The Determinants and Consequences of Political Selection in the Developing World: Evidence from India*

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

Mar 26, 2026

Abstract

We examine the determinants and governance consequences of political selection in low-income democracies. We link administrative data on 1 million villages politicians in Bihar, India with a census of 95 million rural residents, and present three findings. First, we find that politicians are much more educated than the average citizen and their own siblings, but come from elite backgrounds. This strong tradeoff between competence and representativeness contrasts with the inclusive meritocracy Dal Bó et al. (2017) observe in Sweden. Second, we show that this tradeoff shapes governance: using a close election regression discontinuity design, we find that villages governed by highly competent politicians implement skill-intensive programs more effectively, but redistribute less to disadvantaged groups and align public public spending with elites' preferences. Third, we show that political reservations weaken the tradeoff between competence and representativeness. Using a population threshold RD design, we find that female and disadvantaged caste candidates come from less elite households within their own communities but are more positively selected on competence. Our results demonstrate that political selection in developing countries has distinct patterns which shape governance, but that institutional reforms can meaningfully shape who becomes a politician.

*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

Political leaders matter. Even mature democracies with strong institutions can be misgoverned if they select poor leaders. In advanced democracies like Sweden, politicians tend to be both competent and socioeconomically representative of the society at large (Dal Bó et al., 2017). Yet most polities around the world—even among developed countries—are not like Sweden: they are more unequal, more diverse and have a shorter history of democracy. We have surprisingly little evidence on who becomes a politician in low-income democracies—even though selection is arguably more important in settings where institutions impose fewer constraints on incumbents. We also have a limited understanding of the consequences of political selection and how widely-enacted policies to broaden political representation—like decentralisation and affirmative action—shape the political class.

In this paper, we study the causes and consequences of political selection in local government in Bihar, a large Indian state that accounts for 3.5% of the global population living in a democracy. Bihar provides an excellent context to examine how political selection operates in a low-income democracy. High levels of inequality concentrate educational and economic opportunities among unrepresentative elites. Deep social and ethnic divisions may lead voters to prioritise co-ethnic representation over candidate competence. Finally, democratic institutions are often less mature, and mechanisms such as political parties that can facilitate positive and inclusive selection may be weak or absent. These features may sharpen the tradeoff between competence and representativeness.

We begin by describing political selection patterns in rural Bihar[1]. Rich administrative data on the universe of citizens (95 million) and local politicians (nearly 1 million) in rural Bihar allows us to measure both the competence and representativeness of candidates in village elections.

We measure competence by years of formal education. While imperfect, four reasons support its use here. First, it is available for all 97 million individuals in our setting. Second, using a close-election RD design and survey data, we show that more educated politicians have greater knowledge of government programmes and wider professional networks – both crucial for local governance. Third, a forced-choice conjoint experiment with 1100 citizens reveals that citizens view more educated candidates as more likely to contest and win elections. Fourth, citizens rank education among the top three attributes they consider when voting, above prior experience and wealth.

We first show that politicians are positively selected from the citizenry. Candidates for political office have 2.6 more years of schooling (0.67 SD) than the average citizen. Politicians are also 0.49 SD more educated than citizens in their birth cohort, 0.47 SD more educated than citizens in their subcaste, and 0.34 SD more educated than other men in

[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.

their household.[2] Moreover, competence is correlated with political success: winners are more competent than losing candidates, and candidates for higher-level offices are also more educated.

Politicians could be more educated than the average citizen because they are positively selected or simply because they come from elite social backgrounds. It is true that politicians come from more advantaged social groups. However, to distinguish between these explanations and understand whether politicians are representative of the population, we match politicians to their fathers and compare the educational qualifications of politicians' fathers relative to the average citizen's father. We find that politicians come from families that are more elite than the average household. Politicians' fathers have 0.5 SD higher levels of education and higher wealth than others in their birth cohort, caste and neighbourhood. Nevertheless, politicians are positively selected (0.18 SD more educated) relative to their siblings. This suggests that political selection in Bihar is meritocratic (as in Sweden), but not as inclusive and rather elite-biased. We see this clear tradeoff between competence and representativeness across most of Bihar's 8000-odd village governments.

The competence-representativeness tradeoff may be starker in our context than in the advanced economies previously studied (Dal Bó et al., 2017), for three reasons. First, Bihar is substantially more unequal than Sweden. In highly unequal polities, restricted educational opportunities may generate elite-biased political selection. We find that politicians are more educated and less representative in more unequal villages. Second, group identity plays a central role in politics, particularly in developing countries. Candidates from larger groups (e.g. religious groups, sub-castes) are both less competent and from less elite backgrounds, consistent with workhorse models (Dal Bó and Finan, 2018). Yet this relationship is non-linear: marginal increases in group size reduce competence among majority-group candidates while raising it among minority-group candidates, likely reflecting the non-linear relationship between group size and the probability of victory. Third, parties are far less institutionalised in most developing countries than in advanced economies. In our setting, parties are present in some polities but absent in others; we show that politicians are more competent where parties exist, and party-affiliated politicians are particularly positively selected.

Having documented a robust competence-representativeness tradeoff in political selection, we investigate its implications for governance. We collect data on citizen preferences and government plans to measure how closely planned public spending aligns with citizen priorities. Using a close-election RD design, we show that GPs governed by more educated leaders allocate less planned expenditure and fewer development projects to historically marginalized Scheduled Caste (SC) citizens' priorities. Moving beyond plans, we test whether competence also shapes realised service delivery. More educated leaders improve implementation of skill-intensive schemes–such as piped water and road construction–but deliver less redistribution: they provide fewer NREGS benefits to SC citizens. Together, these results show that the competence-representativeness tradeoff

[2] We initially focus on unreserved seats which are contested and won overwhelmingly by men. We discuss later how female reservations impact selection patterns.

has tangible consequences for how local governments operate.

Finally, we evaluate a common policy that countries around the world have implemented to improve and broaden political selection—affirmative action for marginalised groups. Many countries introduce political reservations ("quotas") for members of disadvantaged groups. To evaluate the impacts of reservations, we exploit a natural experiment where we compare polities that had "just enough" SC citizens to be reserved for SCs against polities that had "just too few" SCs. Using this variation, we find that political reservations result in less competent but more representative candidates. Interestingly, this is true not just because SCs are less educated than other groups, but more so because SC politicians are more representative of their community than upper-caste politicians are of upper-caste citizens. Our evidence suggests that reservations reduce the competence-representativeness tradeoff, as the representativeness gains generally outweighs the competence loss.

Second, we study the impact of political reservations for women. Using a regression discontinuity design that exploits the reservation algorithm – such that similar GPs are quasi-randomly assigned either a female or male leader – we show that reservations go beyond improving women's representation, inducing a broader change in the selection and social origins of leaders. Reserved female leaders are less educated than unreserved leaders, though they remain more educated than the average female citizen. They also come from less elite households, meaning reservations improve representativeness along both gender and socioeconomic dimensions. Since proxy governance via husbands is prevalent in this setting (Amar et al., 2025), we use the SECC to identify and assess the husbands of reserved leaders directly – and find that even they are less educated than the average unreserved candidate. Overall, political reservations heighten the competence-representativeness tradeoff: they improve representativeness along both gender and socioeconomic dimensions, but at the cost of lower competence, even when accounting for proxy governance.

Contribution to Literature. The paper contributes to two strands of literature. Primarily, we add to a growing literature on political selection, first surveyed by Besley (2005) and more recently by Dal Bó and Finan (2018) and Gulzar (2021). Existing research has described patterns of political selection in rich countries. In a seminal contribution, Dal Bó et al. (2017) studied who becomes a politician in Sweden, exploiting rich Swedish administrative data to illustrate selection patterns and documenting that Sweden is an "inclusive meritocracy". Thompson et al. (2019) studies the historical US, also finding positive selection but less inclusivity in social backgrounds. We show that competence and representativeness have a much starker tradeoff in developing countries.

Turning to developing countries, two recent studies from Pakistan and Nepal provide evidence on important aspects of political selection. While Gulzar and Khan (2021) focus on how to motivate politicians with prosocial motivations to contest local elections, Gulzar et al. (2021) show how providing information to party leaders about voter preferences improves candidate selection and the overall quality of representation. Our paper

is closest to the ongoing project in Nepal by Bhusal et al. (2023) in terms of method and approach. While they compare representativeness in political selection under monarchy and a new democratic framework, we study the representativeness-competence tradeoff in a relatively more stable local government system. We also develop a model to explain our results, evaluate the impact of two common approaches to broaden political selection and study the economic consequences of more competent politicians.

Second, we contribute to the literature on political reservations. Prior work has documented that gender quotas improve representation of women, by changing policy preferences (Chattopadhyay and Duflo, 2004) and breaking stereotypes (Beaman et al., 2012). These effects persist even after quotas end (Bhavnani, 2009, 2017). We go beyond prior work by showing that quotas also broaden representation along axes other than gender, because selection patterns for male and female leaders differ. Specifically, female politicians are selected from more representative households.

The rest of the paper is organised as follows. Section 2 describes the background and context. Section 3 presents descriptive evidence on political selection patterns. Section 4 presents empirical evidence on the determinants of political selection. Section 5 presents results on the economic consequences of political selection. Section 6 concludes.

2 Background and Context

2.1 Bihar

Population, Poverty and Inequality: Bihar is among the most densely populated and poorest parts of the world. With a population of over 130 million, nearly 90% of whom resided in rural areas in 2011, Bihar lags behind the rest of India on a whole host of indicators. As per the National Family Health Survey - 5, conducted between 2019-21, Bihar trails the national average with respect to poverty rates, infant mortality and schooling outcomes. Bihar is a largely agrarian state, with deep inequities between a small fraction of large land-owning households and households of those engaged in casual wage work. In 2009-10, the land value of a landlord household – defined as those living on rent or solely involved in farming supervision – was 33 times greater than that of an agricultural labour household (Swaminathan and Nagbhushan, 2022).

Caste relations: Caste is deeply entrenched and is fundamental to the organisation of social, economic and political lives of the people of the state. Only 11% of the marriages in Bihar cross caste boundaries and nearly 47% of respondents report that someone in their household practices untouchability (Desai and Vanneman, 2015). The sub-castes at the bottom of the asset wealth distribution are also those at the bottom of the caste hierarchy, suggesting the continuing prominence of caste in determining economic trajectories for a large number of Bihar's citizens (Sharan, 2023). Bihar's thousands of sub-castes can be broadly divided into 5 main groupings: the Scheduled Castes (17%) comprising extremely marginalised sub-castes all historically considered "untouchable"; Scheduled

Tribes (1.7%), who come from tribal communities and are also relatively disadvantaged; Extremely Backward Castes (EBCs), a group of disadvantaged subcastes carved out of the larger, more dominant Other Backward Castes (OBCs) grouping. Finally, at the top of the caste hierarchy sit the general castes.

Recent Developments: Bihar went through a dramatic phase of great social churning, a breakdown of law and order and consequent phase of stalled growth around the turn of the century (Witsoe, 2013). Since then, partly owing to a low base and partly to progressive economic and social policies and support from the central government, Bihar turned around to emerge as one of the country's fastest growing states. Yet, in absolute terms, Bihar contributes a disproportionately small share to India's national GDP.

2.2 Local Governance

Bihar is a largely rural polity. It has under 300 urban local bodies and over 8000 rural local bodies. In this paper, we focus on the latter. The rural local bodies are called Gram Panchayats (GPs). The population of a GP is roughly 13000.

Each GP is further divided into administrative units called wards. In 2021, there were about 13.6 wards for every GP, summing to about 110,000 wards in the state. A ward is, therefore, extremely local, comprising about 1000 persons or 200-250 households.

Both wards and GPs have directly elected heads, called "ward members" and "GP heads" (locally known as Mukhiyas), respectively. GP heads are the most important local leaders in the GP: they have direct access to the state's pool of resources meant for development and are chiefly responsible for implementing development projects and welfare schemes.

Ward members, on the other hand, are far less powerful and, up until 2016, had no direct access to the state's funds. Their role is limited to representing their constituents in the GP meetings and village meetings (called "Gram Sabhas"). GP heads may rely on ward members to implement schemes and projects. On paper, ward members are to their wards what GP heads are to the GP: the persons chiefly responsible for overseeing development in their constituencies. However, the absence of formal financial powers makes the equation lopsided in favour of GP heads.

The elections for GP heads and ward member are often hotly contested. In 2016, as per our data, 94,645 (268,591) candidates contested GP head (ward member) elections. The corresponding numbers for 2021 are 63,284 and 481,993. The average GP head in 2016 had 7.17 years of education, compared to 4.77 for ward members.

Local Elections: Local elections are conducted by the State Election Commission of Bihar. Voting percentage in Bihar's local elections are much higher than in state and national elections. The national elections of 2019 saw a voter turnout rate of 57.3% and state election of 2020 saw a turnout of 57.29%. In contrast, Bihar's 2021 Panchayat elections, held in the wake of the devastating second wave of the COVID-19 pandemic in India, saw a

turnout of over 60%.[3]

Reservation: Since 2006, Bihar has reserved seats in favour of minorities for both the GP head and ward member positions. About 45% seats are reserved at both the GP head and ward member levels for women, 17% for Scheduled Castes (SCs), about 20% for Extremely Backward Castes (EBCs) and under 2% for Scheduled Tribes (STs). Elections are held every 5 years. A constituency's reservation status switches every 10 years. For both the 2016 and 2021 election years – the focus of this paper – the reservation status of a constituency was fixed.

2.2.1 GP Plans and Development

Local governance in India under the 73rd Constitutional Amendment institutionalises deliberative decision-making as a core feature of rural policymaking. The Gram Sabha – a legally mandated assembly of all eligible adult residents within a Gram Panchayat – provides a formal forum for public discussion of development priorities and accountability of elected representatives, creating a structured space for participatory governance and collective choice over village affairs. Empirical work on village assemblies shows that when Gram Sabhas are functioning and convened regularly, deliberative interactions can reflect a broad set of citizen preferences and bring diverse voices into local decision-making, thereby strengthening transparency and the responsiveness of policy agendas to median voter interests in the community (Parthasarathy et al., 2019). Research on decentralised governance also emphasises the constitutional and institutional grounding of these assemblies within the Panchayati Raj system, highlighting both their potential and the challenges of effective participation in contexts marked by social and economic heterogeneity (Ban et al., 2012).

In Bihar, as in many other Indian states, these deliberative processes are operationalised through the Gram Panchayat Development Plan (GPDP), the principal annual planning document for a Panchayat's development agenda. The GPDP is designed to be a comprehensive, participatory plan that lists all projects the Panchayat intends to undertake in the coming year along with estimated costs and proposed sources of funding, reflecting the outcomes of consultations at the ward level and in special Gram Sabha meetings. Under recent policy initiatives such as the People's Plan Campaign, Gram Sabhas are convened specifically to mobilise community input, prioritise locally identified needs, and ensure inclusive participation – including efforts to involve vulnerable groups – before the plan is finalised and uploaded to the *eGramSwaraj* platform for transparency and monitoring. Following Gram Sabha approval, block and district officials review and sanction projects for implementation, integrating GPDP priorities with state and central funding streams. As such, the GPDP not only functions as a comprehensive overview of intended development activities within a Panchayat but also institutionalises the link between grassroots deliberation and formal planning and budgeting in India's decentralised governance sys-

[3] Bihar's SEC has been a pioneer in the use of technology: the 2021 elections used electronic voting machines (EVMs), a first in local body elections in the country.

tem.

