [12.904]	0.069 (0.232)	0.768	5,097
Number mixed team stations within village	1.210 [2.178]	1.309 [2.287]	0.043 (0.083)	0.607	3,231
Number mixed team stations in neighboring villages	4.688 [3.908]	4.829 [4.015]	-0.040 (0.136)	0.768	3,216
Total stations within village	3.686 [5.551]	3.812 [5.868]	0.088 (0.212)	0.676	3,231
Total stations in neighboring villages	14.259 [10.694]	14.479 [10.544]	0.065 (0.369)	0.861	3,216

Notes: Columns (1) and (2) report variable means with standard deviations in brackets for homogeneous and mixed officer teams. Column (3) reports the coefficient from an OLS regression where the listed outcome is regressed on an indicator for polling station mixed team composition and column (4) reports the associated p-value. Also included are sub-constituency and number of officer fixed effects. Neighbor stations are polling stations within the same building/compound of a given polling station. Stations with 0.25 and within 0.25-0.75km are non-neighbor stations within 0.25km and 0.25-0.75km of a given polling station, respectively. Numbers of stations within a village and in neighboring villages are the numbers of non-neighbor polling stations within the same village as a given polling station and in villages adjacent to a given polling station's village, respectively. Sample is restricted to those polling stations matched to the dataset of polling station GPS locations. Village-related outcomes further exclude stations in villages which are in the top 1 percent of the distribution in terms of number of polling stations contained within, or their neighboring villages.*Significant at 10% level **5% level ***1% level.

Table A4. Cross-position balance

	Presiding officer (1)	Polling officer 1 (2)	Polling officer 2 (3)	Polling officer 3 (4)	Polling officer 4 (5)
Muslim/Yadav presiding officer		-0.006	0.006	-0.004	-0.016
		(0.014)	(0.014)	(0.014)	(0.030)
Muslim/Yadav polling officer 1	-0.005		-0.004	-0.019	-0.015
	(0.013)		(0.013)	(0.013)	(0.027)
Muslim/Yadav polling officer 2	0.006	-0.004		0.014	-0.009
	(0.014)	(0.014)		(0.015)	(0.027)
Muslim/Yadav polling officer 3	-0.003	-0.018	0.012		-0.020
	(0.013)	(0.012)	(0.014)		(0.029)
Muslim/Yadav polling officer 4	-0.014	-0.013	-0.009	-0.017	
	(0.026)	(0.030)	(0.031)	(0.025)	
Observations	5,561	5,561	5,561	5,523	1,178

Notes: Each column reports coefficients from an OLS regression where the outcome is Muslim/Yadav assignment to the specified position, and is regressed on dummies for Muslim/Yadav assignment to the other polling officer team positions specified in table. Additionally included are sub-constituency and number of officer fixed effects. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table A5. Position- and number-specific impacts on voting outcomes

	Ln votes RJD (1)	Ln votes BJP (2)	Vote share margin RJD-BJP (3)	Ln total votes (4)
Panel A. Position				
Muslim/Yadav presiding officer	-0.006	-0.017	0.007	-0.012
	(0.052)	(0.043)	(0.020)	(0.018)
Muslim/Yadav polling officer 1	0.088*	-0.013	0.031*	0.017
	(0.050)	(0.037)	(0.019)	(0.011)
Muslim/Yadav polling officer 2	0.050	-0.064	0.021	0.002
	(0.052)	(0.044)	(0.020)	(0.014)
Muslim/Yadav polling officer 3	0.054	-0.086**	0.037*	0.000
	(0.050)	(0.040)	(0.019)	(0.019)
Muslim/Yadav polling officer 4	0.111	-0.001	0.041	0.001
	(0.189)	(0.170)	(0.087)	(0.033)
F-test p-value: equality of coeffs.	0.731	0.612	0.824	0.614
Observations	5,276	5,290	5,293	5,293
Panel B. Number				
Any Muslim/Yadav officer	0.055*	-0.046**	0.027**	0.002
	(0.028)	(0.022)	(0.011)	(0.008)
Multiple Muslim/Yadav officers	-0.061	0.040	-0.024	-0.010
	(0.061)	(0.053)	(0.024)	(0.018)
Observations	5,535	5,535	5,549	5,549

Notes: All columns in Panel A report OLS estimates from regressions at the polling station level of the listed variable on indicators for Muslim/Yadav presence in each polling party position, conditional on there being 1 or fewer total MY officers at the polling station. All columns in Panel B report OLS estimates from regressions at the polling station level of the listed variable on indicators for the degree of Muslim/Yadav presence. Additionally included in all regressions are sub-constituency and number of officer fixed effects and controls for the log number of registered voters at the polling station and the Muslim/Yadav share of registered voters *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table A6. Heterogeneity in impacts of team composition by electorate composition

