Political Selection in an Unequal Democracy: The Tradeoff Between Competence and Representation*

Apurva Bamezai† Siddharth George‡ M. R. Sharan § Borui Sun ¶

May 26, 2026

Abstract

Who enters politics in low-income democracies? We describe political selection patterns in Bihar, India, by linking 1 million candidates in village elections to population-wide administrative data on 97 million citizens. We document a strong tradeoff between politicians' competence and representativeness. Politicians are strongly positively selected on competence: candidates are 0.67 SD more educated than citizens and 0.18 SD more educated than their own siblings. Yet politicians also hail from privileged social groups and families: their fathers are 0.24 SD more educated than the average citizen. This competence-representativeness tradeoff is sharper in unequal, diverse villages with limited party presence. Using a close election regression discontinuity design, we show that the tradeoff matters for governance. Highly competent village leaders are better informed about governance processes and best practices, and implement skill-intensive projects more effectively. But they align public spending more closely with the priorities of advantaged groups, worsening representation for marginalised citizens. Finally, we show that reserving positions for marginalised groups reshapes political selection over time. In the short run, reservations for Scheduled Castes and women raises candidates' representativeness while reducing competence. In the long run, however, reservations improve political selection on both margins, as a more competent and socially representative pool of unreserved citizens enters politics once quotas are removed. Our findings highlight a fundamental tension between capable and representative government in low-income democracies, while showing that institutional design can change who enters politics and mitigate this tradeoff.

*We thank Martin Mattson, Fred Finan, Saad Gulzar, Guo Xu and Jing Cai for feedback. We thank seminar participants at the Midwest International Economic Development Conference, Indian Statistical Institute (Delhi), NUS Singapore, Georgetown University, Delhi School of Economics and PDRI-DevLab at University of Pennsylvania for some excellent comments.

†Kellogg Institute for International Studies, University of Notre Dame

‡Department of Economics, National University of Singapore

§Department of Agricultural and Resource Economics, University of Maryland

¶Department of Agricultural and Resource Economics, University of Maryland

1 Introduction

Effective democracies deliver competent governance and policies that reflect citizens' concerns. How does the composition of the political class affect governance, and which citizens should enter politics to foster competent, responsive democratic government? One view is that the political class should comprise the most capable and competent citizens—a *natural aristoi* who possess the human capital required to navigate complex governance challenges (Jefferson, 1813; Besley, 2005). An alternative view is that elected leaders should be broadly representative of the citizenry, so they can understand and reflect citizens' priorities in policy choices. These visions of political leadership appear compatible in rich, egalitarian democracies: Swedish politicians are more able than the citizens they represent, yet are drawn from nearly every social stratum (Dal Bó et al., 2017). In poor or unequal societies, however, where individuals from privileged backgrounds have more opportunities to acquire human capital, there may be a stark tradeoff between having capable and representative leaders. Yet we have limited evidence on who becomes a politician in low-income democracies, how politicians' competence and representativeness shape governance, and whether institutional design can improve political selection on these dimensions.

In this paper, we examine the determinants and consequences of political selection in the developing world. We focus on the Indian state of Bihar, whose 130 million citizens constitute 3.5% of the global population living in a democracy. Bihar experiences similar challenges to sustaining an effective, responsive democracy as many other low-income democracies. It is deeply unequal: while only 6%of adults have completed high school, privileged families and social groups, such as men and upper castes, have much greater access to education. The most competent politicians are thus often socioeconomic elites. Society is divided along religious lines, so voters must often choose between supporting their own group's and the most competent candidate. Moreover, democratic institutions that might facilitate positive and inclusive selection in mature democracies—such as political parties—are weak. These features bring the potential tradeoff between politicians' competence and representativeness into sharp focus.

We begin by describing patterns of political selection in rural Bihar, whose 37,000 villages are organised into approximately 8,000 local polities known as Gram Panchayats (GPs).[1] Each GP comprises roughly 13,000 citizens, and is governed by an elected president and an elected council comprising 12-14 ward members. Detailed administrative data enable us to link over 1 million candidates in GP elections between 2016 and 2021 to a population census covering over 97 million rural citizens. Using this linked citizen-candidate data—to our knowledge, the largest such database for a developing country—we construct proxies of politicians' competence and representativeness by comparing the education and family background of candidates against those of citizens.

[1]We restrict our focus to rural politicians for two main reasons. First, Bihar remains an overwhelmingly rural state, with only 11% of the population residing in urban areas according to the 2011 Census. Second, the institutional architecture, mandates, and functioning of local bodies differ markedly between rural and urban areas, making it analytically problematic to study them jointly.

Educational attainment is our primary measure of politicians' competence. Beyond maintaining consistency with Dal Bó et al. (2017), we show that education is a relevant proxy in our setting. We find that educated GP leaders have greater knowledge of the governance issues they oversee. Education is also valued by citizens: in an original survey and conjoint experiment (N = 1,147), we show that citizens rank education as an important politician attribute and are much more likely to vote for highly educated candidates. Following Dal Bó et al. (2017), we identify the socioeconomic status of a politician's family using his or her father's education, which—unlike wealth or occupation—is determined before one's entry into politics. The fathers of elite politicians are significantly more educated than the average citizen, while those of representative politicians are similar to the average citizen.

Using these measures, we find that local politicians in Bihar are highly competent but drawn from an unrepresentative elite. Focusing on open (unreserved) seats, we find that candidates have 2.6 more years of schooling (0.67 SD) than the average citizen. While the average citizen is literate and 6% have completed high school, the average candidate has completed primary school and 18% are high school graduates. This gap partly reflects politicians' elite origins: candidates are overwhelmingly male—who are significantly more educated than women in rural Bihar—and are drawn from more educated birth cohorts and castes. Even within these groups, politicians come disproportionately from elite families: politicians' fathers are on average 0.24 SD more educated than the average citizen, and are 0.28 SD more educated than other men in their cohort and caste.

However, these privileged backgrounds only partly explain politicians' educational advantage. Even after restricting attention to more comparable groups of citizens, we find that male politicians are 0.47 SD more educated than other men in the same birth cohort and caste. Most strikingly, politicians are also substantially more educated than their own siblings, with male candidates attaining 0.18 SD more years of schooling than their brothers. This gap between candidates and comparable citizens is present throughout the parental education distribution. Politicians thus appear to be highly positively selected on competence even as they hail from elite families that are unrepresentative of the broader citizenry.

Moreover, we observe a clear tradeoff between politicians' competence and their representativeness. More competent politicians—those more educated in absolute terms or relative to their cohort and caste—tend to come from more elite families. The tradeoff also heightens as one progresses in politics: candidates for GP president are more competent and less representative than those contesting ward member elections, and across both levels, winners are more competent and more elite than losing candidates.

The competence-representativeness tradeoff is also sharper along three important dimensions that characterise Bihar and many other low-income democracies. First, it widens with local inequality: politicians are more educated and less representative in more unequal villages. Second, group size matters. Candidates from larger religious groups are both more competent and less representative. However, each pattern is driven by different groups: the rise in competence comes from minority Muslim candidates, while the

decline in representativeness comes from majority Hindu candidates. This asymmetry is consistent with the logic of Banerjee and Pande (2007): when a group is so large that it is nearly assured victory, the marginal return to competence falls; when it is small and unlikely to win, only the most competent candidates enter. Third, political parties improve selection. Though formally banned in GP elections in Bihar, our surveys reveal that about a fifth of local politicians report a party affiliation. Where parties operate, we find that politicians are substantially more competent but inconclusive evidence that they are more elite. Together, these results suggest that these common features of low-income democracies—economic inequality, ethnic divisions and weak parties—may impede the formation of a political class that is both competent and representative.

Having documented the stark tradeoff between selecting competent and representative politicians, we examine its consequences for governance. We estimate the causal effect of politicians' competence on governance using a close-election regression discontinuity (RD) design. Focusing on elections between candidates with above and below median education, we compare governance outcomes in jurisdictions where more educated candidates narrowly win versus lose. Leveraging an original survey of 7,200 elected ward members, we show that more educated leaders are more knowledgeable about scheme implementation, better informed about governance best practices, and have stronger ties to local bureaucrats. This greater capacity enables more effective delivery of skill-intensive projects: piped water connections and road construction projects—the two most important infrastructure projects during our study period—progress faster in wards where educated leaders narrowly win. Competent politicians thus appear to perform better on governance tasks that benefit from their stronger knowledge and networks.

However, more competent leaders are on average less representative, and we find that they perform worse at reflecting marginalised citizens' preferences in public spending. We elicit citizens' priorities through an original survey of nearly 9,800 citizens in over 2,000 GPs. Then, using administrative records on development plans, we examine the extent to which citizens' concerns are reflected in GPs' planned expenditures. Applying an analogous close election RD design, we show that GPs governed by more educated leaders allocate less spending and fewer development projects to the priorities of historically marginalised Scheduled Caste (SC) citizens, while spending more on public goods valued by advantaged groups. These findings highlight the governance implications of an educated but elite-dominated political class: skill-intensive tasks are performed more competently, but collective choices are less likely to represent the concerns of citizens, especially those from less influential groups.

In the final part of the paper, we explore whether institutional design can weaken the competence-representativeness tradeoff. We analyse the short- and long-term impacts of reserving seats for disadvantaged groups, a widely used tool for broadening political representation in India and around the world (Pande, 2003; Chattopadhyay and Duflo, 2004; Besley et al., 2017). Bihar reserves 17% of GP head positions for SCs and 45% for women, and each GP's reservation status is determined by a population-based formula. These rules generate quasi-random variation in reservation exposure: GPs just above the

SC population threshold, for instance, are 80 percentage points more likely to have an SC president (Kumar and Sharan, 2025). We exploit this variation using an RD design, comparing the selection of politicians in GPs just above and below the cutoff.

In the short run, reserving seats for SCs lowers politicians' competence but raises their representativeness. This tradeoff stems from the substantial across-caste disparity in educational attainment: SC politicians are drawn from much less educated and elite castes than those of unreserved politicians. Relative to their own group, however, SC and unreserved politicians are just as positively selected: both SC and unreserved candidates for GP president are 0.85-0.95 SD more educated than the average citizen of their caste. Yet SC candidates are far more representative: while the fathers of unreserved politicians are 0.38 SD more educated than the average citizen of their caste, the fathers of SC candidates have similar education to the average SC citizen. SC reservations thus improve within-group selection, though the competence-representativeness tradeoff still appears due to large across-group inequalities.

In the long run, however, SC reservations raise both the competence and the representativeness of the political class. Focusing on GPs that were unreserved in 2016 and 2021, we compare GPs that were reserved for SC leaders between 2006-16 against those with no reservation during this 10 year period. In GPs that were previously reserved for SCs, candidates are similarly competent and more representative, while winners are 0.21 SD more competent and 0.33 SD more representative. This improvement in political selection is driven by two forces: first, highly educated SC citizens contest elections at higher rates after reservations are abolished; second, highly educated—but less elite—upper caste citizens are more likely to enter politics. We conjecture that reserving seats for SCs exposes citizens to less elite leaders, inducing a more representative set of non-SC citizens to run for office in the longer term.

Reserving seats for women generates broadly similar patterns. In the short run, reservations lower politicians' competence while increasing their representativeness: candidates in female reserved GPs are 0.79 SD less educated than in unreserved GPs, reflecting the stark gender inequalities in educational attainment. Interestingly, female reservations also induce a more representative set of households to contest elections: female politicians come from less elite families than their male counterparts, and their husbands are 0.15 SD less educated than male politicians. The education gap between female candidates and the average female citizen is only 0.19 SD, much smaller than the analogous 0.65 SD gap between male candidates and citizens. These within-group selection dynamics heighten the competence-representativeness tradeoff in the short run.

By contrast, in the long term, gender reservations—like SC reservations—relax the tradeoff between competence and representativeness. As before, we focus on GPs that were unreserved in 2016 and 2021 and compare GPs that were reserved for a female leader between 2006-16 against those which were not. Candidates in previously reserved GPs are 0.14 SD more competent and similarly representative. These effects are primarily driven by changes in the pool of male candidates: reservations for women influenced which men entered politics in the medium term, and ultimately improved political selection.

Our paper contributes to a growing literature on political selection, reviewed by Besley (2005), Dal Bó and Finan (2018) and Gulzar (2021). Closest to our work, Dal Bó et al. (2017) show that Sweden's political class constitutes an "inclusive meritocracy": politicians are strongly positively selected on ability but are drawn from a broad socioeconomic base. Recent work documents similar patterns in Finland (Jokela et al., 2025) and Denmark (Dahlgaard and Pedersen, 2025). By contrast, political selection appears less inclusive in more unequal settings. Feigenbaum et al. (2026) find that politicians in the 19th and early 20th century United States were highly competent but come from elite families, while Bhusal et al. (2026) show that traditional elites dominated Nepal's political class until the Maoist revolution pushed parties to nominate candidates from historically excluded indigenous groups.[2]

We complement this work by providing large-scale evidence on who becomes a politician in a contemporary low-income democracy. Bihar's politicians are highly positively selected on competence—even relative to their own siblings—but hail from privileged families and social groups. This generates a sharp tradeoff between competence and representativeness. Moreover, this tradeoff varies with economic inequality, diversity, and party presence—forces that may help explain why patterns of political selection differ from those of advanced democracies.

We also show that the competence-representativeness tradeoff in political selection has important consequences for governance. Competent politicians possess greater knowledge of governance processes and best practices, and implement skill-intensive projects more effectively. At the same time, they align public spending more closely with the priorities of advantaged groups, worsening representation for marginalised citizens. The composition of the political class thus shapes both the capacity of local governments and whose preferences they reflect. Our findings provide a new perspective on the mixed impacts of educated leaders (Besley et al., 2011; Carnes and Lupu, 2016; Martinez-Bravo, 2017; Lahoti and Sahoo, 2020): since the returns to political competence depend on the governance task, greater capacity may not translate into more responsive government.

Finally, we contribute to a large literature on political reservations and, more broadly, on how institutions shape political selection. Existing work shows that quotas increase the representation of disadvantaged groups (Bhavnani, 2009; Iyer et al., 2012; O'Connell, 2020) and affect downstream policy choices and governance outcomes (Chattopadhyay and Duflo, 2004; Pande, 2003; Dunning and Nilekani, 2013; Chin and Prakash, 2011; Gulzar et al., 2020). We show that reservations reshape the pool of candidates over time. In the short run, reservations generate a stark competence-representativeness tradeoff: they draw candidates from marginalised groups and, within those groups, from less elite families, further democratising political entry. However, in the longer run, reservations improve political selection on both margins, as more socially representative yet still highly competent citizens from the unreserved group enter politics once quotas are removed.

The rest of the paper is organised as follows. Section 2 outlines the background and con-

[2] Chaudhuri et al. (2025) use lab-in-the-field data on village-level politicians in West Bengal in India to argue that competence and integrity can pull in different directions.

text. Section 3 describes the data. Section 4 describes political selection patterns. Section 5 explains how these patterns vary with institutional features. Section 6 discusses the governance implications of political selection, and Section 7 examines how institutional design can influence political selection. Section 8 concludes.

2 Background and Context

Bihar is a large, poor, agrarian state in eastern India whose political and economic features make it a natural laboratory for studying political selection in developing democracies. Three features in particular – its scale and representativeness of the developing-world norm, its combination of high inequality and salient group identity, and its decentralized local government with extensive reservations – motivate our focus on this setting. We describe each in turn, along with the institutional details needed to interpret our empirical analysis.

2.1 Bihar

Scale and representativeness. With a population of over 130 million – roughly 3.5% of the world's population living in a democracy – Bihar is larger than most countries. It is overwhelmingly rural (nearly 90% in 2011) and agrarian, and its democratic institutions, while stable, are younger and less consolidated than those in advanced economies. These features make Bihar broadly representative of the conditions under which most of the world's voters select their leaders, and where we have the least systematic evidence on who enters politics.

Inequality. Bihar is a highly unequal agrarian economy. In 2009-10, the land value of a landlord household was 33 times that of an agricultural labour household (Swaminathan and Nagbhushan, 2022). Such inequality concentrates educational and economic opportunities among a narrow elite, creating precisely the conditions under which the competence-representativeness tradeoff in political selection is likely to be most acute.

Caste and group identity. Caste is central to social and political life. Only 11% of marriages cross caste boundaries (Desai and Vanneman, 2015), and caste hierarchy maps closely onto economic hierarchy: the sub-castes at the bottom of the asset distribution are also those at the bottom of the caste order (Sharan, 2023). Bihar's thousands of sub-castes fall into five broad groupings: the Scheduled Castes (SCs, 17%), historically considered "untouchable"; Scheduled Tribes (STs, 1.7%); Extremely Backward Castes (EBCs, ∼35%), carved out of the larger Other Backward Castes (OBCs) grouping; and, at the top, the general castes. The salience of group identity – and the wide variation in group sizes across GPs – provides the variation we exploit to study how demographic composition shapes selection.

2.2 Local Governance

Bihar has over 8,000 rural local bodies called Gram Panchayats (GPs), each with a population of roughly 13,000. Each GP is divided into wards of about 1,000 persons. Both units have directly elected heads – GP heads (locally, *Mukhiyas*) and ward members – but GP heads are substantially more powerful, with direct access to state development funds and primary responsibility for implementing welfare schemes. Ward members, by contrast, had no direct access to state funds until 2016 and primarily represent constituents in GP and village meetings (*Gram Sabhas*). Our main analyses focus on GP heads; we use ward members in supplementary validation exercises.

Local elections are conducted every five years by the State Election Commission of Bihar and are hotly contested, with turnout exceeding 60% in 2021 – higher than in recent state or national elections. High turnout, combined with the large stakes of GP head elections, makes this a meaningful setting in which to study political selection.

2.3 Political Reservations

A defining feature of Indian local democracy – and a central element of our empirical strategy – is political reservation. Since 2006, Bihar has reserved GP head and ward member seats for historically disadvantaged groups: approximately 45% for women, 17% for SCs, 20% for EBCs, and under 2% for STs. Reservation status is determined by a population-based assignment algorithm (described in detail in Section 7), which generates quasi-random variation in the identity of leaders across otherwise similar GPs.

Two features of the reservation regime shape how we interpret our results. First, a constituency's reservation status is fixed for two full election cycles (10 years). For both the 2016 and 2021 elections that we study, reservation status was assigned under the same rule and remained unchanged across the two cycles. This allows us to pool candidates from both elections without conflating selection patterns with changes in the reservation regime, and means that the candidates and voters we observe have had time to adjust to the reservation status of their GP – our estimates reflect steady-state selection rather than transitional dynamics. Second, the overlapping caste and gender components of the algorithm generate independent sources of quasi-random variation, enabling us to separately identify the effects of SC reservation and gender reservation (Section 7) on who enters politics.

2.4 GP Development Plans

Under the 73rd Constitutional Amendment, local governance in India institutionalises deliberative planning. In Bihar, this is operationalised through the Gram Panchayat Development Plan (GPDP) – an annual document listing all projects the GP intends to undertake, with estimated costs and funding sources, based on consultations at the ward and

Gram Sabha levels. The GPDP generates detailed administrative data on planned spending by project type, which we combine with a survey of citizen preferences to measure how closely leader choices align with what citizens want (Section 6).

3 Data and Measurement

Studying political selection requires three things that are rarely available together: data on candidates *and* non-candidates, a measure of competence, and a way to benchmark representativeness. Our setting offers all three. This section describes the administrative and survey data we assemble, and defends our measurement choices for competence (years of education) and representativeness (father's education). This paper uses a wide range of datasets – Appendix Table A2 describes all our data in detail.

3.1 Data and Matching

We combine two large datasets: the Socioeconomic Caste Census (SECC) of 2011-12 and candidate-level records from the Bihar State Election Commission for the 2016 and 2021 GP head and ward member elections.

Socioeconomic Caste Census (SECC). The SECC is a household census covering 97 million individuals across 19 million rural Bihar households. For every citizen, we observe gender, broad caste category, age, occupation, and education, along with household-level information on dwelling characteristics, employment, income sources, asset ownership, and land holdings. Crucially, we also observe who each individual's parents are and who else lives in their household, which allows us to link candidates to their families.

Candidate data. The Bihar State Election Commission publishes records of all candidates contesting GP head and ward member elections. Across 2016 and 2021, we observe 909,570 candidate-election records, with data on names, father's or husband's name, age, education, caste category, and votes received.[3]

Matching. We match candidates to citizens in the SECC within the same GP using a string-matching algorithm based on name, father's or husband's name, caste category, and gender. We restrict to matches with a similarity score of at least 0.95 to limit false matches; results are robust to alternative thresholds.

Our overall match rate for our primary population of interest – unreserved GP head candidates – is 60.2%, which compares favourably to other census name-matching work (Abramitzky et al., 2021; Fouka et al., 2022). Match rates are highest for 2016 male GP head candidates and lowest for 2021 women candidates, reflecting two patterns: the SECC is more outdated relative to the 2021 election, and women's records are harder to match because marriage-related migration moves women across villages and because

[3]These are not unique candidates, since individuals can contest across both elections.

candidate records do not distinguish father's from husband's name for female candidates. On average, matched candidates are somewhat older than unmatched ones; other differences are modest. Appendix E provides the full breakdown.

3.2 Measuring Competence: Education

Measuring politician competence is inherently challenging, as direct measures of underlying ability are rarely available. We use years of formal education as our primary proxy. Education is standard in this literature – used in studies of politicians in Sweden (Dal Bó et al., 2017), Denmark (Dahlgaard and Pedersen, 2024), and the United States (Thompson et al., 2019) – and has been linked to politicians' policy performance across a range of outcomes, including growth, public goods provision, and inflation control (Besley et al., 2011; Martinez-Bravo, 2017; Göhlmann and Vaubel, 2007; Sørensen, 2023). It is also, practically, the only competence-relevant variable available for the full pool of candidates (Gulzar, 2021). Education has also been criticised as a competence proxy, particularly in unequal societies where it may reflect elite background rather than ability (Carnes and Lupu, 2016; Curto-Grau and Gallego, 2023; Lahoti and Sahoo, 2020; Dal Bó et al., 2017). We take this concern seriously and address it in two ways. First, we compare politicians to their own siblings in Section 4 to isolate within-family variation. Second, we mount three original empirical exercises – a close-election regression discontinuity design (RD) on politician knowledge, a parallel RD on political networks, and a conjoint experiment on voter perceptions – to validate education as a measure of governance-relevant competence in our specific context.

Voter Perceptions (Survey Evidence): For education to matter for selection, voters must also treat it as a signal of competence. We draw on an original in-person survey of 1,105 respondents across 48 GPs in Bihar, designed specifically to elicit how citizens evaluate political candidates (Appendix G). When asked what qualities they value in a Panchayat candidate, a majority placed education among the top three, ranking it above household wealth, caste, or prior experience (Figure 5).

Voter Perceptions and Vote Choice (Conjoint Experiment):A forced-choice conjoint experiment embedded in the same survey confirms this with randomised variation: holding other candidate attributes constant – including caste, family wealth indicators, and political lineage[4] – respondents consistently perceive more educated candidates as more likely to run and win elections, and are themselves more likely to vote for them. For instance, they perceive a B.A.-holder as 15.5 percentage points more likely to win a Panchayat election than an otherwise-identical 5th-grade pass candidate, and are 19.8 percentage points more likely to say they would vote for the B.A.-holder themselves – large and precisely estimated effects (Figure 6).