3 Overcoming Empirical Challenges

As documented in Dal Bó et al. (2017), there are three key data limitations that any comprehensive study of political selection must overcome. Below, we outline each challenge and its empirical implications; we then discuss how our setting addresses each in the subsections that follow.

Challenge I: Tracking the Selection Pipeline Studying political selection requires observing not only elected politicians, but also non-elected candidates and individuals who never enter politics. In practice, data are typically limited to elected politicians, making it difficult to disentangle selection into politics from selection among candidates.

Challenge II: Quantifying Political Competence A key difficulty is finding a measure of competence that is both meaningful and universally comparable across candidates.

Challenge III: Benchmarking Representativeness Assessing whether politicians represent the broader population requires detailed intergenerational data – such as parental background – which is rarely available. Without such data, it becomes harder to understand if individuals' ability are a function of their abilities or their background.

3.1 Identifying the Selection Pipeline: Data Linkage and Entry

To address the challenge of observing both entrants and non-entrants into politics, we construct a dataset that links electoral records to population-wide administrative data. Specifically, we combine two large datasets: the Socioeconomic Caste Census (SECC) of 2011–12 and candidate-level data on GP head and ward member elections in 2016 and 2021 from the Bihar State Election Commission. This linkage allows us to identify elected politicians, non-elected candidates, and individuals who never enter politics.

Population-wide Administrative Data from the Socioeconomic Caste Census(SECC): The SECC is a census of households across India. We use SECC data for rural Bihar, which contains administrative data on the universe of individuals in rural Bihar. The data includes 97 million individuals from 19 million households. Importantly, for every citizen, we have data on who their parents are and who else lives in their household.

At the within-household level, the survey captured basic information on members of the household including gender, broad caste category, age, type of occupation and education. At the household level, the survey collected data on: type of dwelling including number of rooms, type of wall and roof; employment and income characteristics including whether household has a member having a government job and main source of household income; asset ownership (vehicle, fridge, mechanical agricultural equipment etc); details on land-owned.

Candidate Data from the Bihar State Election Commission: The Bihar State Election Commission uploads data on all candidates contesting GP head and ward member elections for 2016 and 2021 elections. There are 909,570 candidates across the two elections.[4] We have data on candidate full names, father/husband names, age, education, caste category and votes polled.

Citizen-Candidate Matching We match candidates to citizens in the SECC to recover their family and socioeconomic backgrounds. We restrict matches to individuals within the same GP and implement a string-matching algorithm based on name, father/husband name, caste category, and gender. The algorithm assigns a similarity score to each potential match, and to ensure match quality, we restrict the sample to matches with a similarity score of at least 0.95. This threshold implies a low likelihood of false matches (inclusion error), while allowing for minor spelling and transcription differences. At the same time, this procedure may generate some exclusion error, as not all candidates can be confidently matched. Our results are robust to alternative thresholds.

Our overall match rate for our primary population of interest, unreserved GP head candidates, is around 60.2%. This overall match rate is relatively high and compares favourably with other work matching census data on names in other contexts (Abramitzky et al., 2021; Fouka et al., 2022). The match rate by category is given in Table A19 and Table A20. Our match is highest for male GP head candidates for 2016 and lowest for women GP head and ward member candidates in 2021. 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. Appendix section B describes the matching in more detail.

3.2 Quantifying Politician Quality: Education as a Proxy for Competence

Measuring politician competence is inherently challenging, as direct measures of underlying ability are typically unavailable. We therefore use years of formal education as our primary proxy. While widely used in the literature, its validity is not universally accepted. In this subsection, we begin by discussing the advantages and limitations of this approach, and then provide empirical evidence from both the policy implementation perspective and citizens' perceptions of competence, showing that education is a parsimonious and meaningful proxy in our context.

Education is a natural starting point because much of the existing literature on the backgrounds and performance of politicians uses it as a proxy for competence. Using this metric, studies find that politicians tend to be more educated than the general population in countries like Sweden (Dal Bó et al., 2017), Denmark (Dahlgaard et al., 2020) and the United States (Thompson et al., 2019).

In the existing literature, there are three main arguments in favour of education as competence. First, it correlates with earnings and civic engagement, indicating that educa-

[4] These are not unique candidates - since candidates could repeat across the two elections.

tion expands individuals' skills and ability as well as inculcates civic values (Besley and Reynal-Querol, 2011; Besley et al., 2005). Second, politicians' education level has been shown to affect performance on actual policy outcomes: studies have provided empirical evidence for education being associated with higher GDP growth (Besley et al., 2011), higher efficiency on certain indicators (Sørensen, 2023), increase in the provision of local public goods (Martinez-Bravo, 2017), effective inflation control (Göhlmann and Vaubel, 2007), better performance in providing local public goods and lower abuse of political power (Besley, 2005). The third and a more practical advantage of using education as competence comes from the fact that this is often the only data available for the entire candidate pool (Gulzar, 2021).

Education as a proxy for competence, however, is not universally accepted in the literature. Some studies find weak or null links between education and political performance, or show heterogeneous effects (Carnes and Lupu, 2016; Curto-Grau and Gallego, 2021; Dreher et al., 2009; Lahoti and Sahoo, 2020). Even when affirmative action policies such as reservation of seats for historically marginalized groups such as women and lower castes crowds in candidates who obviously have lower levels of educational attainment, there is no compromise with decision-making quality (Duflo, 2005). Other studies warn that education may simply reflect elite membership, rather than ability, particularly in unequal societies where educational attainment is tightly correlated with caste, gender, or wealth (Dal Bó et al., 2017). In other words, more educated politicians may come from more elite or wealthier households and, therefore, be less representative of the underlying population. In such cases, education might reinforce elite capture, rather than serve as a channel for positive selection.[5]

We take these concerns seriously and pursue three strategies to address them, particularly in the absence of any measures of cognitive, leadership or direct earnings in our context. First, we examine whether education predicts politicians' knowledge of scheme implementation – a key indicator of the capacity to govern. Second, we show that more educated politicians have broader political networks and stronger party connections, which facilitate access to higher-level officials and political support. Third, we show that citizens themselves treat education as a salient indicator of competence in our context.

Education and knowledge of governance. A central reason we use education as a measure of competence is that, in our context, the ability to implement policies depends critically on politicians' knowledge of rules, procedures, and scheme-specific requirements (Bamezai et al., 2024). Implementation outcomes are ultimately a joint function of competence and preferences: while preferences determine which projects politicians prioritise, competence determines whether and how those priorities are translated into action.

[5]To address this concern, researchers compare politicians with their parents, siblings and other elite professionals. For example, Dal Bó et al. (2017) find that elected politicians have higher cognitive, leadership, and earnings scores than their siblings, indicating that ability, rather than family background, is the key selection criterion. Similarly, Dahlgaard et al. (2020) find that politicians in Denmark have higher incomes and better education compared to their non-politician siblings, suggesting social mobility into a political career is relatively high.

To validate education as a proxy for this implementation capacity, we use a close-election regression discontinuity design and find that more educated politicians possess substantially greater knowledge of scheme implementation across a range of domains, including housing, infrastructure, and service delivery (Table A5).

Education and state networks. A second dimension along which education may reflect competence is through politicians' ability to build and leverage political networks. In our context, effective governance often requires coordination with higher-level officials and access to administrative and party resources, which can shape both the feasibility and effectiveness of policy implementation.

Using the same close-election RD design applied to ward members, we find that more educated politicians have broader political networks and stronger party connections (Table A6). In particular, they are more likely to have contact information (e.g., personal phone numbers) for higher-level government officials–such as the Block Development Officer, Block Panchayat Raj officials, and the GP head–facilitating communication and access to administrative support. They are also more likely to receive party support and to be actively involved in party activities, suggesting greater integration within the political system.

Together, this evidence establishes a direct link between education and a key indicator of governance – the knowledge of how to go things and the ability to navigate administrative processes and get things done – abstracting from politicians' distributive preferences or policy priorities.Further details on the close-election RD design, validity, and robustness checks are provided in the Appendix **sec:close-election-ward**.

Voters see education as a measure of competence. Next, for education to matter for political selection, however, voters must also perceive it as a relevant indicator of competence. We, therefore, next examine whether citizens in rural Bihar value education when evaluating potential Panchayat candidates.

We show evidence on citizen preferences for education. In a survey of over 1,100 respondents across 48 GPs in Bihar (see survey details in the Appendix F), we asked citizens what qualities they valued most in a Panchayat candidate. As shown in Figure 6, a majority of respondents identified education as one of the top three attributes they look for when voting - ranking above other traits such as household wealth, caste, or prior political experience and performance.

To probe these preferences more rigorously, we conducted a forced-choice conjoint experiment embedded in the same in-person survey (see Appendix F). Each respondent was shown paired profiles of two hypothetical individuals residing in a village similar to their own. The profiles varied randomly across seven attributes, one of which was the individual's education level. Figure 7 presents the Average Marginal Component Effects (AMCE) of the seven attributes from the conjoint experiment. The results reveal a strong, monotonic pattern: conditional on having similar background characteristics as

wealth and occupations, individual profiles that were randomly assigned higher educational qualifications were perceived as both more likely to contest and more likely to win Panchayat elections. Relative to a 5th-grade pass candidate, a 10th-grade pass candidate was viewed as 5.75 percentage points more likely to contest ($p < 0.000$), and a B.A. holder as 13.41 percentage points more likely to contest ($p < 0.000$). Conditional on both contesting, the corresponding effects for perceived likelihood of winning were even higher – 6.13 percentage points for a 10th-grade pass ($p < 0.000$) and 15.52 percentage points for a B.A. pass ($p < 0.000$).[6] These effects are large, statistically significant, and consistent across subgroups. They suggest that education is not only a credential but a salient signal of competence in the eyes of voters across rural Bihar.

Taken together, we show that education is associated with politicians' knowledge of administrative procedures – a key indicator of implementation capacity – and is also valued by citizens as a legitimate indicator of a candidate's ability. By capturing both the capacity to govern and voter-relevant signals of competence, education provides a parsimonious and grounded – both theoretically and empirically – proxy for competence in rural Bihar.

3.3 Benchmarking Social Representativeness: An Intergenerational Approach

We follow (Dal Bó et al., 2017) in using the education of fathers as a measure of the representativeness of male candidates. Since women typically leave their natal villages upon marriage, we are unable to match most adult female candidates with their fathers. No such issue arises for men. Given that male candidates' own education and wealth outcomes are often correlated, we use the relative education of their fathers – compared to other males in the same birth cohort – as a more appropriate proxy for the candidates' representativeness. The measure allows us to assess whether candidates emerge from households that were historically typical or disadvantaged in terms of educational access.

In our data, we are able to match 41.13% of male candidates to their fathers using the 2011 Census village rolls. All representativeness analyses are restricted to this sample. Panel B of table A14 presents covariate balance tests comparing matched and unmatched candidates. On average, matched candidates are slightly more educated (by 0.89 to 1.42 years) and more likely to belong to non-SC, non-ST castes than those for whom no father match was found.

While father's education is our primary measure of representativeness, we also use the conjoint experiment discussed in Section 3.2 to examine whether voters prefer their own type. In addition to education, the profiles in the conjoint experiment varied by household caste category (Upper Castes, Backward Castes, Scheduled Castes). We find modest but consistent evidence that respondents are more likely to favour candidates who share their own caste category (Figure 8).

[6] We find similar results when the respondents are asked a third question "If these two individuals were to contest an election, who would you vote for?" but we do not include those here because of suspected social desirability bias.

These effects are smaller than those for education but suggest that voters value representativeness along caste dimensions as well. We view these findings as complementary evidence that representativeness is an actively desired trait from the voters' perspective, particularly in rural India's socially stratified context.

4 Result #1: Who Becomes A Politician?

Building on the data construction and empirical challenges discussed in the previous section, we begin our analysis by documenting the baseline patterns of political selection in rural Bihar. We show that candidates are more educated than the general population, individuals of the same caste and gender, and their own siblings. We also find that they come from more advantaged family backgrounds. These patterns point to a trade-off between competence and representativeness: while politicians are positively selected on education, they are drawn disproportionately from more elite households. This is consistent with a form of "partially exclusive meritocracy."

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' 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 7.2, 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 rural Bihar have completed high school, this number is three times as high for politicians at 18.5%. These patterns are robust to restricting attention to only adult citizens (Table 1).

We find that politicians are more likely to come from highly-educated groups in society. As Table A15 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. However, what is particularly striking is how positively selected politicians are relative to the average citizen within their social groups. 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.

Panel (a) of Figure 5 presents the citizen mobility curve and the candidate selection frontier. The citizen mobility curve describes the relationship between father's education (representativeness) and son's education (competence) in the underlying GP population. The candidate selection frontier, describes the same relationship among individuals who enter politics. The candidate selection frontier lies clearly above the citizen mobility curve, indicating that candidates are more educated than citizens at any given level of parental background. This is another way of showing clear evidence of positive selection into politics on the basis of competence.

Competence and Electoral Performance. In Table A16, we see that (i) candidates are more competent in elections for higher-level political positions and (ii) winners are more competent than losing candidates. For instance, while the average politician is 0.67 SD

more educated than citizens, winners in GP president elections are a full 1.4 SD more educated than the average citizen. This translated to an additional 5.6 years of schooling. 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 in the past among all men was lower than it is 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 do come from advantaged communities and even within these from advantaged families.

However, from the magnitudes, it is also clear that the competence gap between politicians and the average citizen is much larger than the representativeness gap between their family and the average citizen. We depict this difference more clearly in Figures 1 and 4. Thus, although candidates come from more elite backgrounds than the average citizen, we see significant positive selection in terms of educational attainment even conditional on family background.[7]

4.3 Competence-representativeness Trade-off

Figure 2 shows that there appears to be a tradeoff between competence and representativeness. While politicians are on average more competent and come from advantaged families compared to the average citizen, we see that candidates for higher political offices (GP heads) are more competent and from more elite backgrounds than candidates for lower offices (ward council). Similarly, winners are more competent and from more elite families than losing candidates.

Furthermore, we show directly in Panel (a) of figure 5 using data on all the over 500,000

[7]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.

politicians in our sample that more competent politicians are generally less socially representative (i.e. they come from more elite families).

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 "partially exclusive meritocracy" 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 policy characteristics, we augment the baseline specifications in (1) and (2) by interacting the key indicators with a policy 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 above. 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 only on male candidates in caste unreserved GPs. This is for two reasons: first, as before, we want to focus on "open" seats, where members of all castes can technically contest. Moreover, our representativeness results, as indicated previously, can only be ascertained for male members.

5.1 Local Economic and Structural Constraints

We begin by analyzing the role of the local economic and structural environment. As detailed in Table A1, we utilize six measures constructed from the SECC to capture these baseline conditions: asset wealth, land inequality, educational mobility, ethnic fractionalization, and geographic remoteness. This allows us to assess how the tradeoff between competence and representativeness shifts in response to local resource constraints and structural barriers, prior to accounting for the influence of social identity or political parties. Table A2 shows how selection varies for GP head candidates in caste unreserved and

gender unreserved GPs.

We document the following takeaways: first, that political selection does not vary with a GPs poverty status or how remote it is. Second, places with high land inequality may have elite capture: this is reflected in the relatively high coefficient on candidates' father's education in column (5). In unequal GPs, candidates are more educated than the average citizen and come from more elite families. Thus, the competence-representativeness tradeoff is worse in these GPs.

The results also point to the salience of the trade-off between competence and representativeness: across GP features, factors that may improve (worsen) competence of candidates, seem to reduce (increase) their representativeness too.

