Who Becomes a Local Politician? Evidence from Rural India*

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

October 27, 2023

Preliminary and incomplete.
Please do not cite or circulate.

Abstract

Can local democracy in areas of weak state capacity attract competent leaders while simultaneously ensuring adequate representation of disadvantaged groups? Matching census data of 95 million rural residents and nearly 1 million local politicians from Bihar, we uncover the following facts about politicians' competence and representativeness. First, absent political quotas, Bihar's local electoral system comprises a "partially exclusive meritocracy". Politicians are from more elite backgrounds, but among the elites, the more educated contest and win. Our results suggest a trade-off between competence and representativeness, with women, members of disadvantaged castes, lower ranked candidates and lower tiers of government demonstrating less positive selection and less elite backgrounds. Moreover, while selection patterns vary by various village characteristics such as inequality and caste heterogeneity, these do not fundamentally change the "partially exclusive meritocracy" characterisation. Subsequently, we argue that policy intervention can meaningfully influence selection. Firstly, using a difference-in-differences design, we show that a policy move to devolve financial powers attracts a larger and more competent pool of candidates. Secondly, using a fuzzy regression discontinuity design (RDD), we show that gender-based reservation democratizes selection at not just the individual-level, but also at the household level by encouraging entry of candidates with lower household incomes and wealth. Lastly, using a close election RD approach, we explore the influence of leaders' education on policy implementation. We find no systematic relationship. Taken together, our findings highlight the significance of studying the fiercely competitive landscape of local democracy in understanding the causes and consequences of political selection.

*We thank seminar participants at the Delhi School of Economics for some excellent comments.
†University of Pennsylvania
‡National University of Singapore
§University of Maryland
¶University of Maryland

1 Introduction

The composition of a society's political class is a key determinant of its quality of governance and economic development. Yet we still have limited evidence on political selection in developing countries (Gulzar and Khan, 2021; Gulzar et al., 2021). Prior work examines high-income polities like the United States (Thompson et al., 2019) and Sweden (Dal Bó et al., 2017), where political selection is both meritocratic and inclusive. Patterns of political selection might differ in a developing country for several reasons. First, developing countries tend to have much greater economic inequality, increasing the tradeoff between competence and representativeness. Second, developing countries have much higher levels of ethnic diversity and deeper group divisions, so voters may trade off between voting for their own group and the best candidate. Third, democracy is less mature, so institutions that facilitate positive and inclusive selection may not exist.

We study the causes and consequences of political selection in Bihar, a large Indian state, which accounts for 3.5% of the global population living in a democracy. Bihar is an excellent context to test hypothesis about political selection in a low-income democracy. Like many other developing countries, Bihar is highly unequal, divided along caste and religious lines, and parties play a relatively minor role in local politics.

We begin by providing descriptive evidence on political selection in rural Bihar[1], using rich administrative data on the universe of citizens (95 million) and local politicians (nearly 1 million) in rural Bihar. Measuring the competence of politicians by their education, we first show that politicians are positively selected from the citizenry. Candidates for political office tend to have (1 SD) better education compared not only to the general population, but also among citizens from the same cohort, community (same caste, religion or surname), and even relative to other men in their household (0.6 SD higher).[2] This positive selection is more pronounced for high-performing candidates, such as winners of elections and those seeking higher-level political offices. Furthermore, this trend of positive selection appears to be on the rise over time.

We investigate whether politicians are representative of the population by matching politicians to their fathers and observing the educational qualifications of politicians' fathers relative to the average 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 when compared to their respective cohorts, communities, neighborhoods, and the general citizenry. This suggests that politics in Bihar is meritocratic (as in Sweden), but an exclusive rather than an inclusive one.

One way to get around this "partially exclusive meritocracy" is to introduce political

[1]We focus on rural politicians because Bihar remains an overwhelmingly rural society – as per Census 2011, only 11% of citizens live in urban areas. Moreover, the nature of local bodies, their roles and responsibilities vary considerably between urban and rural areas making the study of them jointly somewhat tricky.

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

reservations ("quotas") for members of disadvantaged groups. We find that candidates in reserved seats are less positively selected than those in unreserved seats, even with respect to their own groups. However, these candidates are more representative of their own groups - as proxied by father's education. This highlights the trade-off between representativeness and competence in our setting.

Next, we explore how patterns of political selection vary with village characteristics, by analysing heterogeneity across Bihar's 8000+ village governments. We find, surprisingly, that the competence and representativeness of politicians are uncorrelated with levels of village poverty and economic development. We find statistically significant but relatively small correlations with ethnic diversity (politicians in villages with lower ethnic fractionalization and higher ethnic polarization are slightly less positively selected and more representative) and land inequality (politicians in unequal villages are more competent and less representative). But overall, the general pattern of political selection is robust – even in the least equal and most diverse village, politicians are positively selected and from advantaged social backgrounds.

In the second part of the paper, we investigate the determinants of political selection. We analyse two natural experiments. First, we study how financial devolution and consequent access to political rents affects selection (Besley, 2005). We exploit the fact that control of funds in a local development scheme was devolved from bureaucrats to ward members (the lowest rung of village politician) in some villages. Using a diff-in-diff design, we show that increased access to financial rents increased the number of candidates in ward elections and raised the competence of the average candidate. This runs counter to the predictions of some models of political selection (Dal Bó and Finan, 2018).[3]

Second, we study the impact of political reservations for women. We show that reservations go beyond improving women's representation and induce a broader change in the selection and social origins of leaders. We find that female leaders are less positively selected (relative to female citizens) than male leaders are (relative to male citizens). However, we also find that female leaders are more representative, not just in terms of their gender, but also in terms of their familial origins: they come from less elite households. Thus, political reservations heighten the competence-representativeness tradeoff, both directly by bringing women (who tend to be less educated) into power and indirectly because selection patterns among female politicians are slightly different to those among men.

In the final part of the paper, we study the consequences of political selection. We estimate the causal effect of electing a more competent (read "more educated relative to citizen") leader. Our primary outcomes are the quality of implementation of social protection and local development schemes. Using a close election regression discontinuity design, we find that more educated politicians have no significant impact on the number of NREGA days worked by citizens or the speed of construction of tap water connections and drains.

[3]Turning to the pool of winners, we find that financial devolution does not change either competence or representativeness. This absence of movement on either representativeness or competence once again indicates that these two factors vary jointly.

The paper contributes to two strands of literature. First, 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.

Turning to developing countries, two recent studies from Pakistan and Nepal provide experimental 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 study the economic consequences of better political selection.

Second, we contribute to the literature on political reservations. Prior work has documented that reservations improve representation of women, by changing policy preferences (Chattopadhyay and Duflo, 2004) and breaking stereotypes (Beaman et al., 2012). These effects persist even after reservations end (Bhavnani, 2009, 2017). We go beyond prior work by showing that reservation also broadens 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

2.2.1 Rural Government

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 "Mukhiyas" respectively. Mukhiyas 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"). Mukhiyas may rely on ward members to implement schemes and projects. On paper, ward members are to their wards what Mukhiyas are to the GP: person chiefly responsible for overseeing development in their constituencies. However, the absence of formal financial powers makes the equation lopsided in favour of the Mukhiyas.

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

2.2.2 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%.[4]

2.2.3 Reservation

Since 2006, Bihar has reserved seats in favour of minorities for both the Mukhiya and ward member positions. About 45% seats are reserved at the Mukhiya and ward levels are reserved 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.

2.2.4 Financial Devolution to Wards

In late 2016, the government of Bihar shifted the responsibility for carrying out two significant water and sanitation initiatives to ward members. These initiatives, known as "Nal Jal" (providing piped water to every household) and "Nali Gali" (building village roads and drains), were central components of the incumbent government's broader development agenda. An approximate budget of $4 billion was allocated for the implementation of these schemes.

[4] 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.

2.3 Overcoming Empirical Challenges

As documented in Dal Bó et al. (2017), there are 3 data limitations that any comprehensive study of political selection needs to overcome. Below we list them and describe how our setting allows us to account for each of these issues:

- *Data on (i) elected politicians, (ii) non-elected politicians and (iii) those who never contested elections:* In our setting, we use data on the near-universe of local politicians obtained from Bihar's State Election Commission for the rural local body polls of 2016 and 2021. We then combine it with a census – the socioeconomic caste census (SECC) – to obtain data on all citizens in rural Bihar (see section 3.1 for a discussion of these datasets and the matching process).
- *A good measure of quality of selected politicians:* We have data on education of all citizens and politicians from our sources and, following standard practice in the literature, use this as our main metric of competence. Section 2.4 has a longer discussion on the use of this as a metric of competence.[5]
- *Intergenerational information to assess "representativeness" of politicians:* we match citizens to their fathers in the census to understand intergenerational information to understand the backgrounds of citizens and local politicians (see section 3.1 for a discussion onrepresentativeness and the matching process).

2.4 Education as Competence

In this paper, we use the number of years of formal education as the key measure of competence. Much of the existing literature on the backgrounds and performance of politicians employs education as a proxy of political 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).

There are three main arguments in favour of education as competence. One, it correlates with earnings and civic engagement, indicating that education 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).

[5] In section 5, we test the empirical validity of this assumption in our context.

Education as a proxy for competence, however, is not universally accepted in the literature. Some research suggests that the number of years of formal education does not lead to any substantial differences in performance on a range of measures (Carnes and Lupu, 2016; Curto-Grau and Gallego, 2021; Dreher et al., 2009) or have heteregeneous effects (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).

Another concern raised by Dal Bó et al. (2017) is the possibility that education might merely reflect elite membership, and thus, it may not be a reliable marker of positive political selection. In other words, more educated politicians may come from more elite or wealthier households and, therefore, be less representative of the underlying population. 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.

In the absence of any measures of cognitive, leadership or earnings in our context, we rely on education as a measure of competence in this paper. However, following Dal Bó et al. (2017) and Dahlgaard et al. (2020), we also compare all politicians in our data with their family members and social groups in order to alleviate some of the concerns highlighted in the literature. Finally, in order to further investigate the links between competence and education, we test in section 5 the causal effects of education on performance in office.

3 Result 1: Who Becomes A Politician?

We begin by describing who, among citizens, enters politics. We ask the following questions: compared to the pool of citizens, how competent are politicians? How representative are they? We then describe how selection varies across a range of GP-level characteristics, from poverty to inequality to ethnic difference.

3.1 Data Sources

We combine two large datasets: the socioeconomic caste census of 2011-12 and political data on Mukhiya and ward candidates in the 2016 and 2021 Panchayat elections obtained from the State Election Comission of Bihar.