[4]The conjoint varied seven attributes independently: education, caste category, gender, behaviour/personality, prior family experience in Panchayat politics, presence of a government worker in the family, and the household's main non-agricultural income source (wage labour, private job, farming, or business). Full design in Appendix G.

Knowledge and Networks: In section Section 6.1.2, we also show that education of political leaders is closely linked to skills required to govern. Using a close-election RD on ward member races, we find that more educated politicians have substantially greater knowledge of how to implement government schemes across housing, infrastructure, and service delivery. They are also more likely to hold the personal contact details of key higher bureaucrats (the Block Development Officer (BDO), Block Panchayat Raj Officer (BPRO) and have stronger links to political parties. Education thus travels with both the procedural knowledge and the relational capital that local governance demands.

3.3 Measuring Representativeness: Father's Education

Benchmarking whether politicians represent the broader population requires a measure that is not itself a product of political selection. A candidate's own education and wealth are endogenous to their path into politics; we therefore follow Dal Bó et al. (2017) in using *father's* education – specifically, education relative to other males in the same birth cohort – as our primary measure of a candidate's social origins. This allows us to assess whether candidates emerge from households that were historically typical or advantaged in terms of educational access.

Two features of this measure shape our analysis. First, we can match fathers only for male candidates. Women typically leave their natal villages upon marriage, and the candidate records do not consistently identify father's versus husband's name for female candidates. All representativeness analyses using father's education are therefore restricted to male candidates. Second, we match 41.1% of male candidates to their fathers in the 2011 Census village rolls. Matched candidates are slightly more educated (by 0.89 to 1.42 years) and more likely to be non-SC and non-ST than unmatched candidates (Appendix Table A19, Panel B). If anything, this suggests our representativeness sample is modestly tilted toward more elite candidates, which would lead us to *understate* the extent of elite persistence among politicians overall.

The conjoint experiment discussed in Section 3.2 provides complementary evidence that representativeness matters to voters as well as to our analysis: respondents consistently favour candidates from their own caste category (Figure 7). The effects are smaller than for education but robust across subgroups, suggesting that voters in rural Bihar value representativeness as a candidate attribute alongside competence.

4 Result #1: Who Becomes A Politician?

Building on the data construction and measurement choices described in the previous section, we now document the baseline patterns of political selection in rural Bihar. Three findings emerge. First, politicians are strongly positively selected on competence: they are substantially more educated than the average citizen, and this gap persists even within narrow comparison groups defined by gender, age cohort, caste, household, and birth

family. Second, politicians come from more elite households than the average citizen – their fathers are more educated and wealthier than the typical man of the same generation and caste. Third, and most importantly, these two forces are in tension: the most competent candidates are systematically the least representative, producing a clear competence-representativeness tradeoff that recurs across GPs, offices, and outcomes.

Estimation Framework To quantify these selection patterns, we begin with a descriptive comparison of the socioeconomic background of politicians (i.e., candidates in local elections) and other citizens within the same GP. Specifically, we estimate

$$Y_{ig} = \alpha \mathbb{1}\{\text{Candidate}\}_{ig} + \epsilon_{ig}, \tag{1}$$

where Y_{ig} denotes the socioeconomic characteristic of individual i in GP g, and $\mathbb{1}\{\text{Candidate}\}_{ig}$ is an indicator equal to one if individual i ran for an unreserved GP head or Ward member seat in the 2016 or 2021 local elections.

To examine politicians' representativeness, we compare the social origins of candidates, measured using their fathers' education and socioeconomic background, to that of the average citizen within the same GP.

$$Y_{ig} = \beta \mathbb{1}\{\text{Candidate's Father}\}_{ig} + \eta_{ig}, \tag{2}$$

where $\mathbb{1}\{\text{Candidate's Father}\}_{ig}$ is an indicator equal to one if individual i has a child who ran in either the 2016 or 2021 local elections for an unreserved seat.

All dependent variables are standardized within GP so that the coefficients are interpreted relative to the local distribution. The coefficients α and β respectively capture the competence and representativeness level of the candidates. ϵ_{ig} and η_{ig} represent the error terms, and standard errors are clustered at the GP level.

Although India has an extensive system of political reservations, we first present descriptive statistics on unreserved seats, which are technically "open" to all citizens of a GP, across genders and castes. This gives us an understanding of what political selection would look like absent reservations. In Section 7, we exploit quasi-random variation in reservation status to analyse how reservations impact selection.

4.1 Selection on Competence: The Education Gap

Average patterns. We find that politicians are significantly more competent than the average citizen. The average citizen in rural Bihar has 3.5 years of schooling. By contrast, politicians in the same GP obtain 2.7 more years of education, an increase of 77% relative to the mean or a 0.67 SD increase (Table 1, column 1). We can also interpret these patterns in the following manner: while the average citizen is merely literate, the average candidate has completed primary schooling. Moreover, while only 6% of citizens in ru-

ral Bihar have completed high school, this number is three times as high for politicians at 18.5%. This gap is robust to alternative matching thresholds (Figure A2).

We find that politicians are more likely to come from highly-educated groups in society. As Table A4 shows, politicians come from age cohorts (column 1) and caste groups (column 2) with higher education than the average citizen. Defining group based on surname, or gender×age cohort × surname cells, we also find that politicians hail from more advantaged social groups. However, the magnitudes already indicate that much of the positive selection is within-group: politicians are 0.67 SD more educated than the average citizen, but come from groups that are on average 0.36 SD more educated. Next, we examine within-group selection more directly.

Within-group. What is particularly striking is how positively selected politicians are relative to the average citizen within their social groups, and even within families. Candidates in elections are 0.56 SD more educated than the average citizen of their gender, 0.49 SD more educated than other citizens in their gender and cohort (e.g. men in their cohort). Furthermore, column 4 shows that candidates are 0.47 SD more educated than citizens in their gender-cohort-caste (i.e. other men of their caste in their cohort). Column 5 shows that politicians are 0.34 SD more educated than other adult members of their household (e.g. other men living in their household). Column 6 is based on a restricted sample where we can identify the father of the citizen. This is a smaller sample since some fathers may have passed away or live in another village. For this restricted sample, we see that politicians are 0.18 SD more educated than their siblings.

Competence and Electoral Performance. In Table A6, we see that candidates are more competent in elections for higher-level political positions. Panel A focuses on GP heads, while Panel B focuses on lower-tiered ward members. Column (1) of Panel A shows that GP head candidates are 1.09 SD more educated than the average citizen. This difference falls to 0.59 SD for candidates in ward elections.

Next, comparing across columns, we see that competence is linked to position in the electoral races. While the average politician is 0.67 SD more educated than citizens, winners in GP head elections are a full 1.4 SD (column 2 of Panel A) more educated than the average citizen. This translates to an additional 5.6 years of schooling. Within GP head elections, winners are more educated than runners-up, who are in turn more educated than other candidates (Panel A, columns 2-4). Taken together, these results suggest that voters value education of politicians.

4.2 Selection on Representativeness: Elite Persistence

We now turn from competence to the representativeness of candidates. To develop a measure of how elite candidates are, and, conversely, whether their education reflects ability or merely inherited wealth, we compare the characteristics of candidates' fathers

to citizens in the same cohort and caste.

Panel B of Table 1 displays results. The average politician's father is 0.24 SD more educated than the average citizen. One possible issue with this is that the educational differential between candidates' fathers and citizens could be mechanically lower because education levels amonge men in the older generation were lower than they are now. To control for this concern, we include cohort fixed effects. We see that politicians' fathers are more educated than their peers: they are 0.28 SD more educated than other men of their caste in their cohort. Thus, politicians do hail not only from more educated and wealthier social groups, but their fathers are also more educated relative to the average man in this social group. In other words, politicians come from advantaged communities and, within these, from advantaged families.[5] This pattern is robust to alternative matching thresholds as well (A2).

4.3 The Competence-Representativeness Tradeoff

The previous two sections establish that politicians in rural Bihar are both positively selected on competence and drawn from more elite households than the average citizen. The central finding of this section is that these two patterns are in tension: gains on one dimension come systematically at the cost of the other. We document this tradeoff through four complementary lenses – the relative magnitudes of the two gaps, the distributional shape of selection, a visual representation of the candidate selection frontier and the citizen mobility curve, and the direct relationship between competence and representativeness within the pool of politicians themselves.

The competence gap is substantially larger than the representativeness gap. The first piece of evidence is simply a comparison of magnitudes. Figure 1 plots the competence and representativeness coefficients side by side across progressively finer comparison groups. In the baseline specification, candidates are 0.67 SD more educated than the average citizen, while their fathers are only 0.24 SD more educated. This 0.43 SD wedge persists as we add controls: within gender and age cohort, the competence gap is 0.49 SD against a representativeness gap of 0.31 SD; within gender, age, and caste, the gap is 0.47 against 0.28. Even in the tightest comparison – politicians' own households – candidates are more educated than their non-politician siblings by 0.34 SD. Candidates are thus substantially more positively selected than their family backgrounds would predict, but their family backgrounds are themselves advantaged.

[5]While some candidates' fathers live with their sons, 58.65% do not. We could, potentially, assess wealth outcomes for fathers who do not live with sons and compare them with other individuals of their own generation. However, while fathers' current asset wealth could also be influenced by sons' wealth, the same cannot be said for education, since education outcomes are almost always determined before sons are born. Hence the relative difference in fathers' education vis-a-vis sons' education offers the cleanest comparison possible.

Distributional shape of selection. The kernel density plot in Figures 3 shows, across both GP head and ward member candidates, the candidate distribution is shifted substantially to the right of the citizen distribution – the competence gap. Candidates' fathers sit somewhat to the right of the citizen distribution as well – the representativeness gap – but the shift is far smaller. Candidates' siblings fall in-between, somewhat closer to the citizen distribution than to the candidate distribution. The picture is one of political selection that pulls substantially on ability within families, but that draws disproportionately from families that are already advantaged.

The candidate selection frontier and citizen mobility curve. Figure 4 provides the cleanest visual summary. Panel (a) plots two curves. The *citizen mobility curve* (orange) describes the relationship between father's education and son's education in the underlying GP population – a standard intergenerational mobility curve. The *candidate selection frontier* (green) plots the same relationship among individuals who enter politics. The candidate selection frontier lies clearly above the citizen mobility curve across the entire range of parental backgrounds, confirming that candidates are positively selected at every level of family origin.

Panel (b) positions the average unreserved candidate and winner on this frontier, with the citizen mean shown as a reference point. Two features stand out. First, both the average candidate and the average winner sit far to the right on the x-axis, indicating that politicians are drawn disproportionately from the upper end of the parental education distribution – the "elite origins" finding made visual. Second, winners sit further up and further right than the average candidate: electoral competition selects upward on competence, but it also selects toward more elite origins. The tradeoff is not only a feature of who enters politics; it is reinforced by who wins.

The tradeoff within the pool of politicians. Finally, the tradeoff is visible not just between politicians and citizens, but also within the pool of politicians itself. Figure 2 disaggregates selection by office, election outcome, and social group. Across all three dimensions, the same pattern recurs: more competent politicians are also less representative. GP heads – the more powerful office – are 1.09 SD more educated than the average citizen, compared to 0.59 SD for ward members; their fathers are 0.53 SD more educated, compared to 0.19 SD for ward members' fathers. Winners are more competent and more elite than runners-up, who are in turn more competent and more elite than other losing candidates. The same pattern holds by caste: upper-caste candidates are both more educated and from more advantaged family backgrounds than SC candidates.

Taken together, these results establish a competence-representativeness tradeoff as a robust feature of political selection in rural Bihar. Politicians are genuinely positively selected on competence – even against their own siblings – but the selection pool tilts toward more elite origins, and electoral competition reinforces rather than offsets this tilt. The next section asks what features of local settings make this tradeoff sharper or more muted.

5 Result #2: Determinants of Selection: Inequality, Group Size, and Parties

Following the documentation of a competence-representativeness tradeoff in Section 4, we now examine the factors that drive these selection dynamics. A central question is why this tradeoff is more pronounced in certain settings than others. To investigate this, we leverage the substantial variation in local conditions across GPs in Bihar. We focus on three primary dimensions that may shape the observed selection patterns: (i) local economic and structural characteristics, (ii) demographic group identity and size, and (iii) the institutional presence of political parties. By exploring these factors, we aim to identify the conditions under which the barriers to political entry are most restrictive.

Estimation Framework. To examine how selection patterns vary with local features, we augment the baseline specifications in (1) and (2) by interacting the key indicators with a feature, $Feature_g$:

$$Y_{ig} = \alpha_1 \mathbb{1}\{\text{Candidate}\}_{ig} + \alpha_2(\mathbb{1}\{\text{Candidate}\}_{ig} \times Feature_g) + \epsilon_{ig}, \quad (3)$$

$$Y_{ig} = \beta_1 \mathbb{1}\{\text{Candidate's Father}\}_{ig} + \beta_2(\mathbb{1}\{\text{Candidate's Father}\}_{ig} \times Feature_g) + \eta_{ig}. \quad (4)$$

where Y_{ig} denotes the relevant standardized outcome for individual i in GP g, and $Feature_g$ denotes one of the GP-level characteristics described below. The coefficients α_2 and β_2 capture how the competence and representativeness gaps vary with the GP-level feature, respectively. Standard errors are clustered at the GP level. We focus on male candidates in caste-unreserved GPs, for two reasons: first, we want to focus on "open" seats where members of all castes can technically contest; second, our representativeness results can only be ascertained for male candidates, since we cannot reliably match female candidates to their fathers.

5.1 Inequality

In GPs with high land inequality, candidates are more educated than the average citizen *and* come from more elite families – both gaps widen simultaneously. As shown in column 5 of Table A3, the coefficient on candidates' fathers' education is much higher in high-inequality GPs, consistent with elite capture of political entry in settings where opportunity is concentrated among a narrow class.

By contrast, we find no heterogeneity along two other dimensions one might expect to matter. Political selection does not vary with a GP's poverty status (column 1) or geographic remoteness (column 6). Candidates are positively selected on education across poor and wealthy GPs, remote and central GPs alike.

These findings indicate that the competence-representativeness tradeoff is thus not uniform across Bihar: it is especially stark in setting swhere inequality restricts access to the

schooling, wealth, and networks that political entry requires.

5.2 Group Diversity and Size

Group identity is central to politics in developing democracies, and Bihar is no exception. Two features of ethnic composition matter: how diverse a GP is overall, and how large a politician's own ethnic group is within the GP. We examine each in turn.

GP-level diversity. Columns (2) and (3) of Table A3 report how selection varies with two summary measures of sub-caste composition. Ethnic fractionalization is positively correlated with candidate competence and representativeness, while ethnic polarization is negatively correlated with both. One interpretation, following Banerjee and Pande (2007), is that polarization captures settings with two large competing groups – where sure-win candidates face little selection pressure – while fractionalization captures settings with many small groups, each of which must field quality candidates to compete. The coefficients on both indices are small, however, suggesting that while GP-level diversity co-varies with selection, the relationships are modest in magnitude.

Own-group size. A more striking pattern emerges when we examine the size of the politician's *own* ethnic group. Table 2 shows that on average, group size sharpens the tradeoff: competence improves (columns 1 and 2), but representativeness of candidates falls (columns 4 and 5). Crucially, in our strictest common specifications (columns 2 and 5), if group share rises from 0 to 1, competence increases by a smaller amount (0.05 SD) than the fall in representativeness (father's education rises by 0.09 SD).

This, however, masks sharp heterogeneity by group. The representativeness result is driven by the majority (Hindu) group. For Hindu candidates, larger group size leaves competence unchanged but draws candidates from more elite families (father's education rises by 0.07 SD): the tradeoff deepens as the majority group dominates the GP. Minority (Muslim) candidates shape the competence increase with larger group sizes. Larger group size improves Muslim candidate competence (0.13 SD) while leaving representativeness unchanged: the tradeoff eases as the minority group grows.

The most natural explanation once again follows Banerjee and Pande (2007)'s logic. Majority-group politicians are essentially sure winners when their group is very large, reducing the competitive pressure to field capable candidates; elite families instead capture the political pipeline. Minority politicians face almost no chance of winning when their group is too small to clear an electoral threshold; as the group grows, selection pressure rises, and the pool of minority politicians shifts toward more competent candidates. The tradeoff thus depends not just on how diverse a GP is, but on whose perspective one takes – majority politicians face less pressure to be competent in homogeneous settings, while minority politicians benefit from larger group size.

5.3 Political Parties

We now turn to showing that parties may somewhat improve the competence-representativeness tradeoff.

Parties are not formally allowed to contest local elections in Bihar, but it is well documented that they play a substantial informal role (Dunning and Nilekani, 2013). We proxy for party presence using the share of ward members in each block who report receiving party support in the baseline survey of over 7,700 ward members conducted by Bamezai et al. (2024). We then estimate the interaction specifications in (3) and (4), replacing $Feature_g$ with this block-level share of party involvement.

Table 3 presents the results. Parties substantially increase candidate competence (columns 1 and 2). The effect on representativeness (column 4) is negative but smaller in magnitude, and – importantly – it is not robust to the inclusion of gender and caste fixed effects. The competence effect, by contrast, is robust across specifications. We therefore interpret the evidence as showing that party presence shifts the pool of candidates toward higher competence without necessarily tilting it toward more elite households.

This result is consistent with the argument that weak party institutions are one reason the competence-representativeness tradeoff is sharper in developing democracies than in advanced ones. Where parties operate – even informally – they appear to ease the tradeoff by actively identifying and supporting more competent candidates, expanding the pool beyond those who would enter politics through family connections or inherited advantage alone.

5.4 Taking Stock

Together, these results show that the competence-representativeness tradeoff is not a monolithic feature of political selection in rural Bihar. It intensifies where inequality concentrates opportunity and where majority ethnic groups dominate the electorate. It eases where parties are active and where minority groups grow large enough to compete electorally. The tradeoff, in other words, is shaped by the forces commonly seen in developing democracies – and it is responsive to institutional features. These findings motivate the analysis in Section 7, where we examine whether a deliberate policy intervention – political reservation – can reshape the tradeoff further.

6 Result #3: Consequences of Political Selection

Having established a systematic tradeoff between competence and representativeness in political selection, we now evaluate whether this tradeoff also appears in governance outcomes. We examine how leader competence—measured by educational attainment—shapes governance along two dimensions. We estimate the causal effect of electing a more com-

petent leader on skill-intensive and representation-intensive governance tasks.

For competence-intensive tasks, we focus on the implementation of infrastructure projects such as piped water schemes, drains, and village roads, which require coordination with line departments, contractor oversight, and bureaucratic engagement. We also measure political competence more directly, using a primary survey of ward members where we collect information on their knowledge and networks within government. For tasks that privilege representativeness, we examine whose preferences are better represented in the plans made by GPs – marginalized SCs or non-SCs. We also examine person-days generated under the Mahatma Gandhi National Rural Employment Guarantee Act (NREGA), a pro-poor workfare programme that is particularly salient for marginalized groups.

6.1 Knowledge, Networks and Skill-Intensive Outcomes

6.1.1 Data and Identification Strategy

Knowledge and Networks: Our first source of data covering politicians' knowledge and networks is a primary dataset covering 7277 ward members, all surveyed in 2023. In the survey, ward members were asked a series of questions about their roles in office.[6] We measure three outcomes. First, scheme knowledge: whether ward members can correctly enumerate the procedural steps required to implement key local programs (ward accounts, solar lights, social pensions, the Public Distribution System, drains and lanes, and housing). Second, bureaucratic networks: whether the ward member holds direct contact information for the Block Development Officer, the Block Panchayat Raj Officer, and the Panchayat Secretary. Third, political networks: whether they are formally involved in any parties and whether they received any external political support received during the campaign.

Mission Antyodaya Data Our second source of administrative data is the Mission Antyodaya survey (Ministry of Rural Development, 2020) accessed via the SHRUG platform (Asher et al., 2021). The survey is a nationwide village-level assessment conducted by the Government of India to benchmark rural infrastructure and service delivery. It covers the near-universe of Bihar's approximately 37,262 revenue villages and provides detailed information on the availability of core public goods.

During our study period, the Government of Bihar prioritised the expansion of basic infrastructure under the Chief Minister's Saat Nischay ("Seven Resolves") programme, including piped tap water (Nal-Jal) and lanes and drains (Nali-Gali) (Sharan, 2021). Importantly, responsibility for implementing these schemes was devolved to the Panchayati Raj Department, placing elected GP heads at the centre of planning, coordination, and execution. These sectors therefore constitute high-stakes, competence-intensive domains of local governance: successful implementation requires technical oversight, contractor management, and coordination with line departments, rather than merely distributive

[6] This was the baseline survey for the peer learning evaluation studied in (Bamezai et al., 2024).

targeting.

To construct village-level measures of realised infrastructure provision, we create: (i) an indicator for the availability of piped tap water, (ii) an index of road availability, (iii) an index of drainage system availability and (iv) an overall index of infrastructure. Together, these outcomes allow us to test whether leader competence translates into improved performance in infrastructure sectors that demand administrative capacity and bureaucratic engagement.

Identification Strategy To estimate the causal effect of leader competence on skill-intensive (knowledge and network) outcomes, we implement a close-election regression discontinuity (RD) design focusing on GP Head (Ward Member) elections held in 2016 (2021). As discussed earlier, leader competence is proxied by formal education. Identification comes from GPs (wards) in which a more educated candidate narrowly defeats a less educated candidate, or vice versa.

Our analysis is restricted to close elections in which the top two candidates differ along the education margin, with one candidate having education above the primary level and the other having education below the primary level. We define treatment as an indicator for whether the winning GP head (Ward member) has education above the primary school level. The running variable is the margin of victory between the more educated (above primary school) and less educated (primary school or below) candidate.

We estimate local linear RD specifications of the form:

$$Y_g = \mu + \theta T_g + \gamma_1 MoV_g + \gamma_2 (T_g \times MoV_g) + \varepsilon_g, \tag{5}$$

where Y_g denotes the realised development outcome in GP (ward) g, and $f(\cdot)$ is a linear function of the running variable, allowed to differ on either side of the cutoff at $MoV_g = 0$.

All specifications are estimated using local linear RD methods with Calonico–Cattaneo–Titiunik (Calonico et al., 2019) optimal bandwidth selection and robust bias-corrected confidence intervals. Standard errors are clustered at the GP level.

6.1.2 Results

Knowledge and Networks. Table 5 reports the results. Column 1 shows that in GPs where a more competent politician narrowly won, the incumbent leader is more knowledgeable about governance: across schemes, more competent leaders know 0.679 SD more of the key steps required to implement the schemes under their charge. Column 2 shows that competent politicians also have stronger networks within the local state: the average likelihood of possessing the phone numbers of the GP Secretary, the Block Panchayati Raj Officer (BPRO), and the Block Development Officer (BDO) increases by 38%. Column 3 shows that connections to political parties rise by 80%.

Together, these results show that more competent politicians are better equipped with the prerequisites for implementing skill-intensive development schemes: knowledge, contact with bureaucrats, and party ties.

