# Does Affirmative Action Worsen Quality? Theory and Evidence to the Contrary from Elections*

Rikhil R. Bhavnani[†] Alba Huidobro[‡] Soledad Artiz Prillaman[§]

October 23, 2024

**Abstract**

Affirmative action improves the representation of women and minorities, but critics worry that it is at odds with meritocracy. We argue that quotas can improve quality under conditions of discrimination, as quota recipients are held to a higher standard despite facing structural inequalities that make meeting these standards difficult. The net effect of quotas on observable proxies for quality – qualifications – therefore depends on the degrees of selection and structural discrimination. We test our argument by examining the effects of electoral quotas on politicians' education and quality in India. Using two censuses covering more than 40 million residents and 13 states, we show that randomly and quasi-randomly assigned quota politicians have lower average education than non-quota politicians but the same or higher quality. We further provide evidence of both voter and structural discrimination. Our results show that quotas can both enhance the representativeness and quality of politicians.

*We thank Stuart Turnbull-Dugarte, Rajeshwari Majumdar, Tine Paulsen and participants at Leiden University, EPSA, APSA and MPSA for comments on previous versions of the paper, and Diego Tocre and Armelle Grondin for yeoman's work with the data analysis. Human subjects research in this article was reviewed and approved by the UW-Madison IRB (submission ID 2022-1281) and the Stanford IRB (submission IDs 67578 and 68998). Bhavnani acknowledges the support of the University of Wisconsin–Madison Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation. Prillaman acknowledges the support of the Stanford King Center on Development.

[†]Professor and Glenn B. and Cleone Orr Hawkins Chair, Department of Political Science, University of Wisconsin–Madison. Email: bhavnani@wisc.edu.

[‡]Postdoctoral Research Fellow, Chair of Political Behavior, Humboldt-Universität zu Berlin. Email: alba.huidobro@hu-berlin.de.

[§]Assistant Professor, Department of Political Science, Stanford University. Email: soledadp@stanford.edu.

To remedy historical inequalities in political representation, over 100 countries across the world have implemented electoral quotas for historically marginalized groups (Bush, 2011). A substantial literature has demonstrated the positive impacts of these institutions on the representation of targeted minorities and their interests (Chattopadhyay and Duflo, 2004; Bhavnani, 2009; Chauchard, 2017; Clayton and Zetterberg, 2018; Brulé, 2020; Weeks, 2022; Chaudhuri et al., 2024). Yet mass media often fears that this increased representation comes at the cost of politician quality. When identity is mandatorily privileged, merit and competence may be sacrificed. Such anti-meritocratic concerns about affirmative action are not unique to politics and are raised in many settings, including school admissions and corporate board selection.

Evidence of a representation-quality trade-off is limited (Gulzar, 2021), partly due to measurement and data availability issues. Candidate quality is inherently difficult to measure. As a result, the literature (and, as we argue, voters) frequently rely on candidate qualifications to proxy for candidate quality. Existing research suggests that the introduction of quotas either does not affect politician qualifications (Murray, 2010; O'Brien, 2012; Weeks and Baldez, 2015) or improves candidate qualifications overall, and specifically for those from historically over-represented groups (Baltrunaite et al., 2014; Besley et al., 2017). We argue that it is important to theoretically differentiate qualifications and quality (Carreri and Payson, 2021) and that this literature may underestimate the true effect of quotas on politicians' qualifications and, especially, quality. If quotas emerge in contexts with historic identity-based inequalities, then one would need to remove the structural inequalities that generate the quotas in the first place to determine whether quotas, as opposed to the underlying inequalities they seek to address, impact quality. This is particularly important in more unequal societies where historical inequities in access to the resources needed to develop and demonstrate quality remain stark, i.e., where inequality of opportunity breeds inequality of qualifications.

We theorize that voter discrimination leads marginalized candidates to be held to a higher standard than dominant group candidates. The more exacting standards voters place on marginalized candidates, even when protected via quotas, increase rather than reduce the quality (and qualifications) demanded of elected politicians. Quotas are therefore expected to increase politician quality and qualifications under voter discrimination.[1] However, structural discrimination[2] leads to a poorer distribution of qualifications (observable proxies for quality) in the population for

[1] Teele et al. (2018) evaluate whether hypothetical women candidates are held to a higher standard on qualifications than male candidates in the U.S. and find no evidence of a double standard using conjoint experiments.

[2] Throughout this paper, we use the term structural discrimination to refer to the disproportionate under-supply of state-provided qualifications to marginalized groups. Our theory and evidence are identical if this phenomenon is referred to as structural inequality. We choose to term this as discrimination because these historic and state-driven inequalities have fallen along the lines of caste and gender in India, for which substantial discrimination by state actors has been documented (Hanna and Linden, 2012; Mosse, 2018; Purohit, 2022).

marginalized groups relative to dominant groups. Lower qualifications among marginalized groups in the population inevitably reduce the supply of high-qualification quota candidates, which in turn reduces the average qualifications of elected politicians. Under structural discrimination, quotas therefore decrease politician qualifications with unclear impacts on quality.

The net effect of quotas on politician quality and qualifications is therefore jointly determined by the degrees of voter discrimination and structural discrimination. In other words, poor overall levels of qualifications counteract the effects of discrimination at the ballot box. Our theory leads to the hypothesis that quotas can improve politicians' quality under discrimination, even in cases where their qualifications may be the same or lower. It also reveals the roots of the under-representation of marginalized candidates, even absent quotas.

To test our theory, we consider the context of local elections in rural India, where more than three million elected representatives oversee local governance and service delivery across India's villages. The Constitution of India mandates the representation of women and historically marginalized caste groups in local elections through reserved seats, institutionalizing the largest affirmative action policy in the world. These reservation-based quotas have been shown to increase both the descriptive and substantive representation of marginalized groups (Chattopadhyay and Duflo, 2004; Bhavnani, 2009; Dunning and Nilekani, 2013; Gulzar et al., 2020), even as some studies argue that this comes at the cost of candidate qualifications (Bamezai et al., 2024; Ban and Rao, 2008). In fear of precisely this trade-off, several states, including Haryana and Rajasthan, have implemented mandatory education minimums for political candidates, highlighting the policy importance of thoroughly understanding this trade-off.

We challenge these concerns by leveraging a unique opportunity and two census datasets. The unique opportunity to examine the causal effect of quotas is presented by the random and quasi-random assignment of reservations for women and marginalized castes across village governments, and the unique datasets are complete population census data for more than 40 million people across 6,000 village governments. The random or quasi-random assignment of quotas across local governments allows us to recover causal estimates of the effects of quotas. The census data allows us to create credible new measures of candidate *quality*, including both a Mincer residual and a comparison of the educational qualifications of politicians and the underlying population distributions that they are drawn from. We merge census data from 6,000 Gram Panchayats (GPs or village governments) in one state and from 166 GPs across 13 states with data on more than 12,000 local politicians (GP chairpersons or Sarpanch) to comprehensively investigate political selection.

Our analysis confirms our theory and reveals several novel aspects of political selection in the world's largest democracy. Politicians, including women and marginalized caste groups, are positively selected from the population on qualifications, which is to say that they are substantially more educated than the people they represent. Women and minority politicians are not only positively selected vis-a-vis their identity groups but are also positively selected with respect to dominant groups, that is, upper-caste men. That said, female and lower-caste politicians have less education on average than upper-caste male politicians, reflecting the representation-quality concern raised by quota opponents.