Citizen Data from the Socioeconomic Caste Census(SECC): The SECC is a nationwide census of households across India. As the name suggests, the census aimed to document the socioeconomic status of various caste groups in the country. The SECC rural dataset for Bihar contains data on 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 households.

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.

We create the following GP-level variables from the SECC:

1. **GP Asset Index**: We create an asset index for every household based on 6 binary asset indicators found in the SECC dataset. We focus on ownership of the most common assets from the data: land; type of the roof (concrete or not) and wall (whether made of burnt brick or concrete) of the main dwelling room of the house structure; whether the house has 4 or more rooms; whether the household has a phone; whether the household owns a vehicle. Each of these assets is owned by at least 10 % of the population. We then calculate the average value of the asset index.

2. **Ethnic Fractionalisation:** We create an ethnic fractionalisation index for every GP by computing one minus the Herfindahl Index of the last names of the household head: $\text{FRACT}_g = 1 - \sum_{n=1}^{N} s_{ng}^2$, where s_{ng} is the share of population with last name n in GP g and N is the total number of unqiue last names in the GP.

3. **Ethnic Polarisation Index**: We create an ethnic polarisation index for every GP based on Reynal-Querol (2002), which takes the form: $\text{POLAR}_g = 1 - \sum_{n=1}^{N}(0.5 - s_{ng})^2 s_{ng}/0.25$.

4. **Educational Mobility:** We calculate, for each GP, the educational mobility of households by comparing education of older cohorts with that of younger cohorts. Specifically, we look at the percentage of children between the age of 14 and 18 who completed primary education within each GP, while restricting our sample to families with parental education level below primary.

5. **Land Gini:** We calculate, for each GP, the extent of land inequality between households using the gini coefficient of total land owned by each household.

6. **Remoteness:** We calculate the remoteness of the GP using census-derived measures of the average distance of villages in the GP for the district headquarters.

Candidate Data from the Bihar State Election Commission: The Bihar State Election Commission uploads data on all candidates contesting the Mukhiya and ward elections for 2016 and 2021 elections. There are 909,570 candidates across the two elections.[6] We have data on candidate full names, father/husband names, age, education, caste category and votes polled.

Matching We match citizens to candidates in the following manner: first, conditional on being in the same GP, we match based on name, father/husband name, caste category and gender using a string-matching algorithm. For every match, the algorithm spits out a match score. Our analysis only restricts attention to match scores that are at least 0.95. This score has a very low inclusion error i.e. it is very unlikely that a match in our dataset is not a true match. On the other hand, there could be some exclusion errors: this is best exemplified by the fact that we do not "match" a significant number of candidates. Our results are robust to increasing/decreasing the match score.

Our overall match rate for our primary population of interest, unreserved Mukhiya 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 18 and Table 19. Our match is highest for the Mukhiya male candidates for 2016 and lowest for women Mukhiya/ward candidates in 2021. 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: Mukhiya male matched and unmatched candidates have similar education levels; female matched candidates (at both ward and Mukhiya 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.

Measuring Representativeness We follow (Dal Bó et al., 2017) to use the education of fathers as a measure of representativeness of male candidates. Since women leave their villages upon marriage, we cannot match most adult women and their fathers. However, no such issues exist for men. Now, male candidates' education and wealth outcomes may

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

be correlated, so a better measure of representativeness of candidates will be the relative education of their fathers compared to other males in their age cohorts.

In our data, we are able to find 41.13% of fathers for our candidates. Our representativeness results are restricted to this sample. Table 20 and Table 21 respectively present the balance tests for candidates whose fathers we can find vis-a-vis those we cannot find. We see that on average, we find fathers' of slightly better educated politicians, with 0.89-1.42 more years of education than their non-matched counterparts. They are also more likely to be from non-SC and non-ST castes.

3.2 Empirical Strategy

We start with a simple regression design to present empirical facts on how electoral candidates differ from the general citizenry across a variety of socio-economic dimensions, including education, employment, income, and asset ownership. Specifically, we estimate

$$Y_{ig} = \alpha_1 \mathbb{1}\{Candidate\}_{ig} + \epsilon_{ig} \tag{1}$$

$$Y_{ig} = \beta_1 \mathbb{1}\{CandidateFather\}_{ig} + \eta_{ig} \tag{2}$$

where Y_{ig} represents the socio-economic background for individual i in GP g. All dependent variables are standardized at GP-level to allow for a fair comparison between candidates and their local peers. $\mathbb{1}\{Candidate\}_{ig}$ is a binary indicator for candidacy in any of the 2016 and 2021 local elections. As an individual's socio-economic status is significantly influenced by their family's background, to examine the candidates' social origins, we also look at the characteristics of candidates' fathers using indicator $\mathbb{1}\{CandidateFather\}_{ig}$, which equals 1 if any children of candidate i participated in local elections. The coefficient α and β respectively describe candidates competence and representativeness level. Standard errors ϵ_{ig} and η_{ig} are clustered at GP level.

3.3 Results

3.3.1 Competence Overall

Our first set of results is restricted to candidates contesting unreserved seats at the ward and Mukhiya levels. We begin with these because these seats are technically "open" to all members in a GP, across genders and castes.

We begin by showing that politicians are more competent – as proxied by education. The average politician is more educated than the average citizen in their own GP. Citizens in Bihar on average have 3.46 years of schooling, while politicians within the same GP obtain 2.67 (0.67 SD) more years of education (Table 22). This 77% higher levels of education can also be interpreted in the following manner: the average citizen is literate, but the average candidate has completed primary schooling. Moreover, while about 6% of citizens in our context have completed high school, this number is three times as high for politicians at 18.5%.

This is true if we compare education levels to all citizens or only adult citizens (Table 2). These results are robust to including a series of fixed effects: gender, caste and age cohort (Table 3). Candidates appear to have more education than peers from their own backgrounds.

We then show that politicians also come from wealthier households. A household that produces a politician is more likely to have someone who has a government or salaried job, pays income tax and self-reports own income to be "high"; the average wealth score of politician-households is also higher (0.2 SD) and these households are more likely to report owning land (0.125 SD) and agricultural equipment (Table 4).

3.3.2 Competence By Groups

We now investigate if selection varies by tier of local government, by vote-rank and across caste and gender.

Tier of Government and Vote Rank: In Table 5, we see that Mukhiya winners are the most positively selected politicians: they are a full 1.4 SD more educated than the average citizen. This translated to an additional 5.6 years of schooling. Ward candidates are the less positively selected among all candidates in unreserved seats. On average, ward candidate has about 2.35 (0.59 SD) years more education than citizens. Moreover, higher ranked politicians seem to be more educated, suggesting that voters value education of politicians.

Caste and Gender: We now turn to see who are the most and least positively selected politicians by caste and gender. Panel A of Figure 1 has the results: while most types of politicians are positively selected on education with respect to the population, the differential reduces as one goes down the caste/gender hierarchy. Only SC candidates at the ward level and female candidates at the ward level seem to have lower education levels than the underlying population in their GPs. However, this could merely be a reflection of the lower educational attainments of low-castes and women overall.

Panel B of Figure 1 shows us how selection varies across politician-types with respect to citizens of their own group. For instance, here, instead of comparing women candidates with all citizens, we only compare with other women. Here too, we see that that unreserved candidates seem to be most positively selected with respect to the underlying population. While the relative selection of low castes and women seems to improve, suggesting strongly that some of the patterns we saw previously merely indicated lower underlying achievements of marginalised groups, the hierarchy in coefficients does not change. Finally, this change in comparison group makes SC ward candidates positively selected on education. However, we do not see a similar sign-flip for women ward candidates: they continue to be negatively selected, even when compared to other women in the GP, though the negative coefficient is much smaller than in the previous figure.

3.3.3 Representativeness

We now turn from competence, as proxied by education, to representativeness of candidates. In the previous section, we show that candidates are not only more educated than the citizens they represent, but also wealthier. This doesn't necessarily mean that candidates come from more elite backgrounds, since current wealth scores could be correlated with their own education.

To develop a measure of how elite candidates are, and, conversely, whether their education reflects ability or merely inherited wealth, we compare education levels of candidates citizens with their fathers and see if the difference is bigger or smaller than for comparable citizens. If candidates' fathers are relatively less (more) educated than citizens than candidates are with respect to citizens, then this indicates that candidates come from more (less) representative backgrounds than their own education levels suggest. Thus, this would suggest positive (negative) selection of candidates.

Tables 6 display the results. While Mukhya (Ward) candidates have a full 1.088 (0.592) SD more education than the average citizen, candidate fathers have about half as much of an advantage over their peers. This suggests while candidates come from more elite backgrounds than the average citizen, some of the positive selection of candidates is driven by candidates' own ability.[7]

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, in Table 7, we show that education differential in SD units is lower for fathers than candidates even when we put in age-cohort fixed effects. The other columns show that the differential is lower even when we control for gender, caste and household fixed effects.

Panel A of Figure 2 shows density plots for standardized years of schooling for 4 groups: all citizens, Mukhiya candidates, Mukhiya candidates' fathers, Mukhiya candidates' siblings. While it is not surprising that candidates are more educated than all citizens and their fathers, the fact that they are more likely to be educated than their siblings indicates that there is positive selection on education of candidates even within their household.

[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.

3.3.4 Trade-off between competence and representativeness

Panel B of Figure 2 shows density plots for ward candidates. The positive selection of candidates on education continues to hold here, albeit the differences between candidates and others is more muted. Indeed, the pattern for ward candiates vis-a-vis Mukhiya candidates suggests a trade-off between representativeness and competence: ward candidates are less educated (see Table 5, column 1) and more representative of the underlying population than Mukhiya candidates.

Another place the trade-off is salient is when we look at competence and representativeness across politicians of different groups: politicians from marginalised caste groups are more representative of the underlying populations (Panel B Figure 3), but the relative education levels are also lower (Panel B of Figure 1).

4 Result 2: What Affects Who Becomes A Politician?

4.1 Selection and GP-level covariates

We begin by describing how political selection varies with a host of GP-level covariates. We rely on the 6 GP-level variables we create from the SECC described in section 3.1: GP aset index, ethnic fractionalisation, educational mobility, land inequality and remoteness.