Skill-Intensive Public Goods: These advantages translate into better implementation. Column 4 shows that GPs that narrowly elected more educated heads have better infrastructure: the infrastructure index increases by 0.10 SD (Column 4), driven by piped water access (Column 5) and roads (Column 6). Panels (a) and (b) of Figure A4 shows that our RD results are robust to a range of bandwidths.[7]

Overall, our results point to clear advantages of having a competent leader: they possess the prerequisites for implementing skill-intensive projects and capitalize on them to deliver better outcomes on the ground.

6.2 Competence and Alignment with Citizen Preferences

Having established that competence improves skill-intensive projects, we turn to representativeness. To measure the impacts on how representative GP policies are, we turn to the GP plans. Each GP prepares an annual Gram Panchayat Development Plan (GPDP), jointly determined by the village council and citizens. These plans reflect prospective local government spending for the subsequent year. To measure representativeness of plans, we collect data on planned projects and caste-wise (SC/non-SC) citizen preferences for those projects across GPs.

6.2.1 Data and Identification Strategy

GPDP Plan Data: We scrape all the data from the 2024-25 plan. This covers all GPs and includes, for each planned development project, the estimated cost of the project. This includes roads, drains, piped water, solar street-lights and so on. We identify project-type using descriptions, so as to match each project with citizen preferences.

Citizen Preferences. We measure citizen preferences using a phone survey of approximately 10,000 respondents across 10 districts in Bihar.[8] Each respondent was asked to name up to two schemes on which they would like the local government to increase spending. Thus, for every GP and scheme, this allows us to calculate the share of citizens who prefer spending on that scheme.

Identification Strategy. We estimate the causal effects of leader competence on development outcomes using a regression discontinuity design (RD) that exploits close GP-level electoral races. Competence is proxied by years of education of the winning leader. The

[7] Panel A of Table A17 shows that our results are similar when we change the definition of competence to being above high school. Panel B shows that while the knowledge and network results hold when we change the definition of competence to being above graduate, the infrastructure results, while positive, are no longer significant.

[8] These citizens were interviewed as part of the endline survey for (Bamezai et al., 2024).

identifying variation comes from elections in which a more educated candidate narrowly defeats a less educated candidate, or vice versa.

Our analysis focuses on GP head elections in 2021 in which the top two candidates differ in whether they have education above the primary level. Like seen in the previous section, we restrict the sample to GPs where one of the top two candidates (either the winner or the runner-up) has education above primary, while the other has education below primary.

Running Variable and Treatment. Let g index GPs. For each GP, the running variable is defined as the vote-share margin between the above-primary and below-primary candidates:

$$MoV_g = VS_g^{above} - VS_g^{below},$$

where VS_g^{above} and VS_g^{below} denote the vote shares of the above-primary and below-primary candidates in GP g, respectively. Positive values of MoV_g indicate that the above-primary candidate wins the election.

Treatment is defined as an indicator equal to one if the above-primary candidate wins the election:

$$T_g = \mathbb{1}\{MoV_g > 0\}.$$

In robustness checks (A16), we re-estimate the RD using alternative education cutoffs, including comparisons between candidates with high school/graduate education and those without.

Empirical Specification. Let s index schemes, and let Y_{gs} denote the development outcome for scheme s in GP g. For each GP-scheme pair, we observe citizen demand for that scheme, measured as the share of citizens in GP g who report preferring scheme s, denoted $Pref_{gs}$. We estimate the following local linear RD specification:

$$\begin{aligned} Y_{gs} = \mu_s + \theta_1 Pref_{gs} + \theta_2 T_g + \theta_3 (T_g \times Pref_{gs}) \\ + \gamma_0 f(MoV_g) + \gamma_1 (T_g \times f(MoV_g)) + \varepsilon_{gs}. \end{aligned} \tag{6}$$

where μ_s are scheme fixed effects and $f(\cdot)$ is a flexible function of the running variable. The function $f(\cdot)$ is allowed to differ on either side of the cutoff at $MoV_g = 0$.

All specifications are estimated using local linear regression discontinuity methods. We use the Calonico–Cattaneo–Titiunik (Calonico et al., 2019) optimal bandwidth selector. Standard errors are clustered at the GP level. We show that GPs above and below the threshold are balanced in Table A14.

6.2.2 Results

Table 4 presents results isolating the effects of narrowly electing a more educated GP Head on plan outcomes. In Columns (1) and (2), the coefficients on citizen preferences are large and statistically significant. These estimates indicate that, on average, an increase in citizen support for a public good in GPs quasi-randomly assigned a less competent leader leads to substantial increases in both planned spending (a 10% increase in support is associated with 61% more spending relative to the mean) and project counts. Electing a more educated leader, however, does not lead to systematically better or worse overall alignment between spending and aggregate citizen preferences. The interaction effects are small and statistically indistinguishable from zero.

These overall RD estimates, however, mask a sharp divergence by social group (Columns 3–6): in particular, competence worsens alignment with SC preferences, while improving alignment with non-SC preferences. Column (3) shows that a 10% increase in SC citizens preferring a public good results in an increase in spending of about INR 195,323. However, this responsiveness is reduced by about INR 73,010 in GPs where the leader is more competent, a decline of roughly 37%. Column (4) shows a similar pattern for project counts, with 41% fewer projects aligned with SC preferences in GPs quasi-randomly assigned a more educated leader.

Columns (5) and (6) display the opposite pattern for non-SC citizens: spending increases by 19% along non-SC citizens' preferences in GPs assigned a more educated leader.

Figure A3 shows that our main results are robust to a range of alternate bandwidths: across every bandwidth we test, SC spending is worse represented by more competent leaders and non-SC spending better. [9]

Taken together, these results provide direct evidence of a competence–representativeness trade-off in planning decisions. While citizen preferences strongly shape GP plans overall, electing a more educated leader does not increase aggregate responsiveness to citizen demand. Instead, competence shifts whose preferences are translated into allocations. Under more educated leadership, responsiveness to the preferences of marginalized SC citizens declines, while responsiveness to non-SC preferences increases.

7 Result #4: Policies to Improve Political Selection

Having shown that the competence–representativeness tradeoff has consequences for governance, we next ask whether institutional design can weaken it. We examine political reservations for SCs and women, distinguishing their effects while seats are reserved from those that persist after seats reopen. This allows us to assess whether reservations

[9] Table A16 shows that patterns are broadly true even we change the definition of more competent to above High School/Graduate: more competent leaders remain considerably better at representing non-SC preferences and worse at representing SC preferences, though some results are no longer statistically significant.

reshape political selection beyond the period in which they restrict eligibility for office.

In the short run, reservations broaden access to office for groups with historically lower educational attainment, potentially increasing representativeness at the cost of competence. However, their effects also depend on who enters from within these groups: do reservations primarily bring in relatively elite members, or attract highly competent citizens from more representative backgrounds? In the longer run, we ask whether exposure to reservations changes who contests and wins once seats are open and eligibility restrictions are removed.

7.1 SC Reservation

Since 2006, Bihar has reserved slightly over a sixth of GP head positions for SCs. Reservation lasts for ten years. We study candidates and winners in the 2016 and 2021 elections, distinguishing the effects of reservation during the current, 2016–2026 cycle from those of reservation during the preceding, 2006–2016 cycle.

7.1.1 Identification Strategy

Our empirical strategy exploits the population-based algorithm used to assign SC reservations, following Kumar and Sharan (2025). Within each administrative block, which contains an average of 15.7 GPs, the algorithm ranks GPs by their SC population and reserves head positions above a block-specific threshold. Crossing the cutoff increases the probability of reservation by approximately 80 percentage points. For the 2016–2026 cycle, the State Election Commission first excludes GPs reserved in the preceding cycle and reapplies the population-based algorithm to the remaining GPs.[10]

For the *short-run* analysis, we compare GPs narrowly above and below the cutoff for the 2016–2026 cycle, among GPs eligible for reservation in that cycle. Reservation during 2016–2026 is the treatment, and the outcomes are the characteristics of candidates and winners in 2016 and 2021. This design measures political selection while eligibility for office is restricted.

For the *long-run* analysis, we return to the cutoff governing reservation during 2006–2016. We restrict attention to GPs unreserved in the 2016 and 2021 elections and compare those narrowly above and below this earlier cutoff. The treatment is prior SC reservation during 2006–2016; outcomes are again measured in 2016 and 2021. Thus, we use the historical assignment threshold to study political selection after seats reopen, rather than using the current threshold to instrument for past reservation.

We estimate the two designs separately using the same fuzzy RD specification. Let $c \in \{2006, 2016\}$ denote the start of the relevant reservation cycle, and let $Reserved^{c}_{gb}$ indicate whether GP g in block b was reserved for SCs during that cycle. Define the running vari-

[10] Section F.1 describes the assignment rule in detail.

able as $R^c_{gb} = SCPop^c_{gb} - T^c_b$, where $SCPop^c_{gb}$ is the SC population used for assignment in cycle c and T^c_b is its block-specific cutoff. The instrument is $Z^c_{gb} = \mathbb{1}\{R^c_{gb} \geq 0\}$. For candidate i observed in election year t, we estimate:

$$Reserved^c_{gb} = \gamma^c_0 + \gamma^c_1 Z^c_{gb} + \gamma^c_2 R^c_{gb} + \gamma^c_3 R^c_{gb} Z^c_{gb} + \mu^c_b + \lambda^c_t + \eta^c_{igbt}, \tag{7}$$

$$Y_{igbt} = \beta^c_0 + \beta^c_1 Reserved^c_{gb} + \beta^c_2 R^c_{gb} + \beta^c_3 R^c_{gb} Z^c_{gb} + \mu^c_b + \lambda^c_t + \epsilon^c_{igbt}. \tag{8}$$

The first stage instruments reservation status with assignment above the relevant cutoff. The specification allows separate linear slopes on either side and includes block fixed effects μ^c_b and election-year fixed effects λ^c_t. The coefficient β^{2016}_1 captures the short-run effect of current reservation, while β^{2006}_1 captures the longer-run effect of prior reservation in the currently unreserved sample. Outcomes are observed at the candidate level; the winner analysis restricts the sample to elected candidates. Standard errors are clustered at the GP level.

7.1.2 Threats to Validity

Gender reservation imbalance. The assignment algorithm also generates an imbalance in gender reservation across the SC cutoff, as discussed by Kumar and Sharan (2025). They address this either by controlling for gender reservation or by excluding female-reserved seats. Our main short-run specification follows the latter approach. It also permits comparisons of fathers' education, which we cannot generally observe for female candidates. Our findings are robust to including female-reserved seats and controlling for gender reservation (See Table A7).

Comparability around the cutoff. Identification requires potential outcomes to vary smoothly across the relevant assignment threshold. Table A11 reports baseline balance for the broader sample, while Table A12 reports balance after excluding female-reserved seats. Figure A4 reports density tests, and Figure A5 examines sensitivity of the short-run estimates to alternative bandwidths.

For the long-run design, comparability must hold around the historical cutoff within the currently unreserved sample. This restriction requires particular care because prior reservation affects eligibility for reservation in the subsequent cycle. The interpretation of the long-run estimates therefore additionally requires that restricting the sample to currently unreserved GPs does not introduce discontinuous differences in potential outcomes at the earlier cutoff.

7.1.3 Short-Run Effects

SC reservation lowers candidates' competence but increases their representativeness. Panel A of Table 6 shows that candidates in SC-reserved GPs are 0.63 SD less educated than

those in unreserved GPs, although they remain substantially more educated than the average citizen. They also come from markedly less elite families: their fathers' education falls by 0.74 SD, from above to below the population average. SC reservation also brings in poorer households that are less likely to contain a government employee or an income-tax payer (Cols. 2–4).

Much of the educational gap reflects the different social groups from which candidates are drawn. When we benchmark candidates against citizens of their own caste, SC and unreserved candidates are nearly as positively selected: their educational advantages over their respective communities are approximately 0.85 and 0.95 SD (Panel B). Yet their social origins remain quite different. Unreserved candidates' fathers are 0.38 SD more educated than their own-group benchmark, while SC candidates' fathers have education levels close to the average in their community. SC reservation thus draws highly educated members of a disadvantaged group into politics without concentrating entry among its most elite families. The short-run aggregate tradeoff coexists with substantially more inclusive selection within caste.

These broad patterns carry through to elected leaders (Table A8). SC winners remain strongly positively selected on education and come from much less elite backgrounds than unreserved winners. However, their within-caste educational advantage is smaller than that of unreserved winners, with a larger gap than we observe between candidates.

7.1.4 Long-Run Effects

After seats reopen, previously reserved GPs attract candidates who are similarly competent but come from more representative backgrounds (Table 7, Panel A). Candidate education is essentially unchanged, while fathers' education falls by 0.25 SD. Among winners, fathers' education falls by 0.33 SD, and the estimated effect on own education is positive at 0.21 SD, although imprecisely estimated (Panel B). SC reservations thus draw candidates from less elite families even after seats reopen, without the reduction in competence observed in the short run.

The household indicators show a more mixed pattern: candidate households are more likely to contain government employees and tax payers, despite candidates coming from less educated families. Among winners, however, household wealth also declines.

Two changes in entry help explain these results. Kumar and Sharan (2025) show that SCs continue to run a Highly educated SC citizens are more likely to contest after reservations end. Moreover, past reservations could reshape participation beyond the group they directly target. Table A10 shows that exposure to less elite SC leaders encourages a more representative set of non-SC citizens to enter politics.

7.2 Gender Reservation

We next examine whether reservations for women generate similar changes in political selection, first while seats are reserved and then after they reopen.

7.2.1 Identification Strategy

Gender reservation is assigned within each block and caste-reservation category, with up to half of the positions in each category reserved for women. GPs are ranked by SC or ST population within those reserved categories, by total population within the EBC category, and by non-SC/ST population among caste-unreserved GPs. The threshold is the midpoint between the relevant populations of the last GP assigned to female reservation and the first GP that misses out. Section F.2 provides the details.

We apply the same distinction between current and historical assignment as in the SC analysis. For the short-run estimates, the treatment is female reservation during 2016–2026, instrumented by the corresponding assignment cutoff. The main results currently focus on caste-unreserved GPs, so the running variable is non-SC/ST population relative to the female-reservation cutoff in that category. For the long-run estimates, the treatment is female reservation during 2006–2016, instrumented by the cutoff governing that earlier assignment. Outcomes are measured in the 2016 and 2021 elections among GPs that are then unreserved.

Formally, let $FemaleReserved^c_{gb}$ indicate female reservation in cycle c. Let Q^c_{gb} denote the relevant assignment population minus its block- and caste-category-specific cutoff, and define $W^c_{gb} = \mathbb{1}\{Q^c_{gb} \geq 0\}$. For each cycle, we estimate:

$$\begin{aligned} FemaleReserved^c_{gb} = \gamma^c_0 + \gamma^c_1 W^c_{gb} + \gamma^c_2 Q^c_{gb} + \gamma^c_3 Q^c_{gb} W^c_{gb} \\ + X^{c\prime}_{gb}\delta^c + \mu^c_b + \lambda^c_t + \eta^c_{igbt}, \end{aligned} \tag{9}$$

$$\begin{aligned} Y_{igbt} = \beta^c_0 + \beta^c_1 FemaleReserved^c_{gb} + \beta^c_2 Q^c_{gb} + \beta^c_3 Q^c_{gb} W^c_{gb} \\ + X^{c\prime}_{gb}\theta^c + \mu^c_b + \lambda^c_t + \epsilon^c_{igbt}. \end{aligned} \tag{10}$$

Where specifications pool caste-reservation categories, X^c_{gb} includes indicators for those categories; these controls are unnecessary in the caste-unreserved-only specification. We estimate fuzzy RDs using outcome-specific optimal bandwidths following Calonico et al. (2019), include block and election-year fixed effects, and cluster standard errors at the GP level. The running variable is the population used by the assignment algorithm, which differs across caste-reservation categories, rather than female population itself.

7.2.2 Threats to Validity

As in the SC design, identification requires comparability on either side of the relevant cutoff. Table A13 reports baseline balance tests and Figure A4 reports density tests. Most covariates show no statistically detectable differences. These diagnostics must correspond to the sample and assignment cycle used in each specification. In particular, the long-run design requires comparability around the historical cutoff after restricting to GPs unreserved in the subsequent cycle.

7.2.3 Short-Run Effects

Gender reservation generates a similar short-run tradeoff. In caste-unreserved GPs, reservation reduces candidate education by 0.79 SD (Table 8, Panel A). Unlike the SC results, a substantial gap remains within group: relative to citizens of their own gender, female candidates have an educational advantage of approximately 0.53 SD, compared with 0.94 SD for unreserved candidates. Women entering politics are therefore positively selected on education, but less strongly so than their male counterparts.

Reservation also broadens the pool of households entering politics. Because married women typically move to their husbands' villages, we cannot generally identify their fathers and instead examine their current households. Female candidates' households have lower wealth and are less likely to pay income or professional taxes, although they remain more advantaged than the average household. Given the prevalence of proxy governance by husbands (Amar et al., 2025), we also examine husbands' education: they are 0.15 SD less educated than unreserved candidates (Col. 3). Reservation therefore changes which households contest, even when we consider the men who may exercise authority on elected women's behalf.

Among winners, education also falls sharply, but the changes in household wealth, government employment and tax payment are not statistically significant (Panel B). The evidence that gender reservation brings in less elite households is therefore clearer for entry than for election to office. However, when we look at seats across all castes (not only caste-unreserved ones), we see a clear trend: both winners and candidate-households are less elite (Table A18).

7.2.4 Long-Run Effects

As with SC reservation, the longer-run effects of gender reservation differ from its contemporaneous effects. Candidate education rises by 0.15 SD and winner education by 0.27 SD, both statistically significant (Table 9). Neither candidates nor winners show a statistically detectable change in fathers' education. Past gender reservation thus raises competence without a corresponding increase in the eliteness of family backgrounds.

These gains primarily reflect changes in the pool of male candidates, who account for

most entrants after gender reservations end. As in the SC case, reservations influence political selection beyond the group whose eligibility they directly expand. The improvements in competence after seats reopen suggest that the effects of reservations extend beyond the short-run tradeoff observed while quotas apply.

8 Conclusion

This paper documents a tradeoff between competence and representativeness in political selection in a developing democracy. Linking administrative records on 97 million rural residents of Bihar to nearly one million candidates contesting local elections, we show that politicians are strongly positively selected on education but come from disproportionately elite families. In unreserved seats, this tradeoff sharpens as candidates progress to higher office and as voters select winners from the candidate pool. It also has consequences for governance: more educated leaders implement skill-intensive projects more effectively, but align public spending more closely with advantaged groups' priorities and less closely with those of marginalised citizens.

Political reservations reshape this tradeoff over time. In the short run, reservations for SCs and women bring in less educated candidates from more representative backgrounds. For SCs, much of the educational gap reflects inequalities across caste groups: SC candidates are nearly as positively selected on education within their own group as unreserved candidates are within theirs, yet come from substantially less elite families. Under gender reservation, a larger educational gap remains within group, although reservations also broaden the pool of households entering politics. The longer-run effects differ from these contemporaneous patterns. After SC-reserved seats reopen, candidates and winners come from less elite families without a detectable reduction in education. Past gender reservation raises the education of both candidates and winners without a detectable increase in fathers' education. These changes extend beyond the groups directly targeted by quotas, as reservations also influence which non-SC citizens and men subsequently enter politics. Reservations can therefore improve political selection beyond the period in which they restrict eligibility for office.

These findings raise further questions about how institutions shape political entry and governance. What explains the changes in participation after reservations end, particularly among citizens outside the groups they target? Do the longer-run improvements in political selection translate into more capable and responsive government? Training and peer learning may offer complementary ways to improve competence among leaders from less elite backgrounds. Extending the analysis to corruption, rent-seeking, and other dimensions of state performance would also clarify the broader consequences of the competence–representativeness tradeoff.

References

Abramitzky, R., L. Boustan, K. Eriksson, J. Feigenbaum, and S. Pérez (2021): "Automated linking of historical data," *Journal of Economic Literature*, 59, 865–918.

Amar, P., A. Bamezai, and R. Kumar (2025): "Mere Proxies or Genuine Leaders? Women in Local Politics in the Context of Gender Quotas," SSRN Working Paper No. 5200785.

Asher, S., T. Lunt, R. Matsuura, and P. Novosad (2021): "Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in India using the shrug open data platform," *The World Bank Economic Review*, 35.

Bamezai, A., S. George, S. Hari, and M. R. Sharan (2024): "Learning to Govern: The Impact of Politicians' Peer Networks," Stage-1 (Pre-Results) Acceptance, Journal of Development Economics. SSRN Working Paper No. 4664964.

Banerjee, A. V. and R. Pande (2007): "Parochial politics: Ethnic preferences and politician corruption," .

Besley, T. (2005): "Political Selection," *Journal of Economic Perspectives*, 19, 43–60.

Besley, T., O. Folke, T. Persson, and J. Rickne (2017): "Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden," *American economic review*, 107, 2204–2242.

Besley, T., J. G. Montalvo, and M. Reynal-Querol (2011): "Do Educated Leaders Matter?" *The Economic Journal*, 121, F205–F227.

Bhavnani, R. R. (2009): "Do Electoral Quotas Work after They Are Withdrawn? Evidence from a Natural Experiment in India," *American Political Science Review*, 103, 23–35.

Bhusal, B., M. Callen, R. Pande, S. A. Prillaman, D. Singhania, and A. Subedi (2026): "From People's War to People's Rule: Rebel Governance and the Foundations of Inclusive Democracy," Tech. rep., National Bureau of Economic Research.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2019): "Regression discontinuity designs using covariates," *Review of Economics and Statistics*, 101, 442–451.

Carnes, N. and N. Lupu (2016): "What Good Is a College Degree? Education and Leader Quality Reconsidered," *The Journal of Politics*, 78, 35–49.

Chattopadhyay, R. and E. Duflo (2004): "Women as policy makers: Evidence from a randomized policy experiment in India," *Econometrica*, 72, 1409–1443.

Chaudhuri, A., V. Iversen, F. R. Jensenius, and P. Maitra (2025): "Selecting the 'Best'? Competing Dimensions of Politician Quality in the Developing World," *British Journal of Political Science*, 55, e81.

Chin, A. and N. Prakash (2011): "The redistributive effects of political reservation for minorities: Evidence from India," *Journal of development Economics*, 96, 265–277.

Curto-Grau, M. and A. Gallego (2023): "How Do the Educated Govern? Evidence from Spanish Municipalities," in *Unequal Democracies: Public Policy, Responsiveness, and Redistribution in an Era of Rising Economic Inequality*, ed. by N. Lupu and J. Pontusson, Cambridge University Press.

Dahlgaard, J. O. and R. T. Pedersen (2024): "Inclusive Meritocracy: Ability and Descriptive Representation Among Danish Politicians," *Political Science Research and Methods*, 1–24, forthcoming/Online first.

——— (2025): "Inclusive meritocracy: ability and descriptive representation among Danish politicians," *Political Science Research and Methods*, 13, 567–590.

Dal Bó, E. and F. Finan (2018): "Progress and Perspectives in the Study of Political Selection," *Annual Review of Economics*, 10, 541–575.