Contrary to this concern, we show that while average education levels are lower for marginalized politicians, these politicians are of the same or higher quality, on average, as upper-caste male politicians. Using a latent measure of quality (residuals from a Mincer earnings regression), we show that women and minority politicians have the same average quality as upper-caste male politicians. Further, women and minority politicians' are relatively more positively selected on education than upper caste male politicians, i.e., they are held to a higher standard. Essentially, only the most educated women and lower castes are elected to office. These results are in line with expectations of voter discrimination—a higher *quality* standard is applied to the marginalized.

We provide two additional pieces of evidence for voter discrimination on candidate qualifications. First, we use a candidate choice experiment with just over 1,200 randomly sampled voters to demonstrate that voters hold reserved seat candidates to a higher standard of education than upper-caste candidates, even when controlling for political experience. Voters require higher qualifications in hypothetical choices in quota seat contests in much the same way we observe the requirement of higher relative education in quota seat contests in actual contests. Second, we evaluate heterogeneity in political selection based on demography. If voter discrimination is rooted in relatively more limited knowledge of the quality of marginalized candidates, then we would expect places where the marginalized are in the minority to more strongly hold marginalized candidates to a higher standard. Our results align exactly with these expectations: the higher standard imposed on marginalized candidates exists only where minorities are truly in the minority.

We also provide suggestive evidence for structural discrimination by showing that the lower average education levels of marginalized politicians primarily stem from regions characterized by limited educational supply. Looking at heterogeneity in political selection based on the underlying supply of education in the population, we find that lower-caste men and women politicians in areas with a relatively larger pool of educated minority citizens have the same or higher average education than upper-caste male politicians in areas with the same pool of educated upper-caste men. These findings are in line with the expectation that structural discrimination explains the overall negative difference in education levels between marginalized and dominant group politicians.

Finally, we consider an alternative explanation of our findings: differential costs to contestation. By examining reserved seats, our analysis inherently controls for differential entry costs as only the marginalized can contest. The elections we study are also nonpartisan, eliminating concerns of party discrimination in political selection (Fujiwara et al., 2024; Auerbach and Ziegfeld, 2020). Furthermore, we show that available labor market opportunities do not explain our results and that candidate entry is as common in quota contests.

Our study deepens our understanding of the consequences of quotas—a common institutional form of affirmative action in politics, education, and the workforce. Whereas the first-generation literature on quotas generally showed that they improve representation while they are in place and once they are withdrawn (Chattopadhyay and Duflo, 2004; Bhavnani, 2009), we now examine whether such improvements come at the cost of politician quality. We highlight the fact that since quotas frequently emerge in discriminatory systems, quota recipients usually have less access to the resources that enable the development of formal qualifications, which they disproportionately need as a signal of quality. Last-mile discrimination by voters reverses this pattern, as marginalized group candidates are held to a higher standard than others (see also Desai et al. 2024). This helps to explain evidence demonstrating that quota politicians can leverage these positions to move into higher levels of governance (Karekurve-Ramachandra, 2023; Goyal, 2024a). In the Indian context, our study suggests that an important constraint in improving politician education is not the propensity of voters to discriminate (which we show they do in striking ways) but rather the state, which has yet to rectify the deeply unequal levels of education across historically marginalized groups.

We also further the literature on the challenges to improved representation of women and minorities without quotas. In particular, a large literature on the quota-free American context has examined the reasons for women's severe under-representation and the constraints that women face in contesting and winning elections. This literature demonstrates that female politicians tend to be of higher qualifications and quality than male politicians (Anzia and Berry, 2011; Bauer, 2020) but with mixed evidence on the role of voter discrimination (Schwarz and Coppock, 2022). Most argue that this quality differential is due to the higher costs to women for running for office (Fox and Lawless, 2004) while others suggest voter discrimination cannot be ruled out (Ashworth et al., 2024). Recent work documents that women are held to higher standards (Bauer, 2020) and placed in double binds (Teele et al., 2018). The Indian context that we study allows us to control for differential running costs, thereby allowing us to document voter discrimination more explicitly. Our findings also demonstrate that these higher standards and double binds exist even in the presence of institutional protections.

Finally, we contribute to the growing literature on political selection (Dal Bó and Finan, 2018) by theorizing how patterns of political selection can vary across space. This builds on past work largely focused on how qualifications shape political selection (Murray, 2010; O'Brien, 2012; Weeks and Baldez, 2015; Baltrunaite et al., 2014) and suggests that these analyses may underestimate the impact of quotas on political quality (as an exception, see Besley et al. 2017). By focusing on a context with deep histories of inequality, we expose how these histories constrain political representation despite giving rise to institutions meant to ensure inclusive representation.

# 1 A Theory of Politician Qualifications and Quality

## 1.1 Qualifications, Quality, and Performance

We first define three dimensions of political selection: qualifications, quality, and performance. We define politician quality as the hard-to-observe capability of an individual to represent the interests of their constituents in political office. Quality encapsulates both the innate talents and aptitudes of an individual and their learned skills. This definition does not impose a common standard over quality. In some instances, the successful representation of citizen's interests would involve efficient and high-quality public service delivery. In others, it will be the delivery of clientelistic patronage. Quality, by this definition, is simply the capability of a politician to do what the people who voted for them want. Given this, we assume that voters seek to maximize politician quality.

Most studies of politician quality actually evaluate politician qualifications or performance (Dal Bó et al. 2017 and Besley et al. 2017 are exceptions). We define politician qualifications as the easily observable attributes that imply that an individual is high quality. Qualifications, by this definition, are simply the politician characteristics expected by voters to strongly correlate with quality. Qualifications, however, often intersect with socioeconomic status (and therefore identity), and so are imperfect predictors of quality.

Additionally, politician quality is often inferred from politician performance (Anzia and Berry, 2011; Das et al., 2023). We define politicians' performance as the observed outcomes and accomplishments of political processes with regard to constituent interests. Whereas quality is the capability of a politician to execute constituent interests, performance is the actual execution of those interests. While performance is a function of politicians' quality, it is also a function of the many complex political processes that governance entails. Bureaucracies, institutions, and other opportunity structures constrain a politician's *ability* to accomplish their goals (despite whatever their *capability*) (Purohit, 2022). If such opportunity structures themselves entail discrimination against marginalized groups, then the impacts of affirmative action on performance may substantially differ from the impact of affirmative action on quality. This paper focuses on the effects of quotas on politician qualifications and quality and leaves other work to consider the complex relationship between quotas and performance.

## 1.2 Political Selection Under Affirmative Action

To theorize how affirmative action affects politician quality, imagine candidates from two groups—dominant and marginalized—indexed by $D$ and $M$. In the Indian context, we think of dominant groups as men and upper castes and the marginalized as women and lower castes.

We have assumed that voters wish to elect high-quality candidates, but candidate quality is imperfectly observable. When information about candidate quality is unknown or uncertain, voters use politicians' observable qualifications ($E$) to imperfectly proxy for their quality ($Q$). We further assume that the average voter has less information about marginalized candidates than dominant candidates, which aligns with existing work on the network centrality of dominant groups (Cruz, 2019; Prillaman, 2023).That said, because of the density of ties within groups, marginalized group voters will have better information about the quality of marginalized group candidates than dominant group voters (Larson and Lewis, 2017).