4.1.1 Empirical Strategy

We run the following empirical specification:

$$Edu_{ig} = \beta_1[\mathbb{1}\{Candidacy\}_{ig} \times Feature_g] + \beta_2\mathbb{1}\{Candidacy\}_{ig} + \epsilon_{ig} \tag{3}$$

where Edu_{ig} is citizen i's years of education standardized at GP g. The indicator $\mathbb{1}\{Candidacy\}_{ig}$ is 1 if citizen i competes for unreserved Mukhiya (Ward) seats in either 2016 or 2021 election. $Feature_g$ represents one of the 6 GP-level variables. ϵ_{ig} is an error term clustered at the GP level. Our coefficient of interest is β_1.

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.

4.1.2 Results

Table 8[8] shows us how selection varies for Mukhiya 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, as columns (2) and (3) 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. Finally, 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).

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 8 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.

4.2 Causes of Political Selection

We now move to describing how political selection is influenced by policy. In particular, we focus on two policies: first, the move to devolve financial powers to ward members in 2016 and second, gender reservation for the post of Mukhiya.

4.3 Financial Devolution

In late 2016, about six months after they were elected, the 2016 cohort of ward members, in a surprise move, were tasked with implementing two key WAS public goods. This was the first time in the history of Bihar that ward members had direct access to the state's pool of resources: every ward was given its own bank account to which funds were transferred. Of the two WAS public goods, the lanes and drains projects were devolved to all wards, while the piped water projects were not. In particular, for about 31,372 wards that the state deemed had polluted groundwater, implementation was handled by

[8] The equivalent table for wards is Table 23. We do not focus on these because we do not have ward level covariates.

the Public Health and Engineering Department (PHED), a parallel bureaucratic wing of the state. This created variation in amount of devolution of funds to wards, often *within* the same GP: while 63,446 wards got, on average, \$35700 (we refer to these as "treated wards"), 31,372 wards got, on average, \$9800 (we refer to these as "control" wards). Put differently, treated wards had a budget 3.5x of control wards.

From the Government of Bihar, we obtain the list of wards where the water was deemed to be polluted. These wards are our control wards, where financial devolution was relatively lower. Our 2016 data on ward members can be seen as baseline data: candidates contesting the ward elections in 2016 had no inclination that financial powers would be devolved to them. These are, therefore, pre-treatment candidates. Candidates in the 2021 ward elections are post-treatment candidates, contesting with the clear knowledge of how much funds were devolved to wards in the previous five years.

4.3.1 Empirical Strategy

To investigate how access to financial resources affected political selection, we use a difference-in-differences identification strategy. The estimating equation:

$$\begin{aligned} Y_{iwgy} =& \beta_1\{\mathbb{1}\{FinancialAccess\}_{wg} \times \mathbb{1}\{y=2021\}\} \\ &+ \beta_2\mathbb{1}\{FinancialAccess\}_{wg} + \beta_3\mathbb{1}\{y=2021\} + X'_{iwgy} + \gamma_g + \epsilon_{iwgy} \end{aligned} \tag{4}$$

where Y_{iwgy} is the outcome of interest for candidate i competing in election year y at ward w of GP g. The indicator $FinancialAccess_{wg}$ is 1 if ward w in GP g is one of the 63,446 wards that received access to state fund in late 2016 for piped water projects implementation. X'_{iwgy} is a vector of individual-level demographic controls (gender, age and caste) to improve precision. GP-level fixed effects are also included to control for all time-invariant differences across GPs. Standard errors ϵ_{iwgy} are clustered at Ward-level.

Our identification relies on the assumption that the groundwater pollution is random and hence candidates from treated and control wards should share similar characteristics prior to the announcement of financial devolution. This assumption is supported by balance test shown in Table 24.

4.3.2 Results

We now proceed to analyse how financial devolution affects selection. Table 9 indicates the results. We document four results: first, devolution increases the number of candidates running; second, turning to quality, the average education of treated candidates marginally increases and they appear to come from marginally wealthier households; third, turning to winners in Panel B, we see that the winner pool is slightly wealthier and has somewhat higher amounts of education, but this latter effect is statistically not significant; fourth, in Table 10, we show that the representativeness of candidates and winners doesn't change in a major way in treated wards.

Overall, our results indicate that the main way in which financial devolution affects political entry is through increasing the number of candidates contesting. The effects on candidate quality and representativeness is small and there is an even more muted effect on the pool of winners. These results could also be interpreted to be informative on the competence-representativeness trade-off: when there is no impact on competence, there seems to be no real impact on representativeness either.

4.4 Gender Reservation

We turn to how gender reservation affects political selection. The effect of gender reservation on the pool of candidates contesting elections and the characteristics of winners has been studied previously. Women candidates in gender-reserved seats are typically less educated and less experienced than the average candidate in non-gender reserved seats (Chattopadhyay and Duflo (2004), Afridi et al. (2017)). We first test to see if these findings replicate in our context too.

Yet, our setting allows us to do more: since we know which households candidates belong to, we can study if gender reservation changes the type of households contesting elections. We already know that gender reservation democratises the candidate pool, but does it also do the same to the candidate household pool? *Ex ante*, it is unclear whether this would occur: if female candidates are mere proxies to other male candidates, then reservation should not change the pool of households contesting elections. On the other hand, if this is not the case, there are reasons to believe that women's agency varies with their social location in the village: for instance, if caste norms disallow women from certain groups to contest elections, then this should reflect in the type of households contesting elections in gender-reserved seats. Indeed, this is what (Cassan and Vandewalle, 2021) find. They show that gender reservation crowds in low-caste households and this, in line, aligns government functioning with low-caste citizen preferences.

We extend previous results along two dimensions: first, turning to candidate households, because Bihar's reservation algorithm reserves seats for women within caste, we can ask how gender reservation change selection of households in caste-reserved and caste-unreserved GPs. The caste unreserved results are a parallel to (Cassan and Vandewalle, 2021), but the caste-reserved seats are different entirely; second, our measures of selection extend beyond caste, showing results on education, asset wealth and occupation of the household.

4.4.1 Empirical 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 (a collection of about 15 GPs), the governments first determines which GPs should have seats reserved for SC, ST, OBC or none. Subsequently, within each grouping, the government reserves 50% of seats for women. The details are presented in the Appendix (section 9), 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.[9] Thus, population thresholds exist for each of these groups too.

Our running variable is simply the distance of the GP's relevant population from the threshold population. The threshold population is the mean of the relevant population of the last GP to be reserved for women (call it GP "1") and the first GP to not be reserved (call it GP "0"):

$$\begin{aligned} Running_{gb} &= WomenPop_{gb} - ThresholdPop_b \\ &= WomenPop_{gb} - (\frac{WomenPop_{1b} + WomenPop_{0b}}{2}) \end{aligned} \tag{5}$$

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 variables specification:

$$\begin{aligned} Reserved_{gb} =& \gamma_1 \mathbb{1}\{Running_{gb} > 0\} + \gamma_2(Running_{gb}) + \\ & \gamma_3 \mathbb{1}\{Running_{gb} > 0\} \times (Running_{gb}) + X'_{gb} + \zeta_b + \eta_{gb} \end{aligned} \tag{6}$$

$$\begin{aligned} Y_{gb} =& \beta_1 Reserved_{gb} + \beta_2(Running_{gb}) + \\ & \beta_3(Running_{gb}) \times Reserved_{gb} + X'_{gb} + \phi_b + \epsilon_{gb} \end{aligned} \tag{7}$$

where Y_{gb} is the outcome of interest in GP g of block b. The indicator $Reserved_{gb}$is a binary indicator that equals 1 if GP g is reserved for female candidates. Block level fixed effects ζ_b and ϕ_b are included as the cutoff thresholds vary across blocks. 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. Standard errors ϵ_{gb} and η_{gb} are clustered at block 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.

[9]The Panchayati Raj Act specifies that *up to* 50% of seats can be reserved for women. Thus, if there are 3 reserved seats for SCs, only one will be reserved for women, since reserving two for women will exceed the 50% threshold. This rule applies for reservation for women across STs, OBCs and open caste seats too. This mechanically implies that fewer than 50% of seats are reserved for women overall.

We also report results within separate caste reservation groups.

4.4.2 Results

Individual Level: We begin by showing that reservation changes the pool of candidates (Table 25): fewer candidates contest overall, but the composition of the candidate pool changes entirely: an additional 5.1 female candidates enter the race in gender reserved seats compared to unreserved counterparts.

We then test to see how candidate education outcomes are affected by reservation: unsurprisingly, as Table 11 shows, candidates in gender-reserved GPs are much less educated than their unreserved counterparts. The coefficient on the RD estimate indicates a 0.63 SD reduction in years of schooling for candidates and 0.89 SD reduction in years of schooling for winners. This translates to 2.51 fewer years for the average candidate and 3.6 fewer years for winners.

We also see that candidates and winners in gender-reserved seats are worse educated than women who contest and win in unreserved seats. This is not entirely unexpected: women candidates and winners in non-gender reserved seats are likely to be particularly positively selected since they face high costs of entry and likely have to overcome patriarchal norms to contest.

Household Level: We then study how gender reservation changes household-level selection. We focus on 3 groups of variables: income, wealth and education (Table 12 and 13). Our results indicate that gender reservation results in the political entry of households with lower income (0.1 SD less likely to self-report that their income levels are high, i.e above 10,000 INR a month), wealth (0.9 SD reduction in wealth scores) and education (0.4 SD reduction in the education levels of the highest member). In other words, our results indicate that gender reservation democratizes the pool of candidate households.

The differences between households of winners in gender reserved and unreserved seats is even more stark (Table 13). In absolute terms, the gap between reserved and unreserved winners is twice as large as the corresponding gap between reserved and unreserved candidates.

Analysis by Caste Reservation Status: Table 26 breaks down the results by caste reservation status. While for both caste unreserved and OBC-reserved seats, we do find that candidate households have lower education levels (as measured by the highest education years of any member in the household), we do not find any significant impacts on household income or wealth. The coefficients are negative, but not significant. On the other hand, for SC reserved seats, gender reservation clearly results in the entry of lower income/wealth status households. Thus the democratisation of households due to gender reservation is driven by the most marginalised caste groups.

Our results point to some interesting trends: first, that gender reservation democratizes the pool of candidate households, encouraging participation of members from households lower in the income and asset wealth distribution; second, these effects are largest in GPs that are reserved for SCs and the effects are relatively small (and insignificant) for caste-unreserved GPs. This suggests that democratisation occurs most when gender and caste-reservations intersect. Finally, there appears to be a tradeoff here too between representativeness and competence (as measured by education): gender reservation not only brings in candidates who are less educated, but also allows households as a whole that are less educated to contest. Thus, at the household level, reservation improves representatives, but lowers competence.[10]

5 Result: Does Education = Competence?

Throughout this paper, as we explained in section 2.4, we use education as a proxy for competence. We now ask: is this empirically true in our context? More broadly, when does education imply competence and when does it not?