The main takeaway from Table A2 is that the positive selection story is true across very different types of villages: while selection varies to some extent by village characteristics, politicians tend to always remain positively selected on education in the overall population. In addition, we continue to see selection of elite households into politics. Thus, the competence-representativeness tradeoff remains stark.

5.2 Demographic Composition and the Returns to Group Size

We now explore the relationship between diversity and political selection. As columns (2) and (3) in table A2 show, sub-caste heterogeneity matters: ethnic polarization is negatively correlated with selection, but ethnic fractionalisation is positively correlated with selection. There are many explanations for these. For instance, ethnic polarisation is highest when there are two large sub-caste groups: previous work has suggested that large groups could see worse selection (Banerjee and Pande, 2007). On the other hand, since greater fractionalisation implies smaller groups, this could result in better selection. Note that the coefficients on both these indices are small, suggesting that while sub-caste heterogeneity could co-vary with selection, the magnitude of these relationships may be modest.

Next, we investigate the relationship between group size and the selection of politicians. Table 2 presents results. On average, we find a weak relationship between group size and political selection, but this masks heterogeneity. Among the majority Hindu population, politicians from numerically larger social groups in a GP are weakly less competent. However, they are from more elite families. Thus, the competence-representativeness tradeoff worsens as majority groups get larger.

For minority Muslim politicians, however, we find that as group size increases, competence actually improves, while representativeness is unchanged. Thus, for minority politicians, the competence-representativeness tradeoff improves as minority groups get larger. Perhaps the best explanation for this is Banerjee and Pande (2007)'s theory that majority group politicians are essentially sure-win candidates when their group is very large, while minorities have almost no chance when their group is too small.

5.3 Institutional Factors: The Role of Political Parties

We now turn to whether parties affect selection. We begin by noting that parties are not formally allowed in local elections in Bihar. However, it is well documented that they do play an informal role in the state (Dunning and Nilekani, 2013). We use data from the baseline survey of (Bamezai et al., 2024) to collect data on party presence across blocks in Bihar. Bamezai et al. survey over 7700 ward members across the state and ask if their campaigns were supported by parties. We calculate share of candidates in a block who report being supported by a party and use it as a proxy for party presence across the state. We now test if candidates in areas with greater party presence differ in their competence and representativeness. We estimate a regression analogous to equations (3) and (4) by interacting the candidate and father-of-candidate indicators with the block-level share of party involvement.

Our results are shown in Tables 2. The main takeaway from the tables are that parties increase competence, but reduce representativeness of candidates contesting local elections.

6 Result #3: Consequences of Political Selection

Having established a systematic tradeoff between competence and representativeness in political selection, we now evaluate how these patterns translate into governance outcomes. We examine how leader competence—measured by educational attainment—shapes governance along two dimensions. First, we study whether more competent leaders better align GP development plans with citizen preferences. Second, we assess whether competence differentially affects performance across policy types, comparing skill-intensive infrastructure projects with pro-poor welfare schemes.

To isolate these effects, we employ a close-election regression discontinuity (RD) design comparing winners who narrowly defeated less educated opponents. We find that while higher competence significantly improves the implementation of infrastructure—such as roads and tap water delivery—it simultaneously reduces alignment with the preferences of disadvantaged groups, resulting in fewer welfare projects and NREGS benefits for Scheduled Caste (SC) households.

6.1 Competence and Alignment with Citizen Preferences

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.

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

We combine election data with administrative information on GP Development Plans (GPDPs). Outcomes are measured at the GP $\times$ scheme level and include (i) the number of schemes proposed in GP plans and (ii) planned spending on schemes. These outcomes capture allocation decisions made by GP leadership during the post-election planning period.

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. Specifically, 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, we re-estimate the RD using alternative education cutoffs, including

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

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:

$$\begin{aligned} Y_{gs} = \mu_s + \theta_1 Pref_{gs} + \theta_2 T_g + \theta_3 (T_g \times Pref_{gs}) \\ + \gamma_0 f(MoV_g) + \gamma_1 (T_g \times f(MoV_g)) + \varepsilon_{gs}. \end{aligned} \tag{5}$$

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, and report robust confidence intervals. Standard errors are clustered at the GP level. We show that GPs above and below the threshold are balanced in Table **??**.

6.1.2 Results

Table 6 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.

Finally, Table A13 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.

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.

6.2 Competence and Policy Implementation Outcomes

Having shown that more competent leaders produce GP development plans that are less aligned with the preferences of marginalized SC citizens, we next examine whether this competence-representativeness trade-off extends to realised implementation. We distinguish between governance tasks that primarily require administrative and technical competence and those that depend more directly on representativeness and distributive responsiveness.

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. As documented earlier (Table A5), more educated leaders demonstrate greater administrative knowledge and broader professional networks outside their GPs—precisely the capacities we argue constitute competence. *Ex ante*, such advantages should improve performance in infrastructure provision.

For tasks that privilege representativeness, we 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. Our preferred measure of representational responsiveness is the number of SC person-days generated under NREGA, which captures the extent to which leaders convert the preferences of marginalized SC citizens into realised programme benefits.

6.2.1 Data and Identification Strategy

NREGA MIS: We draw on administrative data from the NREGA Management Information System (MIS), which reports the number of person-days generated at the GP level, both in aggregate and disaggregated by social group. Our primary outcomes are total person-days generated and person-days generated for Scheduled Castes. We restrict attention to person-days generated between 2016 and 2021, corresponding to the tenure of GP heads elected in 2016.

Mision Antyodaya Data Our second source of administrative data is the Mission Antyodaya survey, a nationwide village-level assessment conducted by the Government of India to benchmark rural infrastructure and service delivery. 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.

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 targeting.

We use Mission Antyodaya data to construct village-level measures of realised infrastructure provision. Specifically, we create: (i) an indicator for the availability of piped tap water, (ii) an index of road availability, and (iii) an index of drainage system availability. 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 realised development outcomes, we implement a close-election regression discontinuity (RD) design focusing on GP Head elections held in 2016. As discussed earlier, leader competence is proxied by formal education. Identification comes from GPs 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 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:

$$Y_g = \mu + \theta T_g + \gamma_1 MoV_g + \gamma_2 (T_g \times MoV_g) + \varepsilon_g, \tag{6}$$

where Y_g denotes the realised development outcome in GP 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.

6.2.2 Results

Consistent with earlier results, Table 5 shows that electing a more educated GP Head worsens outcomes for SCs. While overall NREGA work-days do not change significantly (Column (1)), SC work-days decline by roughly 616 days (Column (2)), corresponding to a 15% reduction relative to mean SC employment under NREGA. This pattern is consistent with greater social and economic distance between more competent leaders and the average SC voter, whose preferences place a higher weight on NREGA implementation.

In contrast, more competent leaders perform better on two of the three skill-intensive outcomes we measure: tap water and road indices increase significantly, with effect sizes on the order of one-fifth of a standard deviation (Columns (4) and (5)). We, however, find no effects on drains (Column (6)).

These gains suggest that competence enhances performance in domains where technical capacity and coordination are paramount, even as it weakens responsiveness in policy areas that are most salient to marginalised voters. Together, these findings indicate that competence improves governance where skills matter, but does not uniformly enhance distributive outcomes. Instead, the trade-off between competence and representativeness identified earlier is reflected not only in plans, but also in realised development outcomes.

7 Result #4: Policies to Improve Political Selection

In the previous sections, we described the competence-representativeness tradeoff among local politicians in Bihar. Citizens value both attributes, and both attributes affect governance: more competent (and thus less representative) politicians are better at representing elites and undertaking complex development tasks, but are worse at pro-poor programs or representing low-caste preferences.

In open seats, as we have seen, leaders tend to be competent but elite: in other words, the tradeoff skews towards competence and away from representativeness. A natural follow-up question is whether policy interventions that boost representativeness decrease competence, and whether, if they do, the tradeoff is "worth it".

We focus on a widely used policy to improve representativeness: affirmative action through political reservations. Affirmative action policies in the form of explicit quotas for disadvantaged groups – called "reservation policies" or simply "reservations" – have been a common tool to increase representativeness among the political class in various countries and are present at various levels in India (Pande, 2003).

Specifically, we causally estimate how reservations for (i) Scheduled Castes (SCs) and (ii) women affect the competence–representativeness trade-off among political leaders. A large literature studies the downstream policy consequences of political reservations (Pande, 2003; Duflo, 2005; Chin and Prakash, 2011), including in Bihar (Dunning and Nilekani, 2013; Kumar and Sharan, 2025). By contrast, far less is known about how reser-

vations shape political entry itself – and, in particular, how they alter the balance between leader competence and representativeness.

Reservations mechanically increase representativeness by expanding political access to historically excluded groups. At the same time, given longstanding disparities in access to education, wealth, and social capital between women(/SCs) and men(/non-SCs), one might *ex ante* expect a decline in average competence among elected leaders under reservation.

Two deeper questions, however, are less straightforward *ex ante*. First, how large is the competence trade-off relative to the gain in representativeness? That is, does reservation substantially erode competence, or does it modestly adjust the composition of leaders while achieving large representational gains? Second – and more importantly from a policy perspective – how do competence and representativeness shift *within* the reserved group? In other words, does reservation primarily enable entry by minorities who are relatively elite within their own group, or does it broaden access to include those who are closer to the average citizen within that group?

This within-group question is critical for understanding the distributional and normative implications of affirmative action. If reservation selects disproportionately from the upper tail of the minority group's education and wealth distribution, it may increase descriptive representation without substantially narrowing social distance between leaders and constituents. By contrast, if it expands entry among less-elite minorities, it may fundamentally alter whose preferences are articulated and acted upon. Empirically addressing this question requires measuring competence and representativeness relative to the full distribution of citizens within each group, an exercise that is typically infeasible. As shown earlier, however, our data allow us to construct precisely such measures.

7.1 SC reservation

Since 2006, Bihar reserves slightly over a sixth of the GP head seats in favour of SCs. Each GP is reserved for a period of 10 years. We study the effects on political entry in reserved and unreserved GPs in the 2016 and 2021 elections.

7.1.1 Identification Strategy

Our empirical strategy exploits the population-based algorithm used to assign reservations for GP head positions for SCs. This algorithm has been used to estimate the causal effects of reservation in Bihar elsewhere (Kumar and Sharan, 2025). In each administrative block – a cluster of roughly 15–16 GPs – the algorithm ranks GPs by their SC population and reserves head positions for those above a block-specific threshold. As a result, GPs just above this threshold are substantially more likely to be reserved for an SC head than those just below it: in practice, crossing the cutoff increases the probability of reservation by about 80 percentage points.

This rule governs reservations in the 2006–2016 cycle. For the subsequent 2016–2026 cycle, the State Election Commission first excludes all GPs that were reserved in the previous cycle and then reapplies the same population-based algorithm to the remaining GPs. Our analysis focuses on GPs lying close to the relevant population thresholds in each cycle. Comparing outcomes for GPs narrowly above and below these cutoffs allows us to causally estimate the effects of SC reservation using a fuzzy regression discontinuity design.[9]

We estimate a fuzzy regression discontinuity design. Formally, for candidate i in GP g and block b, with SC population $SCPop_{gb}$, we specify:

$$\begin{aligned} Reserved_{gb} = \gamma_0 + \gamma_1 \mathbb{1}\{SCPop_{gb} \geq T_b\} + \gamma_2 (SCPop_{gb} - T_b) \\ + \gamma_3 (SCPop_{gb} - T_b)\, \mathbb{1}\{SCPop_{gb} \geq T_b\} + \mu_b + \lambda_t + \eta_{gb}. \end{aligned} \tag{7}$$

$$\begin{aligned} Y_{igb} = \beta_0 + \beta_1 Reserved_{gb} + \beta_2 (SCPop_{gb} - T_b) \\ + \beta_3 (SCPop_{gb} - T_b)\, \mathbb{1}\{SCPop_{gb} \geq T_b\} + \mu_b + \lambda_t + \epsilon_{igb}. \end{aligned} \tag{8}$$

Here T_b denotes the block-specific SC population cutoff, $lambda_t$ denotes election-year fixed effects, and μ_b denotes block fixed effects. The unit of observation is the candidate, while the treatment indicator varies at the GP level. Standard errors are clustered at the GP level.

7.1.2 Threats to Validity

Gender Reservation Imbalance: One feature of the RD design – discussed in detail in (Kumar and Sharan, 2025) – is that gender reservation is imbalanced across the RD threshold. This imbalance is a mechanical artefact of the reservation algorithm. (Kumar and Sharan, 2025) address this issue in two ways: (a) by controlling for whether a seat is gender reserved, and (b) by dropping gender-reserved seats on either side of the threshold.

Our main specification for competence follows approach (a). For representativeness, we follow approach (b), mainly because this choice has the additional advantage that we do not observe fathers' education for women candidates, which prevents us from constructing comparable measures of representativeness for women leaders. Table A7 shows balance across the RD threshold.

7.1.3 Results

Candidate Selection Frontier We begin by examining, visually, how reservation affects where the average candidate sits relative to the Citizen Mobility Curve and the Candidate Selection Frontier. Panel (a) of Figure 5 plots both curves: the lower (orange) line

[9] The reservation rule has been described in more detail in Section C.1.

traces the relationship between father's education and son's education in the underlying GP population – the Citizen Mobility Curve – while the upper (green) line traces the same relationship among political candidates – the Candidate Selection Frontier. Because political entry involves positive selection, the Candidate Selection Frontier lies above the Citizen Mobility Curve: at any given level of parental background, candidates are more educated than citizens.

Panel (b) indicates where the average candidate and winner is on the frontier. The unreserved candidate is positioned far to the right on the x-axis – coming from a privileged background – and sits high on the y-axis, reflecting strong positive selection into politics. Electoral competition reinforces this: unreserved winners are more competent than the candidates from whom they are drawn, but less representative, consistent with the competence-representativeness trade-off documented earlier.

SC candidates present a strikingly different picture. They come from substantially more representative backgrounds than their unreserved counterparts, lying near the left tail of the father's education distribution. Relative to the Citizen Mobility Curve, SC candidates sit above it, indicating positive selection relative to the GP population. Relative to the Candidate Selection Frontier, however, they lie below it – suggesting that, at comparable levels of parental background, SC candidates appear somewhat less competent than the typical political entrant.

A natural concern is that this gap reflects differences in underlying mobility regimes rather than weaker selection into politics. If educational mobility is more constrained within SC communities, SC individuals will have lower educational attainment for any given parental background relative to the pooled population. Benchmarking against a frontier estimated over all castes will then understate their competence within their own social group.

Panel (c) addresses this directly by re-estimating both curves after standardising education within caste $\times$ GP, so that all comparisons are made relative to others from the same caste group in the same village. Two findings emerge. First, SC candidates remain more representative than non-SC candidates even within their respective caste groups, confirming that reservation draws candidates from especially representative social backgrounds. Second, once mobility differences are accounted for, SC candidates move above the Candidate Selection Frontier. The apparent competence gap in the pooled comparison therefore largely reflects differences in mobility regimes across caste groups, not weaker selection into politics among SCs.

The pattern for election *winners* differs strikingly across seat types. Among non-SC seats, winners are more competent but less representative than the candidates from whom they emerge – the familiar trade-off. Among SC seats, the pattern reverses: SC winners are simultaneously more representative *and* more competent than the SC candidates from whom they are drawn. Electoral selection within reserved seats thus moves in both directions at once. Reservation therefore substantially increases descriptive representation while still producing leaders who are positively selected relative to the pool of potential

entrants – and the two goals need not be in tension.