We further assume that qualifications positively affect quality for all people, but this relationship may be stronger for the marginalized. Qualifications affect the skill-based component of quality by providing formal skills (e.g., technical knowledge) and informal resources (e.g., networks). Given historical inequities, marginalized people often have fewer alternatives to develop formal skills and informal resources outside of formal qualifications. As a result, qualifications can have a larger impact on the quality of marginalized candidates than dominant candidates. Qualifications may, therefore, serve as a stronger signal of quality for marginalized groups. We assume that the distribution of innate quality is the same for marginalized and dominant people.

Electoral quotas constrain political representation by increasing the presence of marginalized candidates and/or representatives. We consider the effects of reserved seats – where only marginalized candidates can run for office – on politician quality. Since electoral quotas are introduced to correct historical wrongs, we argue that quotas are usually introduced in the context of two forms of discrimination: selection discrimination by voters[3] and structural discrimination by the state. We outline our theoretical expectations over $E$ and $Q$ under voter and structural discrimination in Table 1 and describe each set of expectations below.

Voter discrimination pertains to the *demand for candidates* based on identity and is the underselection of marginalized candidates relative to dominant candidates given identical quality ($V_D|Q > V_M|Q$) (Ashworth et al., 2024). Voters may discriminate against marginalized candidates for at least two reasons (Teele et al., 2018). First, they might genuinely prefer discriminating against the marginalized (taste-based discrimination). Second, voters may hold and act on negative

[3] We follow existing work and assume a context without parties. This assumption is valid in rural India, as electoral contests are nonpartisan.

stereotypes about the quality of marginalized candidates (statistical discrimination—see Beaman et al. 2009; Chauchard 2017). Less and more uncertain information about marginalized candidates' quality heightens statistical discrimination (Anzia and Bernhard, 2022), where the marginalized are held to higher standards to compensate for uncertainty and assumed lower quality (Teele et al., 2018).

Structural discrimination pertains to the *supply of candidates* and reflects historic inequities in the state provision of resources that enable the acquisition of qualifications. Structural discrimination causes the mean qualifications of the marginalized to be lower than that of the dominant ($\overline{E}_D > \overline{E}_M$). Put simply, structural discrimination leads to inequality of opportunity. Quotas are often introduced to remedy the very inequalities in the supply of public goods that inhibit the acquisition of qualifications.

With only marginalized candidates, the information environment in reserved-seat contests is worse than in open-seat contests. Voters in these elections are, therefore, more likely to rely on stereotypes and the individuating information that qualifications provide. The constraints on candidate selection imply that the difference in likelihood that a voter votes for a high-qualifications candidate and a low-qualifications candidate in a reserved seat race ($R$) is larger than the same difference for candidates in open seat races ($O$) ($V_{R_{EHigh}} - V_{R_{ELow}} > V_{O_{EHigh}} - V_{O_{ELow}}$). Essentially, candidates in quota seats will be held to a higher qualifications standard than candidates in open seats.

This higher qualifications standard could result from voter discrimination: voters demand more qualifications from reserved seat candidates to compensate for their uncertainty over the quality of these candidates or because they have a distaste for these candidates. Internalized beliefs that marginalized candidates are of lower innate quality lead voters to demand stronger signals of quality (i.e., higher qualifications) of these candidates to compensate. If this positive qualifications penalty results from voters' statistical discrimination, reducing uncertainty over candidate quality, such as through greater information, will reduce the qualifications gap between reserved and open seats. Voter discrimination implies that reserved-seat politicians will be of higher quality than open-seat politicians (the top right cell of Table 1).[4]

The presence of structural discrimination complicates how quotas are expected to impact politician qualifications and quality. With structural discrimination, where the average qualifications in the

[4] It is also possible that, when looking at any one qualification, a higher qualifications standard could result from that qualification more strongly signaling the quality of marginalized candidates. In such a case, the higher qualification standard aligns with voters' preferences for high-quality candidates and the use of that qualification as an appropriate signal. In such instances, representatives in reserved seats are expected to have higher qualifications but no difference in quality relative to representatives in open seats (the top left cell of Table 1) That said, when multiple qualifications are considered, differential signaling across qualifications would suggest that marginalized candidates are held to a higher standard on one qualification, whereas dominant candidates are held to a higher standard on another, averaging to no difference in aggregate qualifications.

population are lower for the marginalized, the same average quality will likely open lower average qualifications for marginalized candidates. Structural discrimination is, thus, expected to lower the average qualifications of the marginalized. The effect of structural discrimination on candidate quality depends on the strength of the relationship between qualifications and quality. If this relationship is weak or limited, then we would not expect the average quality in reserved seats to be below that of open seats, even in the presence of a qualifications differential (the bottom left cell of Table 1).

With both structural and voter discrimination, expectations become less clear. In such cases, the impact of reservations on politician qualifications is conditional on the levels of discrimination. When structural discrimination severely limits the supply of qualified marginalized candidates, the impact of affirmative action on qualifications may be negative (i.e., the average qualifications of marginalized candidates are less than that of dominant candidates). When structural discrimination is less severe, we expect that marginalized candidates' qualifications are equal to or greater than those of privileged group candidates under voter discrimination. In both cases, marginalized candidates' quality is expected, because of voter discrimination, to be greater than or equal to that of dominant candidates (the bottom right cell of Table 1). The combination of structural and voter discrimination enables the reconciliation of the qualifications-representation trade-off that concerns critics of affirmative action and the presence of higher-quality marginalized representatives.[5]

**Table 1.** Expectations over Politician Qualifications and Quality under Discrimination

| | | **Voter Discrimination** | |
|---|---|---|---|
| | | No | Yes |
| **Structural Discrimination** | No | $\bar{E}_R \geq \bar{E}_O$<br>$\bar{Q}_R = \bar{Q}_O$ | $\bar{E}_R > \bar{E}_O$<br>$\bar{Q}_R > \bar{Q}_O$ |
| | Yes | $\bar{E}_R < \bar{E}_O$<br>$\bar{Q}_R \leq \bar{Q}_O$ | $\bar{E}_R >< \bar{E}_O$<br>$\bar{Q}_R \geq \bar{Q}_O$ |

This yields a set of expectations over comparative statics based on the levels of discrimination, as outlined in Table 1. First, the difference between average politician quality in reserved and open seats is expected to be decreasing in the level of voter discrimination. Second, the average level of reserved seat politicians' qualifications is expected to be decreasing in the level of structural discrimination. In contexts with no voter discrimination but with structural discrimination, we expect the impact of reservations on politicians' qualifications to be negative and the impact on quality to be null or negative. In contexts with no structural discrimination but with voter

[5] Note that our theory predicts an over-representation of dominant representatives in open seats. The presence of voter discrimination and structural discrimination, all else equal, will cause the underselection of marginalized candidates and unequal and disproportionate representation as is found in non-quota contexts (Anzia and Berry, 2011).

discrimination, we expect the impact of reservations on politicians' qualifications and quality to be positive.[6]

## 2 Background, Data, and Empirical Strategy

### 2.1 Local Elections in India

We explore political selection in the context of local elections in rural India. While India has had local governments since the 1950s, it was only in 1992, with the passage of the 73rd and 74th constitutional amendments, that Indian states were mandated to conduct elections at the local level. Local governments hold substantial power in the decision-making over and execution of local service delivery and are comprised of a chairperson (known as the Sarpanch) and a council. The 1992 constitutional amendments additionally institutionalized a system of quotas to protect the representation of historically marginalized groups. Seats are reserved for Scheduled Castes (SCs) and Scheduled Tribes (STs) in proportion to their population share, and then at least one-third of seats are reserved for women. Since the reservation status of seats changes across elections, there is high electoral turnover.