	Ln votes RJD (1)	Ln votes BJP (2)	Vote share margin RJD-BJP (3)	Ln total votes (4)
Mixed team	0.038	-0.023	0.018	0.003
	(0.037)	(0.027)	(0.014)	(0.010)
Mixed team * Muslim/Yadav registered voter %	0.057	-0.131	0.039	-0.014
	(0.158)	(0.166)	(0.053)	(0.039)
Muslim/Yadav registered voter %	0.031***	-0.029***	0.015***	-0.000*
	(0.001)	(0.001)	(0.000)	(0.000)
Observations	5,535	5,549	5,552	5,552

Notes: Each column reports OLS estimates from regressions at the polling station level of the listed outcome on indicators for mixed team composition, interacted with a continuous measure of the polling station level proportion of registered voters that are Muslim or Yadav. Additionally included are sub-constituency and number of officer fixed effects and a control for log total registered voters. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table A7. Cross-station spillovers - extended range

	Ln votes RJD	Ln votes BJP	Vote share margin RJD-BJP	Ln total votes
	(1)	(2)	(3)	(4)
Panel A. Buffer radius				
Mixed team composition	0.062**	-0.060***	0.033**	-0.001
	(0.028)	(0.022)	(0.011)	(0.008)
Number mixed team neighbor stations	0.024	-0.049**	0.026***	-0.001
	(0.026)	(0.020)	(0.010)	(0.008)
Number mixed team stations within 0.25km	0.012	0.014	-0.006	-0.004
	(0.026)	(0.014)	(0.009)	(0.007)
Number mixed team stations within 0.25-0.75km	-0.006	0.005	-0.003	0.000
	(0.009)	(0.008)	(0.004)	(0.002)
Total neighbor stations	-0.028**	0.028***	-0.021***	-0.011***
	(0.013)	(0.010)	(0.005)	(0.003)
Total stations within 0.25km	-0.028**	0.007	-0.007**	-0.006**
	(0.011)	(0.006)	(0.003)	(0.003)
Total stations within 0.25-0.75km	0.001	0.007***	-0.003*	0.001
	(0.004)	(0.003)	(0.001)	(0.001)
Observations	5,074	5,087	5,090	5,090
Number locations	3,307	3,307	3,307	3,307
Panel B. Village boundaries				
Mixed team composition	0.083**	-0.070**	0.037**	0.010
	(0.038)	(0.031)	(0.015)	(0.010)
Number mixed team neighbor stations	0.069	-0.108***	0.048**	0.011
	(0.054)	(0.041)	(0.019)	(0.011)
Number mixed team stations within village	-0.017	-0.044	0.010	-0.003
	(0.039)	(0.029)	(0.016)	(0.008)
Number mixed team stations in neighboring villages	0.008	-0.020*	0.009	0.004
	(0.013)	(0.011)	(0.006)	(0.003)
Total neighbor stations	-0.033	0.102***	-0.044***	-0.001
	(0.033)	(0.026)	(0.012)	(0.010)
Total stations within village	0.016	0.016	-0.004	-0.000
	(0.019)	(0.013)	(0.007)	(0.003)
Total stations in neighboring villages	0.000	0.011**	-0.004*	0.000
	(0.005)	(0.004)	(0.002)	(0.001)
Observations	3,196	3,210	3,212	3,212
Number villages	1,247	1,247	1,247	1,247

Notes: Each column within a panel reports OLS estimates from a regression at the polling station level of the listed variable on an indicator for mixed team composition. Each regression includes sub-constituency and number of officer fixed effects and controls for log total registered voters and share Muslim/Yadav registered voters. Neighbor stations are those within the same building/compound of a given polling station. Stations within 0.25 and 0.25-0.75km are non-neighbor stations within the stated distance of a given polling station. Numbers of stations within a village and in neighboring villages are the numbers of non-neighbor polling stations within the same village as a given station and in villages adjacent to a given station's village. Panel A is restricted to stations matched to the dataset of station GPS locations. Panel B further excludes stations in the top 1 percent of villages in terms of number of stations contained within, or their neighboring villages. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table A8. Type specific impacts of officer identity on voting outcomes