To understand the effects of education of the winning candidate on policy outcomes, we need the following: (a) a shock that exogenously varies the education levels of the winning candidate; (b) an outcome that reflects competence.

Outcome Measures: We focus on 3 sets of outcome measures, each reflecting somewhat different metrics of competence.

Ward Member Knowledge: Using data from a survey with over 3700 randomly sampled incumbent ward members across 10 districts in Bihar, we measure knowledge. We ask ward members about the steps involved in implementation of six key schemes. These are: (i) opening the "ward account", (ii) implementation of the housing scheme (AWAS), (iii) lanes and drains (NALIGALI), (iv) obtaining a ration card (PDS), (v) obtaining pension benefits and (vi) installing solar lights. We then create a "knowledge index" that is the mean of the standardized step counts for each of the individual schemes. This index is our main metric of ward members' knowledge.

Mukhiya Scheme Implementation: Next, using the universe of Mukhiya candidates contesting in the 2016 elections, we focus on 2 main outcomes: persondays generated under the Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS; henceforth "NREGA"). This is a pro-poor scheme with documented excess demand: more persondays could, therefore, signal more competent Mukhiyas. However, given the disproportionate benefits to women and members of disadvantaged castes, elite Mukhiyas – even if otherwise competent – may not prefer to implement the scheme in spirit for fear of distributing benefits to those unlike them.

To assuage these concerns, we also look at another scheme: the WAS schemes mentioned above. Even though ward members were to implement these schemes, Mukhiyas played a

[10] In the appendix section 10, we study the consequences of gender reservation on implementation of NREGA projects and WAS projects. We find largely null results. This suggests that the competence and representativeness tradeoff may cancel each other out.

prominent role because the funds flowed through them to the wards and they often helped ward members navigate the bureaucracy. These WAS schemes, as mentioned previously, were an extremely important aspect of the incumbent state government's development platform for the period 2016-21.

Taken together, we have outcomes across both tiers of government – ward and GP – and across 3 different metrics: knowledge, performance on an important welfare program, implementation of key local infrastructure projects.

5.1 Data Sources

Knowledge: We use a primary survey of over 3700 candidates to measure knowledge of ward members. These ward members were elected in late 2021 and the survey collects data between 12-18 months after they came to power.

NREGA Data: We scrape data on the the MIS (www.nrega.nic.in) for the period 2016-2022. Our main outcome variables are: total annual persondays generated and persondays generated for SCs/STs/women. For the cohort of Mukhiyas elected in 2016, we focus on outcomes from 2016-2021.

Politician Data: Our data on politicians is the same as used before. Since we have education data from two sources - the SEC data and the SECC data. While we use the SEC data in our main specifications, we show that all our results are robust to using either data source.

5.2 Empirical Strategy: Close Election RD

To use a close election RD design, we first restrict our sample to GPs/wards where winner and runner in the 2016 Mukhiya/ward election possess different levels of education background. For Mukhiya candidates, we focus on races between two candidates, one below primary and one above primary. For the ward races, for reasons of power, our main table shows results for races where one candidate has above high school education and the other has below high school.

For each GP/ward, only the person with education level above primary/high school is kept in the data. Using this sample, the effects of having a more educationally competent candidate on development outcome is then estimated using the following specification:

$$\begin{aligned} Y_g = &\beta_1 \mathbb{1}\{WinElection\}_g + \beta_2 VictoryMargin_g + \\ &\beta_3 VictoryMargin_g \times \mathbb{1}\{WinElection\}_g + \epsilon_g \end{aligned} \tag{8}$$

where Y_g is the development outcome of interest for GP g. Indicator $\mathbb{1}\{WinElection\}_g$ equals 1 if the candidates with above primary education won the 2016 Mukhiya election of GP g. The running variable $VictoryMargin_g$ represents the margin of victory of the winner/runner with above primary education in GP g.

At the outset, we recognize that such "politician characteristic RD designs", while useful, are not ideal. In particular, as Marshall (2022), among others, has pointed out, the treatment effects are picking up not just the effects of education, but also politician features that make high-educated and low-educated politicians run in close elections.

5.3 Results

Knowledge: Table 15 presents the impacts of having a high -school educated ward member win a close election on knowledge outcomes. We see the knowledge index - which is the average of a series of standardized variables - improves by 0.61 units. Columns (2) - (7) break down the improvements by scheme and we see that there is some heterogeneity on how being more educated translates to changes in knowledge: while more educated ward members know how to open ward accounts, implement the housing and solar light schemes, the effects are smaller for the village lanes and drains and pension schemes. Overall, the results indicate that education improves ward member knowledge of scheme implementation. However, does this translate to improve scheme outcomes? We test for this below.

Scheme Outcomes: We focus on close races between candidates who have had different levels of education: above primary education and below primary education. We first show that most covariates are balanced in Table 27.

Next, we study if higher educated Mukhiya candidates improve development outcomes. As Table 14 shows, we do not find any large and significant effects of education of the winner on outcomes: if anything, education marginally reduces NREGA outcomes, but may marginally improve village projects. Both these results are largely insignificant. Table 28 breaks down the results by reservation status of GP. Once more, we do not find distinct patterns by gender/caste reservation status of GP. Overall, the resutls indicate that education does not strongly predict performance in office.

Taken together, the close election RD results indicate that while more education implies better knowledge of scheme implementation for ward members. However, we cannot reject the null that education levels of candidates are uncorrelated with actual scheme outcomes. These schemes require somewhat different skill-sets to implement effectively. Yet, neither of these seem to be causally affected by the education levels of the Mukhiya candidate.

6 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 more educated than others.

This paper also suggests a trade-off between representativeness and competence. Policies aimed at empowering disadvantaged groups - like political reservations for women or members of marginalized caste groups – increase representativeness, not merely of the candidates, but also crowd in a more diverse and less elite set of households. However, such policies also decrease the relative educational levels of these candidates (even when compared to other individuals within their own group). Such a trade-off is also visible when comparing lower and higher-tiers of local government: higher-tiered candidates are more competent and less representative than lower-tiered ones.

While this paper shows the existence of the trade-off between representativeness and competence, our results call into question the salience of this trade-off in predicting performance in office. More educated politicians may have better knowledge regarding schemes, but this does not necessarily translate to improved development outcomes. Higher educated politicians have a marginal effect on performance along two key rural development programs: the NREGA and laying of village lanes and piped water projects.

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.

Afridi, F., V. Iversen, and M. R. Sharan (2017): "Women political leaders, corruption, and learning: Evidence from a large public program in India," *Economic Development and Cultural Change*, 66, 1–30.

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.

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.

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.*

Lahoti, R. and S. Sahoo (2020): "Are educated leaders good for education? Evidence from India," *Journal of Economic Behavior & Organization*, 176, 42–62.

Marshall, J. (2022): "Can close election regression discontinuity designs identify effects of winning politician characteristics?" *American Journal of Political Science.*

Martinez-Bravo, M. (2017): "The local political economy effects of school construction in Indonesia," *American Economic Journal: Applied Economics*, 9, 256–289.

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*.

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.

7 Tables

Table 1: 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

Focusing on the GP and ward elections of 2016-2021, this table shows how characteristics of those politicians we match in the SECC data vary from those we do not find. Panel (A) presents the balance tests for candidates while panel (B) presents the balance tests for candidates' fathers. Age is winsorized at top and bottom 1 percentile. Standard errors are clustered GP levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Political selection on education and caste for unreserved seats (adult sample)

	Education				Caste		
	Yrs of Schooling (1)	At least Literate (2)	At least Secondary (3)	At least Higher (4)	SC (5)	ST (6)	Other (7)
Candidates	0.594^{***}	0.428^{***}	0.450^{***}	0.405^{***}	-0.114^{***}	-0.074^{***}	0.132^{***}
	(0.004)	(0.003)	(0.004)	(0.005)	(0.003)	(0.002)	(0.003)
Citizen Mean	3.811	0.544	0.373	0.092	0.157	0.016	0.827
Citizen SD	4.470	0.498	0.484	0.289	0.364	0.125	0.378
Observations	54,498,666	54,498,515	54,498,642	54,497,849	54,471,654	36,585,002	54,475,734

The table describes the political selection in Bihar for electoral candidates (competing in unreserved seats) in terms of education level and caste. Only citizens above the age of 18 and 2016(2021) candidates above the age of 22(27) are included in the sample. The dependent variables are normalized by substracting the population mean and divided by standard deviation respectively within each GP. Respondents highest education levels reported in the SECC survey are re-categorized into 5 groups and these are 'illiterate', 'literate', 'primary and secondary', 'higher secondary' and 'graduate'. Respondents caste statuses are recorded as 'scheduled caste (SC)', 'scheduled tribe (ST)' and 'other'. Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Political selection across various reference groups for unreserved seats

	Years of Schooling					
	Population (1)	Gender (2)	Age Cohort (3)	Caste (4)	Jatin (5)	Household (6)
Candidates	0.673***	0.561***	0.603***	0.651***	0.566***	0.498***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	97,349,648	97,349,648	97,348,523	97,349,648	93,530,932	96,386,588

The table compares the years of education between electoral candidates (competing for unreserved seats) with all citizens, citizens from their gender group, citizens from their age cohort, citizens from their caste group, citizens who share their same surname, and other members of their households. The dependent variable 'years of education' is normalized by substracting the population mean and divided by standard deviation respectively within each GP. In column (2)-(6), we respectively include fixed effects on gender group, age cohort, caste group, surname and household id. Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Political selection on income and wealth for unreserved seats

	Income				Wealth		
	Gov. Job (1)	Salaried Job (2)	Income Tax (3)	Income High (4)	Wealth Score (5)	Ag. Equipment (6)	Total Land (7)
Candidates	0.043***	0.040***	0.028***	0.075***	0.199***	0.111***	0.125***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Citizen Mean	0.057	0.068	0.031	0.080	0.071	0.030	1.828
Citizen SD	0.231	0.253	0.174	0.272	1.434	0.171	9.855
Observations	97,291,132	97,278,538	97,213,207	97,285,243	97,278,480	97,213,179	97,336,899