We choose this context for three reasons. First, the reservation system in India's local government is one of, if not the, largest quota systems in the world. More than three million politicians and roughly 250,000 Gram Panchayat chairpersons hold office at any time, and more than half of these positions are reserved for women, SCs, or STs. This provides ample opportunity to explore the effects of reservations and to understand how these effects vary with other contextual factors. Second, the allocation of reservations is done in random or predictable ways that enable the evaluation of the impact of reservations. For precisely this reason, these elections have been leveraged to build a wide base of knowledge around the impact of reservations on other political outcomes (Chattopadhyay and Duflo, 2004; Bhavnani, 2009; Beaman et al., 2009; Chauchard, 2017; Bhavnani, 2017; Karekurve-Ramachandra and Lee, 2020, 2024; Goyal, 2024b). Third, local elections in India provide an opportunity to remove many of the additional layers of political selection that complicate the estimation of voter discrimination in the selection process. Local elections in India are, in most states, nonpartisan, making elections close to a citizen-candidate model. These elections are also the lowest level in a much more complex pipeline of political offices and so few candidates have previously held other office. The rotational nature of reservations implies

[6] What does this pattern mean for politician performance? Since marginalized politicians elected in seats with quotas are, under voter discrimination, expected to be of higher quality than the largely dominant politicians elected without quotas, we would expect their performance to also be higher, all else equal. However, rarely is all else equal: we expect that marginalized politicians face higher hurdles in implementation than dominant politicians, not least because of potential discrimination. Put another way, additional forms of discrimination constrain the performance of marginalized politicians.

that incumbency is relatively rare, and each electoral contest is essentially an opportunity to select from a new set of candidates.As a result, qualifications and local networks are the primary sources of information about candidate quality. Since these positions typically represent approximately 5,000 people, local information about candidates is possible, though not guaranteed.

## 2.2 Data

We analyze two unique and comprehensive data sources to describe the nature of political selection in rural Indian villages. Both data sources include an individual-level census of the population paired with data on politicians.

### 2.2.1 Odisha Census

We focus on one state in northern India—Odisha—with a diverse population and utilize population data from the 2011-12 Socioeconomic and Caste Census (SECC), which was a census conducted across India to determine eligibility for government programs.[7] These data include measures of education, age, occupation, caste category, and household characteristics for the roughly 40 million people residing in the state in 2011.

We pair these census data with electoral data scraped from the state election commission's website on all elected Gram Panchayat (GP or village government) chairpersons in the two most recent local elections (2017 and 2022).[8] Electoral data included only the names of the more than 6,500 elected GP chairpersons (discussed in Appendix B). We undertook a fuzzy matching process to merge the politician data with the SECC data. We limited the sample of potential matches based on the eligibility criteria for contesting office – being over the age of 21 and having no more than two children. We conducted a name-based fuzzy merge using algorithms designed for transliterated Indian names, validating matches for accuracy. We discarded all matches where more than one resident of the GP shared the name of the politician, as we had no mechanism to adjudicate the match.[9] In total, we matched 48% of politicians. This enables us to study political selection dynamics in local governments with more than 11 million adults. Appendix Tables B1 and B2 show that the sample of GPs with a matched politician is not substantively different from those without.

The electoral data only provide information on politicians' identity, not on the reservation status of the electoral contest. We worked with the state election commission to acquire reservation information for the 2022 local elections, but data from the 2017 election were either destroyed or

[7] These data have been studied by others, such as in Asher et al. (2018); Bamezai et al. (2024).
[8] Data on candidates is not available for either election, even after detailed conversations with the state election commission.
[9] 25% of unmatched politicians were because of duplicate names.

unable to be shared. The 2022 reservation data were merged with the census data using GP names, with a near-perfect match rate.

#### 2.2.2 All India Census

To validate whether the results from Odisha travel, we leverage data from the Rural Economic and Demographic Survey.[10] Specifically, we analyze data from the sample frame listing undertaken as part of the larger data collection effort between 2014 and 2016 that enumerated all residents in one village of 166 GPs across 13 states, a voting-age population of nearly 235,000.[11] This listing accompanied a survey of the elected chairpersons of these GPs, which included information on chairperson education and seat reservation status. We append these data sets to compare politicians with the constituents they represent.

### 2.3 Measurement

#### 2.3.1 Measuring Politician Qualifications

In our context, we presume that a key qualification is education. Education has been shown to improve politicians' ability to navigate complex governance structures and to deliver public goods (Besley et al., 2011; Carnes and Lupu, 2016; Jain et al., 2023; Lahoti and Sahoo, 2020). Education is also one of the most visible signals to voters—candidates have to declare their educational qualifications in publicly available affidavits—especially in the context of local elections with low incumbency. We validate the assumption that rural Indian voters prize education in their political selection with the aid of a survey experiment which shows that voters place more emphasis on education than other attributes when selecting candidates (see Appendix **??**). In addition, we show that politician education is positively correlated with performance (see Appendix Table E1).[12]

Both census surveys provide information on education for politicians and the entire population. Education is measured as a categorical variable with the statewide census including seven education categories (illiterate, literate, primary, middle, secondary, higher secondary, and graduate or higher) and the all-India census including four categories (illiterate, primary, secondary, and higher secondary). For ease of interpretation, we convert these data to years of education.

[10] These data were collected by the National Council of Applied Economic Research in Delhi and provided by Andrew Foster.

[11] In large villages, the SEPRI exercise enumerated only a sizable number of households, though we do not know the precise share.

[12] The expected positive correlation between education and performance suggests that our results on the role of qualifications in political selection can also be interpreted as partial evidence on the role of quality.

### 2.3.2 Measuring Politician Quality

Politician quality is notoriously difficult to measure. Quality, as we have argued, is comprised of both an individual's innate aptitudes and the skills they have developed throughout their lifetime. In addition to looking at education, we directly measure quality in two ways: with a Mincer earnings regression model and by capturing relative educational attainment.

First, we measure quality using the residuals from a Mincer earnings regression model and following the approach of Besley et al. (2017). Mechanically, we regress an individual's household wealth (measured using an asset index[13] constructed using factor analysis) on a dense set of individual socioeconomic characteristics (age, education, occupation, household size, household number of workers, household number of children, and the fully saturated interactions of all of these) and village fixed effects separately for each identity group. We extract the residuals from this model (the difference between a person's predicted household wealth and their actual household wealth based on their observed individual characteristics) and use this residual as a measure of quality. Intuitively, this model estimates how all observable characteristics predict household wealth and then, for each individual, measures whether their household wealth level is above or below what their individual characteristics would predict. Highly competent people should have household wealth levels above what their circumstances would suggest, as they "make more" of the opportunities they have been given.[14] Therefore, a large and positive residual is used to indicate a high-quality individual.