Having inspected these patterns visually, we now establish them causally via the RD design outlined above.

7.1.4 Regression Discontinuity Design

The regression discontinuity compares GPs that are similar in all respects except that one has reserved status and the other does not. Table A7 in the appendix confirms that GPs on either side of the threshold are balanced on a host of observables, and Panel (a) of Figure A1 confirms that the McCrary test is satisfied.

Candidates: Compared to all citizens in the GP. Panel A of Table 3 benchmarks SC candidates in reserved GPs against candidates in unreserved GPs, with all outcomes standardised relative to the distribution of all GP citizens. The average candidate in an unreserved GP is approximately 1 SD more educated than the average citizen. Reservation reduces this premium by 0.62 SD, but the average SC candidate in a reserved GP still remains around 0.38 SD more educated than the average citizen – reservation does not eliminate positive selection, it attenuates it.

Despite this competence advantage, SC candidates come from households that closely resemble the average GP household. Unreserved candidates, by contrast, are substantially more elite: their households score around 0.55 SD higher on a PCA-based wealth index, are 0.20 SD more likely to pay income or professional taxes, and are 0.26 SD more likely to contain a salaried government employee (Columns 2–4). Reservation sharply attenuates all of these gaps. If anything, SC candidates are marginally less wealthy than the average citizen.

The clearest evidence on representativeness comes from fathers' education (Column 5). In unreserved GPs, candidates' fathers are on average 0.44 SD more educated than the average citizen. Under reservation, this reverses: fathers of SC candidates are approximately 0.30 SD *less* educated than the average citizen. SC candidates are thus drawn from markedly marginalised household backgrounds, even as they themselves remain more competent than the average GP resident.

Candidates: Compared to citizens within own caste. Panel B of Table 3 benchmarks each group of leaders against citizens from their own caste. Along the competence dimension, SC leaders are strongly positively selected: they are on average 0.85 SD more educated than the average SC citizen, compared to 0.95 SD for unreserved leaders relative to non-SC/ST citizens. The selection premium is large in both cases.

Patterns of representativeness are more nuanced. In unreserved GPs, the average candidate household scores approximately 0.50 SD higher on the wealth index than the average non-SC/ST household. This gap narrows by about 0.16 SD under reservation, indicating

that reservation attenuates – but does not fully eliminate – selection toward relatively elite households within groups. Because household wealth partly reflects leaders' own competence rather than pre-existing advantage, Column 5 focuses on fathers' education as our preferred representativeness measure. In unreserved GPs, candidates' fathers are 0.38 SD more educated than the average non-SC/ST citizen. Under reservation, this advantage is almost entirely eliminated: fathers of SC candidates closely resemble the average member of their own social group.

Taken together, reserved leaders are strongly selected on competence relative to their own group, while remaining far more representative in terms of household background. Within own-group comparisons, reservation produces candidates who are nearly as competent as unreserved leaders but substantially more representative – helping explain why many studies find little evidence of an equity-efficiency trade-off under caste-based reservation (Pande, 2003; Dunning and Nilekani, 2013).

Winners: Compared to all citizens in the GP. Turning from candidates to winners, column (1) of Panel A of Table A3 shows that SC winners are on average 0.88 SD less educated than unreserved winners and come from substantially more marginalised backgrounds (columns (2) - (4)). SC winners are poorer than the average citizen in the GP, unlike their non-SC counterparts who are 0.9 SD higher in wealth score (column 4). Column 5 shows that, on representativeness, SC candidates' fathers are 0.69 SD less educated than the average citizen, compared to +0.64 SD for unreserved winners (β = - 1.33 SDs). Nonetheless SC winners are 0.41 SD more educated than the citizens they represent (column 1; non-SC mean is 1.29, β = -0.88), indicating that even within reserved seats, electoral competition selects positively on competence.

Winners: Compared to citizens within own-caste. Panel B of Table A3 turns to comparisons with own-caste members. Within their own caste group, column 1 shows that SC winners are significantly more educated than the average SC citizen (0.84 SD; non-SC winners are 1.21 SD more educated than non-SC winners; β = -0.375), confirming positive selection on competence. Their households are wealthier than the average SC household, but they are less likely to contain a government employee or an income-tax payer – suggesting they come from households that are relatively advantaged in assets but not in the kinds of institutional connections that typically characterise unreserved political elites.

On representativeness, the results are striking. Column 5 shows that SC winners' fathers are 0.36 SD *less* educated than the average SC citizen (non-SC winners' fathers are 0.56 SD more educated than the average non-SC citizen; β = -0.92) – meaning they are more representative not only than unreserved winners, but than the average SC candidate. Combined with the competence finding, this confirms the pattern visible in Panel (c) of Figure 5: within reserved seats, electoral selection moves simultaneously toward greater competence *and* greater representativeness relative to the SC candidate pool. The usual trade-off between the two does not apply.

7.2 Gender Reservation

Next, we turn to gender-based reservations and their implications for the competence–representativeness trade-off.

We measure competence as before – by years of education. However, unlike in the caste context, we lack direct measures of representativeness for women candidates, since information on fathers' education is unavailable.

We therefore shift focus to household characteristics, asking how the households of women contesting in gender-reserved seats differ from those of candidates contesting in unreserved seats. *Ex ante*, the direction of these differences is ambiguous. If women candidates primarily serve as proxies for male relatives[10], gender reservation may leave the pool of contesting households largely unchanged. If not, variation in women's agency across social groups suggests a different possibility: when social or caste norms restrict political participation for women from certain groups, these constraints should shape which households enter politics under gender reservation.[11]

7.2.1 Identification Strategy

To estimate the effects of gender reservation on political selection, we circumvent the endogeneity of GP reservation status with a fuzzy regression discontinuity (RD) approach. The Bihar government mandates that up to 50% of GP seats must be reserved for women. Within each block, the governments first determines which GPs should have seats reserved for SC, ST, OBC or none. Subsequently, within each grouping, the government reserves up to 50% of seats for women. The details are presented in the Appendix (section C.2), but we offer a stylized explanation below.

For gender reservation within GPs reserved for SC (STs), GPs are arranged in descending order of SC (ST) population and the 50% of GPs are reserved for women. This allows for a clear SC (ST) population threshold, above which GPs are reserved for women. Now, for EBC reserved seats, GPs are arranged in descending order of total population and the top 50% of seats are reserved for women. For caste-unreserved GPs, GPs are arranged in descending order non-SCST population and the top 50% of GPs in the list are reserved for women. Thus, population thresholds exist for each of these groups too.

The running variable is simply the distance of the relevant population to the threshold – which is defined by the mean population of the last GP to be reserved and the first GP that missed out on being reserved. By comparing GPs that narrowly selected for reservation to those that narrowly lost, we are able to capture the causal affects of gender reservation on political selection outcomes using the following two-stage instrumental

[10] (Amar et al., 2025; Heinze et al., 2025; Mori et al., 2025)

[11] Consistent with this view, Cassan and Vandewalle (2021) show that gender reservation crowds in low-caste households and, in turn, aligns government functioning more closely with low-caste citizen preferences.

variables specification:

$$
\begin{aligned}
Reserved_{gb} = \gamma_0 + \gamma_1 \mathbb{1}\{FemalePop_{gb} \geq T_b\} + \gamma_2 (FemalePop_{gb} - T_b) \\
+ \gamma_3 (FemalePop_{gb} - T_b)\, \mathbb{1}\{FemalePop_{gb} \geq T_b\} + X'_{gb} + \mu_b + \lambda_t + \eta_{gb}. \\
Y_{igb} = \beta_0 + \beta_1 Reserved_{gb} + \beta_2 (FemalePop_{gb} - T_b) \\
+ \beta_3 (FemalePop_{gb} - T_b)\, \mathbb{1}\{FemalePop_{gb} \geq T_b\} + X'_{gb} + \mu_b + \lambda_t + \epsilon_{igb}.
\end{aligned} \tag{9}
$$

Here T_b denotes the block-specific Female population cutoff, $lambda_t$ denotes election-year fixed effects, and μ_b denotes block fixed effects. Considering that the thresholds are determined separately for GPs with different caste reservation status, a vector of controls on GP caste reservation X'_{gb} is also included. The unit of observation is the candidate, while the treatment indicator varies at the GP level. Standard errors are clustered at the GP level.

Following Calonico et al. (2019), we estimate a fuzzy RD with optimal bandwidths. We pool all the running variables together to estimate an "overall" effect of gender reservation. We also report results within separate caste reservation groups.

Threats to Validity: The RD design – like all similar designs – assumes that GPs on either side of the threshold are comparable. Table A8 shows balance. Moreover, Figure A1 shows that McCrary test and assures that there is no manipulation of the running variable around the threshold.

7.2.2 Results

Impact on Candidate Pool Table 4 reports the effects of gender reservation on political selection. Panel A compares candidates elected in gender-reserved GPs to candidates elected in unreserved GPs, with all outcomes standardised relative to the distribution of all citizens. Gender reservation is associated with a large decline in the elected candidate's own education: women elected from reserved GPs have, on average, 0.77 SD fewer years of schooling than candidates elected from unreserved GPs. As a result, female leaders are only marginally more educated than the typical citizen in the GP.

Column 2 benchmarks candidates relative to their own gender. The average candidate in unreserved seats is about 0.65 SD more educated than the average male in the GP. The coefficient of -0.458 on female reservation indicates that the average woman candidate is 0.19 SD more educated than the average woman in reserved GPs. Thus, while gender reservation substantially reduces the educational advantage of leaders relative to unreserved candidates and relative to citizens overall, it continues to select women who are meaningfully more educated than other women.

However, focusing solely on the elected woman's own education may overstate the im-

plied loss in effective competence. In our setting, proxy leadership is prevalent and administrative responsibilities are frequently mediated through other household members. If governance capacity is embedded within the household rather than the individual alone, then household-level human capital becomes the more relevant measure of effective competence.

In the Bihar context, Amar et al. (2025) show that the overwhelming majority of proxy governance is done via the husband. Column (3) compares the husband of the woman leader on competence with the governing unreserved leader. The impact of reservation is negative (β = -0.15), suggesting that even husbands are less competent than the governing unreserved candidate.

Columns (4) - (6) of Table 4 show that gender reservation brings in less elite and poorer households: gender-reserved candidate households are less likely to have members holding a government job (-0.08 SD), report paying income tax (-0.07 SD) and have lower PCA-based asset index scores (-0.09 SD). They are, however, still more elite than the average household in the GP, since the estimates do not entirely wipe out the substantial advantage unreserved candidate households have.

Impact on Winners Panel B of Table 4 shows that the patterns documented for candidates largely persist among elected winners. Under gender reservation, the education of the elected GP head falls sharply: Column 1 shows that years of schooling decline by 0.93 SD relative to unreserved leaders and by 0.65 SD even when we restrict comparisons to members of own gender (Column (2)). However, the effects are smaller when we expand the lens to include the husband of the winner: the husband is about 0.36 SD less educated than the winner in unreserved seats (column 3), indicating that the potential proxy winner is still less competent than the unreserved winner, but remains highly positively selected relative to the underlying population (the unreserved winner is 1.1 SD more educated than the average citizen; thus, the husband is still 0.74 SD more educated than the average citizen). Gender reservation also reduces elite household markers (columns 4- 6) – government employment (-0.22 SD), tax payment (-0.17 SD), and wealth (-0.18 SD) – but households remain somewhat more elite than citizens.

Gender reservation by caste Because the gender reservation RD pools across castes, we can examine whether the reduction in competence or the democratisation of the candidate pool is driven by particular caste groups. Table A4 presents these results. The reduction in candidate competence is visible across castes, though it is not statistically significant for SCs — likely reflecting the smaller sample around the SC-gender threshold. The democratisation of household backgrounds, by contrast, is concentrated among SC and unreserved women; the corresponding estimate for OBC women is small and insignificant. No single caste group therefore drives both effects simultaneously, and the overall gender-reservation results reflect broadly overlapping patterns across the three caste-groups.

7.3 Discussion: How Reservation Affects the Competence-Representativeness Tradeoff

Taken together, the evidence points to three commonalities between caste- and gender-based reservation, and one important difference. The first commonality is increased representativeness. Both forms of reservation broaden the candidate pool toward less elite households – in obvious and subtle ways. The obvious way is mechanical: reservation opens seats to marginalised groups that would otherwise be shut out. The subtler finding is that SC candidates are less elite relative to their own group than unreserved candidates are relative to theirs, and that gender reservation draws in women from less elite households.

The second commonality is a reduction in candidate competence. Both forms of reservation reduce the competence of the candidate pool. However, within-group, we see that SCs are more competent conditional on their father's education. Similarly, when we broaden the measure of competence for gender reservation to include husbands' education, the within-group reduction narrows.

The third commonality concerns winners. Across all seat types, winners are more competent than the candidate pool from which they emerge — electoral competition selects upward on competence regardless of reservation status. This holds for unreserved, SC, and women candidates alike.

The key difference between SC and gender reservation emerges on the representativeness of winners. Among SC seats, winners are not only more competent than the SC candidate pool but also more representative — their fathers are less educated than the average SC candidate's father. Electoral selection within reserved SC seats thus moves simultaneously toward greater competence and greater representativeness, dissolving the usual trade-off. Among women's seats, winners appear to come from wealthier households than the average woman candidate, which moves in the opposite direction on representativeness. However, we cannot rule out – even if it seems implausible – that the increased wealth owes to the fact that women winners (and their husbands) are more competent.

How do reservations affect the competence-representativeness trade-off? Taking the candidate selection frontier as the benchmark, SC reservation does not compromise competence: as shown above, once mobility differences are accounted for. SC candidates sit above the Frontier, and SC winners are simultaneously more competent and more representative than the candidate pool they emerge from. Thus SC reservation reduces the competence-representativeness tradeoff, as the gain in representativeness outweighs the decline in competence.

For gender reservation, the picture is less clean. Candidate competence falls, including within own-group comparisons, and we cannot make a definitive statement about whether the gains in representativeness outweigh these reductions. What we can assert — and what distinguishes this analysis from much of the prior literature — is that the representativeness gains from gender reservation operate along two distinct dimensions:

at the individual level, in that women replace men, and at the household level, in that the entering candidates come from less elite backgrounds. Having at-scale data allows us to document both margins with confidence, even if the net welfare implications remain difficult to sign.

8 Conclusion

This paper studies the nature of political selection in local democracies in a developing country context using rich data on incumbent politicians, political aspirants and citizens. It argues that selection can be quite different here than in more developed countries with mature democracies. For one, consistent with the literature, local politicians in developing countries come from more elite backgrounds than the average citizen, suggesting the exclusivity of the space (Bardhan and Mookherjee, 2000). However, contrary to more conventional understandings of these systems, local democracies continue to select politicians who are more "competent", as measured by education. This competence premium of politicians persists even when controlling for politicians' backgrounds. In other words, among the elites, it is the more educated who contest local elections. Voters too clearly show a preference for more educated candidates, with winning candidates in local elections typically being significantly more educated than others.

We also find that the competence-representativeness tradeoff is persistent across a variety of dimensions, including office type (GP Head and Ward members), electoral outcomes (winners, runners-up, and other candidates), caste group, and gender. A natural follow-up question to ask is why this tradeoff emerges in developing settings but not in more advanced democracies such as Sweden (Dal Bó et al., 2017). To explore this, we examine three potential explanatory factors: economic inequality, ethnic group size and dynamics, and party involvement. We find that the tradeoff worsens in contexts of greater economic inequality, as the majority group grows larger, and where party presence is stronger.