Dal Bó, E., F. Finan, O. Folke, T. Persson, and J. Rickne (2017): "Who Becomes a Politician?" *The Quarterly Journal of Economics*, 132, 1877–1914.

Desai, S. and R. Vanneman (2015): "India human development survey-ii (ihds-ii), 2011-12. icpsr36151-v2," *Ann Arbor, MI: Inter-university Consortium for Political and Social Research* [*distributor*], 31.

Dunning, T. and J. Nilekani (2013): "Ethnic quotas and political mobilization: caste, parties, and distribution in Indian village councils," *American political Science review*, 107, 35–56.

Feigenbaum, J. J., A. B. Hall, D. M. Thompson, and J. Yoder (2026): "Fortunate Sons: Elite Political Selection in American History," Tech. rep., National Bureau of Economic Research.

Fouka, V., S. Mazumder, and M. Tabellini (2022): "From immigrants to Americans: Race and assimilation during the Great Migration," *The Review of Economic Studies*, 89, 811–842.

Göhlmann, S. and R. Vaubel (2007): "The Educational and Occupational Background of Central Bankers and Its Effect on Inflation: An Empirical Analysis," *European Economic Review*, 51, 925–941.

Gulzar, S. (2021): "Who Enters Politics and Why?" *Annual Review of Political Science*, 24, 253–275.

Gulzar, S., N. Haas, and B. Pasquale (2020): "Does Political Affirmative Action Work, and for Whom? Theory and Evidence on India's Scheduled Areas," *American Political Science Review*, 114, 1230–1246.

Iyer, L., A. Mani, P. Mishra, and P. Topalova (2012): "The Power of Political Voice: Women's Political Representation and Crime in India," *American Economic Journal: Applied Economics*, 4, 165–193.

Jefferson, T. (1813): "Letter to John Adams, 28 October 1813," in *The Adams–Jefferson Letters: The Complete Correspondence Between Thomas Jefferson and Abigail and John Adams*, ed. by L. J. Cappon, Chapel Hill: University of North Carolina Press, vol. 2, 387–392, reprinted 1959.

Jokela, M., J. Meriläinen, J. Tukiainen, and Å. von Schoultz (2025): "Personality traits and cognitive ability in political selection," *Journal of the European Economic Association*, 23, 1908–1949.

Kumar, C. and M. R. Sharan (2025): "The Short- and Long-Run Distributional Consequences of Political Reservations," R&R, Journal of the European Economic Association.

Lahoti, R. and S. Sahoo (2020): "Are Educated Leaders Good for Education? Evidence from India," *Journal of Economic Behavior & Organization*, 176, 42–62.

Martinez-Bravo, M. (2017): "The Local Political Economy Effects of School Construction in Indonesia," *American Economic Journal: Applied Economics*, 9, 256–289.

Ministry of Rural Development, G. o. I. (2020): "Mission Antyodaya," Accessed via SHRUG https://www.devdatalab.org/shrug_download/, v2.1.pakora.

O'Connell, S. D. (2020): "Can Quotas Increase the Supply of Candidates for Higher-Level Positions? Evidence from Local Government in India," *Review of Economics and Statistics*, 102, 65–78.

Pande, R. (2003): "Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India," *American Economic Review*, 93, 1132–1151.

Sharan, M. R. (2021): *Last Among Equals: Power, Caste and Politics in Bihar's Villages*, Westland Publishers.

——— (2023): "Making the Bihar Caste Census Better," *The India Forum*.

Sørensen, R. J. (2023): "Educated Politicians and Government Efficiency: Evidence from Norwegian Local Government," *Journal of Economic Behavior & Organization*, 210, 163–179.

Swaminathan, M. and S. Nagbhushan (2022): "Wealth Inequality: Evidence from Two Villages in Bihar," *Review of Agrarian Studies*, 12, 66–89.

Thompson, D. M., J. J. Feigenbaum, A. B. Hall, and J. Yoder (2019): "Who Becomes a Member of Congress? Evidence from De-Anonymized Census Data," NBER Working Paper 26156, National Bureau of Economic Research.

9 Tables

Table 1: Political Selection Among Candidates Contesting for Unreserved Seats

	Years of Schooling					
	Baseline	Gender	Gender + Age Cohort	Gender + Age + Caste	Gender + Household	Father
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Competence: Comparison between Citizens and Candidates						
Candidate	0.673***	0.561***	0.492***	0.474***	0.344***	0.185***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.007)
Panel B - Representativeness: Comparison between Citizens and Candidates Fathers						
Candidate's Father	0.242***	0.106***	0.306***	0.282***	0.158***	
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	
Observations	97,349,644	97,349,644	97,348,519	97,348,519	96,385,396	10,856,994

This table reports regression estimates of differences in years of education between individuals who have ever run for Mukhiya or Ward office in unreserved seats and those who have never run for office, using data from the 2016 and 2021 Panchayat elections. Panel A reports differences between candidates and citizens, while Panel B reports differences between candidates' fathers and citizens. The dependent variable, years of schooling, is standardized within GP. Columns (2)–(6) sequentially include fixed effects for gender (2); gender and age cohort (3); gender, age cohort, and caste (4); gender and household (5); and father (6). Standard errors are clustered at GP-level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Heterogeneous Selection by Religious Group Size

	Competence		Representativeness	
	Baseline (1)	Gender + Age + Caste (2)	Baseline (3)	Gender + Age + Caste (4)
Panel A: All Population				
Candidate	0.620***	0.424***		
	(0.020)	(0.019)		
Candidates X Group Share	0.056**	0.046**		
	(0.022)	(0.022)		
Candidate Father			0.123***	0.194***
			(0.018)	(0.018)
Candidates Father X Group Share			0.133***	0.092***
			(0.021)	(0.021)
Observations	94,770,823	94,769,723	94,770,823	94,769,723
Panel B: Hindu Population				
Candidate	0.676***	0.420***		
	(0.033)	(0.032)		
Candidates X Group Share	-0.004	0.033		
	(0.037)	(0.035)		
Candidate Father			0.077**	0.067**
			(0.032)	(0.032)
Candidates Father X Group Share			0.186***	0.226***
			(0.036)	(0.035)
Observations	79,978,946	79,978,012	79,978,946	79,978,012
Panel C: Muslim Population				
Candidate	0.499***	0.382***		
	(0.024)	(0.024)		
Candidates X Group Share	0.128**	0.128**		
	(0.053)	(0.053)		
Candidate Father			0.124***	0.220***
			(0.023)	(0.023)
Candidates Father X Group Share			0.004	0.006
			(0.049)	(0.049)
Observations	14,791,877	14,791,711	14,791,877	14,791,711

This table reports heterogeneity in political selection by ethnic group size, using data from the 2016 and 2021 Gram Panchayat elections. Panel A pools all population and interacts an indicator for ever standing for office with the own-group population share at the GP level. Panels B and C restrict the sample to Hindu and Muslim individuals, respectively, and estimate the analogous specification using each group's own population share. Columns (1)–(2) report differences between candidates and citizens; columns (3)–(4) report differences between candidates' fathers and citizens. Column (2) and (4) include gender, age-cohort, and caste fixed effects. All dependent variables are standardized within GP. Standard errors, clustered at the GP level, are reported in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Heterogeneous Selection by Party Affiliation

	Competence		Representativeness	
	Baseline (1)	Gender + Age + Caste (2)	Baseline (3)	Gender + Age + Caste (4)
Candidate	0.574***	0.394***		
	(0.026)	(0.025)		
Candidates X Party Involv.	0.463***	0.402***		
	(0.118)	(0.113)		
Candidate Father			0.197***	0.250***
			(0.025)	(0.024)
Candidates Father X Party Involv.			0.189*	0.132
			(0.111)	(0.111)
Observations	26,083,033	26,082,888	26,083,033	26,082,888

This table reports heterogeneity in political selection by party involvement, using data from the 2016 and 2021 Gram Panchayat elections. We examine heterogeneity by interacting an indicator for ever standing for office with the share of candidates in the block supported by any political party. Columns (1)–(2) report differences between candidates and citizens; columns (3)–(4) report differences between candidates' fathers and citizens. Columns (2) and (4) include gender, age-cohort, and caste fixed effects. All dependent variables are standardized within GP. Standard errors, clustered at the GP level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Alignment of Public Spending with Citizen Preferences

	Overall		SC Pref.		Non-SC Pref.	
	Spending (1)	Projects (2)	Spending (3)	Projects (4)	Spending (5)	Projects (6)
Above Primary	-4136.251	-0.127	-33495.772	-0.061	-1121.092	-0.140
	(40561.287)	(0.127)	(64144.157)	(0.216)	(41303.792)	(0.130)
Pref.	2790806.395***	7.246***	1953232.499***	5.684***	2423916.218***	6.229***
	(208218.051)	(0.515)	(242999.214)	(0.685)	(192146.563)	(0.484)
Above Primary × Pref.	160525.064	0.170	-730101.491**	-2.337***	452260.136*	0.916
	(277697.717)	(0.720)	(315195.252)	(0.861)	(258905.339)	(0.680)
Dep. Mean	451465.14	1.29	471101.52	1.34	449939.43	1.29
Bandwidth	0.12	0.17	0.12	0.17	0.12	0.17
Observations	5,145	6,735	2,430	3,045	5,025	6,585

This table reports the heterogeneous effects of electing a more educated GP heads—defined as having above-primary education—on government scheme outcomes by local citizen preferences. Estimates are obtained from a close-election regression discontinuity design. Citizen preferences are measured across a comprehensive set of local public goods and welfare schemes, including water and sanitation, infrastructure, and social protection. We control for scheme-specific measures of citizen preferences and include interactions between these measures and the treatment indicator. The outcome variables are scheme-level expenditures (*Spending*) and the number of projects implemented (*Projects*). Columns (1)–(2) use preferences of the overall population, Columns (3)–(4) use preferences of the Scheduled Caste population, and Columns (5)–(6) use preferences of the non–Scheduled Caste population. All specifications include scheme fixed effects. Standard errors are clustered at the running-variable level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Political competence and development outcomes

	Information	Network		Project Implementation			
	Scheme Knowledge (1)	Bureaucratic Network (2)	Party Support (3)	Infra. Aggregated (4)	Tap Water (5)	Road Index (6)	Drainage Index) (7)
Winner	0.679***	0.173***	0.092***	0.103*	0.160*	0.174*	-0.009
	(0.180)	(0.045)	(0.030)	(0.060)	(0.092)	(0.099)	(0.097)
Dep. Mean	1.95	0.45	0.11	0.00	0.00	0.02	-0.02
Bandwidth	0.08	0.08	0.08	0.13	0.17	0.12	0.13
Observations	708	697	734	1,515	1,799	1,422	1,482

This table reports the effects of electing a more educated politician—defined as having above-primary education—on three dimensions of governance: information, network, and project implementation. Estimates are obtained from a close-election regression discontinuity design, restricted to constituencies in which the top two candidates' education levels straddle the above-primary threshold. The treatment variable (*Above Primary*) equals one if the winning candidate has above-primary education. Columns (1)–(3) use a survey of male Ward Members elected in the 2021 Ward elections. *Scheme Knowledge* (1) is the average of standardized scores measuring each Ward Member's ability to recount the implementation steps for six government schemes (ward account, AWAS housing, drains, PDS, pensions, and solar lights), matched to a predetermined checklist. *Bureaucratic Network* (2) is the average of indicators for whether the respondent knows the phone numbers of the Panchayat Secretary, BPRO, and BDO. *Party Support* (3) is the average of indicators for any party involvement and receipt of political support.Columns (4)–(7) use GP-level data from the 2016 Mukhiya elections. *Infra. Aggregated* (4) is the average of three indicators for the availability of piped tap water (5), road availability (6), and drainage systems (7), each also shown separately. Standard errors in Columns (4)–(7) are clustered at the running-variable level. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Effects of SC Reservation on Selection for GP Head Candidates

	Yrs of Schooling	Representativeness			
	Candidate Edu Yrs	Gov. Job	Income Tax	Wealth Score	Father Edu Yrs
	(1)	(2)	(3)	(4)	(5)
Panel A: Comparison to all citizen					
SC RSVN	-0.626***	-0.283***	-0.321***	-0.639***	-0.744***
	(0.051)	(0.051)	(0.049)	(0.044)	(0.051)
Dep. Mean	1.01	0.26	0.20	0.55	0.44
Bandwidth	141.557	150.272	143.460	141.613	159.372
Observations	10,768	11,430	10,863	10,761	6,381
Panel B: Comparison to citizen of own group					
SC RSVN	-0.105**	-0.152***	-0.174***	-0.161***	-0.358***
	(0.052)	(0.051)	(0.039)	(0.044)	(0.053)
Dep. Mean	0.95	0.25	0.18	0.50	0.38
Bandwidth	140.548	150.184	351.059	148.457	160.452
Observations	10,742	11,318	19,910	11,293	6,412

This table reports the effects of Scheduled Caste (SC) reservation on political selection outcomes for GP Head candidates in the 2016 and 2021 Panchayat elections. Estimates are obtained from a regression discontinuity design around the population threshold for SC reservation. Columns (1)–(5) report results for the following candidate-level outcomes: standardized years of education (1); an indicator for whether any household member holds a salaried government job (2); an indicator for whether any household member pays income or professional taxes (3); a PCA-based wealth index constructed using land ownership, housing characteristics, and asset ownership (4); and the candidate's father's years of education (5). Panel A reports outcomes standardized relative to all citizens within the GP, while Panel B reports outcomes standardized relative to individuals from the same caste within the GP. The sample is restricted to seats unreserved for women. All specifications include block fixed effects and election-year controls. Standard errors are clustered at the GP level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Long-Run Effects of SC Reservation on Political Selection

	Yrs of Schooling	Representativeness			
	Candidate Edu Yrs	Gov. Job	Income Tax	Wealth Score	Father Edu Yrs
	(1)	(2)	(3)	(4)	(5)
Panel A: Candidates					
SC RSVN	-0.014	0.265***	0.105*	0.040	-0.246***
	(0.075)	(0.061)	(0.057)	(0.054)	(0.079)
Dep. Mean	1.04	0.28	0.21	0.59	0.47
Bandwidth	239.468	224.086	292.321	255.509	222.937
Observations	7,927	7,551	9,447	8,511	4,631
Panel B: Winners					
SC RSVN	0.211	-0.162	-0.275*	-0.242**	-0.333***
	(0.129)	(0.128)	(0.155)	(0.123)	(0.127)
Dep. Mean	1.30	0.38	0.42	0.93	0.67
Bandwidth	366.625	464.966	426.554	313.725	343.450
Observations	1,045	1,365	1,238	844	643

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Effects of Female Reservation on Political Selection: Caste-Unreserved GPs

	Yrs of Schooling			Representativeness		
	Edu Yrs (All Citizen) (1)	Edu Yrs (Own Group) (2)	Husband Edu Yrs (3)	Gov. Job (4)	Income Tax (5)	Wealth Score (6)
Panel A: Candidates						
Female RSVN	-0.794***	-0.413***	-0.154***	-0.062	-0.085*	-0.094**
	(0.038)	(0.038)	(0.052)	(0.046)	(0.045)	(0.038)
Dep. Mean	1.10	0.94	1.10	0.33	0.25	0.65
Bandwidth	1010.169	1029.815	921.515	918.683	980.925	1039.182
Observations	19,001	19,212	14,367	18,239	18,739	19,304
Panel B: Winners						
Female RSVN	-0.918***	-0.528***	-0.685***	0.031	-0.285	-0.096
	(0.114)	(0.116)	(0.146)	(0.147)	(0.176)	(0.113)
Dep. Mean	1.45	1.27	1.45	0.45	0.51	1.05
Bandwidth	448.804	445.913	436.782	654.905	716.648	445.736
Observations	1,268	1,257	927	1,622	1,682	1,257

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

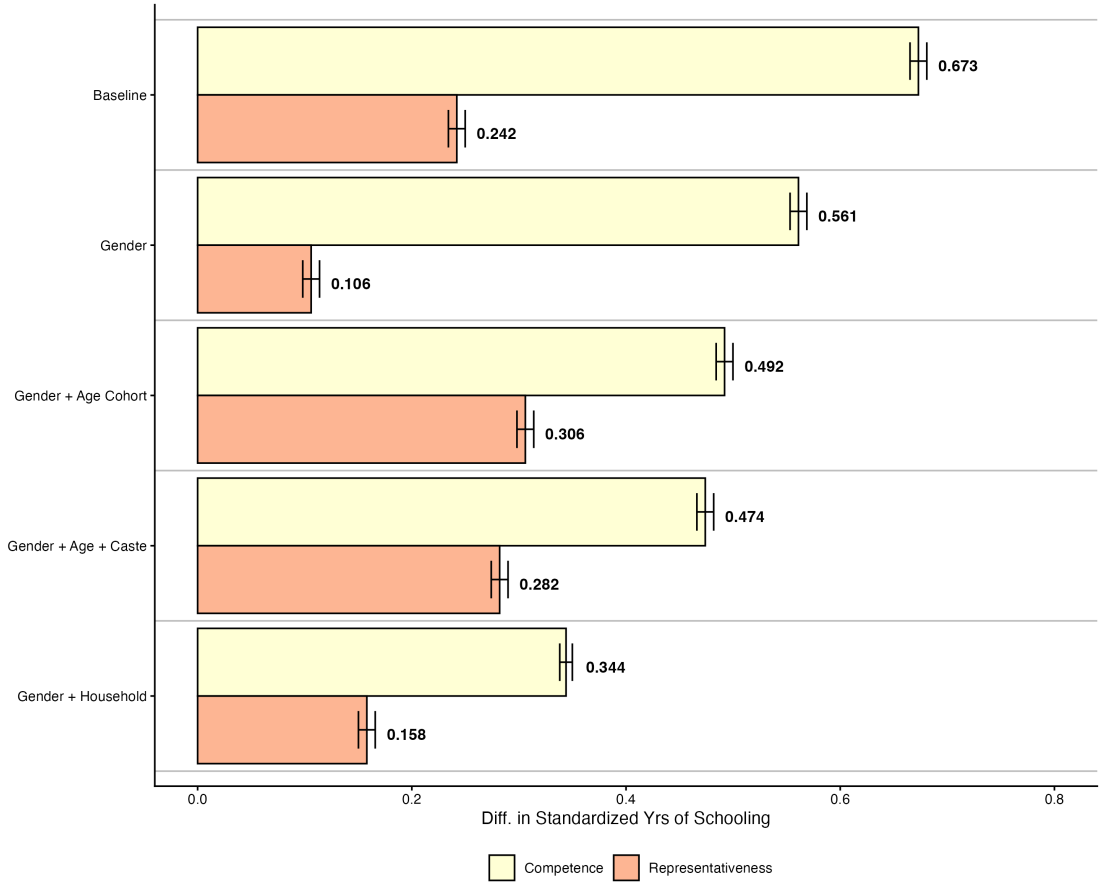
Table 9: Long-Run Effects of Female Reservation on Political Selection

	Yrs of Schooling	Representativeness			
	Candidate Edu Yrs (1)	Gov. Job (2)	Income Tax (3)	Wealth Score (4)	Father Edu Yrs (5)
Panel A: Candidates					
Female RSVN	0.148***	0.033	0.129***	0.004	0.008
	(0.036)	(0.036)	(0.038)	(0.033)	(0.042)
Dep. Mean	0.86	0.24	0.16	0.39	0.32
Bandwidth	335.081	390.293	304.099	428.821	330.813
Observations	20,889	23,093	19,382	24,463	12,132
Panel B: Winners					
Female RSVN	0.273***	0.446***	0.140	0.047	-0.094
	(0.090)	(0.104)	(0.114)	(0.087)	(0.100)
Dep. Mean	1.19	0.34	0.37	0.75	0.52
Bandwidth	538.687	628.051	582.590	490.005	312.417
Observations	2,497	2,740	2,615	2,359	1,138

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

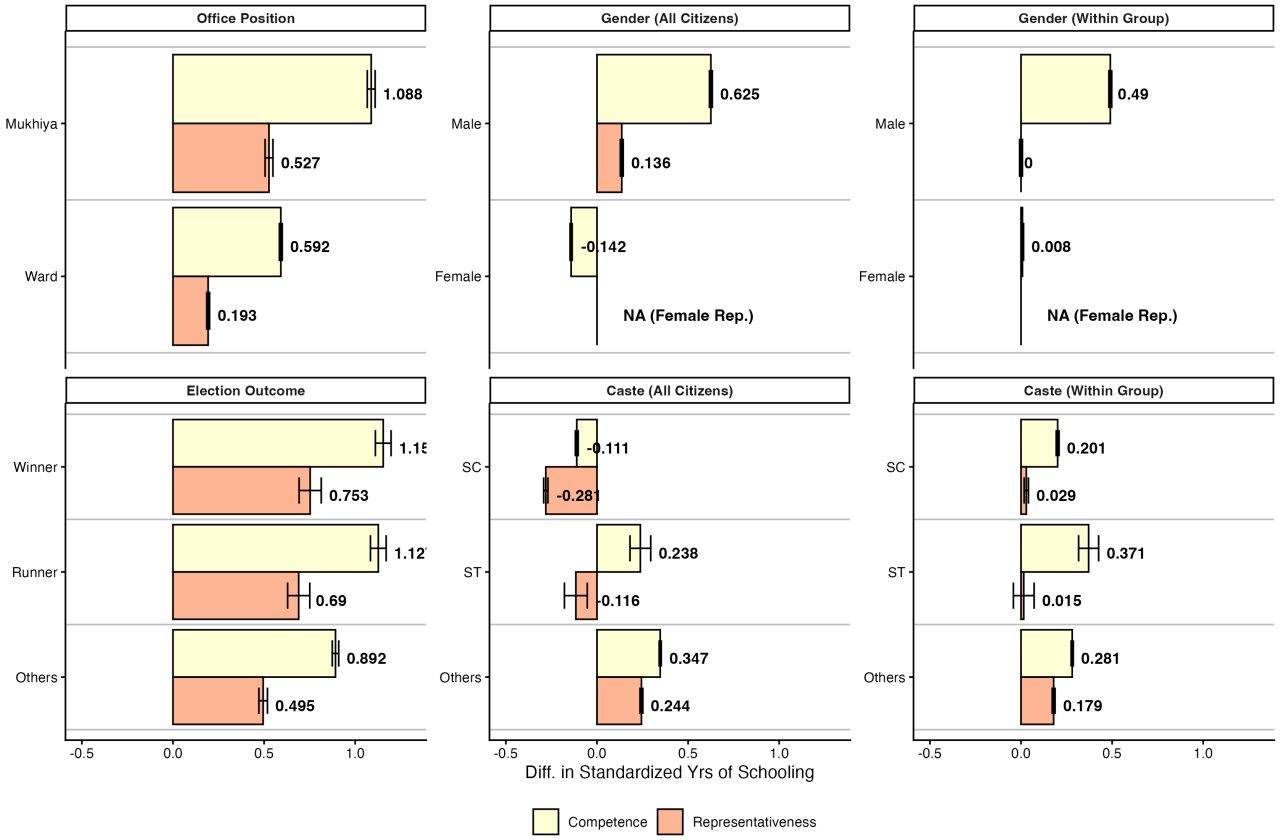
10 Figures

Figure 1: Political Selection: Competence and Representativeness



Notes: This figure visualizes the coefficients from the regressions reported in Table 1. Each bar represents the average difference in standardized years of education between candidates and the general population (Competence, yellow bars), and between candidates' fathers and the general population (Representativeness, orange bars). The "Baseline" specification includes GP fixed effects. Subsequent rows progressively add fixed effects for gender, age cohort, and caste group. The final row ("Gender + Household") utilizes household fixed effects to compare candidates to their own siblings. Error bars represent 95% confidence intervals; standard errors are clustered at the GP level.