	Ln votes RJD (1)	Ln votes BJP (2)	Vote share margin RJD-BJP (3)	Ln total votes (4)
Any Muslim officer	0.051*	-0.034	0.023**	0.011
	(0.030)	(0.024)	(0.012)	(0.008)
Any Yadav officer	0.070	-0.100**	0.044**	-0.033
	(0.057)	(0.044)	(0.022)	(0.025)
Muslim/Yadav registered voter %	0.031***	-0.030***	0.015***	-0.000**
	(0.001)	(0.001)	(0.000)	(0.000)
Ln total electors	1.007***	1.198***	-0.069***	0.934***
	(0.062)	(0.050)	(0.024)	(0.019)
Control Mean [SD]	4.451	5.143	-0.181	6.180
	[1.198]	[0.969]	[0.452]	[0.402]
Observations	5,276	5,290	5,293	5,293

Notes: All columns report OLS estimates from regressions at the polling station level of the listed variable on indicators for Muslim and Yadav presence, conditional on there being 1 or fewer total Muslim/Yadav officers at the polling station. Additionally included are sub-constituency and number of officer fixed effects and controls for Muslim/Yadav share of registered voters and log total registered voters. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Table A9. Back-of-the-envelope calculation details

The total estimated effect on the RJD-BJP vote share margin of shifting to a mixed composition polling team is the sum of the within-station effect and the cross-station spillover effect multiplied by the number of neighbor polling stations, adjusting for the sub-constituency level of voter identity card coverage, ID_c. Using available sub-constituency-level administrative data for the entire state of Bihar, I calculate the average number of neighbors for a polling station in each sub-constituency, N_c. Taking the coefficients from a modified version of equation (2) allowing for heterogeneity by identity card coverage, estimated on the sample districts for which I possess officer assignment information:

$$Y_{pc} = \mu_c + \theta_o + \beta Mixed_{pc} + \gamma T_{pc} + \phi N_{pc} + \beta_2 \left[Mixed_{pc} * ID_c\right] + \gamma_2 \left[T_{pc} * ID_c\right] + \phi_2 \left[N_{pc} * ID_c\right] + \mathbf{X'_{pc}}\lambda + \epsilon_{pc}$$

the impact of a change of magnitude, X, in the proportion of mixed polling stations in a sub-constituency can be estimated as $X * [(\beta + \gamma * N_c) + (\beta_2 + \gamma_2 * N_c) * ID_c]$. While I do not observe the actual baseline proportion of mixed teams outside of my sample area, the value of X needed to change the outcome of the race between the RJD and BJP coalitions can be calculated using the formula above together with the constituency level margins of victory. When calculating impacts at the parliamentary constituency level, I take a weighted average (based on number of polling stations) across the sub-constituencies within that parliamentary constituency. The impacts of alternative team composition scenarios can then be assessed based on the range within which one assumes the baseline proportion of mixed team polling stations in each constituency falls. I assume that the baseline proportion in all sub-constituencies is the same as that in the observable sample, 0.324.

Table A10. Cross-election impacts of randomized officer team composition on voting outcomes

	Ln votes RJD 2015 (1)	Ln votes BJP 2015 (2)	Vote share margin RJD-BJP 2015 (3)	Ln total votes 2015 (4)
Mixed team	-0.016	-0.012	0.001	0.004
	(0.043)	(0.043)	(0.020)	(0.012)
Number mixed team neighbor stations	0.005	-0.067	0.035	0.025
	(0.058)	(0.058)	(0.025)	(0.015)
Total number neighbor stations	-0.107***	-0.011	-0.045***	-0.054***
	(0.036)	(0.043)	(0.015)	(0.009)
Muslim/Yadav elector %	0.020***	-0.026***	0.013***	0.000
	(0.001)	(0.002)	(0.000)	(0.000)
Ln total electors	0.969***	0.858***	0.015	0.819***
	(0.146)	(0.147)	(0.062)	(0.047)
Observations	1,335	1,332	1,335	1,335
Number locations	1,008	1,007	1,008	1,008
Homogeneous team mean [SD]	5.096	5.106	-0.010	6.243
	[0.908]	[0.973]	[0.433]	[0.321]

Notes: All columns report OLS estimates from regressions at the polling station level of the listed variable from the 2015 elections on an indicator for mixed team composition and variables for the numbers of total and mixed composition team neighboring polling stations in the 2014 elections. Additionally included are AC and number of officer fixed effects and controls for Muslim/Yadav share of registered voters and log total registered voters from 2014. Neighbor stations are polling stations within the same location (building/compound) as a given polling station. Standard errors clustered at the location level. The sample is restricted to the district where only minor changes were made to the polling station locations between the 2014 and 2015 elections. *Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.