One might still question, however, whether this observed tradeoff translates into meaningful differences in actual policy outcomes. To address this, we leverage a close-election regression discontinuity design that compares villages led by more competent GP heads against those led by less competent ones within a narrow margin of victory. The results are policy-relevant and striking: more competent, elite politicians are less likely to be responsive to the demands of disadvantaged groups. Specifically, scheme implementation is less aligned with SC population preferences under elite leadership. Moreover, while villages led by more competent, elite politicians exhibit improved outcomes in skill-intensive programs — such as road construction and tap water delivery — they implement fewer welfare projects and deliver fewer NREGS benefits to SC households.

Lastly, we attempt to propose a feasible solution to this tradeoff by evaluating the effects of reservation for minority groups, one of the most popular institutional interventions in developing democracies. While reservations do not fully close the competence-representativeness gap, which is likely due to persistent mobility barriers, we find that

they can meaningfully shift the tradeoff frontier: relative to their own group, more competent and representative SC candidates now have greater opportunities to enter and contest. Moreover, winners from SC-reserved seats are even more positively selected on competence, while simultaneously coming from less elite and more representative backgrounds.

These findings have broader implications for the design of democratic institutions in developing countries where similar tradeoffs may arise. While reservation policies are often evaluated by examining the politician pool in isolation, our findings offer a more nuanced view by benchmarking candidate competence and elite background relative to the general citizen population. Such policies can serve as an effective institutional lever for broadening political representation without sacrificing quality — a lesson potentially relevant for developing democracies grappling with similar tradeoffs.

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 5200785, SSRN.

Bamezai, A., S. George, S. Hari, and R. S. Mamidipudi (2024): "Learning to Govern: The Impact of Politicians' Peer Networks," *Available at SSRN 4664964*.

Ban, R., S. Jha, and V. Rao (2012): "Who has voice in a deliberative democracy? Evidence from transcripts of village parliaments in south India," *Journal of development economics*, 99, 428–438.

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

Bardhan, P. and D. Mookherjee (2000): "Capture and governance at local and national levels," *American economic review*, 90, 135–139.

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012): "Female leadership raises aspirations and educational attainment for girls: A policy experiment in India," *science*, 335, 582–586.

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

Besley, T., J. G. Montalvo, and M. Reynal-Querol (2011): "Do educated leaders matter?" *The Economic Journal*, 121, F205–227.

Besley, T., R. Pande, and V. Rao (2005): "The Political Economy of Gram Panchayats in South India: Drawing Policy Implications from a Research Project on Democratically Elected Village Governments', World Bank," .

Besley, T. and M. Reynal-Querol (2011): "Do democracies select more educated leaders?" *American political science review*, 105, 552–566.

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.

——— (2017): "Do the effects of temporary ethnic group quotas persist? Evidence from India," *American Economic Journal: Applied Economics*, 9, 105–123.

Bhusal, B., M. Callen, R. Pande, S. Prillaman, and D. Singhania (2023): "Does Revolution Work? Post-revolutionary evolution of Nepal's political classes," .

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.

Cassan, G. and L. Vandewalle (2021): "Identities and public policies: Unexpected effects of political reservations for women in india," *World Development*, 143, 105408.

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

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 (2021): "How do the educated govern? Evidence from Spanish mayors," *Unpublished manuscript*.

Dahlgaard, J. O., R. T. Pedersen, and L. Pedersen (2020): "Must Voters Choose Between Political Competence and Descriptive Representation?" *Unpublished manuscript*.

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.

Dreher, A., M. J. Lamla, S. M. Lein, and F. Somogyi (2009): "The impact of political leaders' profession and education on reforms," *Journal of comparative economics*, 37, 169–193.

Duflo, E. (2005): "Why political reservations?" *Journal of the European Economic Association*, 3, 668–678.

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.

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., Z. Hai, and B. K. Paudel (2021): "Information, candidate selection, and the quality of representation: Evidence from Nepal," *The Journal of Politics*, 83, 1511–1528.

Gulzar, S. and M. Y. Khan (2021): "'Good Politicians': Experimental Evidence on Motivations for Political Candidacy and Government Performance," *Available at SSRN 3826067*.

Heinze, A. R., R. E. Brulé, and S. Chauchard (2025): "Who actually governs? Gender inequality and political representation in rural India," *The Journal of Politics*, 87, 818–822.

Kumar, C. and M. R. Sharan (2025): "The Short- and Long-Run Distributional Consequences of Political Reservation," Working paper.

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.

Mori, Y., R. Manjula, D. Rajasekhar, J. Goto, and T. Kurosaki (2025): "Who becomes a politician in a gender quota system? Evidence from India," *World Development*, 188, 106911.

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

Parthasarathy, R., V. Rao, and N. Palaniswamy (2019): "Deliberative democracy in an unequal world: A text-as-data study of south India's village assemblies," *American Political Science Review*, 113, 623–640.

Reynal-Querol, M. (2002): "Ethnicity, political systems, and civil wars," *Journal of Conflict Resolution*, 46, 29–54.

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

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

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," Tech. rep., National Bureau of Economic Research.

Witsoe, J. (2013): *Democracy against development: Lower-caste politics and political modernity in postcolonial India*, University of Chicago Press.

9 Tables

Table 1: Political Selection Among Candidates Contesting for Unreserved Seats

	Years of Schooling					
	Baseline (1)	Gender (2)	Gender + Age Cohort (3)	Gender + Age + Caste (4)	Gender + Household (5)	Father (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***	-0.135***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.018)
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 Party Affiliation and Ethic Group Size

	Candidate Edu.			Father Edu.	
	Baseline (1)	Gender + Age + Caste (2)	Father (3)	Baseline (4)	Gender + Age + Caste (5)
Panel A: Party Affiliation					
Candidate	0.574***	0.392***	0.149***		
	(0.026)	(0.025)	(0.031)		
Candidate X Party Involv.	0.461***	0.398***	0.145		
	(0.118)	(0.113)	(0.138)		
Candidate Father				0.198***	0.249***
				(0.025)	(0.024)
Candidate Father X Party Involv.				0.190*	0.132
				(0.111)	(0.111)
Observations	25,439,618	25,439,474	2,894,850	25,439,618	25,439,474
Panel B: Hindu Population					
Candidate	0.676***	0.420***	0.254***		
	(0.033)	(0.032)	(0.052)		
Candidate X Group Share	-0.005	0.033	-0.074		
	(0.037)	(0.035)	(0.058)		
Candidate Father				0.078**	0.068**
				(0.032)	(0.032)
Candidate Father X Group Share				0.186***	0.225***
				(0.036)	(0.035)
Observations	79,978,946	79,978,012	9,155,314	79,978,946	79,978,012
Panel C: Muslim Population					
Candidate	0.499***	0.382***	0.119***		
	(0.024)	(0.024)	(0.035)		
Candidate X Group Share	0.128**	0.128**	0.045		
	(0.053)	(0.053)	(0.077)		
Candidate Father				0.123***	0.220***
				(0.023)	(0.023)
Candidate Father X Group Share				0.005	0.007
				(0.049)	(0.049)
Observations	14,791,877	14,791,711	1,517,817	14,791,877	14,791,711

This table reports heterogeneity in political selection by ethnic group size and party involvement, using data from the 2016 and 2021 Panchayat elections. Panel A examines heterogeneity of party involvement by interacting the binary indicator for ever running for office with the share of candidates in the block supported by any political party. Panels B and C examine heterogeneity by ethnic group size. Panel B restricts the sample to Hindu individuals and interacts the candidacy indicator with the Hindu population share at the GP-level. Panel C restricts the sample to Muslim individuals and conducts the analogous analysis using the Muslim population share. Columns (1)–(3) report differences between candidates and citizens, while Columns (4)–(5) report differences between candidates' fathers and citizens. Columns (2) and (5) include fixed effects for gender, age cohort, and caste, while Column (3) includes father fixed effects. All dependent variables are standardized within GP. Standard errors are clustered at GP-level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Effects of SC Reservation on Selection for Mukhiya 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 Mukhiya 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 4: Effects of Female Reservation on Political Selection for Mukhiya Candidates

	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	-0.765***	-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 Mukhiya 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. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table 5: Effects of Leader Competence on Local Governance Outcomes

	Distributive			Skill-intensive		
	NREGA Work-Days (1)	SC Work-Days (2)	SC (%) Work-Days (3)	Tap Water (4)	Road Index (5)	Drainage Index (6)
Above Primary	-1193.774	-616.843*	-0.022	0.160*	0.174*	-0.009
	(1266.620)	(326.622)	(0.015)	(0.092)	(0.099)	(0.097)
Dep. Mean	15841.96	2457.75	0.17	0.00	0.02	-0.02
Bandwidth	0.15	0.13	0.13	0.17	0.12	0.13
Observations	1,726	1,553	1,554	1,799	1,422	1,482

This table reports the effects of electing a more educated Mukhiya—defined as having above-primary education—on government scheme 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 NREGA distributive outcomes: average annual total workdays generated (1), average annual workdays generated for Scheduled Castes (2), and the share of workdays generated for Scheduled Castes (3). Columns (4)–(6) report effects on skill-intensive infrastructure outcomes, including an indicator for the availability of piped tap water (4), an index of road availability (5), and an index of drainage system availability (6). 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 running-variable level and reported in parentheses. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

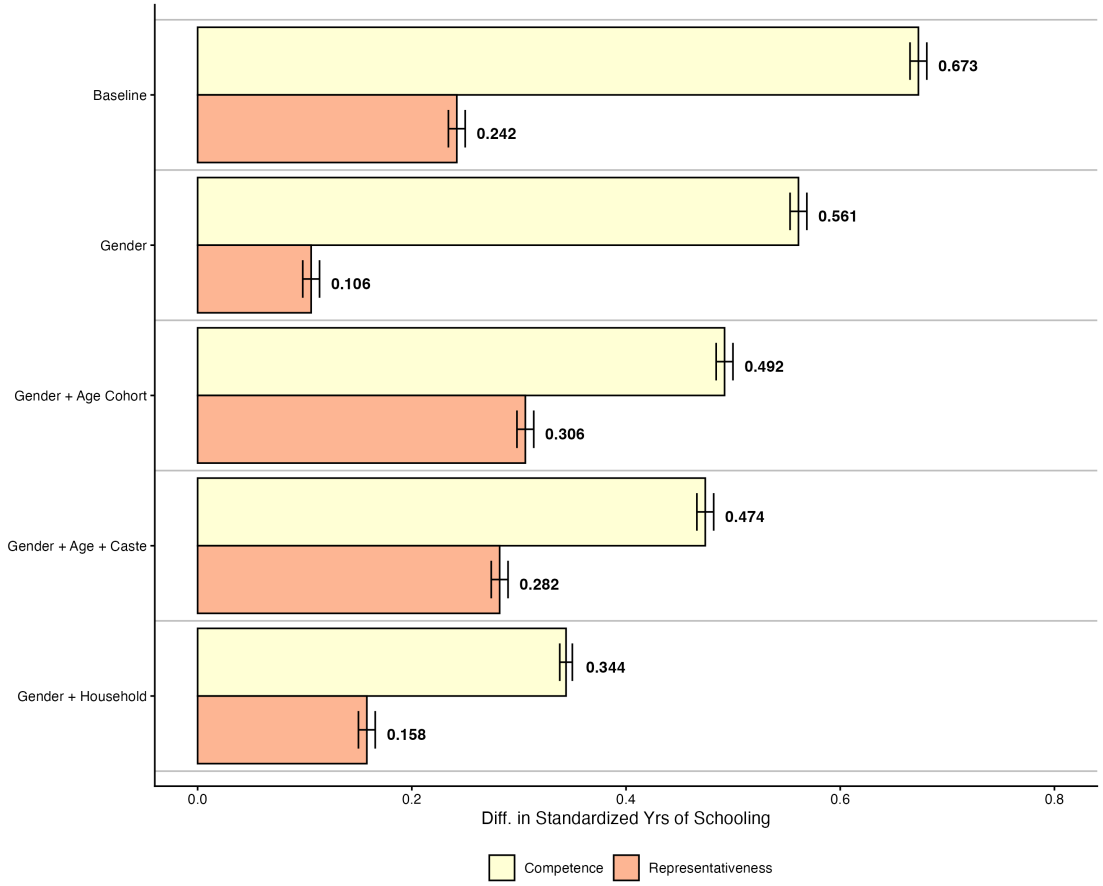
Table 6: Alignment of Public Spending with Citizen Preferences

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

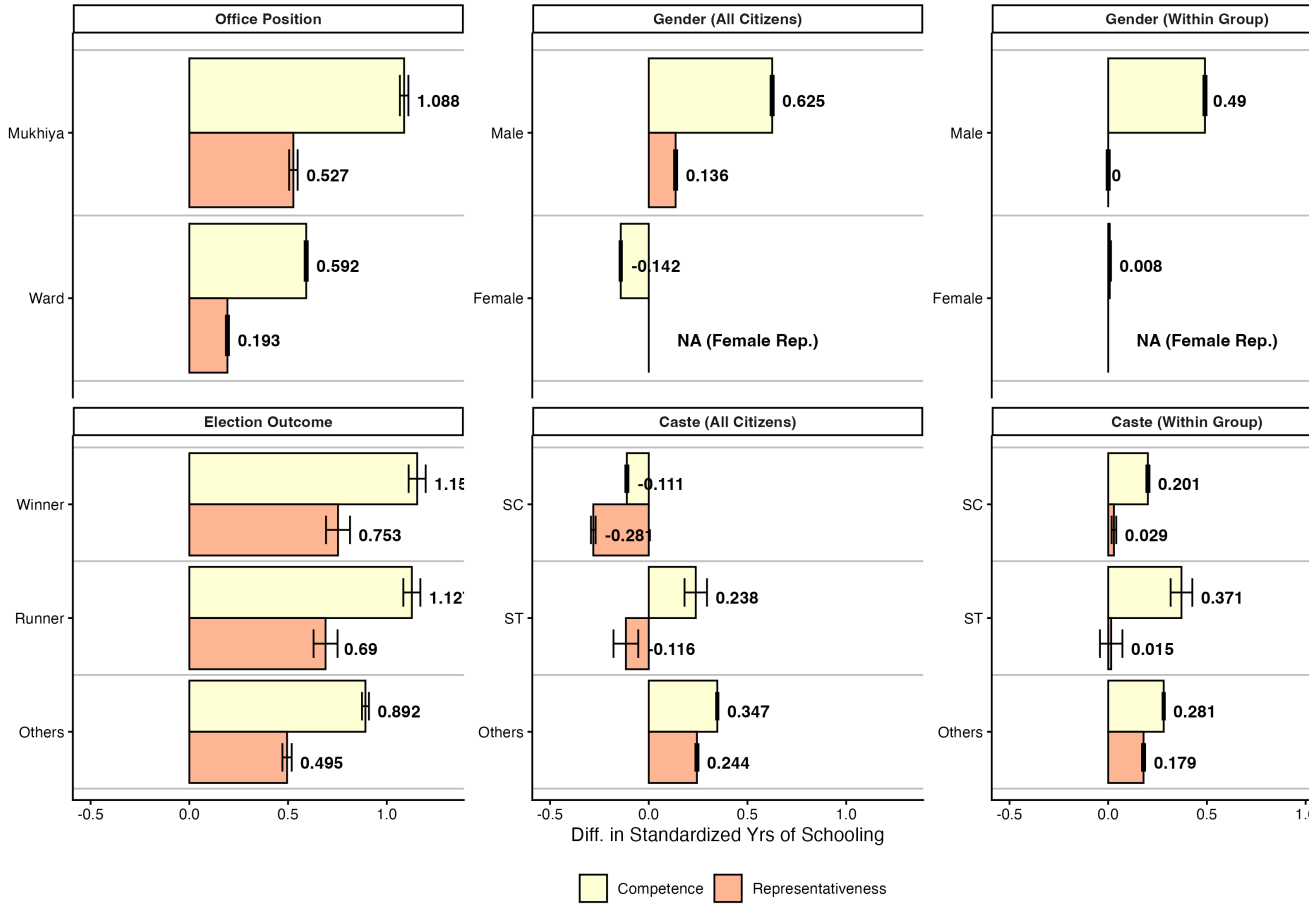
10 Figures

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.