The table describes the political selection in Bihar for electoral candidates (competing for unreserved seats) in terms of income and wealth level. The dependent variable is a binary indicator for whether any household member has a government job (column 1), has a salaried job (column 2), pay income tax (column 3), has monthly income greater than 5,000 Rs. (column 4), owns mechanized wheeler agricultural equipment (column 6). The depenent variable 'Wealth Score' (column 5) is a PCA score constructed by using respondents land ownership, home wall condition, roof material, number of rooms, phone ownership and vehicle ownership. The depdendent variable 'Total Land' (column 7) is the total number of land owned by the household, top-coded at the value of 120. All dependent variables are normalized by substracting the population mean and divided by standard deviation respectively within each GP. The dependent variables are. Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Political selection and relationship with 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 6: Representativeness trade-off on Education and Caste

	Education				Caste		
	Yrs of Schooling	At least Literate	At least Secondary	At least Higher	SC	ST	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
*Panel A - **Mukhiya**: Competence - Comparison between Citizens and Candidates*							
Candidates	1.088***	0.538***	0.712***	1.089***	-0.238***	-0.082***	0.254***
	(0.011)	(0.006)	(0.008)	(0.015)	(0.005)	(0.004)	(0.005)
*Panel B - **Mukhiya**: Representativeness - Comparison between Citizens and Candidate Fathers*							
Candidates Father	0.438***	0.277***	0.323***	0.350***	-0.415***	-0.041***	0.420***
	(0.011)	(0.009)	(0.009)	(0.013)	(0.007)	(0.004)	(0.007)
*Panel C - **Ward**: Competence - Comparison between Citizens and Candidates*							
Candidates	0.592***	0.367***	0.430***	0.418***	-0.121***	-0.072***	0.139***
	(0.004)	(0.003)	(0.004)	(0.005)	(0.003)	(0.002)	(0.004)
*Panel D - **Ward**: Representativeness - Comparison between Citizens and Candidate Fathers*							
Candidates Father	0.147***	0.144***	0.130***	0.038***	-0.691***	-0.061***	0.703***
	(0.007)	(0.006)	(0.006)	(0.007)	(0.009)	(0.005)	(0.009)
Citizen Mean	3.461	0.559	0.361	0.059	0.166	0.023	0.818
Citizen SD	3.983	0.496	0.480	0.236	0.372	0.151	0.386
Observations	97,349,645	97,349,371	97,349,612	97,348,291	97,300,487	65,208,192	97,308,381

The table describes the trade-off between competence and representativness in Bihar for Mukhiya/Ward (Panel (A-B)/(C-D)) male candidates (competing for unreserved seats) in terms of education level and caste. The dependent variables are normalized by subtracting the population mean and divided by standard deviation respectively within each GP. Respondents highest education levels reported in the SECC survey are re-categorized into 5 groups and these are 'illiterate', 'literate', 'primary and secondary', 'higher secondary' and 'graduate'. Respondents caste statuses are recorded as 'scheduled caste (SC)', 'scheduled tribe (ST)' and 'other'. Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Representativeness trade-off across various reference groups

	Years of Schooling					
	Population (1)	Gender (2)	Age Cohort (3)	Caste (4)	Jatin (5)	Household (6)
*Panel A - **Mukhiya**: Competence - Comparison between Citizens and Candidates*						
Candidates	1.088***	0.975***	1.046***	1.054***	0.911***	0.631***
	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)	(0.009)
*Panel B - **Mukhiya**: Representativeness - Comparison between Citizens and Candidate Fathers*						
Candidates Father	0.527***	0.392***	0.743***	0.484***	0.352***	0.082***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)
*Panel C - **Ward**: Competence - Comparison between Citizens and Candidates*						
Candidates	0.592***	0.480***	0.516***	0.572***	0.499***	0.472***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
*Panel D - **Ward**: Representativeness - Comparison between Citizens and Candidate Fathers*						
Candidates Father	0.193***	0.058***	0.385***	0.162***	0.095***	-0.021***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	97,349,645	97,349,645	97,348,520	97,349,645	93,530,798	96,386,585

The table compares the years of education between Mukhiya/Ward male candidates (panel A/C) of unreserved seats as well as the fathers of Mukhiya/Ward male candidates (Panel B/D) across all citizens, citizens from their gender group, citizens from their age cohort, citizens from their caste group, citizens who share their same surname, and other members of their households. The dependent variable 'years of education' is normalized by substracting the population mean and divided by standard deviation respectively within each GP. In column (2)-(6), we respectively include fixed effects on gender group, age cohort, caste group, surname and household id. Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Political selection and relationship with 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						
Candidates	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						
Candidates 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

The table shows how do politicial selection varies by village characteristics. The dependent variable is years of schooling, normalized by substracting the population mean and divided by standard deviation respectively within each GP. In column (1)-(6), we respectively interact a binary indicator for Mukhiya male candidates with a Panchayat-level standardized PCA wealth score constructed by using respondents land ownership, home wall condition, roof material, number of rooms, phone ownership and vehicle ownership (column 1), a Panchayat-level standardized ethnic fractionalization index constructed by taking one minus the sum of squares of the share of population for each Jatin (column 2), a Panchayat-level standardized ethnic polarization index constructed by replicating the Reynal-Querol index from J of Conflict Resolution, 2002 (column 3), a Panchayat-level education mobility index constructed by taking the fraction of children in the age between 14-18 who completed primary education but their parents did not (column 4), a (total) land inequality Gini-index (column 5) and the distance from the Panchayat to the district head quarter (column 6) . Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Effects of Financial Incentives on Political Selection

	Education				Wealth			Participation
	Yrs of Schooling	At least Literate	At least Secondary	At least Higher	Gov. Job	Total Land	Wealth Score	N. of Candidates
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Candidates Outcomes							
Financial Access	-0.080*	-0.016**	-0.012**	-0.000	-0.001	0.154	-0.008	-0.042**
	(0.041)	(0.006)	(0.005)	(0.003)	(0.002)	(0.124)	(0.015)	(0.017)
2022 Election	0.173***	0.015***	0.021***	0.007***	0.002**	-0.073**	0.013	1.769***
	(0.025)	(0.004)	(0.003)	(0.001)	(0.001)	(0.033)	(0.009)	(0.011)
Finance X 2022	0.040*	0.006	0.004	0.000	0.001	0.112***	0.012	0.086***
	(0.023)	(0.004)	(0.003)	(0.001)	(0.001)	(0.031)	(0.009)	(0.012)
Control Mean	4.011	0.579	0.390	0.090	0.041	1.639	0.095	2.905
Observations	365,721	365,721	365,721	365,721	365,709	365,721	365,454	178,508
	Panel B: Winner Outcomes							
Financial Access	-0.088	-0.016**	-0.012	-0.001	-0.006	0.118	-0.017	
	(0.058)	(0.007)	(0.009)	(0.004)	(0.004)	(0.152)	(0.023)	
2022 Election	0.298***	0.028***	0.037***	0.009***	0.006**	0.136	0.047***	
	(0.048)	(0.005)	(0.005)	(0.003)	(0.002)	(0.089)	(0.012)	
Finance X 2022	0.033	0.002	-0.003	0.004	0.002	-0.060	0.026*	
	(0.031)	(0.006)	(0.005)	(0.003)	(0.003)	(0.087)	(0.014)	
Control Mean	4.209	0.592	0.408	0.101	0.046	1.670	0.125	
Observations	100,627	100,627	100,627	100,627	100,624	100,627	100,560	

This table reports difference-in-differences estimates of the effect of financial access on political entry and selection. Panel (A) reports the effects on candidate fathers, while Panel (B) shows the results for winners. Regressions are estimated according to Equation 4 with GP-level fixed effects and individual-level demographic controls. Standard errors are reported in parentheses and clustered at ward-level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Effects of financial devolution on candidate and winner father outcomes

	Education				Wealth		
	Yrs of Schooling	At least Literate	At least Secondary	At least Higher	Gov. Job	Total Land	Wealth Score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Candidate Father Outcomes							
Financial Access	-0.179*	-0.022*	-0.013	-0.009	0.001	0.172	-0.012
	(0.088)	(0.012)	(0.010)	(0.005)	(0.006)	(0.162)	(0.022)
2022 Election	0.253***	0.026***	0.025***	0.006***	-0.005**	-0.146	-0.009
	(0.032)	(0.004)	(0.003)	(0.002)	(0.002)	(0.088)	(0.017)
Finance X 2022	-0.022	-0.005	-0.004	0.003	0.002	0.050	0.003
	(0.045)	(0.006)	(0.005)	(0.003)	(0.002)	(0.107)	(0.019)
Panel B: Winner Father Outcomes							
Financial Access	-0.049	-0.007	0.006	-0.001	-0.017**	0.314	0.043
	(0.141)	(0.016)	(0.014)	(0.010)	(0.007)	(0.289)	(0.043)
2022 Election	0.427***	0.043***	0.042***	0.014**	-0.006	0.301*	0.028
	(0.070)	(0.009)	(0.009)	(0.006)	(0.007)	(0.160)	(0.034)
Finance X 2022	-0.108	-0.015	-0.018	-0.000	0.010	-0.354**	0.020
	(0.105)	(0.013)	(0.013)	(0.007)	(0.006)	(0.167)	(0.041)
Control Mean	3.668	0.537	0.354	0.069	0.074	2.198	0.296
Observations	36,741	36,741	36,741	36,741	36,740	36,741	36,719

This table reports difference-in-differences estimates of the effect of financial access on political selection (representativeness). Panel (A) reports the effects on candidate fathers, while Panel (B) shows the results for winners. Regressions are estimated according to Equation 4 with GP-level fixed effects and individual-level demographic controls. Standard errors are reported in parentheses and clustered at ward-level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Effects of reservation on candidates & winner competence

	Candidates				Winner			
	Yrs of Schooling (1)	At least Literate (2)	At least Secondary (3)	At least Higher (4)	Yrs of Schooling (5)	At least Literate (6)	At least Secondary (7)	At least Higher (8)
Panel A: Comparison to All Mukhiya Candidates/Winners								
Female RSVN	-0.629***	-0.085***	-0.873***	-4.132***	-0.885***	-0.051**	-1.174***	-6.881***
	(0.031)	(0.017)	(0.047)	(0.362)	(0.061)	(0.021)	(0.092)	(0.694)
Bandwidth	415.83	455.48	541.74	693.40	469.09	773.18	477.30	460.72
Control Mean	0.832	0.966	0.664	3.038	0.999	0.940	0.796	5.455
Observations	2,679	2,821	3,072	3,419	1,973	2,457	2,001	1,964
Panel B: Comparison to Female Mukhiya Candidates/Winners Only								
Female RSVN	-0.137***	-0.029	-0.234***	-0.581	-0.469***	-0.114*	-0.642***	-4.277***
	(0.037)	(0.021)	(0.062)	(0.379)	(0.131)	(0.063)	(0.199)	(1.334)
Bandwidth	1003.44	1016.71	1036.06	1021.97	707.91	666.57	595.92	617.94
Control Mean	0.182	0.857	-0.216	-1.118	0.214	0.881	-0.289	-0.167
Observations	2,950	2,974	2,990	2,976	1,221	1,205	1,155	1,170