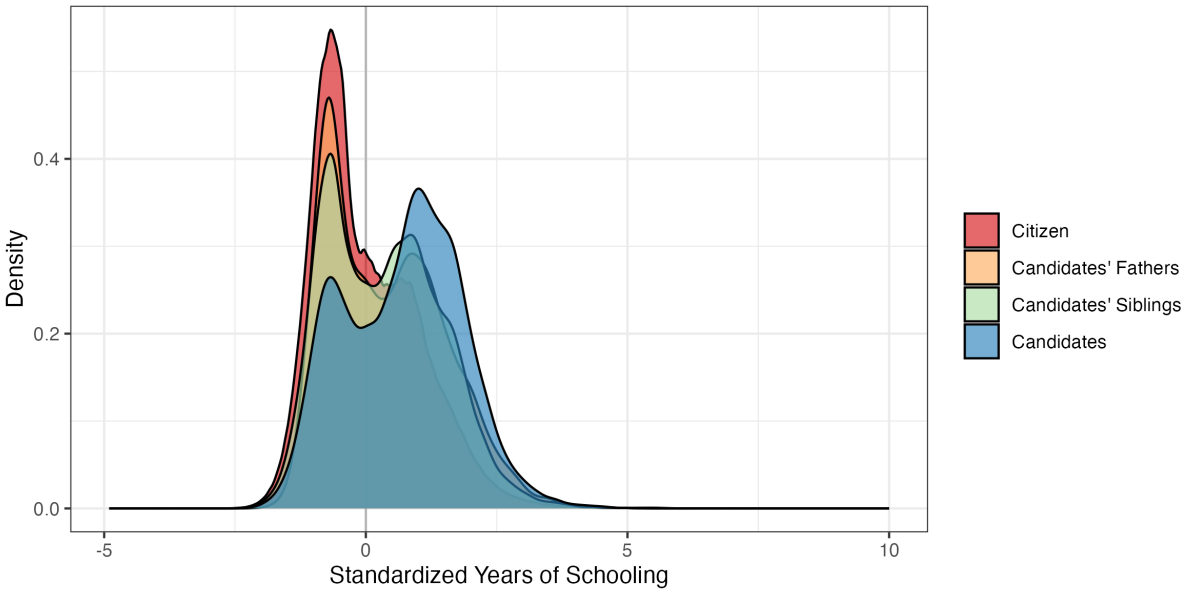
Figure 2: Heterogeneity in Selection



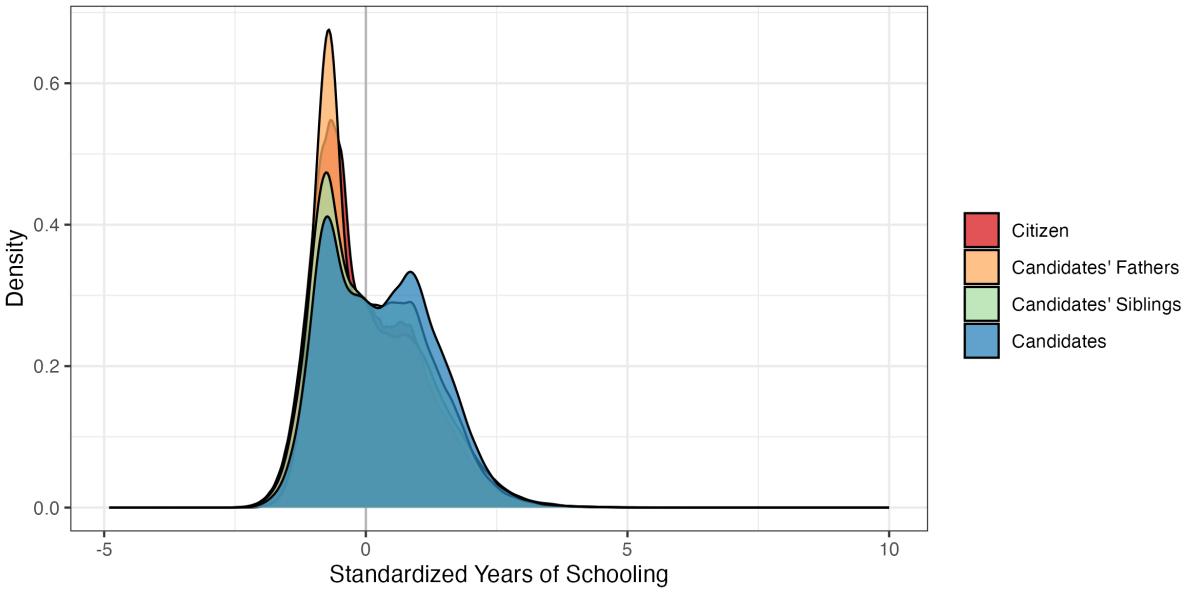
Notes: This figure plots coefficients from the baseline specifications (Equation 1 and **??**) across various sub-groups, with standardized years of education as the dependent variable. Top left compares candidates for GP head and Ward member positions against the general population. Top middle and Top right compare male and female candidates against the full population and their respective gender groups, respectively. Bottom left disaggregates candidates by election outcome (winners, runners-up, and other candidates) relative to the population. Bottom middle and Bottom right compare candidates from different caste categories (SC/ST and Others) against the full population and their own caste groups, respectively. All specifications include GP fixed effects. Error bars represent 95% confidence intervals with standard errors clustered at the GP level.

Figure 3: Candidates, Fathers & Siblings

(a) GP Head



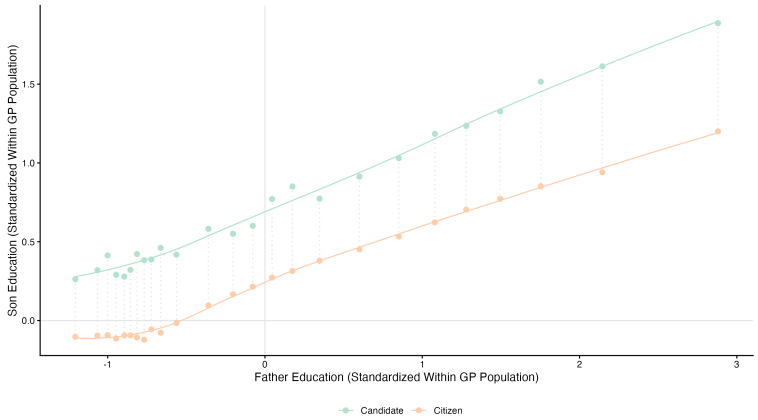
(b) Ward Member



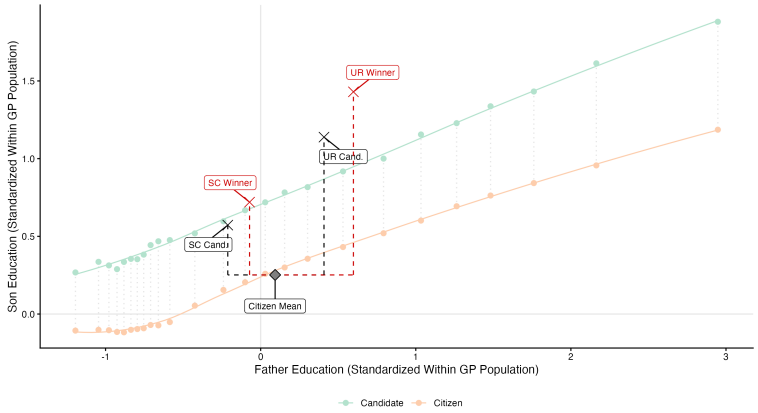
Notes: This figure plots the Kernel density of standardized years of education for candidates, their fathers, their siblings, and the general population. Years of education are standardized by gender, caste, and age cohort within each GP to account for local demographic differences. Panel A restricts the sample to candidates running for the position of GP heads and their respective family members. Panel B reports the same distributions for candidates running for Ward member positions and their families.

Figure 4: Competence-Representativeness Trade-off

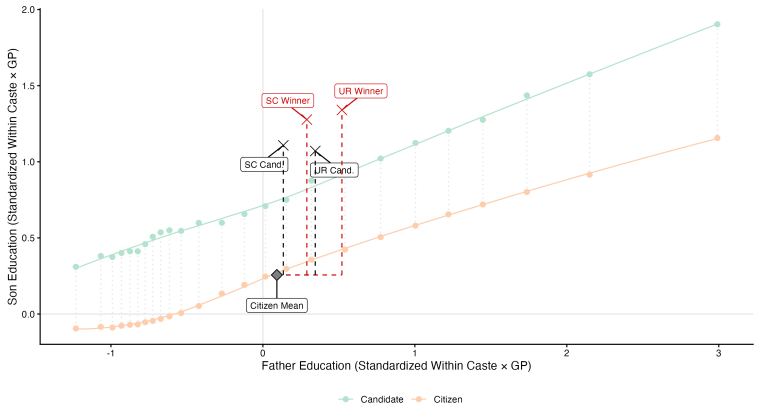
(a) The Candidate Selection Frontier and Citizen Mobility Curve



(b) Selection Frontier by Reservation Status



(c) Selection Frontier by Reservation Status (Caste-Benchmarked)



Notes: This figure plots intergenerational education mobility — father's education against son's education — for candidates and citizens in the SC reservation regression discontinuity sample. Bin scatter points are constructed by dividing the sample into 25 quantile bins based on father's education, with vertical bars indicating the within-bin range and smooth curves fitted separately for candidates and citizens. Panel (a) presents the baseline bin scatter without reference points, with years of education standardized within GP. Panel (b) augments Panel A by adding RD point estimates for mean father's and son's education among unreserved (UR) and SC candidates and winners, with the citizen mean shown as a diamond reference point; dashed lines indicate each group's deviation from the citizen mean along both dimensions. Panel (c) mirrors Panel B but benchmarks years of education within both GP and caste group, allowing comparisons relative to one's own caste peers. Black markers correspond to candidates; red markers correspond to winners.

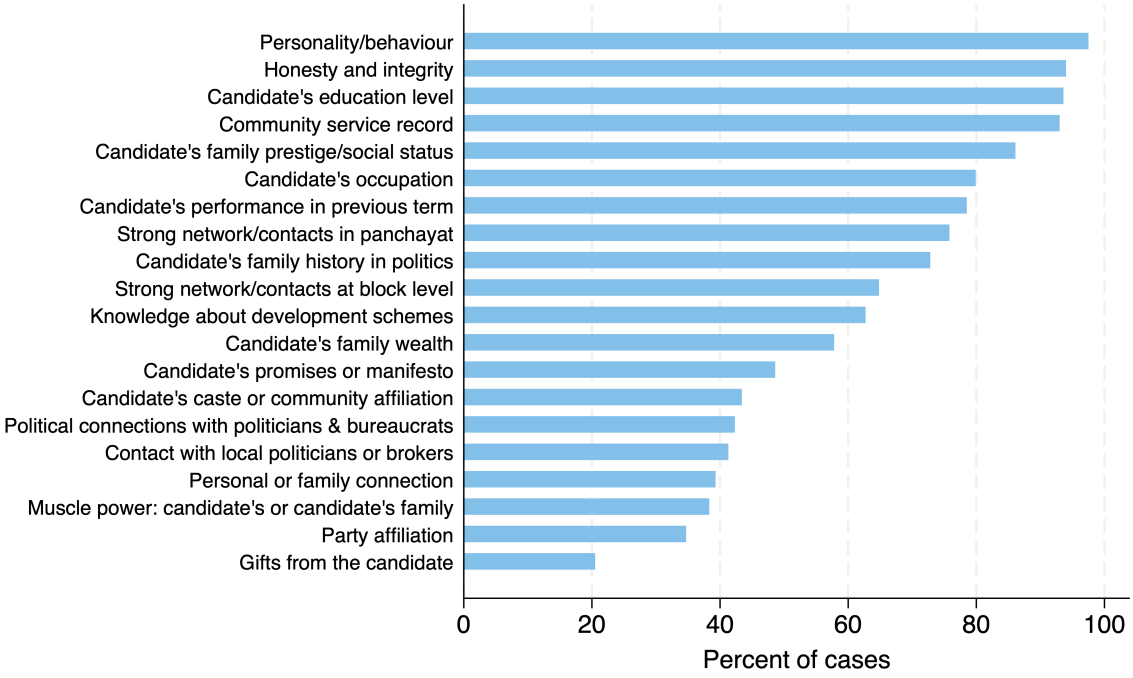
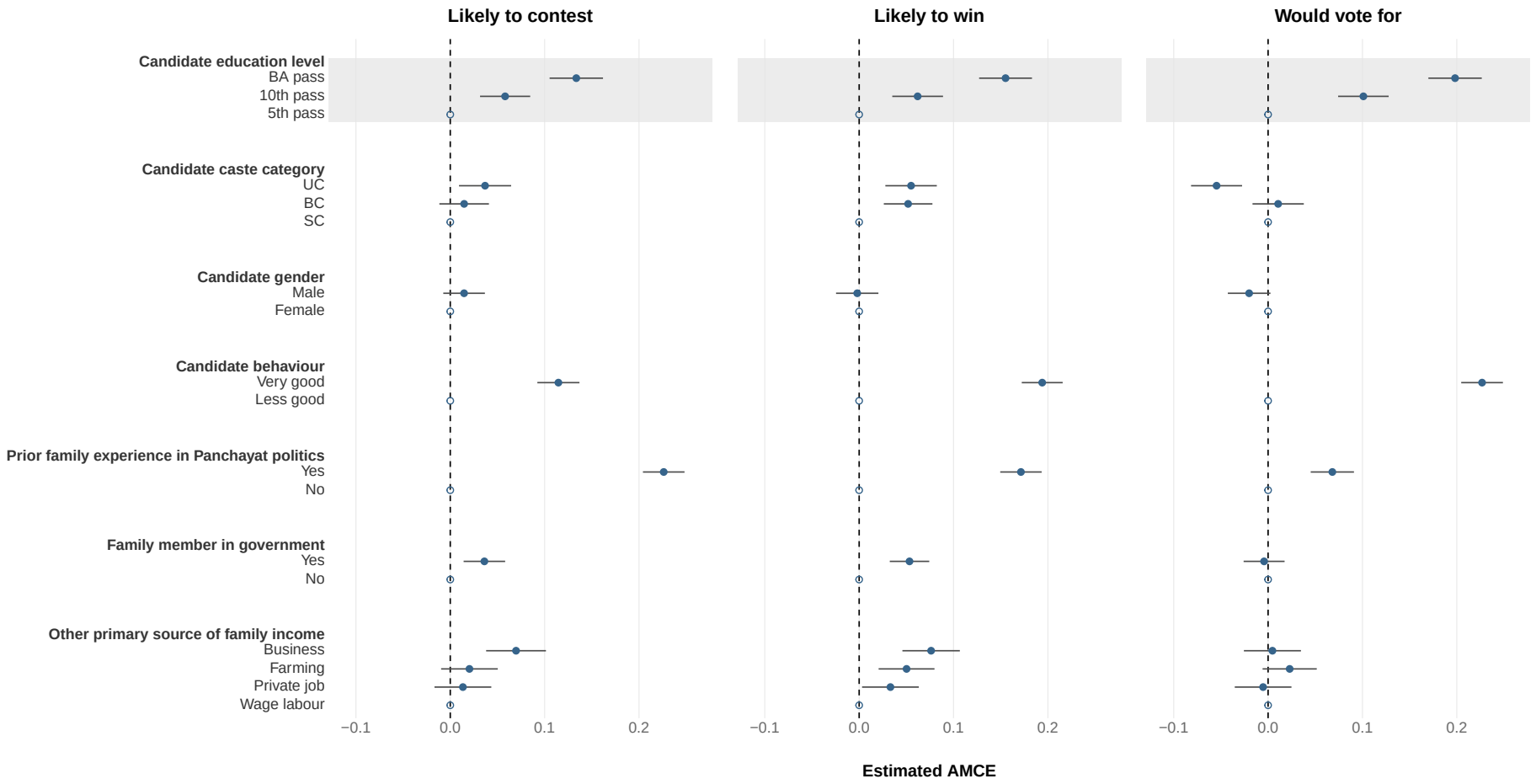


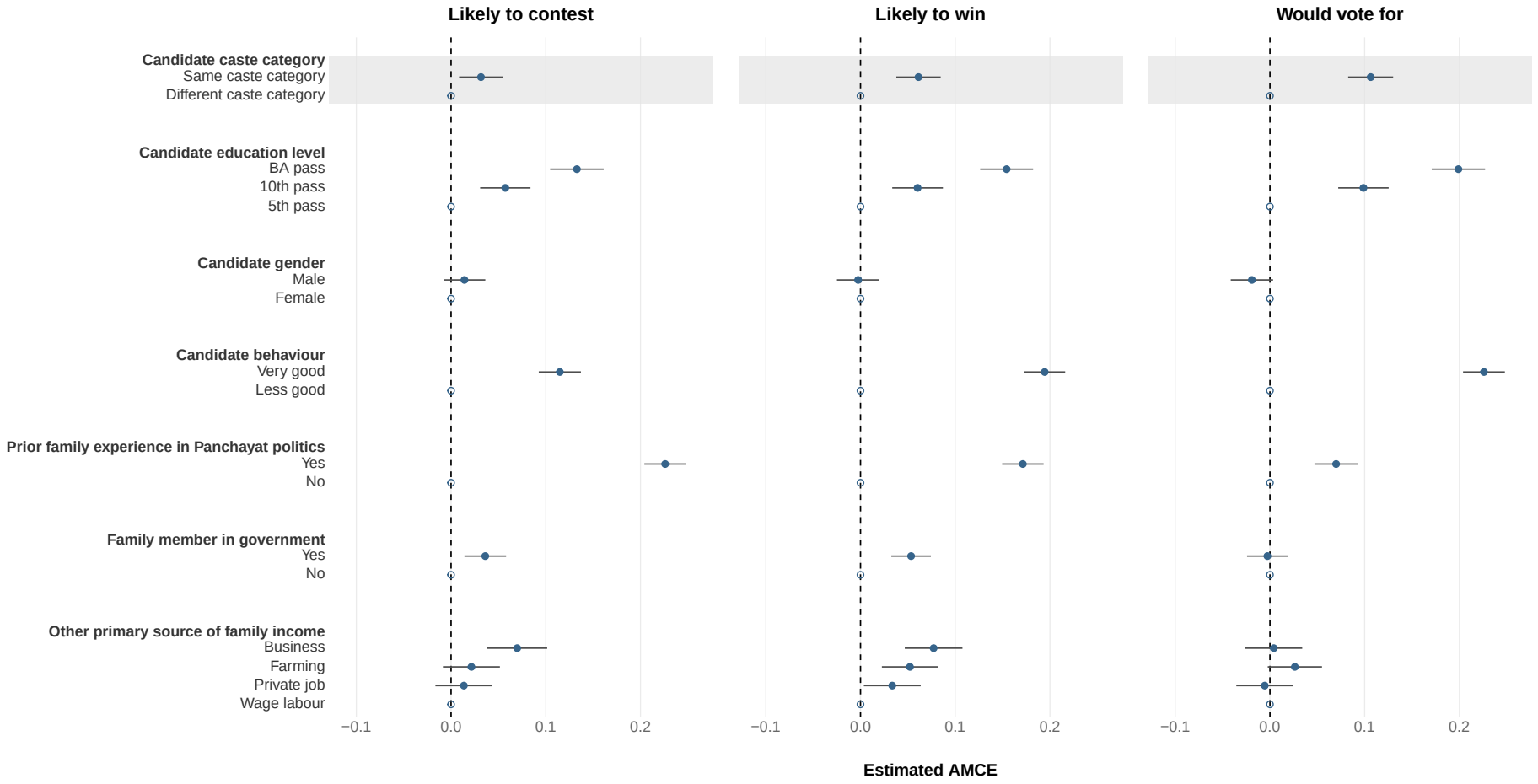
Figure 5: Qualities based on which citizens report evaluating Panchayat candidates: Percent of cases

Figure 6: Voter perceptions of which candidate types are likely to contest and win, and their own vote choice



Notes: Bars represent 95% confidence intervals. Estimates are Average Marginal Component Effects (AMCEs) from a forced−choice conjoint experiment (N = 1,105 respondents; effective N = 7,600 profile evaluations). Source: Bihar Citizen Survey, February 2024.

Figure 7: Voter perceptions of which candidate types are likely to contest and win, and their own vote choice (Candidate caste category recoded as per respondent caste category)



Notes: Bars represent 95% confidence intervals. Estimates are Average Marginal Component Effects (AMCEs) from a forced−choice conjoint experiment (N = 1,105 respondents; effective N = 7,600 profile evaluations). Source: Bihar Citizen Survey, February 2024.

APPENDIX

A Data and Construction

Table A1: Construction of GP-level Variables from SECC

Variable	**Definition and Construction**
GP Asset Index	We construct a household-level asset index using six binary indicators: ownership of land; quality of housing (concrete roof and burnt brick or concrete walls); having four or more rooms; ownership of a phone; and ownership of a vehicle. Each asset is owned by at least 10% of households. The GP-level asset index is the average of this index across households in the GP.
Ethnic Fractionalisation	Measured as one minus the Herfindahl index of last names of household heads within a GP: $\text{FRACT}_g = 1 - \sum_{n=1}^{N} s_{ng}^2$, where s_{ng} is the share of individuals with last name n in GP g, and N is the number of distinct last names.
Ethnic Polarisation	Constructed following **?**: $\text{POLAR}_g = 1 - \sum_{n=1}^{N} (0.5 - s_{ng})^2 s_{ng} / 0.25$, where s_{ng} is the share of individuals with last name n in GP g.
Educational Mobility	For each GP, we measure educational mobility as the share of children aged 14–18 who completed primary education, restricting to households where parents have below primary education.
Land Gini	Measured as the Gini coefficient of total land owned across households within a GP.
Remoteness	Measured using census-derived data on the average distance of villages in the GP from the district headquarters.