Second, we measure quality by calculating the relative position of politicians in the education distribution of all individuals in their identity group in their GP (Dal Bó et al., 2017). Whereas most studies of political selection can only compare average characteristics across politicians, a key advantage to having population data is that we can estimate the degree of positive selection of politicians vis-a-vis their constituents. While average differences in qualifications may reflect structural discrimination (opportunity), the relative position of politicians is more likely to reflect their quality. Intuitively, this is akin to comparing where politicians fall in percentile terms in their groups' population distribution, with politicians who sit more in the tails of their groups' qualifications distribution (regardless of the actual level of their qualifications) seen as being of higher quality than politicians who sit more in the center of their groups' qualifications distribution.

[13] The Odisha census data only include wealth and income measured at the household level. We are, therefore, unable to perform an individual-level Mincer regression, as would be ideal. We use wealth, as opposed to household income, because it allows for more variation across households since household income includes only three categories. We are only able to create this measure for the Odisha sample because the All India survey data do not include measures of earnings or wealth.

[14] An alternative explanation, perhaps particularly for women, is that they select into higher "quality" households than their characteristics would suggest. We believe this is still a measure of quality as these observable characteristics are central in the Indian marriage market ((Afridi et al., 2023)), but the possibility remains that this could upward bias our estimates of female "quality." That said, this concern does not affect our analysis of within-gender selection patterns.

### 2.4 Empirical Strategy

Our empirical aim is to evaluate the impact of quotas on politician quality. A key concern is that quotas may be assigned to places that value certain forms of representation more or select politicians differently. We resolve this endogeneity concern by studying the random or quasi-random assignment of reservations to GPs in rural India. First, the allocation of caste reservations is based on the caste composition of the population. In Odisha, state rules mandate that GPs are ranked by their Scheduled Caste (SC) and Scheduled Tribe (ST) population shares within a block. The GPs with the highest population shares of each are reserved for that group, rotating every two elections in descending order of population shares.[15] To account for this assignment mechanism, we control for block fixed effects,[16] the SC and ST population shares according to the 2011 decennial census (the data used to allocate reservations in our elections), and the squared SC and ST population shares to account for possible nonlinearities in assignment. Our main identifying assumption is that caste reservation assignments across GPs with the same blocks and SC/ST population and squared population shares are as-if random. Our results are robust to excluding the squared levels of SC and ST population shares and to controlling for the interaction between block fixed effects and SC and ST population shares (see Appendix D). Our results are also robust to considering only non-Scheduled Areas.

Second, reservations for women are randomly assigned. According to state rules in Odisha, GPs are ordered alphabetically by their Odia names, and every other GP is reserved for women (separately for caste-based reservation), with rotations every election cycle, allowing for a causal interpretation of their effects.

Our theory suggests that marginalized politicians, even when elected under quotas, will be held to different qualifications and quality standards than dominant politicians. The above identification strategies allow us to estimate the impact of reservations on quality but do not allow us to estimate the impact of politician identity – their caste and gender – unless identity is perfectly responsive to reservations (i.e., the marginalized only run in reserved seats). In the All India data, we observe both whether an elected position was reserved and the identity of the politician. In the Odisha data, we only observe GP reservation status in 2022 and observe politicians' identity in both the 2017 and 2022 elections. The interpretation of results from these different data sources (reservation or identity) in the analysis of the Odisha data represents a trade-off between potential selection bias

[15] The language in the Odisha Gram Panchayat Act is as follows: "The Grama Panchayats in relation to Gramas in which the density of Population of the Scheduled Castes and the Schedule Tribes is higher in the Block shall be reserved by the [District] Collector for the Scheduled Castes and the Scheduled Tribes respectively and shall rotate in the descending order at every two terms of General Election."

[16] We can only include state fixed effects in the analysis of the All India data given the limited sample of GPs in each state.

(as the assignment mechanism of reservations is known, but the assignment mechanism of identity is not) and potential measurement bias (as attribute data was collected five to ten years before the elections).

We consider the potential selection bias from estimating the impact of politician identity instead of GP reservation status by comparing the alignment of the two in the 2022 electoral data: upper-caste men were elected in 83% of open seats, upper-caste women were elected in 89% of open women seats, and SC/ST men were elected in 82% of open SC/ST seats. Simply, reservation status aligned to a very high extent with expected politician identity.[17] The strong correlation between identity and reservation status minimizes concerns of selection bias, and past work has also used identity as a proxy for reservation status (Chattopadhyay and Duflo, 2004). We further confirm that there is no systematic imbalance in the observable characteristics of GPs across reservations and politician identities in Appendix C, except in one dimension unrelated to the assignment of reservations: total population. We control for population size in our analyses to account for this difference.

Data from more recent elections is less prone to measurement bias due to the time difference between when data on representative qualifications was captured (2011/12) and when the representative contested for office (2017 or 2022). Education data for young representatives is likely to be downward-biased the longer the time between the collection of education and electoral data, as these representatives were likely still pursuing their education at the point of data collection. We presume that our education data should be complete for anyone over the age of 22 at the time of census data collection, as this is the age by which one would have attained the highest level of education in our data. Our data show that 9% of representatives elected in 2017 and 16% of representatives elected in 2022 were under 22 in 2012. To minimize measurement bias, we exclude politicians from all analyses who were under the age of 22 in 2012 but demonstrate the robustness of our results to their inclusion in Appendix D.5.

Given these trade-offs, we report estimates using all three sources of data in Odisha: politician identity in 2017 and 2022 and reservation status in 2022. We also report estimates using both politician identity and reservation status in the All India data.

We evaluate our hypotheses using two empirical approaches. First, we compare politicians across GPs to understand the average relationship between politician identity/GP reservation status and

[17] Expected identity is the most dominant identity group that abides by reservation requirements.

politician qualifications and quality. Specifically, we estimate the following model:

$$\begin{aligned}\text{Qualification/Quality}_{p,gp,b} =& \beta_1 \text{Identity/Reservation}_{p,gp,b} + \\ & \beta_2 \text{SC Pop. Share}_{gp,b} + \beta_3 \text{SC Pop. Share}^2_{gp,b} + \\ & \beta_4 \text{ST Pop. Share}_{gp,b} + \beta_5 \text{ST Pop. Share}^2_{gp,b} + \\ & \beta_4 \text{X}_{gp,b} + \gamma_b + \varepsilon_{gp}\end{aligned} \tag{1}$$

where $\text{Qualification/Quality}_{p,gp,d}$ is the education level or the Mincer residual for politician $p$ in GP $gp$. $b$ references blocks in the Odisha data and states in the All India data. $\text{Identity/Reservation}_{i,gp,b}$ is a vector of either the identity of the politician (upper-caste men, upper-caste women, SC/ST men, and SC/ST women)[18] or the reservation status of the GP (open, open women, open SC/ST, SC/ST women) depending on the data source used. $\text{SC Pop. Share}_{gp,b}$ and $\text{ST Pop. Share}_{gp,b}$ are the shares of the population in GP $gp$ that are SC and ST, respectively. $\gamma_b$ are block/state fixed effects to account for the level at which caste reservations are allocated in Odisha (block) or reservation rules are decided across India (state). $\text{X}_{gp,b}$ is a vector of GP-level occupational shares (share of the population in agricultural work, manual work, service work, business, or other employment, committing those out of the labor force) for the identity group of the politician to capture outside labor market opportunities and the total GP population. Standard errors are clustered at the GP level, the unit of reservations.