This table shows the effects of female reservation on candidate and winner educational competence level. Only 2016 candidates are included in the sample. The data is at GP-level. Column (1)-(4) show the results on candidates educational background and column (5)-(8) show the results on winner background. In Panel (A), we compare the average education levels of candidate/winner from reserved GPs to unreserved GPs, while Panel (B) restrict the samples to female candidates/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 12: Effects of Reservation on Candidates Household Characteristics

	Income				Wealth			Edu
	Gov. Job	Salaried Job	Income Tax	Income High	Wealth Score	Ag. Equipment	Total Land	Highest Edu Yrs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Comparison to All Mukhiya Candidates								
Female RSVN	-0.056*	-0.082***	-0.062**	-0.099***	-0.087***	-0.048	-0.021	-0.103***
	(0.030)	(0.027)	(0.028)	(0.029)	(0.025)	(0.033)	(0.028)	(0.025)
Bandwidth	739.80	914.75	951.58	965.69	912.74	840.29	852.99	903.29
Control Mean	0.239	0.219	0.160	0.278	0.387	0.243	0.232	1.437
Observations	6,762	7,330	7,398	7,487	7,328	7,070	7,142	7,307
Panel B: Comparison to Female Mukhiya Candidates Only								
Female RSVN	-0.207***	-0.169***	-0.054	-0.143***	-0.018	-0.010	0.006	-0.037
	(0.057)	(0.051)	(0.050)	(0.046)	(0.039)	(0.053)	(0.038)	(0.043)
Bandwidth	1124.78	1135.15	1116.21	1192.33	1158.27	947.27	1057.29	1147.26
Control Mean	0.373	0.328	0.175	0.238	0.261	0.156	0.117	1.264
Observations	5,540	5,565	5,501	5,651	5,610	5,222	5,436	5,598

This table shows the effects of female reservation on candidates household-level characteristics. Both 2016 and 2021 electoral candidates are included in the sample. We collapse the data into GP level and respectively calculate the average (standardized) outcome variables across all candidates within the same GP. Panel (A) compares average candidate characteristics from female-reserved GPs with average characteristics of candidates from non-reserved GPs, while in Panel (B), the comparison is restricted to only female candidates between reserved and non-reserved GPs. The running variable 'Population Diff.' represents the difference between the GP population and the cutoff population size (the population cutoff line is calculated by taking the average population size between the last reserved GP and the first unreservedGP). 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 13: Effects of Reservation on Winner Household Characteristics

	Income				Wealth			Edu
	Gov. Job	Salaried Job	Income Tax	Income High	Wealth Score	Ag. Equipment	Total Land	Highest Edu Yrs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Comparison to All Mukhiya Winners								
Female RSVN	-0.218***	-0.204***	-0.167**	-0.172**	-0.189***	-0.028	-0.039	-0.329***
	(0.078)	(0.077)	(0.085)	(0.086)	(0.061)	(0.099)	(0.078)	(0.070)
Bandwidth	504.17	510.17	472.99	465.32	582.77	413.86	420.55	394.03
Control Mean	0.317	0.316	0.305	0.521	0.649	0.434	0.508	1.631
Observations	3,343	3,360	3,233	3,211	3,582	2,995	3,027	2,949
Panel B: Comparison to Female Mukhiya Winners								
Female RSVN	-0.547***	-0.729***	-0.101	-0.561***	-0.274**	-0.058	-0.159	-0.126
	(0.180)	(0.194)	(0.193)	(0.167)	(0.111)	(0.147)	(0.132)	(0.136)
Bandwidth	595.96	576.20	601.68	629.47	625.60	622.27	620.18	661.48
Control Mean	0.501	0.462	0.340	0.456	0.483	0.305	0.272	1.476
Observations	1,845	1,818	1,851	1,893	1,890	1,885	1,881	1,931

This table shows the effects of female reservation on winner household-level characteristics. Both 2016 and 2021 electoral candidates are included in the sample. Panel (A) compares winner characteristics from female-reserved GPs to winner characteristics from non-reserved GPs, while in Panel (B), the comparison is restricted to only female winners only between reserved and non-reserved GPs. The running variable 'Population Diff.' represents the difference between the GP population and the cutoff population size (the population cutoff line is calculated by taking the average population size between the last reserved GP and the first unreservedGP). 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 14: Effects of Political Competence on Local Economic Development

	NREGA Work-Day					Village Projects	
	Total Work-Day	Women Work-Day	Men Work-Day	SC Work-Day	Non-SC Work-Day	Total Projects	Total Amounts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Winner	-1186.210	35.753	-766.001	-675.035*	-484.202	0.175	248256.407
	(1400.801)	(843.269)	(686.812)	(358.431)	(1199.134)	(2.875)	(1295441.285)
Bandwidth	0.15	0.16	0.16	0.13	0.16	0.12	0.12
Control Mean	16,195	8,727	7,621	2,513	13,702	21	10,095,678
Observations	1,633	1,686	1,705	1,423	1,693	1,394	1,343

This table shows the effects of having a more (educationally) competent politician on NREGA and village project outcomes. Regressions are estimated according to equation 8. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 15: Effects of Political Competence on Scheme Knowledge and Solutions

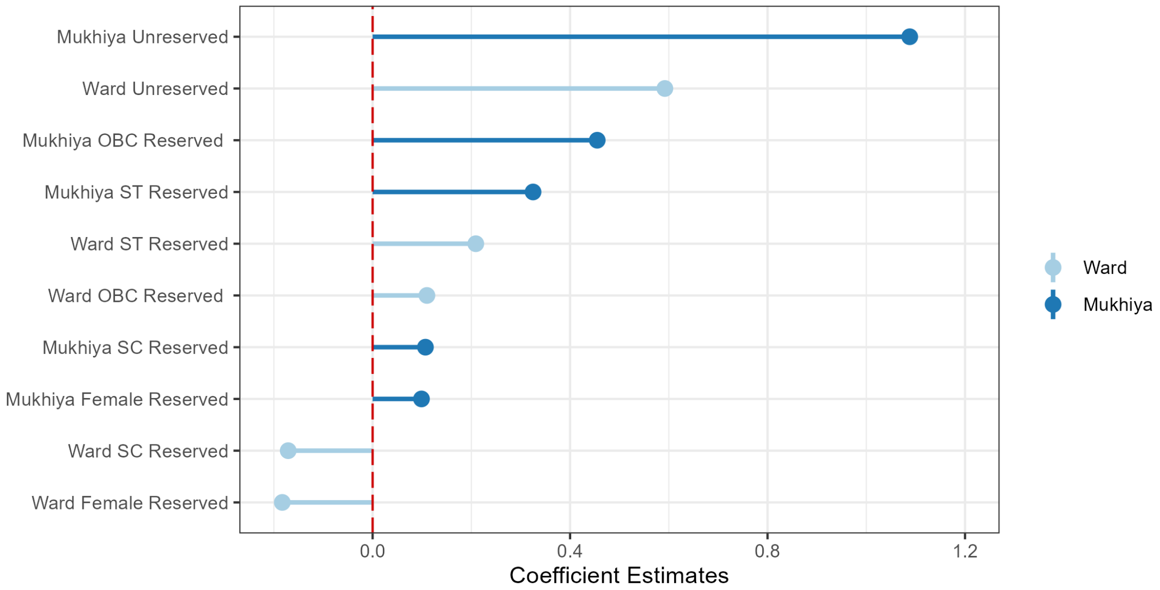
	Knowledge							Solution	
	Avg. (1)	Acct. (2)	AWAS (3)	NALIGALI (4)	PDS (5)	Pension (6)	Solar Light (7)	Total (8)	Avg. (9)
Winner	0.617*	0.712	0.651*	-0.676	-3.017	0.197	1.140**	4.620	4.620
	(0.351)	(0.608)	(0.340)	(0.785)	(2.423)	(1.996)	(0.445)	(6.546)	(6.546)
Bandwidth	0.14	0.13	0.11	0.09	0.04	0.05	0.13	0.70	0.70
Control Mean	1.85	2.51	1.63	1.38	2.19	2.56	1.33	1.80	0.26
Observations	470	451	367	271	52	66	383	133	133

This table shows the effects of having a more (educationally) competent politician on scheme knowledge and solutions. Regressions are estimated according to equation 8 using below and above primary education cutoffs. 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 the sample. Robust tandard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