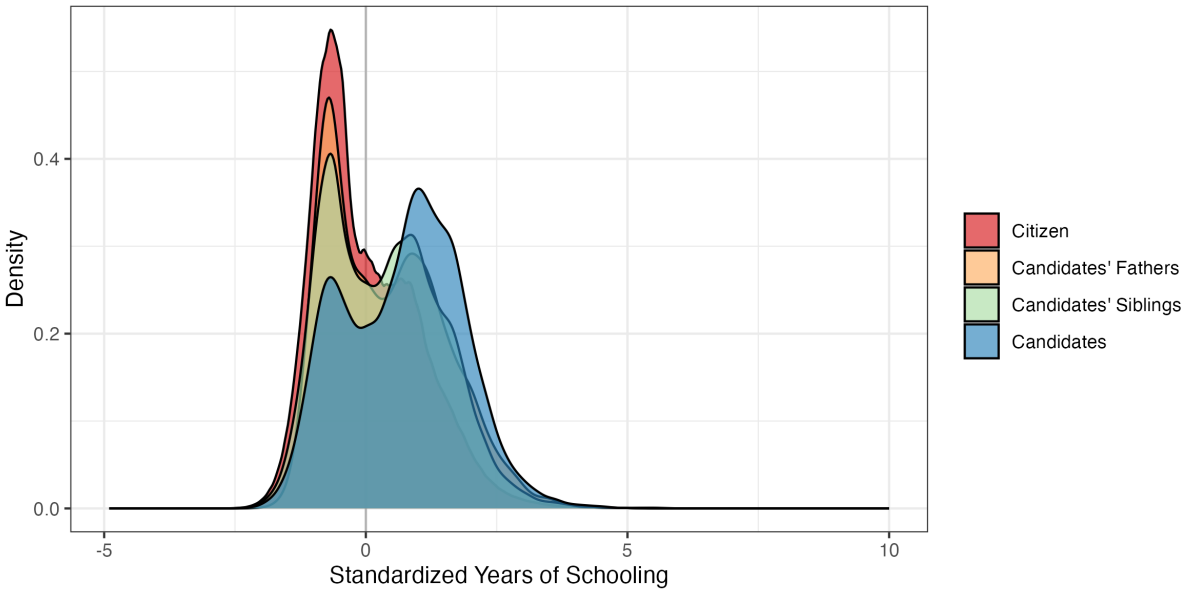
Figure 2: Heterogeneity in Selection



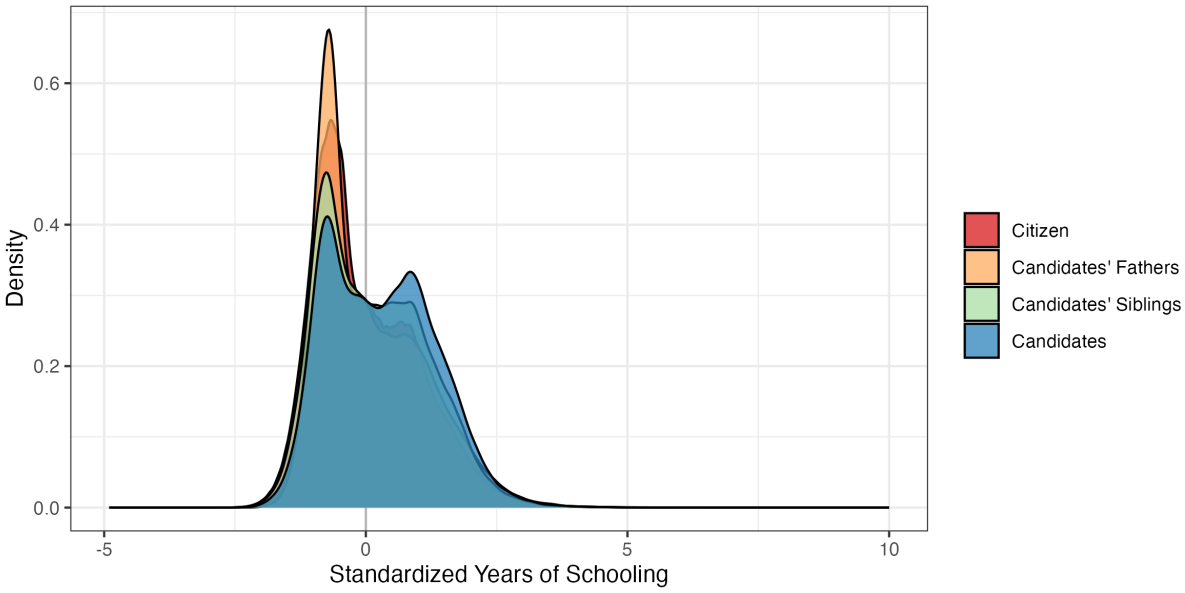
Notes: This figure plots coefficients from the baseline specifications (Equation 1 and 2) across various sub-groups, with standardized years of education as the dependent variable. Top left compares candidates for GP head and Ward member positions against the general population. Top middle and Top right 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



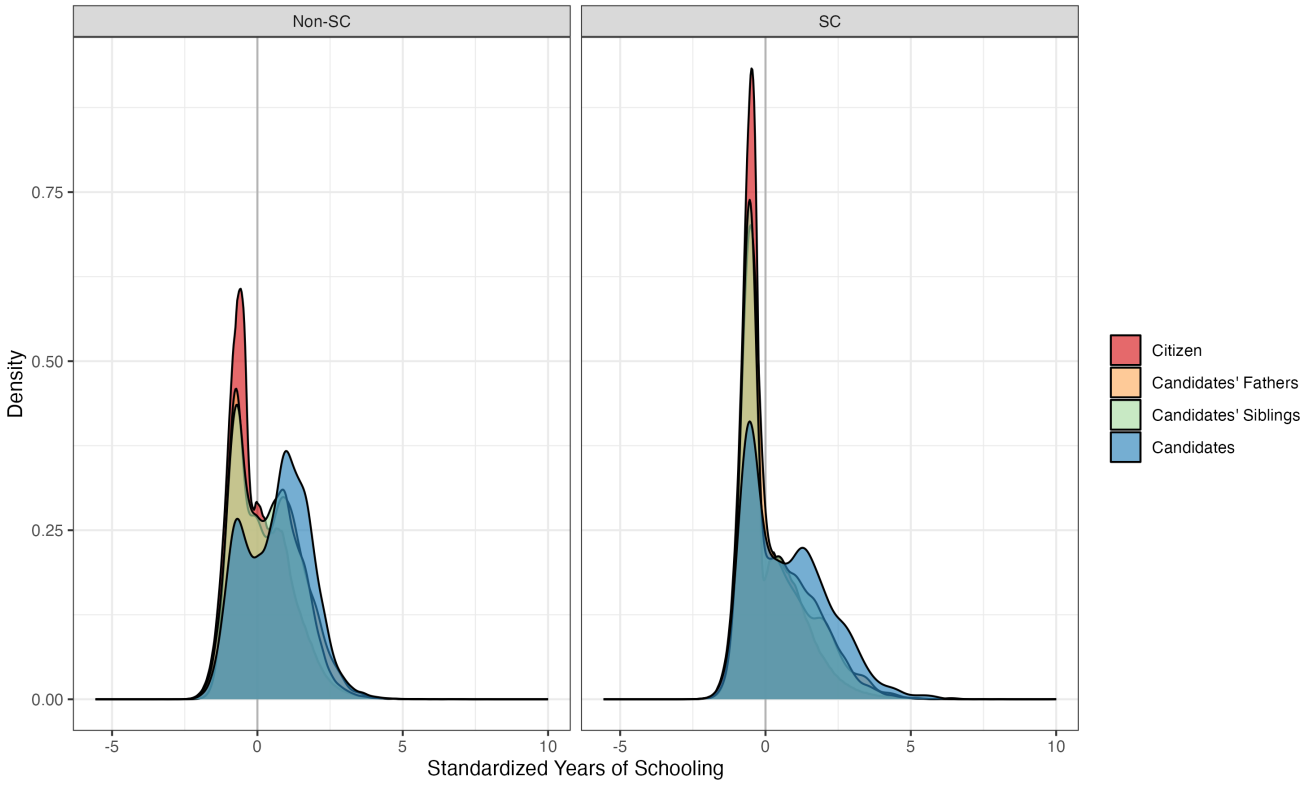
(b) Ward Member



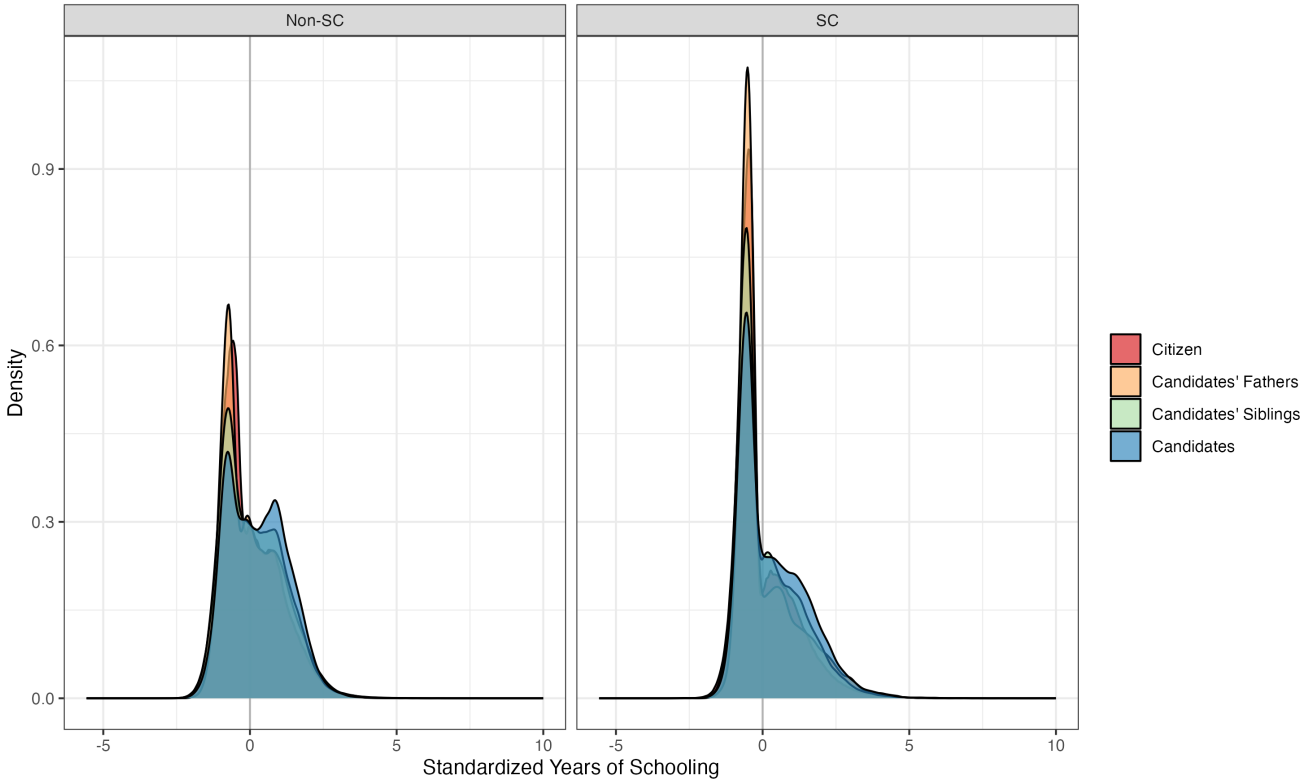
Notes: This figure plots the Kernel density of standardized years of education for candidates, their fathers, their siblings, and the general population. Years of education are standardized by gender, caste, and age cohort within each GP to account for local demographic differences. Panel A restricts 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: Candidates, Fathers & Siblings (By SC)

(a) GP Head



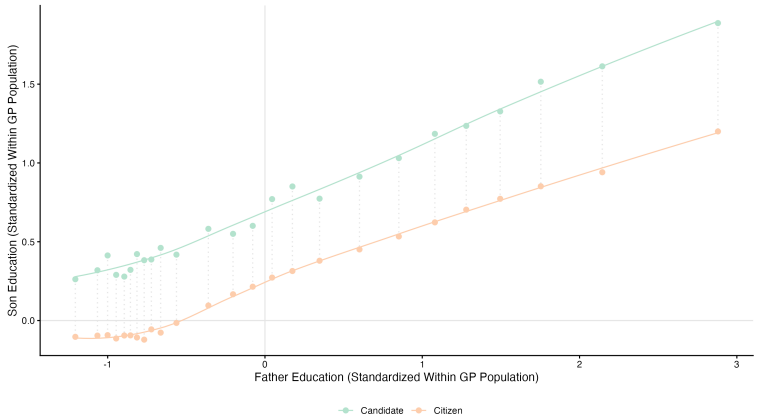
(b) Ward Member



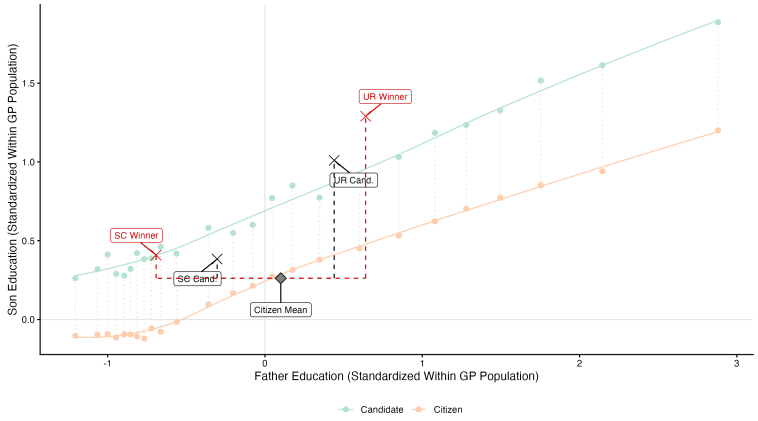
Notes: Notes: This figure plots the Kernel density of standardized years of education for candidates, their fathers, their siblings, and the general population. Years of education are standardized by gender, caste, and age cohort within each GP to account for local demographic differences. Panel A presents the distributions for candidates running for the position of GP heads and their family members, disaggregated by SC (left) and Non-SC (right) status. Panel B provides the same disaggregated distributions for candidates running for Ward member positions. The "General Population" density in each sub-panel serves as the reference group for that specific caste category. Standardized years of education are calculated relative to the full GP-level population. See Section 4.3 for a discussion of how these selection gaps vary across caste and institutional hierarchies.