Second, we compare politicians with their constituents to understand how political selection varies by politician identity/reservation status. Specifically, we estimate the following model:

$$\begin{aligned}\text{Representative}_{i,gp|g} =& \beta_1 \text{Qualification}_{i,gp|g} + \beta_2 \text{Qualification X Identity/Reservation}_{i,gp|g} + \\ & \gamma_{gp} + \varepsilon_{gp}\end{aligned} \tag{2}$$

where $\text{Representative}_{i,gp|g}$ is an indicator equal to 10,000 for person $i$ from group $g$ in GP $gp$ who is the elected chairperson. $\text{Qualification}_{i,gp|g}$ is the education level of individual $i$ from group $g$ in GP $gp$. $\text{Identity/Reservation}_{i,gp|g}$ is a vector of either the identity of the politician or the reservation status of the GP depending on the data source used. We subset the population to only individuals $i$ who are in the same identity group as the politician (in the identity models) or who were eligible for the reservation status (in the reservation models) $g$ to compare politicians to the group they would compete with. $\gamma_{gp}$ are GP fixed effects to compare politicians only to the populations they represent. Standard errors are clustered at the GP level, the unit of reservations.

Our preferred specification in these analyses leverages the set of villages in Odisha for which we have panel data on the identity of the politicians from 2017 and 2022. In these models, the

[18] We combined SC and ST into a single category for ease of interpretation. Our results are robust to their separation.

inclusion of GP fixed effects allows for a *within GP* comparison of patterns of political selection and, therefore, accounts for all GP-specific characteristics that might impact political selection. Put simply, we can observe how, on average, the same GPs make different selection decisions based on the identity of the person elected. Appendix Table D2 confirms substantial rotation in this set of GPs. We report the results from this panel analysis in addition to the analysis for each year separately to demonstrate the robustness of our findings.

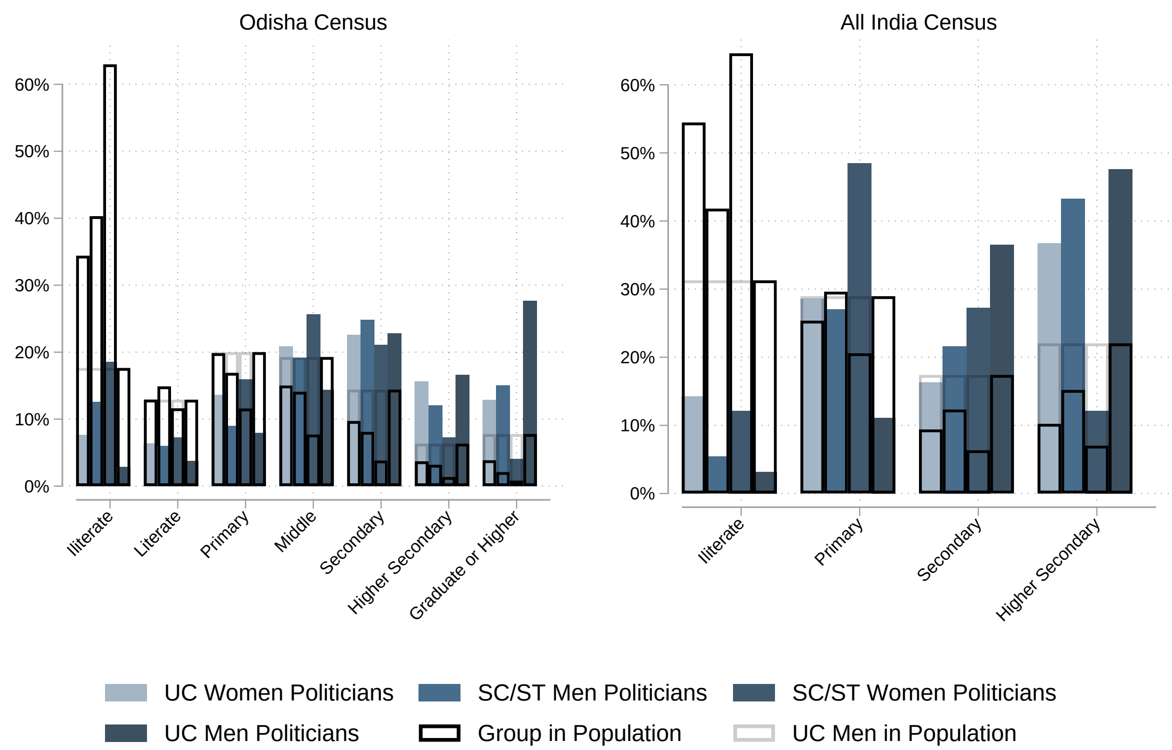
## 3 Political Selection under Electoral Quotas

We begin by examining the aggregate distributions of education in the population and among politicians in Figure 1. The transparent black bars denote the education distribution in the population for each identity group (the light gray bars are the population education distribution for upper-caste men), and the blue bars denote the education distribution of elected politicians by identity group. This figure reveals several facts in both the Odisha and the All India samples. First, the distribution of education among upper-caste male politicians is more right-skewed than that of women and SC/ST politicians. This highlights the concern that many quota opponents raise about upper-caste men having significantly higher education than their quota-elected counterparts. However, what is additionally evident is that politicians of all identity groups are positively selected in terms of education relative to their identity group *and* with respect to upper-caste men. Last, a comparison of the population education distributions across identity groups reveals substantial inequalities: women and minorities have substantially lower levels of education in the population than upper-caste men, suggesting substantial structural discrimination.

To test whether quotas affect politician qualifications and quality, we estimate Specification (1) for years of education and the Mincer residual. Table 2 reports the estimated marginal effect of each identity group/reservation status on politician education and Mincer residual, with upper-caste men/open seats serving as the reference group. The group measure denotes whether the core independent variables are the identity of the politician or the reservation status of the GP.

Simply comparing across politicians, the first six columns of Table 2 show that the average education of women and SC/ST politicians is significantly below that of men. These differences are most acute for women, with upper-caste women politicians/politicians in women-reserved seats having, on average, between 1 and 3 fewer years of education than upper-caste men politicians/politicians in open seats across specifications. This difference is around one year of education for SC/ST men politicians/reserved seats and around four years of education for SC/ST women politicians/reserved seats, highlighting the particular challenge for minority women. The results are similar in size but noisier when looking at the data from across India, which is unsurprising given the smaller sample size and limited coverage of each identity group.

**Figure 1.** Women and Minorities are Positively Selected on Education Relative to their Own Group and Upper Caste Men




*Note*: This figure plots the education distribution for politicians and citizens by their caste and gender. The blue bars depict the distribution of education for each politician identity group, the transparent black bars depict the population distribution for the corresponding identity group, and the transparent grey bars depict the population distribution for upper caste men. Illiterate corresponds to 0 years of education, literate to less than five years of education, primary to having completed at least five years of education, middle to having completed at least eight years of education, secondary to having completed at least ten years of education, higher secondary to having completed at least 12 years of education, and graduate or higher to having completed at least a bachelors degree or more. We exclude politicians who were not 2022 at the time of data collection to ensure accurate education data.