Table A2: Summary of Datasets

Dataset	Description	Analysis
Socioeconomic Caste Census (SECC), 2011–12	A nationally administered census of households conducted by the Government of India. We use the rural Bihar sample, covering approximately 97 million individuals across 19 million households. The data provide rich individual- and household-level socioeconomic information and, crucially, allow us to link individuals to parents and other household members.	All analyses. Primary data source for competence and representativeness comparisons, GP-level covariates, and candidate–citizen matching.
Bihar State Election Commission Data, 2016 & 2021	Administrative records on all candidates contesting GP head and ward member elections. The data cover 909,570 candidates across two election cycles (94,645 GP head and 268,591 ward candidates in 2016; 63,284 and 481,993 in 2021) and include detailed candidate characteristics, including years of education, age, gender, caste, father's name, votes received, and seat reservation status.	All analyses. Provides candidate-level data underlying all selection, close-election RD, and reservation RD results.
Gram Panchayat Development Plan (GPDP), 2024-25	Annual Gram Panchayat Development Plans listing all proposed projects, with estimated costs and funding sources, based on ward- and Gram Sabha-level consultations. Scraped from eGramSwaraj (Ministry of Panchayati Raj), the data cover all GPs and provide project-level information on planned spending (e.g., roads, drains, piped water, solar streetlights). We combine these data with a survey of citizen preferences to measure the alignment between leader choices and citizen demand.	Effects of leader competence on policy alignment with citizen preferences (close-election RD); used jointly with citizen survey.
Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) MIS, 2016–2021	Administrative data scraped from the Management Information System (MIS) reporting employment generated at the GP level under MGNREGA, which guarantees up to 100 days of wage employment per rural household per year. The data include total person-days and person-days by social group, including Scheduled Castes. We focus on outcomes between 2016 and 2021, corresponding to the tenure of GP heads elected in 2016.	Effects of leader competence on distributive policy outcomes (close-election RD).
Mission Antyodaya, 2019–20	Village-level administrative data from the Ministry of Rural Development, Government of India (downloaded from SHRUG platform by Asher et al. (2021)). The survey covers the near-universe of Bihar's approximately 37,262 revenue villages and provides detailed information on the availability of core public goods and services (e.g., infrastructure, basic amenities).	Effects of leader competence on skill-intensive policy outcomes (close-election RD).
Ward Member Survey (Bamezai et al. (2024), Baseline)	Baseline in-person survey conducted as part of a randomized evaluation of politician peer networks in Bihar. Over 7,700 elected ward members were surveyed in person across blocks statewide. The survey measures politicians' knowledge of rules and procedures for implementing government schemes (across domains including housing, infrastructure, and service delivery), their professional networks with higher-level officials (Block Development Officers, Block Panchayati Raj officials, GP heads), party support, and campaign characteristics.	Effects of education on scheme-implementation knowledge and bureaucratic networks (close-election RD on ward members); heterogeneous political selection by party affiliation.
Citizen Survey (Bamezai et al. (2024), Endline)	Endline phone survey conducted as part of a randomized evaluation of politician peer networks in Bihar. Approximately 10,000 citizens across 10 districts were asked to name up to two development schemes on which they would like their local government to increase spending. Responses are aggregated to construct GP×scheme-level preference shares, separately for Scheduled Caste and non-SC populations.	Effects of leader competence on policy alignment with citizen preferences (close-election RD); used jointly with GPDP data.

Notes: SECC and Bihar SEC linked via string-matching (Section 3.1, Appendix B; match rate ≈60.2% for unreserved GP head candidates). Education standardized within GP. Reservation strategies exploit population-based algorithms in Bihar SEC data (Appendix C). See Table A1 for GP-level variable definitions; Appendix F for survey details.

B Additional Results / Extensions

Table A3: Political selection and village characteristics for male GP Heads

	Years of Schooling					
	Wealth Score (1)	Ethnic FRACT. (2)	Ethnic POLAR. (3)	Edu. Mobility (4)	Gini Index Tot. Land (5)	Distance District HQ (6)
Panel A: Competence - Comparison between Citizens and Candidates						
Candidate	1.088***	1.090***	1.091***	1.244***	0.842***	1.105***
	(0.011)	(0.010)	(0.011)	(0.039)	(0.144)	(0.023)
Candidate X Panchayat Feature	0.009	0.034**	-0.033***	-0.327***	0.284*	-0.001
	(0.011)	(0.014)	(0.012)	(0.072)	(0.167)	(0.001)
Observations	97,349,645	97,349,645	97,349,645	97,349,238	97,349,645	97,284,398
Panel B: Representativeness - Comparison between Citizens' and Candidates' Fathers						
Candidate's Father	0.527***	0.532***	0.533***	0.696***	-0.062	0.571***
	(0.011)	(0.011)	(0.011)	(0.038)	(0.131)	(0.022)
Cand. Father X Panchayat Feature	0.017	0.066***	-0.052***	-0.353***	0.682***	-0.001**
	(0.011)	(0.012)	(0.012)	(0.071)	(0.151)	(0.001)
Observations	97,349,645	97,349,645	97,349,645	97,349,238	97,349,645	97,284,398

Notes: This table reports heterogeneity in political selection by village-level characteristics, using data from the 2016 and 2021 Panchayat elections. The dependent variable is years of schooling, standardized within Gram Panchayats. Panel A reports differences between candidates and citizens, while Panel B reports differences between candidates' fathers and citizens. Columns (1)–(6) interact an indicator for ever running for office with the following Gram Panchayat–level characteristics: average household wealth levels (1), ethnic fractionalization (2), ethnic polarization (3), educational mobility (4), land inequality (5), and distance from the Panchayat to the district headquarters (6). Standard errors are clustered at the Gram Panchayat level and reported in parentheses. $^{*}\, p < 0.10$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table A4: Political Selection Candidate's Social Groups

	Avg. Years of Schooling			
	Age Cohort (1)	Caste (2)	Surname (3)	GenderXAgeXSurname (4)
Candidate	0.071***	0.022***	0.108***	0.366***
	(0.001)	(0.000)	(0.001)	(0.002)
Observations	97,349,674	97,349,674	97,349,674	97,349,674

This table reports the extent to which politicians come from more educated social groups relative to the average citizen. The outcome variable in each column is the average years of schooling of the social group to which each individual belongs: age cohort (1), caste group (2), surname group (3), and gender × age cohort × surname group (4). The coefficient on the candidate indicator captures how much more educated the average politician's social group is relative to that of the average citizen. All specifications include GP-level fixed effects. Standard errors are clustered at the GP level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Political competence and development outcomes

	NREGA Work-Days		
	Non SC Total (1)	SC Total (2)	SC (%) (3)
Winner	-529.046	-616.828*	-0.022
	(1070.763)	(326.728)	(0.015)
Dep. Mean	13373.02	2457.75	0.17
Bandwidth	0.16	0.13	0.13
Observations	1,828	1,553	1,554

Notes: This table reports the effects of electing a more educated GP Head on NREGA workday outcomes. Estimates are obtained from a close-election regression discontinuity design using the 2016 Panchayat elections. The treatment variable (*Above Primary*) is a binary indicator equal to one if the winning candidate has above-primary education in the Mukhiya election. Columns (1)-(3) report effects on annual average NREGA workdays generated, decomposed into non-Scheduled Caste total workdays (1), Scheduled Caste total workdays (2), and the Scheduled Caste share of total workdays (3). The sample is restricted to Gram Panchayats in which the top two candidates' education levels straddle the above-primary threshold, with one candidate above and one below primary education. Standard errors are clustered at the Gram Panchayat level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Political selection, level of office and electoral performance

	Years of Education			
	All Candidates	Winner	Runner	Others
	(1)	(2)	(3)	(4)
Panel A: GP Head Candidates				
Candidates	1.088***	1.417***	1.401***	1.099***
	(0.011)	(0.026)	(0.026)	(0.011)
Observations	97,349,648	97,349,648	97,349,648	97,349,648
Panel B: Ward Candidates				
Candidates	0.592***	0.690***	0.610***	0.627***
	(0.004)	(0.007)	(0.007)	(0.005)
Citizen Mean	3.457	3.460	3.460	3.457
Citizen SD	3.981	3.983	3.983	3.981
Observations	97,349,648	97,349,648	97,349,648	97,349,648

The table describes the political selection in Bihar for electoral candidates competing for unreserved seats. Panel A shows the results for GP Head candidates, while panel B shows the results for Ward Candidates. The dependent variables are years of education, normalized by substracting the population mean and divided by standard deviation respectively within each GP. Column (1) presents the difference in years of education between all electoral candidates and citizens. Column (2)-(4) further breaks down the analysis by winner, runner and all other candidates respectively. Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Effects of SC Reservation on Selection for GP Head Candidates (including female)

	Yrs of Schooling	Representativeness			
	Candidate Edu Yrs	Gov. Job	Income Tax	Wealth Score	Father Edu Yrs
	(1)	(2)	(3)	(4)	(5)
Panel A: Comparison to all citizen					
SC RSVN	-0.630***	-0.148***	-0.240***	-0.650***	-0.738***
	(0.039)	(0.034)	(0.038)	(0.033)	(0.051)
Dep. Mean	0.67	0.24	0.19	0.51	0.44
Bandwidth	400.686	613.649	416.315	489.966	161.306
Observations	37,810	51,338	38,785	43,449	6,478
Panel B: Comparison to citizen of own group					
SC RSVN	-0.121***	-0.002	-0.092**	-0.165***	-0.350***
	(0.041)	(0.037)	(0.039)	(0.034)	(0.053)
Dep. Mean	0.62	0.23	0.17	0.45	0.38
Bandwidth	406.693	542.187	486.262	486.586	162.027
Observations	38,124	46,552	41,667	43,203	6,508

This table reports the effects of female reservation on political selection outcomes for GP Head candidates and winners in the 2016 and 2021 Panchayat elections, extending the main analysis to include female-reserved seats. Estimates are obtained from a regression discontinuity design exploiting the population threshold that determines female reservation status. Columns (1)–(3) report competence measures based on standardized years of education benchmarked against: all citizens (1), citizens of the same group (2), and all citizens using husband's education in place of the candidate's own (3). Columns (4)–(6) report eliteness measures: an indicator for whether any household member holds a salaried government job (4); an indicator for whether any household member pays income or professional taxes (5); and a PCA-based wealth index constructed from land ownership, housing characteristics, and asset ownership (6). Panel A reports results for candidates; Panel B reports results for winners. All specifications include GP-level fixed effects and control for election year and GP-level covariates. Standard errors are clustered at the GP level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Effects of Reservation on Political Selection for GP Head Winners

	Yrs of Schooling	Representativeness			
	Candidate Edu Yrs	Gov. Job	Income Tax	Wealth Score	Father Edu Yrs
	(1)	(2)	(3)	(4)	(5)
Panel A: Comparison to all citizen					
SC RSVN	-0.881***	-0.527***	-0.812***	-1.040***	-1.332***
	(0.121)	(0.145)	(0.137)	(0.106)	(0.119)
Dep. Mean	1.29	0.38	0.40	0.92	0.64
Bandwidth	172.877	163.254	261.318	215.440	197.064
Observations	1,094	1,054	1,548	1,339	722
Panel B: Comparison to citizen of own group					
SC RSVN	-0.375***	-0.465***	-0.776***	-0.542***	-0.918***
	(0.125)	(0.155)	(0.161)	(0.135)	(0.134)
Dep. Mean	1.21	0.36	0.38	0.85	0.56
Bandwidth	191.133	173.623	238.573	176.391	201.106
Observations	1,186	1,089	1,373	1,103	742

This table reports the effects of Scheduled Caste (SC) reservation on political selection outcomes for GP Head winners in the 2016 and 2021 Panchayat elections. Estimates are obtained from a regression discontinuity design around the population threshold for SC reservation. Columns (1)–(5) report results for the following candidate-level outcomes: standardized years of education (1); an indicator for whether any household member holds a salaried government job (2); an indicator for whether any household member pays income or professional taxes (3); a PCA-based wealth index constructed using land ownership, housing characteristics, and asset ownership (4); and the candidate's father's years of education (5). Panel A reports outcomes standardized relative to all citizens within the GP, while Panel B reports outcomes standardized relative to individuals from the same caste within the GP. The sample is restricted to seats unreserved for women. All specifications include block fixed effects and election-year controls. Standard errors are clustered at the GP level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Effects of Female Reservation on Political Selection By Caste

	Yrs of Schooling		Representativeness			
	Candidate Edu Yrs	Husband Edu Yrs	Gov. Job	Income Tax	Wealth Score	Highest HH Edu Yrs
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Unreserved						
Female RSVN	-0.794***	-0.154***	-0.062	-0.085*	-0.094**	-0.117***
	(0.038)	(0.052)	(0.046)	(0.045)	(0.038)	(0.033)
Bandwidth	1010.17	921.52	918.68	980.92	1039.18	969.56
Observations	19,001	14,367	18,239	18,739	19,304	30,264
Panel B: SC Reserved						
Female RSVN	-0.738***	-0.101	-0.121**	-0.062	-0.115**	-0.174***
	(0.050)	(0.078)	(0.051)	(0.042)	(0.047)	(0.043)
Bandwidth	157.45	106.49	144.01	161.70	156.72	149.62
Observations	8,072	4,890	7,621	8,206	8,064	11,936
Panel C: OBC Reserved						
Female RSVN	-0.833***	-0.162**	-0.028	-0.029	-0.038	0.011
	(0.047)	(0.064)	(0.047)	(0.041)	(0.051)	(0.043)
Bandwidth	304.09	150.04	308.62	405.56	292.42	266.20
Observations	5,987	2,695	6,046	7,603	5,774	8,846

This table shows the effects of female reservation on candidate competence, in terms of education background. Only 2016 candidates are included in the sample. We collapse the data into GP level and respectively calculate the average (standardized) education level of candidates within the same GP. In column (1)-(4), we are comparing the average education level of candidates from reserved GPs and those from unreserved GPs, while in column (5)-(8), only GPs with at least 1 female candidate running is included, and we are comparing the average education level of female candidates from reserved GPs and female candidates from unreserved GPs. The running variable 'Population Diff.' represents the difference between the Gram Panchayat population and the cutoff population size (the population cutoff line is calculated by taking the average population size between the last reserved Gram Panchayat and the first unreserved Gram Panchayat). Controls include fixed effects on election year. Standard errors are clustered at Block level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Long-Run Effects of SC Reservation on Political Selection: Non-SC Candidates

	Yrs of Schooling	Representativeness			
	Candidate Edu Yrs	Gov. Job	Income Tax	Wealth Score	Father Edu Yrs
	(1)	(2)	(3)	(4)	(5)
Panel A: Candidates					
SC RSVN	-0.058	0.302***	0.199***	-0.005	-0.226***
	(0.080)	(0.066)	(0.062)	(0.058)	(0.081)
Dep. Mean	1.06	0.29	0.22	0.61	0.49
Bandwidth	240.690	208.518	261.312	261.590	213.714
Observations	7,295	6,546	7,983	7,993	4,249
Panel B: Winners					
SC RSVN	0.150	-0.298**	-0.337**	-0.257**	-0.494***
	(0.132)	(0.121)	(0.152)	(0.124)	(0.119)
Dep. Mean	1.30	0.38	0.42	0.94	0.68
Bandwidth	368.587	441.285	445.505	350.624	374.042
Observations	1,015	1,240	1,253	931	696

This table shows the long-run effects of SC reservation on political selection for non-SC Mukhiya candidates. SC candidates are excluded from the sample. Both 2016 and 2021 candidates are included in the sample. Gram Panchayats reserved for SC in the 2016 cycle are dropped, so the comparison is between GPs reserved for SC in the 2006 cycle and GPs never reserved; GPs reserved for women in 2016 are also excluded. Panel A includes all candidates and Panel B restricts the sample to winners. The dependent variables are standardized by subtracting the population mean and dividing by the standard deviation within each GP. Father's education is available for matched male candidates only. The running variable 'Population Diff.' represents the difference between the Gram Panchayat SC population and the cutoff SC population within the block; 2006 SC reservation status is instrumented with an indicator for crossing the cutoff. Controls include fixed effects on block and election year, and 2016 female reservation status. Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C RD Design Validation

Table A11: SC Reservation Balance Test

	Control Mean	RD Estimate	Obs	p-value (FDR q-value)
Paved Road (Census 2001)	1.78	-0.20	4,463	0.16 (1)
Bank Facilities (Census 2001)	0.34	0.03	5,410	0.56 (1)
Bus facilities (Census 2001)	0.54	0.04	5,015	0.65 (1)
Child Welfare Centre (Census 2001)	0.05	0.28	6,111	0 (0)
Distance District HQ	31.88	0.21	4,647	0.76 (1)
Total Educational Facilities (Census 2001)	3.46	-0.06	5,318	0.68 (1)
Education Mobility	0.46	0.01	4,727	0.47 (1)
Ethnic Fractionalization	0.89	0.00	4,774	0.85 (1)
Ethnic Polarization	0.34	-0.01	4,340	0.16 (1)
Medical Facility (Census 2001)	0.88	0.05	5,458	0.58 (1)
Middle School Facility (Census 2001)	1.21	-0.05	4,632	0.62 (1)
Distance Nearest Statutory Town	20.65	1.08	4,906	0.05 (0.6)
Wealth Score	0.00	0.02	4,613	0.82 (1)
Phone Lines (Census 2001)	0.54	0.07	4,831	0.3 (1)
Primary Health Sub Centres (Census 2001)	0.43	0.08	4,510	0.18 (1)
Post Office (Census 2001)	1.03	-0.05	4,590	0.52 (1)
Power Supply (Census 2001)	1.67	0.05	4,133	0.74 (1)
Tank (Census 2001)	0.52	0.05	5,914	0.6 (1)
Tap Water (Census 2001)	0.15	-0.11	5,037	0.01 (0.12)
Mean Village Income (Census 2001)	434,692.63	41,417.58	4,225	0.84 (1)
Mean Village Income (Census 2001)	457,166.91	-2,114,289.00	2,618	0.27 (1)
Well (Census 2001)	0.82	0.05	4,479	0.07 (0.67)
Gini Index (Total Land)	0.87	0.00	3,942	0.84 (1)
Gini Index (Irrigated Land)	0.88	0.00	4,266	0.81 (1)

This table reports balance tests for the SC reservation regression discontinuity design. For each baseline GP-level covariate, we estimate the RD coefficient using the SC population recentered around block-specific reservation thresholds as the running variable, with SC reservation status as the treatment, and block-level fixed effects as covariates. The mean of each covariate is reported for unreserved GPs within the optimal bandwidth. Standard errors are clustered at the GP level. The final column reports sharpened q-values following the Benjamini, Krieger, and Yekutieli (2006) procedure to account for multiple hypothesis testing.

Table A12: SC Reservation Balance Test (Female Unreserved)

	Control Mean	RD Estimate	Obs	p-value (FDR q-value)
Paved Road (Census 2001)	1.73	0.04	2,410	0.78 (1)
Bank Facilities (Census 2001)	0.32	0.05	1,936	0.51 (0.78)
Bus facilities (Census 2001)	0.51	0.18	2,518	0.07 (0.21)
Child Welfare Centre (Census 2001)	0.05	0.06	2,509	0.12 (0.28)
Distance District HQ	32.01	0.57	3,101	0.42 (0.71)
Total Educational Facilities (Census 2001)	3.42	0.02	2,215	0.89 (1)
Education Mobility	0.46	0.00	2,342	0.81 (1)
Ethnic Fractionalization	0.89	0.01	2,560	0.02 (0.09)
Ethnic Polarization	0.34	-0.03	2,892	0 (0.01)
Medical Facility (Census 2001)	0.88	0.05	2,176	0.62 (1)
Middle School Facility (Census 2001)	1.19	0.04	2,441	0.69 (1)
Distance Nearest Statutory Town	20.76	0.92	2,187	0.14 (0.28)
Wealth Score	-0.02	-0.03	2,283	0.7 (1)
Phone Lines (Census 2001)	0.52	0.19	2,401	0.02 (0.09)
Primary Health Sub Centres (Census 2001)	0.44	0.06	2,016	0.42 (0.71)
Post Office (Census 2001)	1.01	0.00	2,174	1 (1)
Power Supply (Census 2001)	1.63	-0.03	2,842	0.86 (1)
Tank (Census 2001)	0.53	-0.16	2,355	0.11 (0.28)
Tap Water (Census 2001)	0.16	-0.13	2,888	0.01 (0.06)
Mean Village Income (Census 2001)	369,680.78	1,218,851.90	1,950	0 (0)
Mean Village Income (Census 2001)	570,399.00	224,651,072.00	1,212	0.45 (0.71)
Well (Census 2001)	0.82	-0.04	2,301	0.74 (1)
Gini Index (Total Land)	0.87	0.01	3,086	0.01 (0.07)
Gini Index (Irrigated Land)	0.88	0.01	2,266	0.08 (0.21)

Table A13: Gender Reservation Balance Test

	Control Mean	RD Estimate	Obs	p-value (FDR q-value)
Paved Road (Census 2001)	1.76	-0.05	5,325	0.55 (1)
Bank Facilities (Census 2001)	0.32	-0.04	5,354	0.27 (1)
Bus facilities (Census 2001)	0.53	-0.12	5,351	0.03 (0.24)
Child Welfare Centre (Census 2001)	0.05	0.03	5,349	0.22 (1)
Distance District HQ	31.89	0.20	5,248	0.64 (1)
Total Educational Facilities (Census 2001)	3.56	-0.11	5,521	0.24 (1)
Education Mobility	0.46	0.00	5,205	0.58 (1)
Ethnic Fractionalization	0.89	0.00	5,263	0.68 (1)
Ethnic Polarization	0.34	0.00	5,304	0.82 (1)
Medical Facility (Census 2001)	0.90	-0.02	5,298	0.72 (1)
Middle School Facility (Census 2001)	1.16	0.04	5,316	0.47 (1)
Distance Nearest Statutory Town	20.64	-0.10	5,253	0.79 (1)
Wealth Score	-0.06	0.06	5,200	0.2 (1)
Phone Lines (Census 2001)	0.51	0.00	5,348	0.93 (1)
Primary Health Sub Centres (Census 2001)	0.44	0.01	5,351	0.81 (1)
Post Office (Census 2001)	0.98	-0.04	5,337	0.4 (1)
Power Supply (Census 2001)	1.70	-0.26	5,305	0 (0.06)
Tank (Census 2001)	0.55	0.02	5,351	0.76 (1)
Tap Water (Census 2001)	0.16	0.00	5,495	0.88 (1)
Mean Village Income (Census 2001)	400,185.47	-17,207.46	5,235	0.86 (1)
Mean Village Income (Census 2001)	490,421.59	32,917.38	5,380	0.8 (1)
Well (Census 2001)	0.83	-0.06	5,488	0 (0)
Gini Index (Total Land)	0.87	0.00	5,204	0.27 (1)
Gini Index (Irrigated Land)	0.88	0.00	5,226	0.21 (1)

This table reports balance tests for the female reservation regression discontinuity design. For each baseline GP-level covariate, we estimate the RD coefficient using the female population recentered around block-specific reservation thresholds as the running variable, with female reservation status as the treatment, and controlling for caste reservation status and block-level fixed effects. The mean of each covariate is reported for non-female-reserved GPs within the optimal bandwidth. Standard errors are clustered at the GP level. The final column reports sharpened q-values following the Benjamini, Krieger, and Yekutieli (2006) procedure to account for multiple hypothesis testing.