8 Figures

Figure 1: Competence by Gender and Caste Reservation

(a) In comparison to general citizenry



(b) In comparison to citizenry of their own group

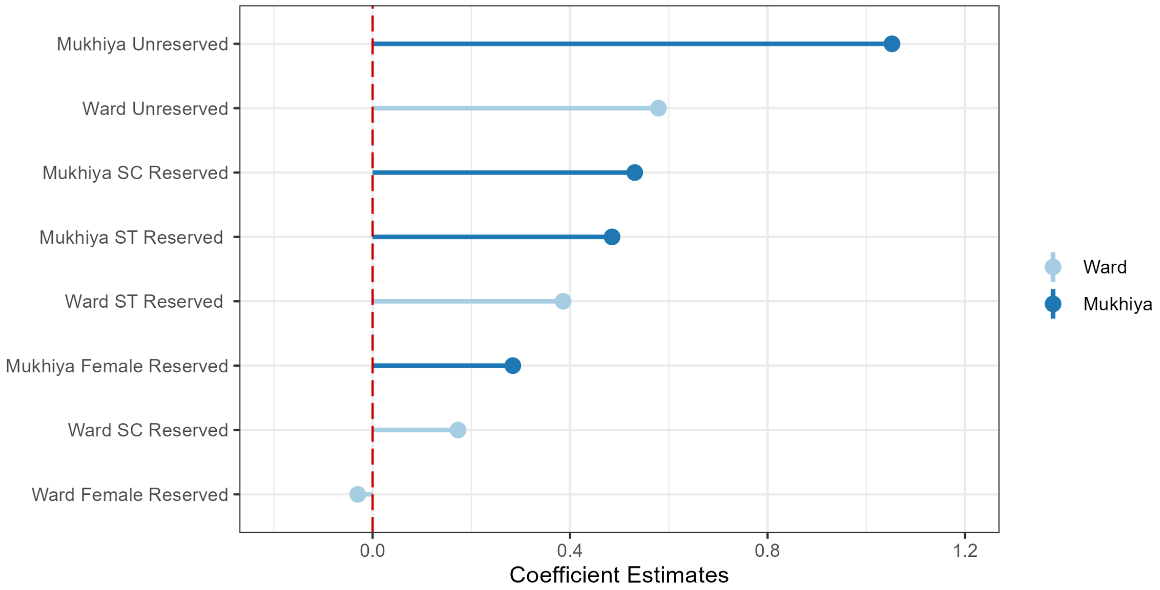
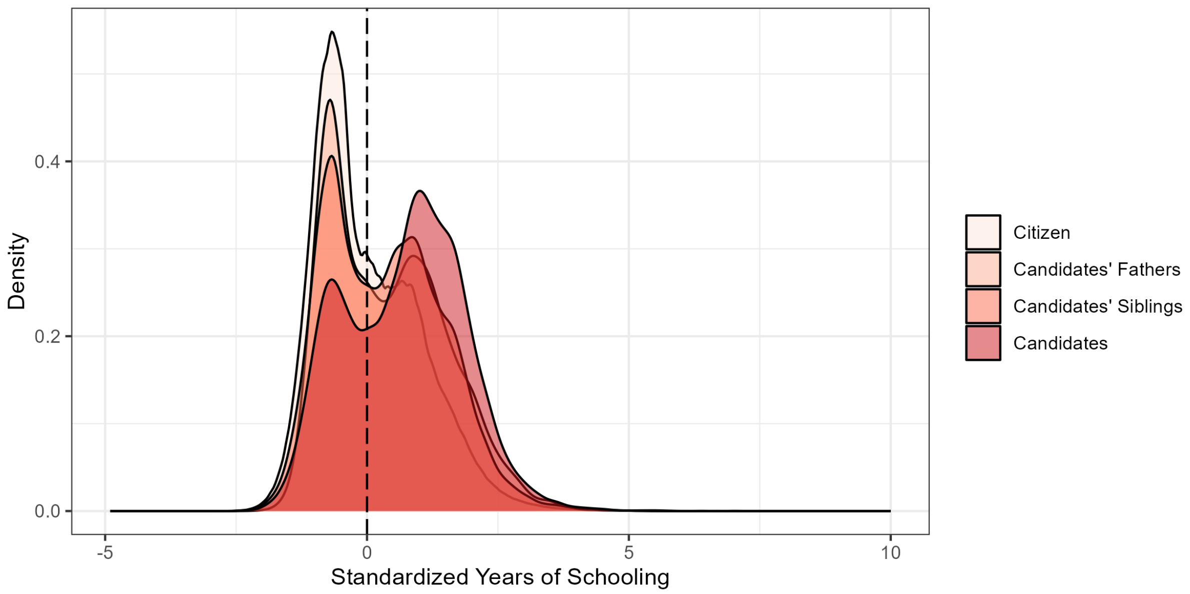


Figure 2: Candidates, Fathers & Siblings

(a) Mukhiya

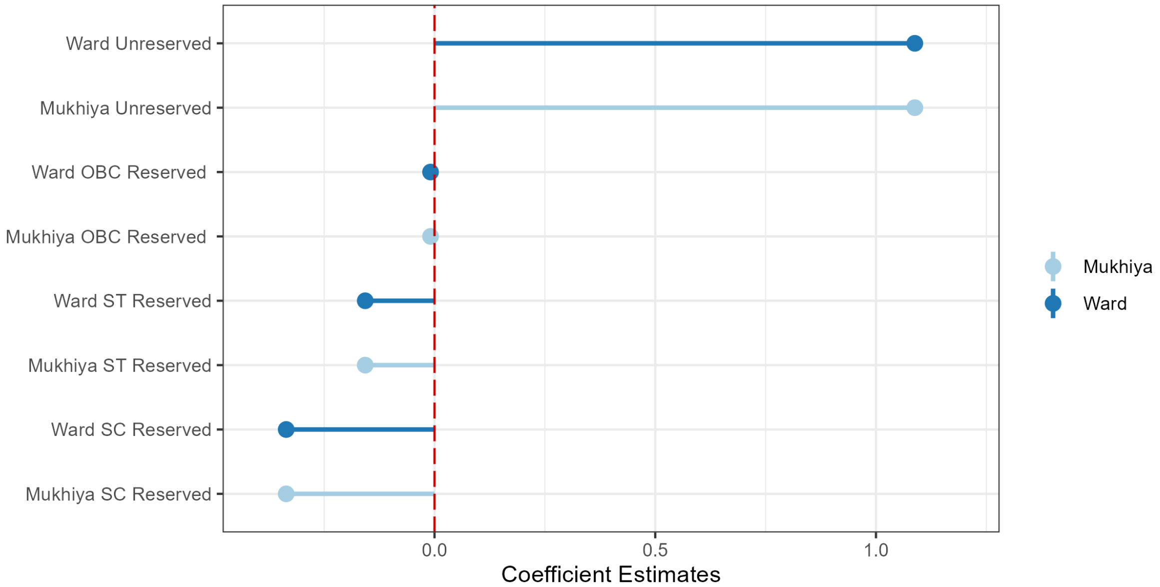


(b) Ward

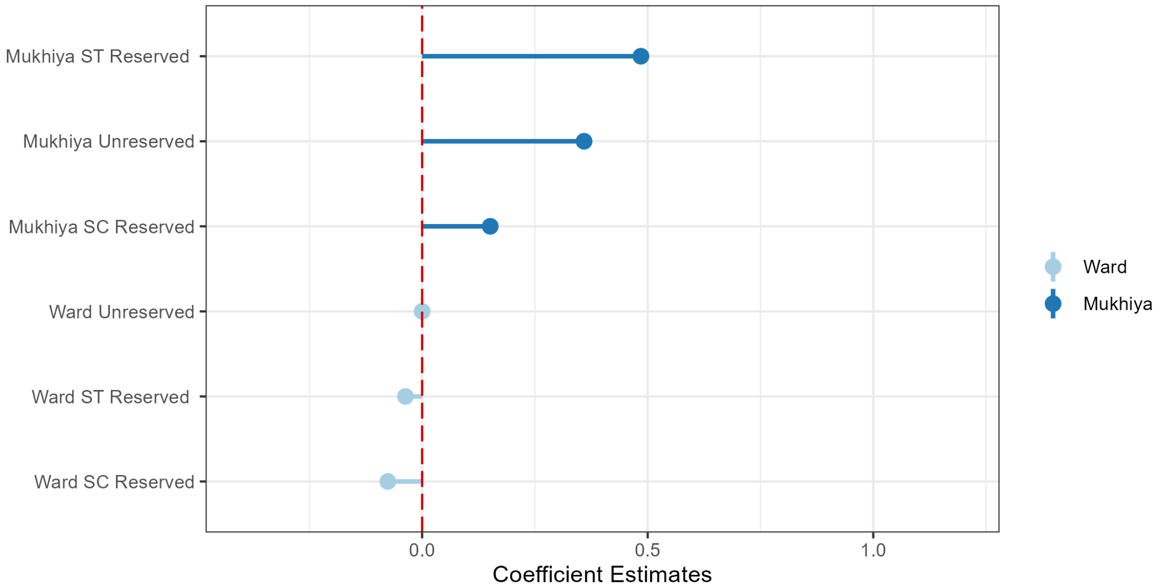
Notes: Years of education is standardized by gender, caste and age cohort within each GP

Figure 3: Representativeness by Gender and Caste Reservation

(a) In comparison to general citizenry



(b) In comparison to citizenry of their own group



9 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).

9.1 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.

9.2 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.

10 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 16 and Table 17 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 16: 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 17: Effects of Reservation on Development Outcomes

	NREGA Work-Day					Village Projects	
	Total Work-Day	Women Work-Day	Men Work-Day	SC Work-Day	Non-SC Work-Day	Total Projects	Total Amounts
	(1)	(2)	(3)	(4)	(5)	(6)	(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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A Appendix

A.1 Tables

Table 18: Balance in observables between matched and non-matched candidates

	Education				Caste			
	Yrs of Schooling	At least Literate	At least Secondary	At least Higher	SC	ST	Other	Age
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(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 19: Balance in observables between matched and non-matched candidates

	Caste			
	SC (1)	ST (2)	Other (3)	Age (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 20: 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 21: 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-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 22: Political selection on education and caste for unreserved seats

	Education				Caste		
	Yrs of Schooling	At least Literate	At least Secondary	At least Higher	SC	ST	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Candidates	0.673***	0.395***	0.476***	0.528***	-0.140***	-0.074***	0.157***
	(0.004)	(0.003)	(0.004)	(0.005)	(0.003)	(0.002)	(0.003)
Citizen Mean	3.456	0.559	0.361	0.059	0.166	0.016	0.818
Citizen SD	3.980	0.497	0.480	0.235	0.372	0.124	0.386
Observations	97,349,648	97,349,374	97,349,615	97,348,294	97,300,490	65,208,195	97,308,384

The table describes the political selection in Bihar for electoral candidates (competing in unreserved seats) in terms of education level and caste. The dependent variables are normalized by substracting the population mean and divided by standard deviation respectively within each GP. Respondents highest education levels reported in the SECC survey are re-categorized into 5 groups and these are 'illiterate', 'literate', 'primary and secondary', 'higher secondary' and 'graduate'. Respondents caste statuses are recorded as 'scheduled caste (SC)', 'scheduled tribe (ST)' and 'other'. Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 23: Political selection and relations with village characteristics for Ward male

	Years of Schooling					
	Wealth Score	Ethnic FRACT.	Ethnic POLAR.	Edu. Mobility	Gini Index Tot. Land	Distance District HQ
Panel A: Competence - Comparison between Citizens and Candidates						
Candidates	0.592***	0.592***	0.592***	0.455***	0.763***	0.600***
	(0.004)	(0.004)	(0.004)	(0.015)	(0.055)	(0.009)
Candidate X Panchayat Feature	0.012***	-0.002	0.007	0.295***	-0.197***	-0.000
	(0.005)	(0.005)	(0.005)	(0.028)	(0.064)	(0.000)
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						
Candidates Father	0.193***	0.193***	0.193***	0.244***	0.095*	0.204***
	(0.004)	(0.004)	(0.004)	(0.015)	(0.055)	(0.009)
Cand. Father X Panchayat Feature	0.015***	0.021***	-0.012***	-0.110***	0.112*	-0.000
	(0.005)	(0.005)	(0.005)	(0.030)	(0.063)	(0.000)
Observations	97,349,645	97,349,645	97,349,645	97,349,238	97,349,645	97,284,398