Figure 5: Competence-Representativeness Trade-off

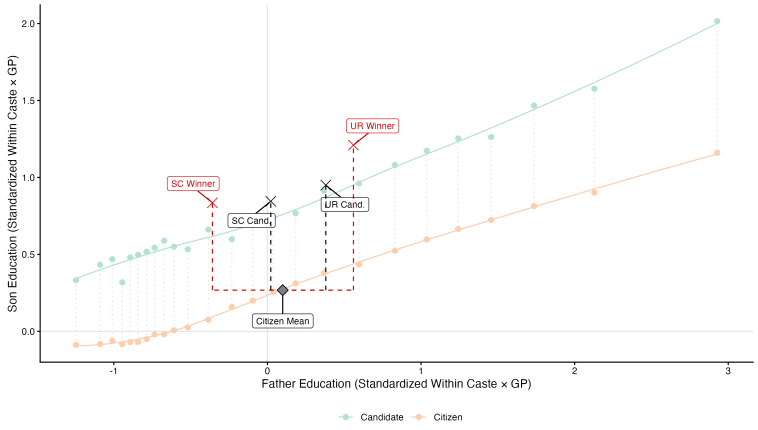
(a) The Candidate Selection Frontier and Citizen Mobility Curve



(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.

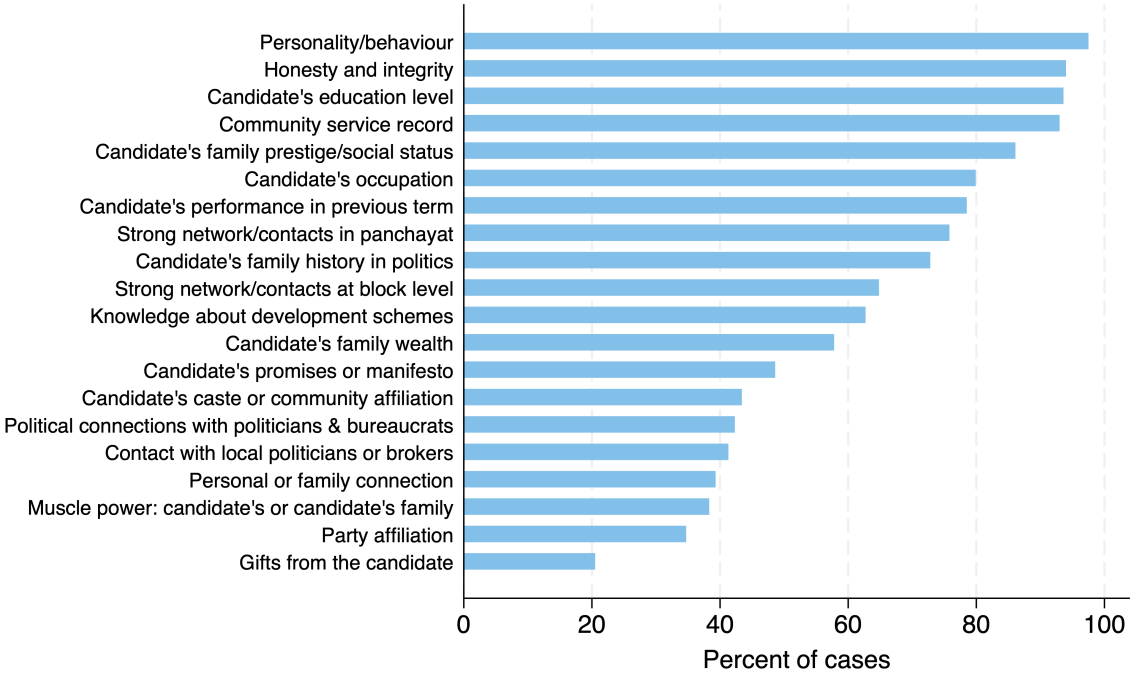
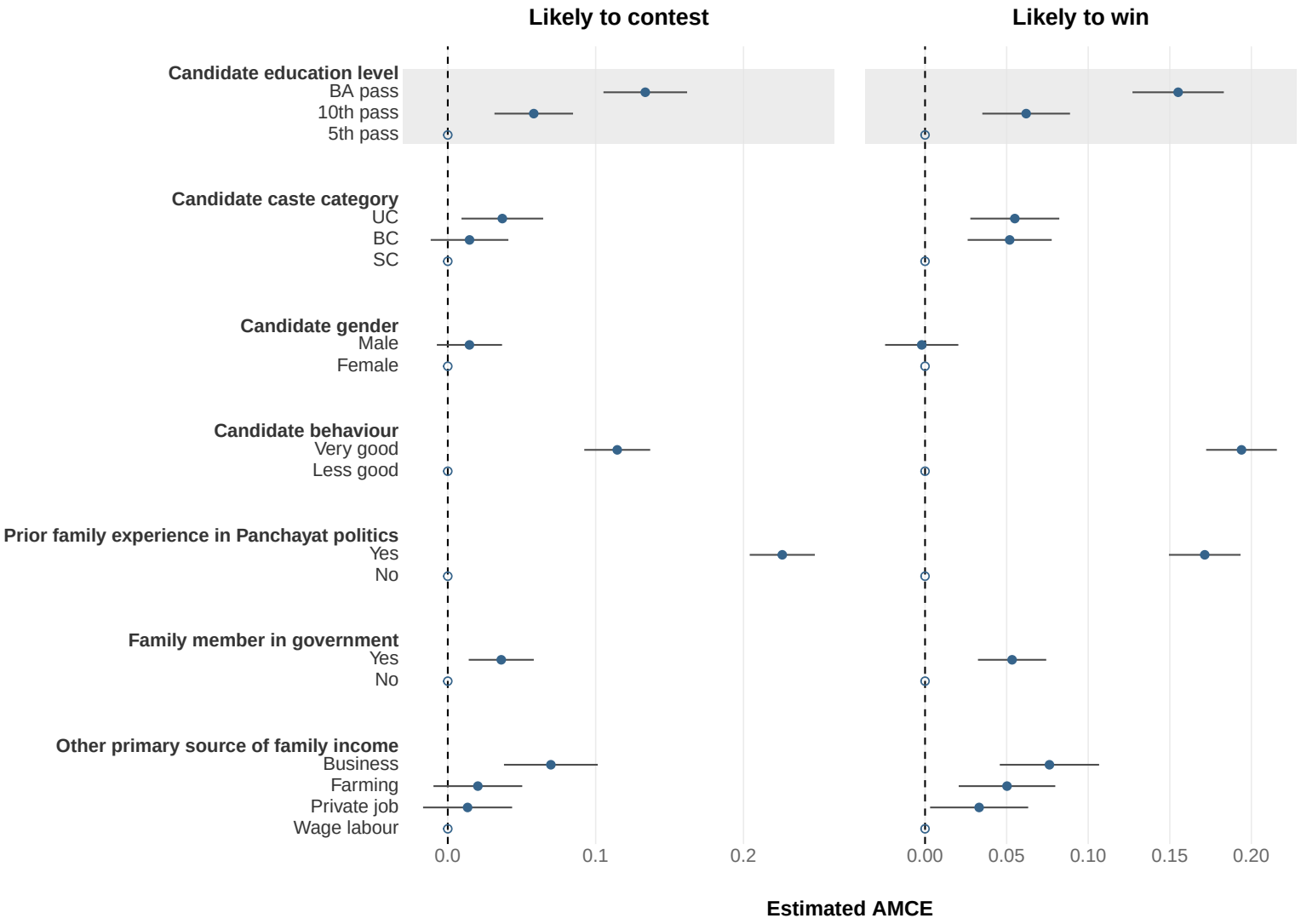


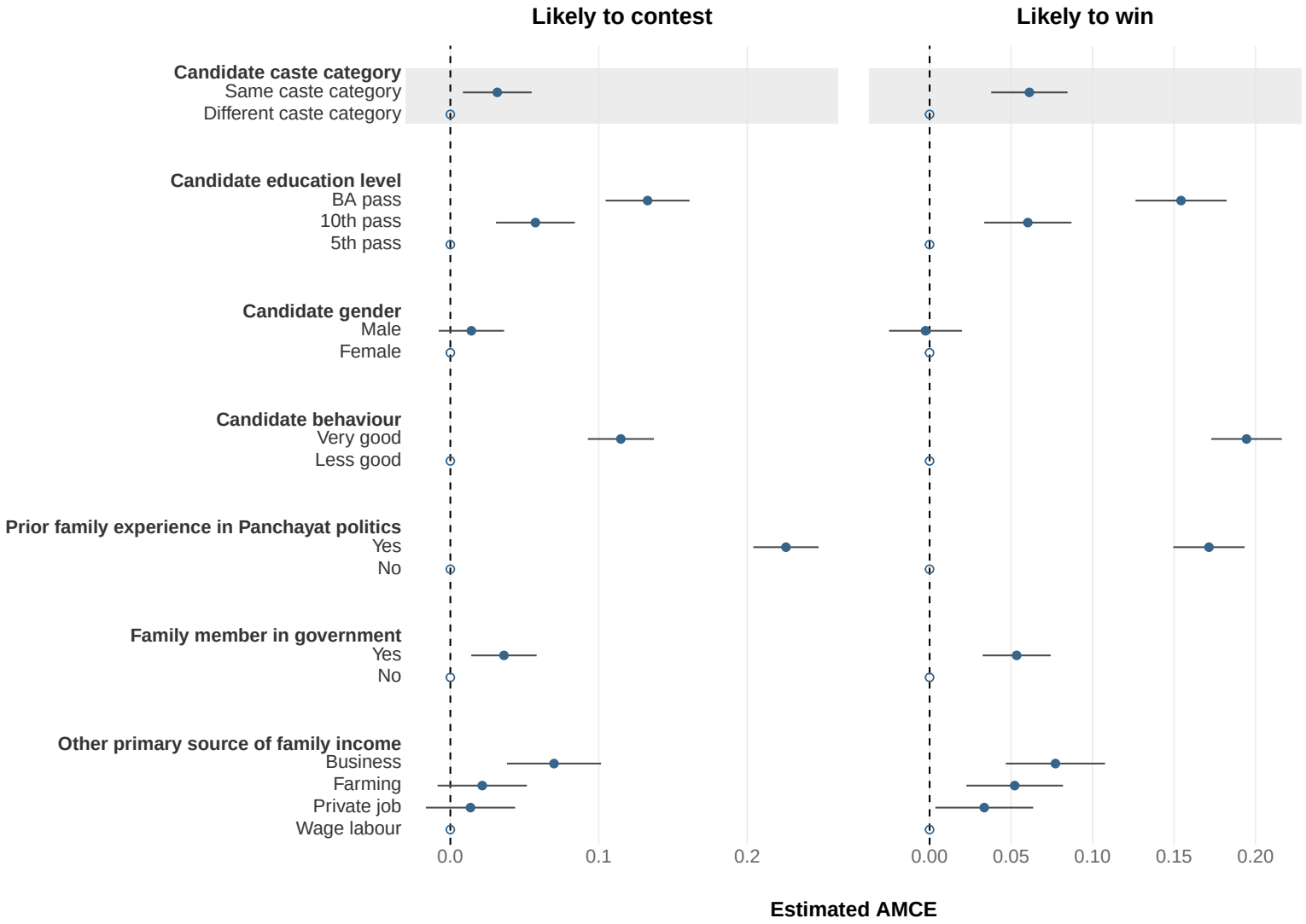
Figure 6: Qualities based on which citizens report evaluating Panchayat candidates: Percent of cases

Figure 7: Voter perceptions of likelihood of candidate type who contest and win



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 8: Voter perceptions of likelihood of candidate type who contest and win (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 Additional Tables & Figures

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 Reynal-Querol (2002): $\text{POLAR}_g = 1 - \sum_{n=1}^{N}(0.5 - s_{ng})^2 s_{ng}/0.25$, 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: Political selection and village characteristics for male Mukhiyas

	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 A3: Effects of Reservation on Political Selection for Mukhiya 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 Mukhiya 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 A4: Effects of Female Reservation on Political Selection By Caste

	Yrs of Schooling		Representativeness			
	Candidate Edu Yrs (1)	Husband Edu Yrs (2)	Gov. Job (3)	Income Tax (4)	Wealth Score (5)	Highest HH Edu Yrs (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 A5: Effects of Political Competence on Scheme Knowledge and Solutions

	Knowledge						
	Avg. (1)	Acct. (2)	AWAS (3)	NALIGALI (4)	PDS (5)	Pension (6)	Solar Light (7)
Above Primary	0.353**	0.541**	0.592***	0.135	-0.533*	-0.188	0.409*
	(0.147)	(0.262)	(0.213)	(0.206)	(0.273)	(0.356)	(0.215)
Bandwidth	0.10	0.11	0.09	0.15	0.12	0.11	0.13
Control Mean	1.95	2.56	1.75	1.35	2.50	2.70	1.51
Observations	911	949	814	1,141	431	388	1,023

This table reports the effects of electing a more educated GP Head—defined as having above-primary education—on politician's scheme Knowledge. Estimates are obtained from a close-election regression discontinuity design using the 2016 ward elections. The treatment variable (*Above Primary*) is a binary indicator equal to one if the winning candidate has above-primary education in the ward elections. Each WM was asked to list down steps to be undertaken to implement government schemes. The outcome variable is simply the number of steps a ward members is able to recount, matched to a predetermined checklist. Column (1) is the average score created using knowledge of the six schemes shown in columns (2)-(7). Column (2) is steps required to open the ward account, column (3) on housing (AWAS), column (4) drains, column (5) is the PDS, column (6) is pensions, column (7) is the solar light schemes. All dependent variables are standardized across GPs. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

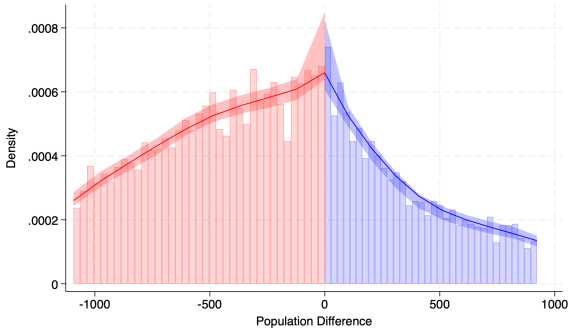
Table A6: Effects of Political Competence on Network and Party Connections

	Network (Contact Info.)				Party	
	Mukhiya (1)	Panchayat Secretary (2)	BPRO (3)	BDO (4)	Party Involvement (5)	Political Support (6)
Above Primary	0.095	0.264**	0.115	0.283**	0.133*	0.054**
	(0.068)	(0.114)	(0.093)	(0.112)	(0.080)	(0.024)
Bandwidth	0.15	0.13	0.18	0.11	0.11	0.07
Control Mean	0.91	0.73	0.28	0.33	0.17	0.05
Observations	1,242	1,100	1,456	951	953	681

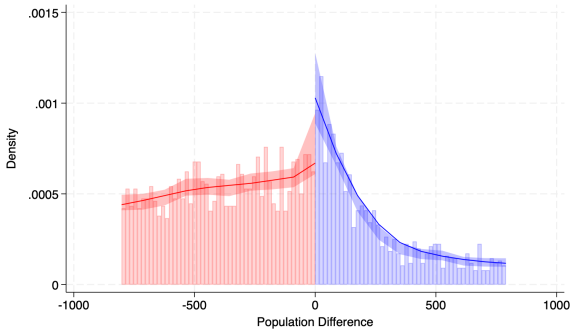
This table reports the effects of electing a more educated Ward members—defined as having above-primary education—on politician's network and party connections. The data uses responses from a survey of male ward members. Estimates are obtained from a close-election regression discontinuity design using the 2021 ward elections. The treatment variable (*Above Primary*) is a binary indicator equal to one if the ward winner has above-primary education in the Ward election. Columns (1)–(4) examine political network outcomes, measured by whether the respondent knows the phone number of the GP Head (1), Panchayat Secretary (2), Block Panchayat Raj Officer (BPRO) (3), and Block Development Officer (BDO) (4). Columns (5)–(6) examine Ward members' party connections, measured by indicators for any party involvement (5) and receipt of political support (6). Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A1: McCrary tests for manipulation of the running variable

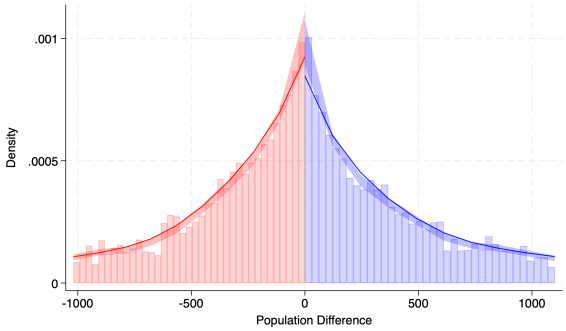
(a) SC Reservation RD (All Seats)



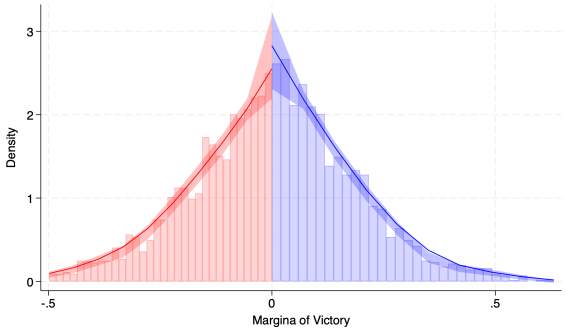
(b) SC Reservation RD (Female-Unreserved Only)



(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.