Table A14: Mukhiya Close-Election Balance Test

	Control Mean	RD Estimate	Obs	p-value (FDR q-value)
Candidate Father Edu	5.30	-0.17	354	0.91 (1)
Paved Road (Census 2001)	1.78	0.08	2,560	0.64 (1)
Bank Facilities (Census 2001)	0.34	0.06	2,544	0.35 (1)
Bus facilities (Census 2001)	0.59	0.08	2,440	0.47 (1)
Child Welfare Centre (Census 2001)	0.06	-0.05	2,394	0.29 (1)
Distance District HQ	32.12	-0.04	2,414	0.98 (1)
Total Educational Facilities (Census 2001)	3.45	-0.14	2,153	0.59 (1)
Education Mobility	0.46	0.01	2,449	0.68 (1)
Ethnic Fractionalization	0.89	0.00	2,092	0.74 (1)
Ethnic Polarization	0.34	0.00	2,037	0.93 (1)
Medical Facility (Census 2001)	0.88	0.22	2,598	0.07 (1)
Middle School Facility (Census 2001)	1.24	-0.07	2,374	0.55 (1)
Distance Nearest Statutory Town	20.40	-0.74	2,180	0.63 (1)
Wealth Score	-0.04	0.09	2,449	0.44 (1)
Phone Lines (Census 2001)	0.54	-0.05	1,910	0.64 (1)
Primary Health Sub Centres (Census 2001)	0.44	0.04	2,240	0.62 (1)
Post Office (Census 2001)	1.04	-0.25	2,167	0.02 (1)
Power Supply (Census 2001)	1.66	-0.38	2,177	0.17 (1)
Tank (Census 2001)	0.52	-0.02	2,227	0.89 (1)
Tap Water (Census 2001)	0.16	-0.11	2,246	0.13 (1)
Mean Village Income (Census 2001)	288,080.00	199,424.03	2,123	0.49 (1)
Mean Village Income (Census 2001)	413,981.06	290,817.31	2,445	0.5 (1)
Well (Census 2001)	0.82	0.06	2,120	0.23 (1)
Gini Index (Total Land)	0.87	0.00	2,476	0.77 (1)
Gini Index (Irrigated Land)	0.88	0.00	2,413	0.82 (1)

This table reports balance tests for the close-election regression discontinuity design. For each baseline GP-level covariate, we estimate the RD coefficient using the margin of victory of the more-educated candidate over the less-educated candidate as the running variable. The mean of each covariate is reported for GPs where the less-educated candidate won within the optimal bandwidth. Standard errors are clustered at the GP level. The final column reports sharpened q-values following the Benjamini, Krieger, and Yekutieli (2006) procedure to account for multiple hypothesis testing.

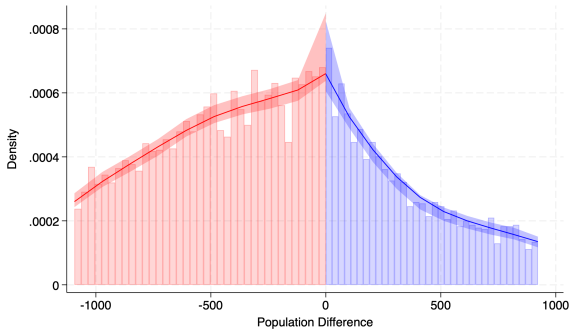
Table A15: Balance Test: Diff in Candidate Characteristics in Close Election - Ward

	N	Dep. Mean	Winner/se
Scheduled Caste	14944	.22	.0156**
			(.00681)
Scheduled Tribe	14944	.01	.00295
			(.00193)
Other Caste	14944	.77	-.0186***
			(.00696)
Woman	14944	.47	-.045***
			(.0082)
Age	14943	38.86	-1.03***
			(.173)
Gov. Job	13038	.10	-.0173
			(.0211)
Salaried Job	13034	.09	-.0122
			(.0206)
Income Tax	13026	.07	-.0285
			(.0203)
Income High	13033	.10	-.0226
			(.0205)
Wealth Score	13037	.27	-.00302
			(.019)
Ag. Equipment	13025	.08	.00598
			(.0221)
Tot. Land	13044	.10	.0178
			(.0191)
Father Edu Yrs	3868	4.10	-.146
			(.145)

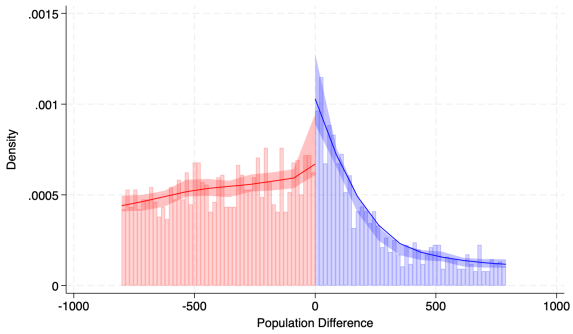
This table shows means of candidates characteristics in column (2) and differences around cutoff in column (3). Only 2021 Ward winners and runners that meet the sample restriction are included (below and above primary education cutoffs). Age is winsorized at top and bottom 1 percentile. Household-level characteristics and father education are only available for matched candidates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A1: McCrary tests for manipulation of the running variable

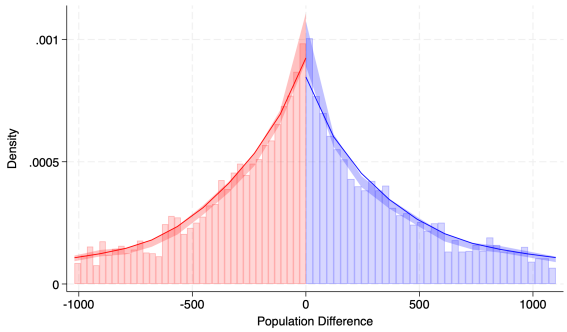
(a) SC Reservation RD (All Seats)



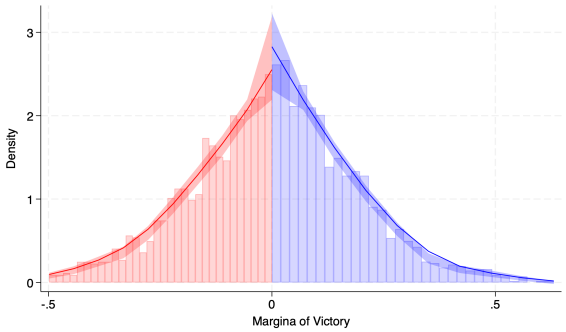
(b) SC Reservation RD (Female-Unreserved Only)



(c) Female Reservation RD



(d) Close-election Education RD



Notes: This figure presents McCrary density tests for manipulation of the running variable for each regression discontinuity design used in the paper. Panel (a) and Panel (b) use the SC population recentered around block-specific reservation thresholds as the running variable; Panel (a) includes all seats while Panel (b) restricts to female-unreserved seats only. Panel (c) uses the female population recentered around block-specific reservation thresholds. Panel (d) uses the margin of victory of the more-educated candidate minus the margin of victory of the less-educated candidate. The construction of each running variable is detailed in the main text.

D Robustness Checks

Figure A2: Main Results by Alternative Matching Rates

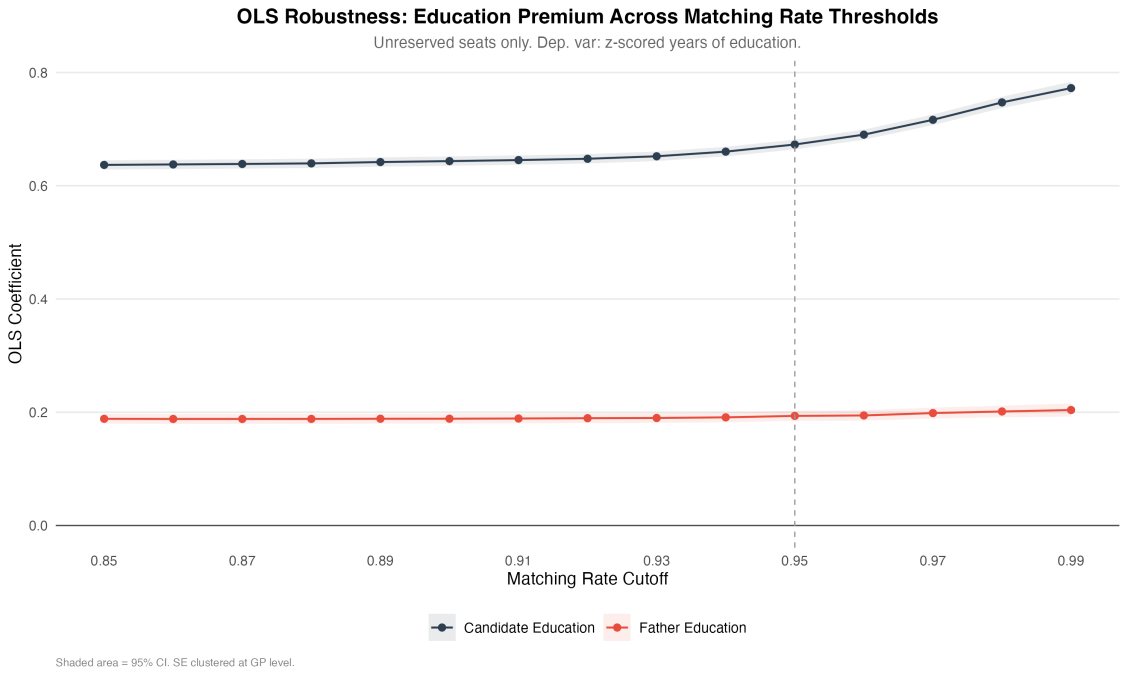


Figure A3: Close-Election Citizen Preference Outcomes with Alternative BandWidth Selection

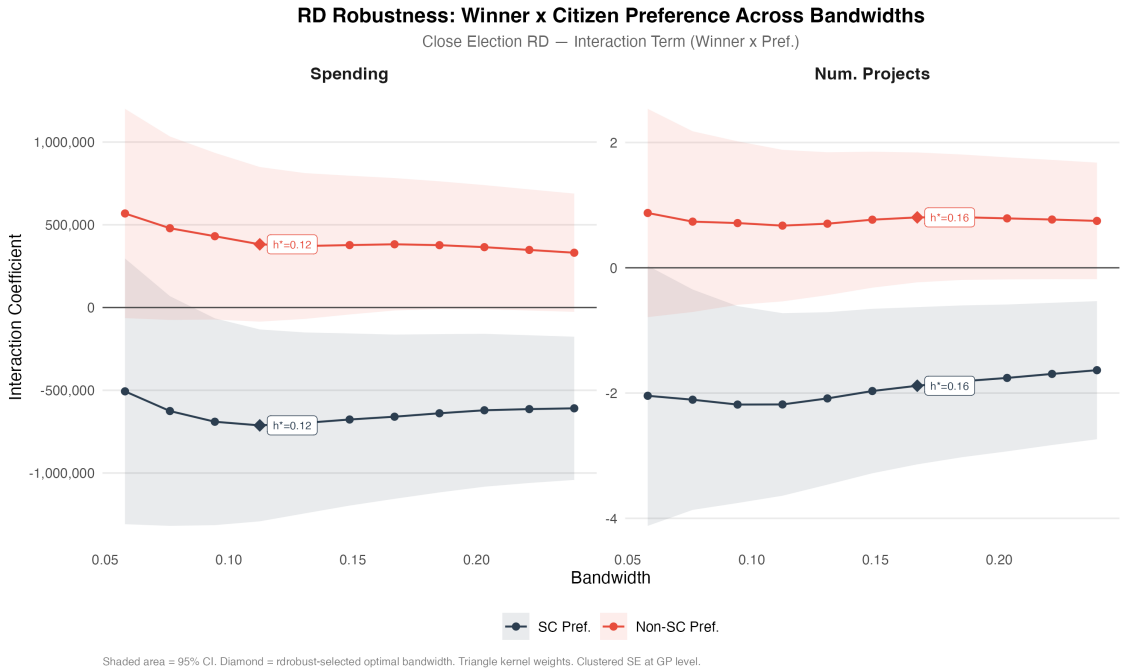
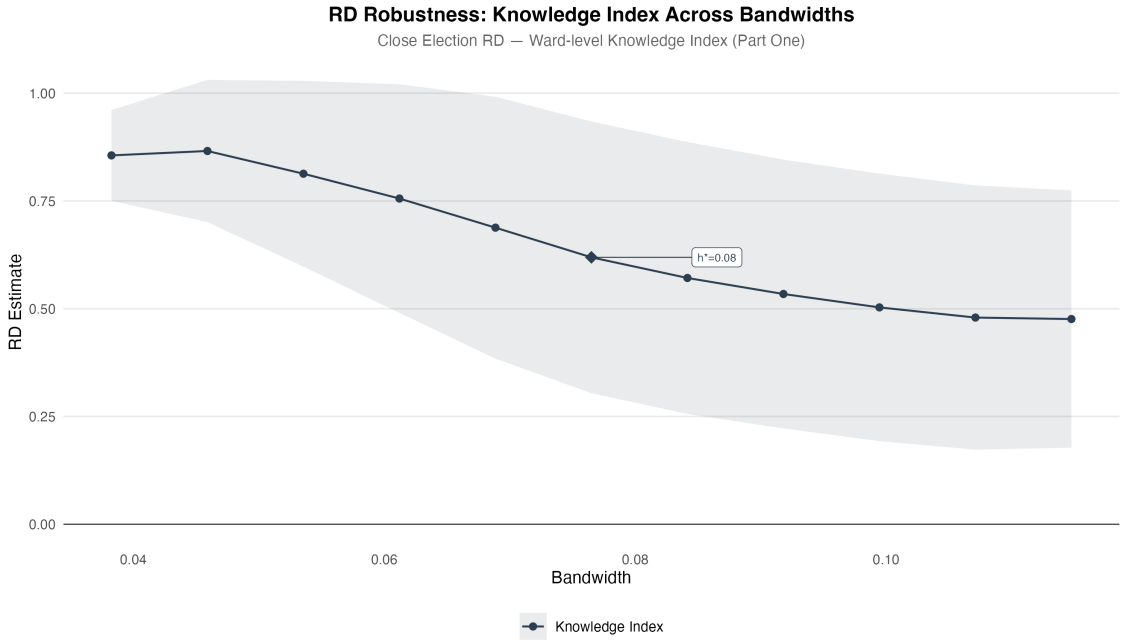


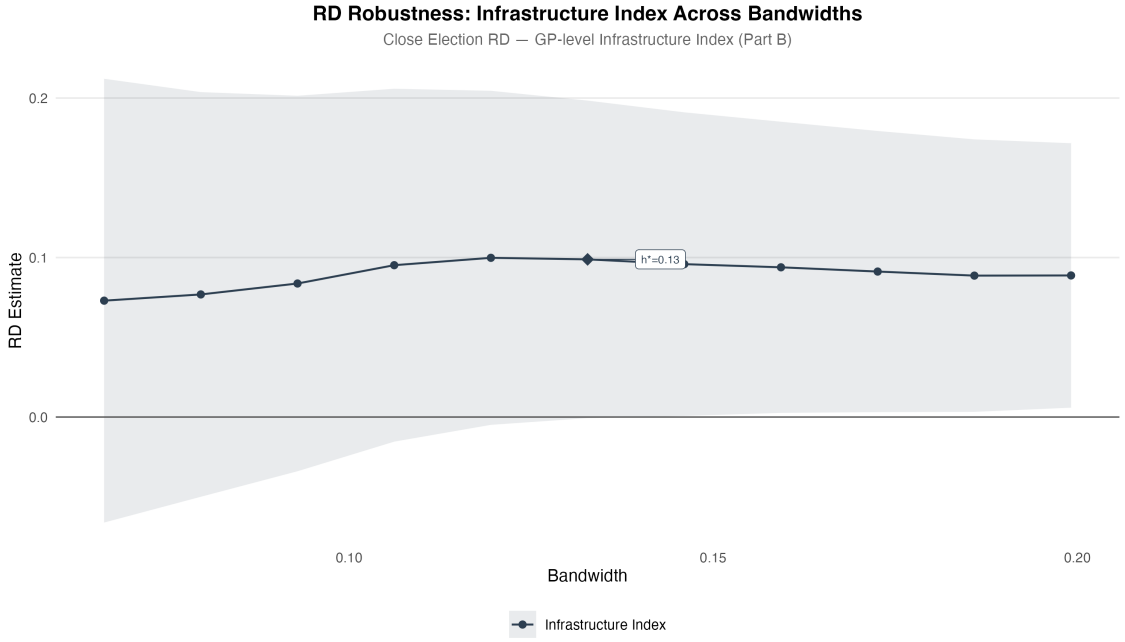
Figure A4: Close-Election Outcomes with Alternative BandWidth Selection

(a) Policy Knowledge



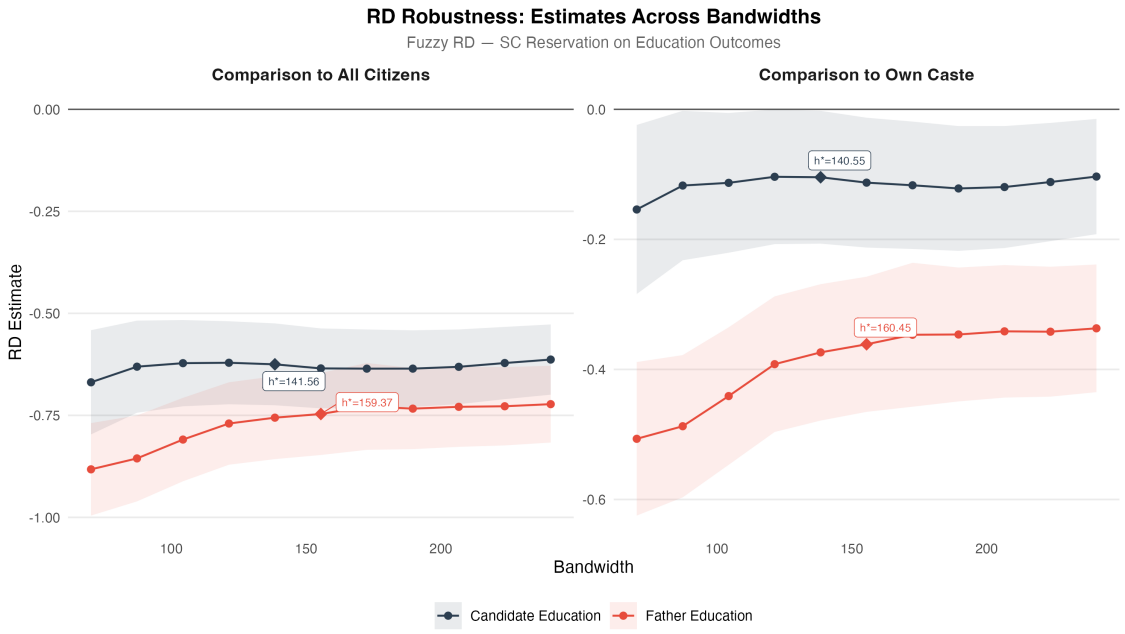
Shaded area = 95% CI. Diamond = rdrobust-selected optimal bandwidth. SE clustered at ward level.

(b) Policy Implementation



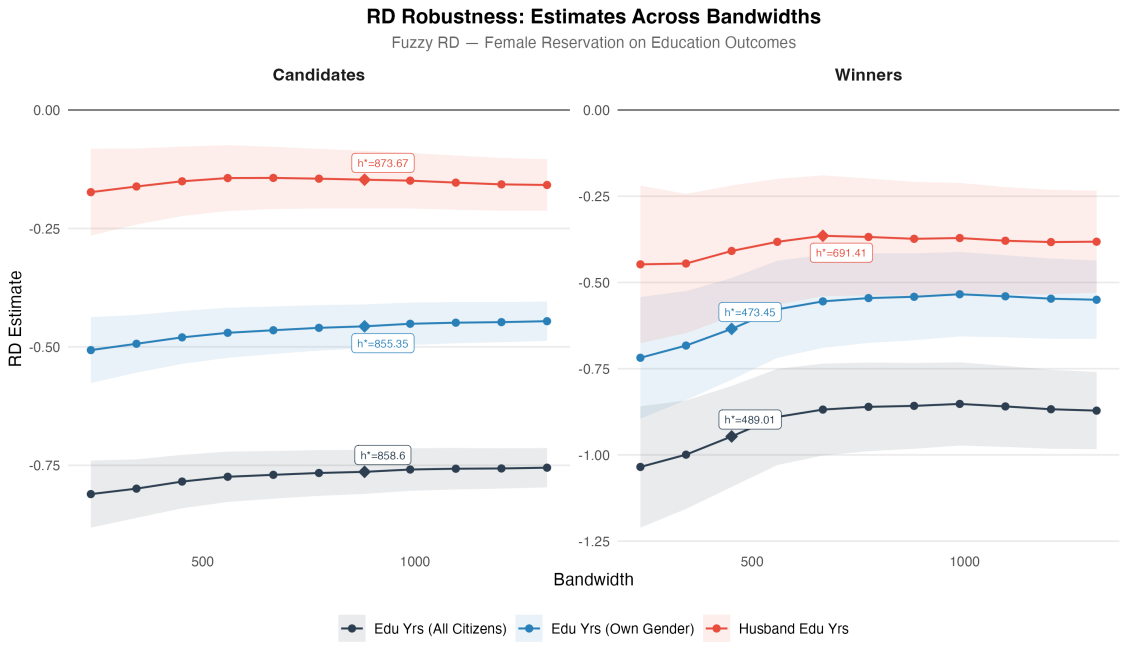
Shaded area = 95% CI. Diamond = rdrobust-selected optimal bandwidth. SE clustered at GP level.

Figure A5: SC Reservation Main Results with Alternative BandWidth Selection



Shaded area = 95% CI. Diamond = rdrobust-selected optimal bandwidth. Sample: Female-unreserved GPs.
Covariates: year, block FE, female reservation.

Figure A6: Female Reservation Main Results with Alternative BandWidth Selection



Shaded area = 95% CI. Diamond = rdrobust-selected optimal bandwidth. Covariates: caste reservation status, year, block FE.