Despite these differences in average qualifications, columns 7-10 of Table 2 show no average difference in our direct measure of quality—the Mincer residual—across groups.[19] In fact, women and SC/ST politicians have somewhat higher Mincer residuals than upper-caste men (as shown by the positive direction of the coefficients), though not significantly so. Contrary to education, which is constrained by the structural discrimination documented in Figure 1, politicians in reserved seats are of no less and occasionally higher quality.

**Table 2.** Women and Minority Politicians Have Lower Education but the Same Innate Quality as Upper Caste Men on Average

| **Dependent Variable:** | Years of Education | | | | | | Mincer Residual | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Sample:** | Odisha Census | | | | All India Census | | Odisha Census | | | |
| **Group Measure:** | Identity Panel | Identity 2017 | Identity 2022 | Reservation 2022 | Identity | Reservation | Identity Panel | Identity 2017 | Identity 2022 | Reservation 2022 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| UC Women/Open Women | -2.43*** | -2.72*** | -3.92*** | -1.08*** | -1.81 | -1.55* | 0.04* | -0.01 | 0.01 | -0.01 |
| | (0.34) | (0.72) | (0.82) | (0.39) | (1.21) | (0.92) | (0.02) | (0.06) | (0.06) | (0.03) |
| SC/ST Men/Open SC/ST | -1.60*** | -1.21*** | -1.02*** | -1.01*** | -1.97 | -2.28** | -0.01 | 0.00 | 0.04* | 0.03 |
| | (0.47) | (0.32) | (0.31) | (0.31) | (1.22) | (1.07) | (0.03) | (0.02) | (0.02) | (0.02) |
| SC/ST Women/SC/ST Women | -3.73*** | -4.13*** | -5.27*** | -3.30*** | -3.53*** | -2.86* | 0.05 | -0.01 | 0.04 | 0.02 |
| | (0.44) | (0.66) | (0.78) | (0.43) | (1.19) | (1.64) | (0.03) | (0.05) | (0.05) | (0.03) |
| E(Y \| UC Men/Open) | 10.87 | 10.61 | 10.72 | 10.71 | 11.02 | 10.48 | 0.07 | 0.08 | 0.07 | 0.08 |
| # Observations | 2,522 | 2,932 | 2,583 | 2,498 | 126 | 157 | 2,522 | 2,932 | 2,583 | 2,498 |
| # GPs | 1,261 | 2,932 | 2,583 | 2,498 | 126 | 157 | 1,261 | 2,932 | 2,583 | 2,498 |
| Fixed Effects | GP | Block | Block | Block | State | State | GP | Block | Block | Block |
| Population Controls | | X | X | X | X | X | | X | X | X |
| Occupation Controls | | X | X | X | | | | X | X | X |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are reported in parentheses. The table reports the relationship between politician identity/GP reservation status and education following specification 1. The sample for each column includes only politicians. Models where the group measure is identity report the relationship between politician identity and education. Models where the group measure is reservation report the relationship between GP reservation status and politician education. For the models using the Odisha Census data, the sample includes the set of GPs where politicians were able to be matched (e.g., 2,932 politicians were matched to census data in 2017; 2,583 politicians were matched to census data in 2022; and 2,498 politicians were matched to GPs with reservation data in 2022). The identity panel models in columns (4) and (10) include the 1,261 GPs where we match politicians in both 2017 and 2022, allowing for a within-GP comparison. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects and controls included as specified. Population controls include SC and ST population shares and their squared values and the total population size. Occupation controls include the share of the population from the politicians' identity group in agriculture, business, service, manual work, and others. The full table with control coefficients is in Appendix Table D1. We exclude politicians who were not 2022 at the time of data collection to ensure accurate education data.

To more definitively estimate the impact of quotas on politician quality, Table 3 reports the estimates from Specification (2), showing politicians' relative degrees of positive selection on education vis-a-vis the population in their constituencies from their identity group. Specifically, Table 3 reports both the coefficient on education (the estimated degree of positive selection for upper-caste men/open seats) and the coefficients on the interactions between identity group/reservation status and education levels (the additional degree of positive/negative selection for each identity group/reservation status relative to upper-caste men/open seats). The dependent variable is an indicator for being a politician (equal to 10,000 for politicians to ease interpretation). The expected value of this in open seats shown at the bottom of the table reveals that, in Odisha, the probability of

[19] We cannot estimate the Mincer residual in the All India sample as we do not have measures of their earnings or assets.

being elected as a politician in an open seat is 1 out of 10,000, and the probability that any specific upper caste man is elected is roughly 1.5 out of 10,000.

First, Table 3 shows that upper-caste men politicians have significantly more education than upper-caste men in the population, i.e., they are positively selected (shown by the positive coefficient on years of education). For each additional year of education, an upper-caste man is roughly two times more likely to be elected as a politician. Second, all columns reveal that women and SC/ST politicians are generally held to a higher standard than upper-caste men politicians, and significantly and consistently more so for SC/ST men and women. Put another way, SC/ST politicians are more positively selected vis-a-vis their groups' education distribution than upper-caste male politicians. Upper-caste women, however, are only somewhat more positively selected than upper-caste men. SC/ST women are the most positively selected identity group, with an additional year of education almost tripling their likelihood of being elected. Importantly, these results hold when considering the identity panel in Odisha (column 1), which shows that marginalized politicians are more positively selected than upper caste male politicians *from the same GP*, ensuring that it is not underlying GP characteristics that explain these patterns of political selection. Additionally, these results are comparable in significance and size for both the Odisha sample and the All India sample, suggesting the generalizability of this finding.

The Odisha census data further allows us to consider positive selection within households. An intra-household comparison allows us to better assess whether the observed positive selection is truly an indicator of quality: Given homogamy, households often have similar opportunities and resources; if voters discern and differentiate among household members, this would suggest that they are selecting on quality as opposed to opportunity or eliteness. Table 4 reports the estimated degree of positive selection on education for politicians vis-a-vis the other members of their households from their identity group (i.e., their eligible household members). Since households are almost always of the same caste category, this amounts to within-gender comparisons.

Table 4 shows that all politicians are significantly more educated than the other members of their household; voters select on education even when considering potential candidates from the same household. SC/ST men and women politicians are even more positively selected vis-a-vis the eligible members of their household, further suggesting that marginalized candidates are held to even higher standards than dominant candidates.