Table A7: 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 A8: 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 A9: Mukhiya Close-Election Balance Test

	Control Mean	RD Estimate	Obs	p-value (FDR q-value)
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 A10: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Effects of Political Competence on Scheme Knowledge and Solutions

	Knowledge						
	Avg. (1)	Acct. (2)	AWAS (3)	NALIGALI (4)	PDS (5)	Pension (6)	Solar Light (7)
Panel A: Above and Below High-School							
Winner	0.175**	0.167	0.144	0.155	0.118	0.078	0.180
	(0.079)	(0.121)	(0.111)	(0.126)	(0.208)	(0.220)	(0.122)
Bandwidth	0.18	0.18	0.21	0.20	0.16	0.15	0.20
Control Mean	2.22	2.86	1.86	1.71	2.78	3.28	1.82
Observations	1,829	1,875	2,034	1,788	669	640	1,800
Panel B: Above and Below Graduate							
Winner	0.192*	0.174	0.312**	-0.041	0.270	0.129	0.199
	(0.101)	(0.146)	(0.149)	(0.182)	(0.250)	(0.276)	(0.155)
Bandwidth	0.16	0.19	0.16	0.14	0.16	0.15	0.18
Control Mean	2.30	2.99	1.90	1.81	2.85	3.45	1.86
Observations	1,581	1,745	1,600	1,273	603	570	1,535

This table reports the effects of electing a more educated Mukhiya—defined as having above- primary education—on politician's scheme Knowledge. Estimates are obtained from a close-election regression discontinuity design using the 2016 Panchayat elections. Panel A reports results for electing a leader with above–high-school education, while Panel B reports results for electing a leader with above-graduate education.The outcome variable in column (1) is the average score created using knowledge of the six categories shown in column (2)-(7). All dependent variables are standardized across GPs. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Effects of Political Competence on Network

	Network (Contact Info.)				Party	
	Mukhiya (1)	Panchayat Secretary (2)	BPRO (3)	BDO (4)	Party Involvement (5)	Political Support (6)
Panel A: Above and Below High-School						
Winner	0.008	0.108**	0.144**	0.171**	0.013	-0.002
	(0.025)	(0.043)	(0.058)	(0.068)	(0.047)	(0.019)
Bandwidth	0.16	0.21	0.18	0.15	0.21	0.18
Control Mean	0.94	0.81	0.35	0.42	0.21	0.03
Observations	1,734	2,039	1,849	1,649	2,044	1,838
Panel B: Above and Below Graduate						
Winner	-0.001	0.170***	0.206**	0.136**	0.096	0.048
	(0.030)	(0.052)	(0.084)	(0.068)	(0.063)	(0.029)
Bandwidth	0.15	0.15	0.13	0.20	0.18	0.17
Control Mean	0.95	0.83	0.37	0.44	0.21	0.03
Observations	1,484	1,449	1,359	1,816	1,721	1,644

This table reports the effects of electing a more educated Ward members—defined as having above- primary education—on politician's scheme Knowledge. Estimates are obtained from a close-election regression discontinuity design using the 2016 Panchayat elections. Panel A reports results for electing a leader with above–high-school education, while Panel B reports results for electing a leader with above-graduate education. Columns (1)–(4) examine political network outcomes, measured by whether the respondent knows the phone number of the Mukhiya (1), Panchayat Secretary (2), Block Panchayat Raj Officer (BPRO) (3), and Block Development Officer (BDO) (4). Columns (5)–(6) examine Ward members' party connections, measured by indicators for any party involvement (5) and receipt of political support (6). Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: 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 Mukhiya 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 A14: 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 Mukhiya 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 A15: 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 A16: Political selection, level of office and electoral performance

	Years of Education			
	All Candidates	Winner	Runner	Others
	(1)	(2)	(3)	(4)
Panel A: Mukhiya 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 Mukhiya 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 A17: Effects of Reservation on Political Selection for Mukhiya 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 Mukhiya 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 A18: Alignment of Public Spending with Citizen Preferences

	h=0.1 (1)	h=0.12 (2)	h=0.14 (3)	h=0.16 (4)	h=0.18 (5)	h=0.2 (6)	h=0.22 (7)	h=0.24 (8)
					Panel A: Spending			
Above Primary	-53467.643	-32811.462	-23326.478	-10295.073	-1927.614	6749.768	19592.841	27333.383
	(68687.004)	(63790.178)	(58966.338)	(54952.579)	(52345.595)	(49908.076)	(47652.561)	(45866.991)
Pref.	1922509.566***	1954973.677***	1977401.571***	1991305.752***	1990778.533***	1987656.046***	1989762.878***	1991783.187***
	(262633.793)	(241900.368)	(226279.557)	(211742.930)	(199854.222)	(190758.616)	(183068.498)	(176444.082)
Above Primary × Pref.	-718105.072**	-728853.350**	-703956.461**	-688214.223**	-664933.438**	-641803.786**	-632540.965**	-626508.929***
	(337664.596)	(313910.803)	(295710.926)	(280047.395)	(267676.103)	(257445.696)	(247568.204)	(239040.118)
Dep. Mean	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Observations	2,190	2,460	2,670	2,970	3,120	3,450	3,645	3,900
					Panel B: Projects			
Above Primary	-0.253	-0.180	-0.162	-0.145	-0.126	-0.092	-0.067	-0.057
	(0.202)	(0.188)	(0.173)	(0.161)	(0.153)	(0.146)	(0.139)	(0.134)
Pref.	5.272***	5.286***	5.265***	5.232***	5.183***	5.157***	5.120***	5.083***
	(0.720)	(0.659)	(0.614)	(0.573)	(0.539)	(0.514)	(0.493)	(0.474)
Above Primary × Pref.	-2.454***	-2.400***	-2.246***	-2.127***	-2.026***	-1.953***	-1.869***	-1.791***
	(0.881)	(0.816)	(0.769)	(0.728)	(0.696)	(0.670)	(0.646)	(0.624)
Dep. Mean	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Observations	2,190	2,460	2,670	2,970	3,120	3,450	3,645	3,900

This table reports robustness checks for the alignment of public spending with SC citizen preferences across a range of bandwidths. Each column corresponds to a different bandwidth, varying from h = 0.10 to h = 0.30 in increments of 0.02, centered around the bandwidth used in the main specification. Within each bandwidth, a triangular kernel is applied by weighting observations by their distance to the cutoff. The running variable is the margin of victory, defined as the vote share of the more-educated candidate minus that of the less-educated candidate. All specifications control for the SC citizen preference measure and its interaction with the winner indicator, as well as a linear polynomial in the running variable interacted with the winner indicator. Panel A reports results for scheme-level expenditures; Panel B reports results for the number of projects implemented. Standard errors are clustered at the GP level and reported in parentheses. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

B 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 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 A20: 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 A21: 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 A22: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C 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).

C.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.

C.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.

D Consequences of Gender Reservation

In this section, we study the consequences of gender reservation on two schemes studied previously in the paper: the implementation of the NREGA and WAS projects. We use the same fuzzy RD strategy described previously to uncover causal effects of gender reservation.

Tables A23 and Table A24 below show the results. Reservation largely has no effects on either set of outcomes. There are marginally fewer WAS projects in reserverd GPs (5% fewer), but amount spent on projects does not change. Moreover, effects do not vary when broken down by caste.

Table A23: Effects of reservation on development outcomes

	NREGA Work-Day					Village Projects	
	Total Work-Day (1)	Women Work-Day (2)	Men Work-Day (3)	SC Work-Day (4)	Non-SC Work-Day (5)	Total Projects (6)	Total Amounts (7)
Panel A: Comparison to All Mukhiya Winners							
Female RSVN	-247.327	-269.372	46.962	-20.010	-218.183	-1.263**	9081.357
	(407.436)	(234.670)	(208.658)	(123.888)	(334.424)	(0.609)	(443534.654)
Bandwidth	863.72	854.82	856.14	841.08	884.65	961.15	1011.50
Control Mean	15679.488	8261.084	7486.388	2868.709	12823.346	20.166	9460208.883
Observations	3,780	3,770	3,772	3,740	3,820	3,688	3,755
Panel B: Comparison to Female Mukhiya Winners							
Female RSVN	-529.502	-321.890	-109.569	71.894	-591.450	0.124	879324.320
	(451.169)	(242.046)	(240.258)	(147.207)	(375.743)	(0.764)	(570562.398)
Bandwidth	1120.21	1162.04	1064.05	1025.67	1112.22	1063.87	984.43
Control Mean	15781.004	8267.845	7547.391	2841.388	12945.996	20.628	9688757.034
Observations	2,763	2,801	2,721	2,694	2,755	2,548	2,493

This table shows the effects of female reservation on development outcomes. Only 2016 candidates are included in the sample. Panel (A) compares the development outcomes of reserved GPs with unreserved GPs, while Panel (B) restrict the comparison to female winners only. 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 A24: Effects of Reservation on Development Outcomes

	NREGA Work-Day					Village Projects	
	Total Work-Day (1)	Women Work-Day (2)	Men Work-Day (3)	SC Work-Day (4)	Non-SC Work-Day (5)	Total Projects (6)	Total Amounts (7)
				Panel A: Unreserved			
Female RSVN	-943.435	-547.817	-326.251	-496.796	-453.359	-1.305	-458969.678
	(932.549)	(523.220)	(472.285)	(303.266)	(772.014)	(1.840)	(886410.549)
Bandwidth	2032.27	2239.09	1918.80	1888.33	2086.56	2084.47	2077.52
Observations	2,472	2,585	2,420	2,407	2,503	2,334	2,329
				Panel B: SC Reserved			
Female RSVN	599.964	-30.915	471.947	302.222	349.957	0.291	869103.828
	(1209.123)	(735.848)	(647.907)	(369.598)	(1049.438)	(2.860)	(1304813.433)
Bandwidth	314.17	313.13	302.15	344.61	303.02	355.09	354.22
Observations	909	907	890	941	891	888	888
				Panel C: ST Reserved			
Female RSVN	-13983.980	-6283.692	-7629.966	-2801.279	-11143.168	29.101	6041102.486
	(9338.292)	(5125.870)	(5162.168)	(2000.369)	(8004.875)	(21.463)	(9830281.099)
Bandwidth	174.65	167.38	177.41	177.45	168.47	206.66	208.86
Observations	21	21	21	21	21	19	19
				Panel D: OBC Reserved			
Female RSVN	-1384.040	-464.378	-735.557	-151.409	-1102.255	-3.176	1350225.934
	(1742.608)	(988.146)	(894.908)	(450.857)	(1551.996)	(3.373)	(2770700.845)
Bandwidth	512.77	525.59	496.82	750.74	507.27	536.98	584.58
Observations	761	773	742	927	754	726	771

This table shows the effects of female reservation on NREGA and village project development outcomes. Only 2016 candidates are included in the sample. Panel (A)-(D) shows the regression results broken down by caste reservation status. 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. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

E Close-Election Regression Discontinuity Design on Politician Knowledge and Network

This appendix describes the close-election regression discontinuity (RD) design used to identify the effects of education on politician knowledge and networks. To support the use of education as a proxy for competence, we examine whether more educated leaders have greater knowledge of scheme implementation as well as broader political networks and stronger party connections. The questions on knowledge and networks were part of the baseline survey data of 7719 WMs, taken from (Bamezai et al., 2024).

We implement an RD design that exploits close electoral races in 2021 ward-level elections.

Following a similar strategy in Section **??** and Section **??**, we restrict our sample to ward-level elections where the top two finishers differ in their attainment of primary education. We estimate:

$$Y_w = \mu + \theta \mathbb{1}\{MoV_w > 0\} + \gamma_1 MoV_g + \gamma_2(\mathbb{1}\{MoV_w > 0\} \times MoV_w) + \varepsilon_w, \tag{10}$$

where Y_w denotes the outcome variable of winning candidate in ward w. The running variable, MoV_w, is the vote share margin between the candidate with above-primary education and the candidate with primary education or below. Treatment indicator equals 1 if the more educated candidate wins ($MoV_w > 0$). The coefficient θ identifies the local average treatment effect of electing a more educated ward member. We also control for if candidates are running for themselves or a family members. Standard errors are clustered at ward-level.

We evaluate the impact of politician's education level on two sets of outcomes. First, we measure institutional and scheme knowledge by assessing whether candidates can correctly identify the administrative steps required to implement various government programs. This includes the setup of ward member accounts, the implementation of solar street lights, and the procedural requirements for giving out social pensions, the Public Distribution System, drains and lanes, and housing schemes. Table A5 documents the results. More educated WMs are able to name 18% more steps than control WMs (Column 1), driven by greater knowledge of housing, solar lights and ward accounts.

Second, we look at for the politicians' political network. This is measured by an indicator for whether the ward member possesses the direct contact information for key block and panchayat officials—specifically the Block Development Officer (BDO), the Block Panchayat Raj Officer (BPRO), the GP head, and the Panchayat Secretary. Additionally, we also capture partisan support through the member's formal involvement with political parties and whether they received external political support during their campaign. Table A6 shows that WMs have wider networks, especially with the key bureaucrats: the GP Secretary and the Block Development Officer. Moreover, they are also more likely to

be linked to political parties.

Lastly, our results are robust to the alternative choice of educational threshold. Tables A11 and A12 show that the estimated effects persist when defining the treatment based on high school completion or university graduation. These results suggest that the benefits of candidate education are not idiosyncratic to the primary school margin.

F 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.

F.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.

F.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.

F.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:

F.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.