Table A16: Alignment of Public Spending with Citizen Preferences

	Overall		SC Pref.		Non-SC Pref.	
	Spending (1)	Projects (2)	Spending (3)	Projects (4)	Spending (5)	Projects (6)
Panel A: Above and Below High-School						
Above High-School	22229.434	0.039	36272.647	0.201	6092.960	0.008
	(38554.704)	(0.133)	(59733.556)	(0.237)	(40969.687)	(0.135)
Pref.	2777468.559***	6.586***	1919622.903***	5.230***	2434939.874***	5.729***
	(184776.156)	(0.445)	(186924.233)	(0.551)	(186258.330)	(0.421)
Above High-School × Pref.	502805.276*	1.205*	-160697.658	-0.381	604732.388**	1.587**
	(292014.105)	(0.665)	(286673.425)	(0.740)	(291916.886)	(0.673)
Dep. Mean	463296.39	1.27	483295.63	1.35	462680.50	1.27
Bandwidth	0.13	0.20	0.13	0.20	0.13	0.20
Observations	5,715	7,500	3,105	3,885	5,580	7,350
Panel B: Above and Below Graduate						
Above Graduate	-10600.803	-0.076	9566.461	0.060	-22857.532	-0.082
	(47293.082)	(0.181)	(73776.396)	(0.352)	(47535.600)	(0.179)
Pref.	2671999.883***	5.936***	1771972.918***	3.847***	2404363.629***	5.327***
	(238370.008)	(0.526)	(219298.312)	(0.461)	(243664.441)	(0.528)
Above Graduate × Pref.	240642.568	1.086	-14210.429	1.196	293036.717	1.134
	(332573.582)	(0.762)	(321151.296)	(0.797)	(331312.018)	(0.744)
Dep. Mean	451632.07	1.21	466443.99	1.20	448337.54	1.19
Bandwidth	0.13	0.16	0.13	0.16	0.13	0.16
Observations	3,285	4,065	1,605	1,905	3,210	3,990

This table reports robustness checks for the heterogeneous effects of electing a more educated GP Head on government scheme outcomes by local citizen preferences, using alternative education thresholds. Panel (a) defines the more-educated candidate as having above-high-school education, while Panel (b) uses above-graduate education as the cutoff. All other specifications follow the main analysis: estimates are obtained from a close-election regression discontinuity design, controlling for scheme-specific measures of citizen preferences and their interactions with the treatment indicator. The outcome variables are scheme-level expenditures (*Spending*) and the number of projects implemented (*Projects*). Columns (1)–(2) use preferences of the overall population, Columns (3)–(4) use preferences of the Scheduled Caste population, and Columns (5)–(6) use preferences of the non–Scheduled Caste population. All specifications include scheme fixed effects. Standard errors are clustered at the running-variable level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A17: Political competence and development outcomes

	Information	Network		Poject Implementation			
	Knowledge (1)	Bureaucratic Network (2)	Party Support (3)	Infra. Aggregated (4)	Tap Water (5)	Road Index (6)	Drainage Index (7)
Panel A: Above and Below High-School							
Winner	0.323***	0.155***	0.032	0.153***	0.116	0.153	0.218**
	(0.110)	(0.033)	(0.055)	(0.057)	(0.089)	(0.094)	(0.089)
Dep. Mean	2.22	0.53	0.12	0.01	0.05	0.03	-0.04
Bandwidth	0.13	0.19	0.12	0.14	0.18	0.13	0.13
Observations	1,450	1,889	1,344	1,611	1,952	1,536	1,573
Panel B: Above and Below Graduate							
Winner	0.374***	0.170***	0.175***	0.059	-0.033	0.110	0.103
	(0.099)	(0.051)	(0.062)	(0.076)	(0.127)	(0.126)	(0.115)
Dep. Mean	2.30	0.55	0.12	0.04	0.08	0.06	-0.02
Bandwidth	0.08	0.19	0.08	0.16	0.15	0.15	0.15
Observations	895	1,757	902	1,100	1,034	1,031	1,056

This table reports the effects of electing a more educated politician on three dimensions of governance—information, network, and project implementation—using alternative education thresholds. Panel A defines the treatment as having education above high school (*Above High-School*), and restricts the sample to constituencies in which the top two candidates straddle this threshold. Panel B uses graduate education (*Above Graduate*) as the threshold, with the sample restricted analogously. Estimates are obtained from a close-election regression discontinuity design, and the treatment variable equals one if the winning candidate has education above the relevant threshold. Columns (1)–(3) use a survey of male Ward Members elected in the 2021 Ward elections. *Scheme Knowledge* (1) is the average of standardized scores measuring each Ward Member's ability to recount the implementation steps for six government schemes (ward account, AWAS housing, drains, PDS, pensions, and solar lights), matched to a predetermined checklist. *Bureaucratic Network* (2) is the average of indicators for whether the respondent knows the phone numbers of the Panchayat Secretary, BPRO, and BDO. *Party Support* (3) is the average of indicators for any party involvement and receipt of political support. Columns (4)–(7) use GP-level data from the 2016 GP Head elections. *Infra. Aggregated* (4) is the average of three indicators for the availability of piped tap water (5), road availability (6), and drainage systems (7), each also shown separately. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A18: Effects of Female Reservation on Political Selection for GP Head Candidates

	Yrs of Schooling			Representativeness		
	Edu Yrs (All Citizen)	Edu Yrs (Own Group)	Husband Edu Yrs	Gov. Job	Income Tax	Wealth Score
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Candidates			
Female RSVN	-0.765***	-0.458***	-0.147***	-0.081***	-0.074***	-0.091***
	(0.024)	(0.024)	(0.031)	(0.026)	(0.024)	(0.022)
Dep. Mean	0.80	0.65	0.80	0.21	0.14	0.32
Bandwidth	858.596	855.351	873.669	799.300	873.669	977.299
Observations	44,583	44,552	36,498	43,369	44,737	46,487
			Panel B: Winners			
Female RSVN	-0.921***	-0.619***	-0.364***	-0.216**	-0.176*	-0.190***
	(0.073)	(0.074)	(0.089)	(0.086)	(0.093)	(0.070)
Dep. Mean	1.11	0.95	1.11	0.32	0.30	0.65
Bandwidth	489.012	473.452	691.411	566.563	539.418	642.694
Observations	3,294	3,249	3,058	3,528	3,456	3,759

This table reports the effects of female reservation on political selection outcomes for GP Head candidates and winners in the 2016 and 2021 Panchayat elections. Estimates are obtained from a regression discontinuity design exploiting the population threshold that determines female reservation status. Columns (1)–(3) report competence measures based on standardized years of education benchmarked against: all citizens (1), citizens of the same group (2), and all citizens using husband's education in place of the candidate's own (3). Columns (4)–(6) report eliteness measures: an indicator for whether any household member holds a salaried government job (4); an indicator for whether any household member pays income or professional taxes (5); and a PCA-based wealth index constructed from land ownership, housing characteristics, and asset ownership (6). Panel A reports results for candidates; Panel B reports results for winners. All specifications include GP-level fixed effects and control for election year and GP-level covariates. Standard errors are clustered at the GP level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Matching

As described in Section 3.1, we match candidates to citizens in the SECC using a string-matching algorithm based on name, father/husband name, caste category, and gender. We restrict to matches with a similarity score of at least 0.95 and achieve an overall match rate of 60.2% for unreserved GP head candidates. Here we provide further details on the matching procedure and robustness to alternative thresholds.

The match rates are lower at the ward level because we do not have citizens' ward information (only GP). Our match rates are also lower for 2021 than 2016 because the SECC is more outdated for the former than the latter. Finally, our match rates are lower for women than men because marriage norms in India make women more mobile, since they leave their villages to be with their husbands. Thus, women who would have married into their husband's GP after the SECC was conducted will be missing from our data. Moreover, for women, the candidate data has either one of the father or husband's name and doesn't specify which one is mentioned, whereas for men the candidate data has only the father's name. The latter makes it more straightforward to match.

On balance, the candidates we find in the SECC look slightly different from the sample of candidates we do not match. In particular, our matched candidates are always somewhat older than our unmatched candidates. Our 2021 matched sample candidates are even older: the average matched candidate is abut 2.5 - 4.5 years older than average unmatched candidate. The corresponding range for 2016 candidates is 0.24 - 2 years. Education-wise, for the 2016 sample of candidates, there is no clear pattern: male GP head matched and unmatched candidates have similar education levels; female matched candidates (at both ward and GP head levels) have marginally lower education levels than their unmatched counterparts, while matched ward male candidates have slightly higher years of education. Caste-wise, we find more SC candidates in the SECC in 2021, but the patterns are more mixed for 2016 candidates.

Table A19: Balance Test of SECC and Electoral Candidates Data Matching

	Education				Caste			
	Yrs of Schooling (1)	At least Literate (2)	At least Secondary (3)	At least Higher (4)	SC (5)	ST (6)	Other (7)	Age (8)
Panel A: Candidates Matching (55.36% candidates matched)								
Matched	-0.021	0.011***	0.003**	-0.012***	0.044***	-0.019***	-0.025***	2.604***
	(0.015)	(0.001)	(0.002)	(0.001)	(0.001)	(0.000)	(0.001)	(0.025)
Control Mean	5.465	0.899	0.376	0.137	0.250	0.010	0.740	41.627
Observations	361,293	363,217	363,217	363,217	908,511	908,511	908,511	907,560
Panel B: Candidates Fathers Matching (41.14% candidates matched)								
Father Matched	1.128***	0.026***	0.101***	0.095***	-0.055***	-0.015***	0.070***	-8.522***
	(0.021)	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)	(0.035)
Control Mean	7.578	0.969	0.579	0.274	0.202	0.013	0.785	36.683
Observations	175,681	176,614	176,614	176,614	433,003	433,003	433,003	432,551

This table compares demographic characteristics of candidates who can be matched to the SECC data with those who cannot, using data from GP Head and Ward Panchayat elections in 2016 and 2021. Panel A presents balance tests for candidates, while Panel B presents balance tests for candidates' fathers. Columns (1)–(4) report education outcomes; Columns (5)–(7) indicate whether the candidate belongs to the Scheduled Caste, Scheduled Tribe, or general category; and Column (8) reports age, winsorized at the top and bottom 1 percent. Standard errors are clustered at the Gram Panchayat level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table A20: Balance in observables between matched and non-matched candidates

	Education				Caste			
	Yrs of Schooling (1)	At least Literate (2)	At least Secondary (3)	At least Higher (4)	SC (5)	ST (6)	Other (7)	Age (8)
Panel A: Mukhiya Man 2016 (63.62% candidates matched)								
Matched	0.056	0.000	0.012**	-0.004	-0.010***	-0.006***	0.016***	0.927***
	(0.048)	(0.002)	(0.005)	(0.005)	(0.003)	(0.001)	(0.003)	(0.120)
Control Mean	8.495	0.973	0.660	0.342	0.248	0.031	0.721	42.643
Observations	46,843	47,178	47,178	47,178	47,178	47,178	47,178	47,130
Panel B: Mukhiya Woman 2016 (54.26% candidates matched)								
Matched	-0.853***	-0.008***	-0.072***	-0.076***	0.005**	-0.004***	-0.001	2.046***
	(0.044)	(0.002)	(0.005)	(0.004)	(0.003)	(0.001)	(0.003)	(0.116)
Control Mean	6.035	0.943	0.410	0.182	0.206	0.016	0.778	38.017
Observations	45,840	46,132	46,132	46,132	46,132	46,132	46,132	46,107
Panel C: Ward Man 2016 (63.83% candidates matched)								
Matched	0.372***	0.018***	0.047***	0.009***	0.008**	-0.017***	0.009**	0.240***
	(0.031)	(0.001)	(0.003)	(0.002)	(0.004)	(0.001)	(0.004)	(0.049)
Control Mean	5.838	0.930	0.409	0.154	0.234	0.030	0.736	41.331
Observations	128,845	129,447	129,447	129,447	129,447	129,447	129,447	129,065
Panel D: Ward Woman 2016 (57.83% candidates matched)								
Matched	-0.177***	0.010***	-0.015***	-0.019***	0.083***	-0.017***	-0.066***	0.537***
	(0.023)	(0.003)	(0.003)	(0.001)	(0.005)	(0.001)	(0.004)	(0.048)
Control Mean	3.665	0.795	0.196	0.054	0.194	0.024	0.783	38.968
Observations	138,446	139,144	139,144	139,144	139,144	139,144	139,144	138,704

Focusing on the GP and ward elections of 2016, this table shows how characteristics of those politicians we match in the SECC data vary from those we do not find. Age is winsorized at top and bottom 1 percentile. Panel (A-B) and (C-D) respectively include GP and Ward level fixed effects. Standard errors are clustered at the same level as the fixed effects.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A21: Balance in observables between matched and non-matched candidates

	Caste			
	SC	ST	Other	Age
	(1)	(2)	(3)	(4)
Panel A: Mukhiya Man 2021 (53.14% candidates matched)				
Matched	0.012***	-0.013***	0.002	4.478***
	(0.003)	(0.001)	(0.003)	(0.153)
Control Mean	0.173	0.030	0.797	42.268
Observations	30,654	30,654	30,654	30,654
Panel B: Mukhiya Woman 2021 (48.7% candidates matched)				
Matched	0.011***	-0.010***	-0.001	3.820***
	(0.003)	(0.001)	(0.003)	(0.146)
Control Mean	0.140	0.017	0.843	38.930
Observations	31,242	31,242	31,242	31,241
Panel C: Ward Man 2021 (51.86% candidates matched)				
Matched	0.051***	-0.018***	-0.033***	4.125***
	(0.004)	(0.001)	(0.004)	(0.056)
Control Mean	0.191	0.032	0.777	38.865
Observations	225,601	225,601	225,601	225,578
Panel D: Ward Woman 2021 (52.3% candidates matched)				
Matched	0.084***	-0.023***	-0.061***	2.501***
	(0.003)	(0.001)	(0.003)	(0.066)
Control Mean	0.164	0.029	0.806	38.507
Observations	256,392	256,392	256,392	256,359

Focusing on the GP and ward elections of 2021, this table shows how characteristics of those politicians we match in the SECC data vary from those we do not find. Age is winsorized at top and bottom 1 percentile. Panel (A-B) and (C-D) respectively include GP and Ward level fixed effects. Standard errors are clustered at the same level as the fixed effects.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A22: Balance in observables between matched and non-matched candidates fathers

	Education				Caste			
	Yrs of Schooling (1)	At least Literate (2)	At least Secondary (3)	At least Higher (4)	SC (5)	ST (6)	Other (7)	Age (8)
Panel A: 2016 Mukhiya (36.52% candidates fathers matched)								
Father Matched	0.889***	0.007***	0.073***	0.096***	-0.022***	-0.004***	0.026***	-8.538***
	(0.043)	(0.001)	(0.004)	(0.005)	(0.002)	(0.001)	(0.002)	(0.107)
Control Mean	9.339	0.980	0.736	0.419	0.223	0.013	0.764	37.987
Observations	46,843	47,178	47,178	47,178	47,178	47,178	47,178	47,130
Panel B: 2016 Ward (37.7% candidates fathers matched)								
Father Matched	1.421***	0.039***	0.137***	0.102***	-0.049***	-0.011***	0.060***	-9.128***
	(0.024)	(0.001)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.062)
Control Mean	6.958	0.965	0.524	0.222	0.208	0.013	0.779	35.823
Observations	128,845	129,447	129,447	129,447	129,447	129,447	129,447	129,065

Focusing on the GP and ward elections of 2016, this table shows how characteristics of those politicians whose fathers we match in the SECC data vary from those whose fathers we do not find. Age is winsorized at top and bottom 1 percentile. Panel (A-B) and (C-D) respectively include GP and Ward level fixed effects. Standard errors are clustered at the same level as the fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A23: Balance in observables between matched and non-matched candidates fathers

	Caste			
	SC (1)	ST (2)	Other (3)	Age (4)
Panel A: 2021 Mukhiya (41.61% candidates fathers matched)				
Matched	-0.015***	-0.009***	0.024***	-7.750***
	(0.003)	(0.001)	(0.003)	(0.138)
Control Mean	0.172	0.015	0.813	40.105
Observations	30,654	30,654	30,654	30,654
Panel B: 2021 Ward (43.78% candidates fathers matched)				
Matched	-0.031***	-0.019***	0.051***	-8.136***
	(0.002)	(0.001)	(0.002)	(0.048)
Control Mean	0.200	0.012	0.788	36.438
Observations	225,601	225,601	225,601	225,578

Focusing on the GP and ward elections of 2021, this table shows how characteristics of those politicians whose fathers we match in the SECC data vary from those whose fathers we do not find. Age is winsorized at top and bottom 1 percentile. Panel A and B respectively include GP and Ward level fixed effects. Standard errors are clustered at the same level as the fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

F Seat reservation rule

Bihar has reservations for SCs, STs, EBCs (*not* OBCs) and women. The rule for these is given below. The crucial takeaway from the sections below is that, within each caste-reserved and caste-unreserved GPs, there exists a population threshold above which all seats are gender-reserved. Note, also, that the rule for 2006 is fixed till 2016 (two election cycles), after which the reservation cycle switches Below, we describe the reservation algorithm for caste/gender for 2016 and 2021 (the relevant period in our study).

F.1 Caste Reservation

- First, based on the proportion of SCs (STs) in the block, the number of GPs to be reserved for SCs (STs) is decided. If there are N_j GPs in block j and θ_j is the proportion of SCs (STs) in block j, then the number of GPs, n_j, to be reserved is

$$n_j = round(\theta_j * N_j, 1)$$

- Let n_{SC} and n_{ST} be the number of GPs to be reserved in block j for SCs and STs, respectively. The number of GPs to be reserved for EBCs is given by

$$n_{EBC} = min(round(0.2 * N_j, 1), round(0.5 * N_j - n_{SC} - n_{ST}, 1))$$

- Now, all GPs are rearranged in the descending order of their non SCST population. The first GP on this truncated list is "blocked". The choice of word is deliberate and conveys an important distinction: the GP is not "reserved", it is merely blocked.
- Now, all unreserved and unblocked GPs are rearranged in descending order of their SC population. The first GP in this further truncated list is now reserved for SCs, unless it has already been reserved for SCs in the 2006-16 reservation cycle. If the latter, then the next GP in the list is selected.
- If there are no STs in the block or n_{ST} is 0 (which is true in 480 of the 534 blocks), then the rule skips to the next step. However, if $n_{ST} > 0$, the rule proceeds by arranging all remaining GPs in descending order of their ST population. The first GP in the list is then reserved for STs (unless it has already been reserved for STs in the previous cycle, in which case, the 2nd GP in the list is picked).
- This algorithm proceeds until the number of GPs reserved for STs = n_{ST} or the number reserved for SCs is n_{SC}. Once, a group hits its quota of reserved GPs, then the rearranging of GPs is no longer done by that group. For instance, if $n_{ST} = 1$, then, in the second round, GPs are no longer rearranged by ST population - instead, the rule proceeds straight to rearranging by non-SCST population.
- The algorithm further proceeds till the second group also hits its quota of reserved GPs. This throws up two sets of GPs, n_{ST} GPs that are reserved for STs and n_{SC} GPs that are reserved for SCs.

- Now, all the unreserved GPs (including the "blocked" ones) are collected and arranged by descending order of GP population.
- The first n_{EBC} GPs in this list is reserved for EBCs.
- Thus, for each block, one can arrive at an SC population cut-off - the SC population of the last GP to be reserved for SCs - below which no GP is reserved. This threshold varies by block. Figure **??** gives the first stage and shows that the first stage is robust to (a) adding block fixed effects (Panel (b)) and (b) dropping all ST/EBC reserved GPs.

F.2 Gender Reservation

- All the n_{SC} seats determined to be reserved for SCs within a block are arranged in descending order of SC population and *up to* 50% of them with the highest populations are reserved for SC women. For instance, if there are $n_{SC} = 3$, then 1 GP is reserved for SC women – the one with the highest SC population. On the other hand, if $n_{SC} = 4$, then 2 GPs are reserved for SC women – the top 2 highest SC population GPs.
- All the n_{ST} seats determined to be reserved for STs within a block are arranged in descending order of ST population and *up to* 50% of them with the highest ST populations are reserved for ST women.
- All the n_{EBC} seats determined to be reserved for EBCs within a block are arranged in descending order of total population and *up to* 50% of them with the highest total populations are reserved for EBC women.
- All the n_{GEN} seats determined to be caste unreserved within a block are arranged in descending order of non-SCST population and *up to* 50% of them with the highest non-SCST populations are reserved for women.

G Citizen Survey and Conjoint Experiment Design and Implementation Details

This appendix provides additional details on the primary survey used to elicit citizen preferences over candidate characteristics, including education, through a conjoint experiment. This citizen survey was conducted as part of Bamezai (2025; FIXME) under the University of Pennsylvania (IRB Protocol #855284). The study was pre-registered on the Open Science Framework (Registry ID: p65hr). Details are available at https://osf.io/p65hr.

G.1 Survey Implementation and Sample Composition

The survey was fielded between February and March 2024 across 48 Gram Panchayats (GPs) located in two districts of Bihar. The study yielded a total of 1,105 completed interviews, comprising 969 citizens, 110 elected Ward Members (WMs), 3 GP heads, and 23 frontline workers (including Anganwadi workers, ASHAs, and other local functionaries). All respondents were included in the conjoint analysis.

G.2 Sampling Strategy

The sampling strategy was designed to capture heterogeneity in political experiences and expectations at a highly local level. GPs in Bihar typically comprise multiple revenue villages and wards, with substantial variation across caste, religion, and socio-economic status. Given this heterogeneity, sampling was conducted at the ward level.

Two districts were selected: Bhojpur and Samastipur. Within each district, three blocks were randomly selected. From each block, eight GPs were randomly sampled, and within each Panchayat, four wards were selected, yielding a total of 192 wards.

Within each ward, enumerators were instructed to complete up to six interviews, subject to time and travel constraints. Households were selected using a standard random-walk procedure. Enumerators first obtained an estimate of the total number of households in the ward from either the elected ward member or a frontline worker, and then selected every k-th household, where k equals the total number of households divided by six and rounded down. If a household refused consent or no eligible respondent was available, the enumerator moved to the next household.

Within selected households, one respondent aged 18–75 was randomly chosen using the survey instrument. Enumerators also interviewed the ward member or a frontline worker in the ward when consent was provided. Approximately 25% of interviews were conducted with female respondents. This share reflects feasibility constraints encountered during piloting, which indicated substantial difficulty among many women respondents – particularly those with lower education levels – in answering questions related to Panchayat elections and candidate evaluation.

G.3 Conjoint Experiment Design

The conjoint experiment was embedded within the survey to elicit respondent preferences over hypothetical Panchayat candidates. Each respondent was shown paired profiles of two hypothetical individuals residing in a village similar to their own. Respondents were explicitly informed that the two profiles were identical in all respects except for the attributes shown, thereby holding representativeness constant.

Candidate profiles varied randomly across seven attributes, one of which was education. All seven attributes are described below:

G.3.1 Conjoint experiment: Other attributes

1. **Education level**: There are three levels: (i) completed 5th grade, (ii) completed 10th grade, and (iii) completed B.A.. Each candidate can be assigned one of the three levels withequal probability (0.33 each).
2. **Gender**: There are two levels: male and female. Each candidate can be assigned one of the two values with equal probability (0.5 each).
3. **Caste category**: There are three levels here: (i) Scheduled caste category (like Paswan, Chamar, Manjhi, etc.), (ii) Backward caste category (like Yadav, Kurmi, etc.), (iii) General caste category (like Bhumihar, Rajput, etc.). Each candidate can be assigned one of the three levels with equal probability (0.33 each).
4. **Behavior/personality ("vyavhaar")**: There are two levels here: (i) very good, and (ii) not very good. Each candidate can be assigned one of the two levels with equal probability (0.5 each).
5. **Previous family experience in local electoral politics**: There are two levels: (i) no such experience, and (ii) someone in the family has been in an elected Panchayat position before (like Mukhiya, Sarpanch, etc.). Each candidate can be assigned one of the two levels with equal probability (0.5 each).
6. **Presence of a government worker in the family**: Therer are two levels: (i) yes, and (ii) no (0.5 each).
7. **Other source of family income**: There are four levels here: (i) Wage labor, (ii) Farming, (iii) Business, (iv) Private job. Each candidates can be assigned one of the four levels with equal probability(0.25 each). The decision to include this attribute was based on my learning that there are barely any households who rely entirely on a government job for their household's subsistence. In rural Bihar, many families will have at least one other source of household income.

For each profile pair, respondents were asked two questions: (a) which individual they believed was more likely to contest a Panchayat election; and (b) conditional on both contesting, which individual was more likely to win the election.

Respondents completed three conjoint tasks if their own education was below 10th grade, and four tasks if they had completed 10th grade or higher. This design yielded approximately 7,600 evaluated profile pairs.