The table shows how do politicial selection varies by village characteristics. The dependent variable is years of schooling, normalized by substracting the population mean and divided by standard deviation respectively within each GP. In column (1)-(6), we respectively interact a binary indicator for Ward male candidates with a Panchayat-level standardized PCA wealth score constructed by using respondents land ownership, home wall condition, roof material, number of rooms, phone ownership and vehicle ownership (column 1), a Panchayat-level standardized ethnic fractionalization index constructed by taking one minus the sum of squares of the share of population for each Jatin (column 2), a Panchayat-level standardized ethnic polarization index constructed by replicating the Reynal-Querol index from J of Conflict Resolution, 2002 (column 3), a Panchayat-level education mobility index constructed by taking the fraction of children in the age between 14-18 who completed primary education but their parents did not (column 4), a (total) land inequality Gini-index (column 5) and the distance from the Panchayat to the district head quarter (column 6). Standard errors are clustered at GP level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 24: Balance on observables across treated and control wards

	N	Var. Mean	Financial Access
Scheduled Caste	246751	.24	.00309
			(.00737)
Scheduled Tribe	246751	.02	-.00194
			(.00176)
Other Caste	246751	.74	-.00115
			(.00716)
Schooling (Yrs)	151390	4.11	-.0944
			(.0649)
At Least Literate	151390	.59	-.0171*
			(.0086)
At Least Primary	151390	.41	-.0128
			(.00792)
At Least Higher	151390	.09	.000791
			(.00417)
Age	245976	40.37	.163
			(.103)
Joint p-value			.26

This table shows means of 2016 candidates characteristics in column (2) and correlation between access to financial resources and outcomes in column (3). Age is winsorized at top and bottom 1 percentile. Only candidates from wards in both 2016 and 2021 elections are included. Controls include candidates gender and UGP fixed effects. Standard errors in parentheses and clustered at ward level. Joint p-value tests equality of all coefficients with zero. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 25: Effects of reservation on female & minority representation

	Candidates Count					Candidates Percentage			
	Tot.	Female	SC	ST	OBC	Female	SC	ST	OBC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Reduced Form									
Population Diff.	-1.056***	4.281***	-0.089	-0.025**	-0.346***	0.744***	0.005	-0.003	0.019
	(0.186)	(0.155)	(0.056)	(0.011)	(0.114)	(0.018)	(0.007)	(0.002)	(0.013)
Bandwidth	1048.05	882.21	1365.66	1061.01	1206.76	664.58	1450.41	1016.97	1291.12
Control Mean	6.738	0.778	0.724	0.097	2.696	0.123	0.088	0.013	0.385
Observations	3,637	3,330	4,141	3,650	3,891	2,837	4,297	3,585	4,014
Panel B: Fuzzy RD									
Female RSVN	-1.172***	5.094***	-0.089	-0.026*	-0.359**	0.890***	0.009	-0.004	0.026
	(0.234)	(0.150)	(0.072)	(0.014)	(0.152)	(0.008)	(0.009)	(0.003)	(0.016)
Bandwidth	817.89	909.09	1050.87	915.55	832.07	764.88	979.51	999.32	983.85
Observations	3,169	3,387	3,643	3,399	3,206	3,039	3,524	3,553	3,529

This table shows the effects of female reservation on female as well as minority representation. 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 block, election year and minority reservation status. Standard errors are clustered at Block level and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 26: Effects of Reservation on Candidate (Household) Competence

	Income				Wealth			Edu
	Gov. Job	Salaried Job	Income Tax	Income High	Wealth Score	Ag. Equipment	Total Land	Highest Edu Yrs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Unreserved								
Female RSVN	-0.024	-0.043	0.048	-0.078	-0.072	-0.093	-0.044	-0.386***
	(0.056)	(0.055)	(0.072)	(0.057)	(0.046)	(0.069)	(0.063)	(0.077)
Bandwidth	2027.40	2043.86	2355.22	2418.35	2420.98	2202.14	2501.16	2268.41
Observations	4,658	4,674	4,947	5,030	5,031	4,824	5,091	4,797
Panel B: SC Reserved								
Female RSVN	-0.135**	-0.113*	-0.119**	-0.141***	-0.133**	0.061	-0.016	-0.432***
	(0.066)	(0.060)	(0.055)	(0.052)	(0.059)	(0.043)	(0.037)	(0.112)
Bandwidth	295.09	334.43	333.59	318.35	332.90	257.68	360.52	242.76
Observations	1,704	1,798	1,796	1,774	1,798	1,595	1,860	1,487
Panel C: ST Reserved								
Female RSVN	0.619	0.589	-0.099	0.899	-0.814	0.384	-0.495	-1.506
	(0.699)	(0.713)	(0.168)	(0.685)	(0.905)	(0.730)	(0.346)	(1.342)
Bandwidth	197.40	198.94	196.81	201.86	197.25	181.49	380.50	174.03
Observations	35	35	35	35	35	35	50	36
Panel D: OBC Reserved								
Female RSVN	-0.044	-0.060	-0.091	0.005	-0.055	-0.072	-0.050	-0.269*
	(0.091)	(0.081)	(0.070)	(0.109)	(0.082)	(0.092)	(0.075)	(0.161)
Bandwidth	444.95	465.72	440.60	443.76	553.35	438.05	445.66	444.12
Observations	1,315	1,337	1,307	1,313	1,471	1,301	1,315	1,291

This table shows the effects of female reservation on candidate competence, in terms of household income and wealth level. Only candidates matched in the SECC data are included in the sample. We collapse the data into GP level and respectively calculate the average (standardized) wealth indicators of candidates within the same GP. The running variable "Population Diff." is calculated according to Equation 5. 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 27: Balance Test: Diff in Candidate Characteristics in Close Election - Mukhiya

	N	Dep. Mean	Winner
Woman	2122	.55	-.0304
			(.022)
Scheduled Tribe	2122	.02	.000229
			(.00617)
Scheduled Caste	2122	.19	-.0298*
			(.0175)
Age	2121	38.93	.493
			(.463)
Female Reserved	2122	.49	-.0248
			(.0221)
Minority Reserved	2122	.37	-.0246
			(.0214)
Father Edu Yrs	311	5.29	-.225
			(.562)

This table shows means of candidates characteristics in column (2) and differences around cutoff in column (3). Only 2016 Mukhiya winners and runners who meet the sample restriction are included (using below and above primary education cutoff). Age is winsorized at top and bottom 1 percentile. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 28: Effects of Political Competence on Local Economic Development

	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: Female Reserved							
Winner	-1858.659	277.563	-1173.982	-715.793*	-1022.510	1.298	-365768.722
	(2004.232)	(1122.331)	(1018.256)	(432.552)	(1762.682)	(3.665)	(1830207.562)
Bandwidth	0.14	0.18	0.14	0.14	0.14	0.16	0.13
Control Mean	16,703	9,015	7,893	2,482	14,248	22	10,533,830
Observations	729	876	737	757	751	803	695
Panel B: Female Unreserved							
Winner	-691.545	-121.286	-521.094	-618.357	-106.768	-1.864	779836.562
	(2053.489)	(1176.794)	(1079.683)	(523.358)	(1850.598)	(4.229)	(1732700.316)
Bandwidth	0.14	0.16	0.14	0.14	0.13	0.10	0.12
Control Mean	15,693	8,443	7,352	2,544	13,162	21	9,662,333
Observations	795	865	775	799	748	642	713
Panel C: Minority Reserved							
Winner	-856.047	-306.299	-488.199	-1029.760*	178.182	-0.752	1519870.431
	(2368.455)	(1364.024)	(1262.272)	(559.770)	(2047.663)	(4.435)	(1862060.656)
Bandwidth	0.18	0.18	0.16	0.15	0.17	0.12	0.18
Control Mean	17,234	9,311	8,120	3,070	14,195	22	10,633,036
Observations	667	677	634	617	652	523	669
Panel D: Minority Unreserved							
Winner	-1435.092	195.166	-839.638	-389.930	-901.255	0.511	22771.413
	(1696.464)	(1019.957)	(832.155)	(420.453)	(1478.967)	(3.709)	(1707816.469)
Bandwidth	0.14	0.16	0.16	0.13	0.15	0.13	0.11
Control Mean	15,559	8,370	7,315	2,173	13,400	21	9,766,572
Observations	962	1,040	1,051	914	1,028	901	805

This table shows the effects of having a more (educationally) competent politician on NREGA and village project outcomes. Panel (A)-(D) respectively presents the regression results breakdown by reservation status. Regressions are estimated according to equation 8. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 29: 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 30: Effects of Political Competence on Scheme Knowledge and Solutions

	Knowledge							Solution	
	Avg. (1)	Acct. (2)	AWAS (3)	NALIGALI (4)	PDS (5)	Pension (6)	Solar Light (7)	Total (8)	Avg. (9)
Panel A: Above and Below High-School									
Winner	-0.064	-0.012	0.086	-0.039	0.037	-0.071	-0.587	2.364	2.364
	(0.274)	(0.391)	(0.413)	(0.369)	(0.432)	(0.525)	(0.498)	(1.772)	(1.772)
Bandwidth	0.11	0.13	0.14	0.16	0.14	0.16	0.11	0.11	0.11
Control Mean	2.14	2.68	1.84	1.60	2.78	3.25	1.73	0.54	0.08
Observations	296	334	348	349	141	165	258	35	35
Panel B: Above and Below Graduate									
Winner	0.244	0.403	0.514	-0.078	-0.444	-0.770	0.269	0.993	0.993
	(0.334)	(0.528)	(0.598)	(0.594)	(0.515)	(0.772)	(0.824)	(0.629)	(0.629)
Bandwidth	0.10	0.12	0.12	0.11	0.12	0.11	0.10	0.15	0.15
Control Mean	2.17	2.65	1.93	1.98	3.20	3.60	1.57	0.67	0.10
Observations	165	181	191	149	69	68	137	29	29

This table shows the effects of having a more (educationally) competent politician on scheme knowledge and solutions. Regressions are estimated according to equation 8. Panel (A) shows the regressions result using below and above high-school education cutoff, while Panel (B) uses below and 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 the sample. Robust tandard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.