**Table 3.** Women and Minority Politicians are Held to a Higher Standard than Upper Caste Men Politicians

| **Dependent Variable:** | Sarpanch (× 10,000) | | | | | |
|---|---|---|---|---|---|---|
| **Sample:** | Odisha Census | | | | All India Census | |
| **Group Measure:** | Identity Panel | Identity 2017 | Identity 2022 | Reservation 2022 | Identity | Reservation |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Education | 1.75*** | 1.94*** | 2.24*** | 1.29*** | 3.74*** | 1.83*** |
| | (0.11) | (0.07) | (0.08) | (0.04) | (0.60) | (0.23) |
| Education × UC Women/Open Women | 0.59*** | 0.31** | 0.21 | 0.68*** | 2.15 | 1.56** |
| | (0.20) | (0.12) | (0.13) | (0.09) | (1.33) | (0.67) |
| Education × SC/ST Men/Open SC/ST | 2.63*** | 2.14*** | 1.93*** | 1.61*** | 5.70** | 2.79** |
| | (0.29) | (0.18) | (0.18) | (0.12) | (2.26) | (1.12) |
| Education × SC/ST Women/SC/ST Women | 4.23*** | 4.31*** | 2.85*** | 2.83*** | 10.66** | 23.59*** |
| | (0.51) | (0.28) | (0.29) | (0.25) | (5.33) | (6.67) |
| E(Y \| UC Men/Open) | 1.40 | 1.60 | 1.20 | 0.83 | 2.64 | 0.93 |
| # Observations | 1057195 | 2907032 | 2561632 | 5323698 | 46,538 | 143,193 |
| # GPs | 970 | 2,933 | 2,587 | 2,749 | 130 | 176 |
| Fixed Effects | GP | GP | GP | GP | GP | GP |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP clustered standard errors are reported in parentheses. The table reports the relationship between identity/eligibility for a GP's reservation status, education, and selection as the Sarpanch following specification 2. The sample includes all elected politicians and the population in their GP from their identity group (identity models) or who were eligible for the reservation (reservation models). The dependent variable equals 10,000 for politicians and 0 for everyone else to aid interpretation. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects included as specified. We exclude people who were not 2022 at the time of data collection to ensure accurate education data.

**Table 4.** Minority Politicians are more Positively Selected vis-a-vis their Households than Upper Caste Men Politicians

| **Dependent Variable:** | Sarpanch (× 1) | | | |
|---|---|---|---|---|
| **Sample:** | Odisha Census | | | |
| **Group Measure:** | Identity Panel (1) | Identity 2017 (2) | Identity 2022 (3) | Reservation 2022 (4) |
| Education | 0.03** | 0.04*** | 0.04*** | 0.03*** |
| | (0.01) | (0.01) | (0.01) | (0.00) |
| Education × UC Women/Open Women | 0.03** | -0.00 | -0.00 | 0.00 |
| | (0.01) | (0.01) | (0.01) | (0.01) |
| Education × SC/ST Men/Open SC/ST | 0.03** | 0.02** | 0.02* | 0.01*** |
| | (0.01) | (0.01) | (0.01) | (0.00) |
| Education × SC/ST Women/SC/ST Women | 0.03* | 0.02* | 0.01 | 0.02** |
| | (0.02) | (0.01) | (0.01) | (0.01) |
| E(Y \| UC Men/Open) | 0.21 | 0.29 | 0.19 | 0.15 |
| # Observations | 2,268 | 6,120 | 5,670 | 9,261 |
| # GPs | 970 | 2,933 | 2,587 | 2,682 |
| Fixed Effects | HH | HH | HH | HH |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP clustered standard errors are reported in parentheses. The table reports the relationship between identity/eligibility for a GP's reservation status, education, and selection as the Sarpanch within the household of the elected Sarpanch following specification 2. The sample includes all elected politicians and the members of their household from their identity group (identity models) or who were eligible for the reservation (reservation models). The dependent variable equals 1 for politicians and 0 for everyone else. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects included as specified. We exclude people who were not 2022 at the time of data collection to ensure accurate education data.

Overall, our data reveal that reserved-seat politicians are less educated on average, have the same Mincer residuals, and are relatively more positively selected with respect to education than open-seat politicians. Since we have argued that politician quality is a function of both innate aptitudes and acquired skills, these results present possibly contradictory findings on the impact of reservations on politician quality. Quota skeptics might argue that the lower average education levels indicate lower average quality. Viewing our results in tandem suggests that such arguments must rely on a narrow definition of quality: that quality is solely the output of qualifications. Broader measures of quality lead to more benign conclusions around the impact of reservations on politician quality. The relative political selection results suggest that quotas improve quality as voters are even more attuned to education when choosing minority politicians. In the following sections, we turn to direct tests of voter and structural discrimination to explain these observed patterns of political selection.

## 4 The Role of Voter Discrimination

Our theory predicts that the relatively higher quality standard imposed on quota politicians derives from voter discrimination. To more directly evaluate this hypothesis, we consider the role of discrimination in the voter selection process in two ways: by estimating whether the positive quality penalty for minorities is conditioned by the information environment (as proxied by demography) and by directly examining voter selection in a candidate choice experiment.

First, we evaluate whether the positive education penalty applied to quota politicians is conditioned by the information environment, which we proxy for with the share of minority group members in the population. If voter discrimination emerges because of poorer information about the quality of quota-elected politicians or because of taste-based discrimination, we would expect discrimination to be lower as the share of the population from minority groups increases, assuming minorities have more information about in-group member quality or less distaste for minorities (Habyarimana et al., 2007). We test this observable implication in Table 5 by amending specification 2 to include the triple interaction between identity/reservation status, education, and the SC and ST shares of the population in a GP. Essentially, we extend Table 3 to include this triple interaction, and we disaggregate SC and ST politicians/reservations to more precisely evaluate the role of demography. The coefficients on the triple interaction represent the difference in political selection in places with higher SC and ST population shares.

Consistent with voter discrimination, Table 5 shows that the additional positive selection penalty on education for SC/ST men and women (and only SC/ST men and women) is declining in the SC/ST population share (as evidenced by the negative and significant coefficient on the triple interaction indicated by the grey shaded rows). In fact, the positive penalty on education for SCs diminishes only as the SC population share increases but is not affected by the ST population share. The same is true for STs. We can compare this to the additional degree of positive selection that these groups face in GPs where they comprise none of the population (shown by the coefficients on the double interaction in rows 2-6), which replicates the findings from Table 3. This comparison reveals that SCs and STs are selected similarly to upper castes when they comprise at least 50-60% of the population.

Second, we provide direct evidence that voters discriminate against reserved-seat candidates by leveraging a candidate choice experiment conducted with 1,251 randomly sampled citizens.[20] This experiment asked eligible voters to select between two candidate profiles contesting for the

[20] This experiment was embedded within a separate study and so was conducted in two districts of Madhya Pradesh. Citizen sampling was done as part of a larger randomized experiment, which focused on women's group members. Of the 1,251 respondents, 1,054 are women randomly sampled from registers of women's groups and 197 are randomly sampled husbands of these women. Citizens were surveyed in person by gender-matched enumerators.

**Table 5.** The Higher Standard for Minorities is Declining in Their Population Share

| **Dependent Variable:** | Sarpanch (× 10,000) | | | |
|---|---|---|---|---|
| **Sample:** | Odisha Census | | | |
| **Group Measure:** | Identity Panel (1) | Identity 2017 (2) | Identity 2022 (3) | Reservation 2022 (4) |
| Education | 1.19*** | 1.09*** | 1.54*** | 0.94*** |
| | (0.25) | (0.17) | (0.19) | (0.09) |
| Education | 0.00 | 0.00 | 0.00 | 0.00 |
| | (.) | (.) | (.) | (.) |
| Education × UC Women/Open Women | 0.41 | 0.30 | -0.12 | 0.70*** |
| | (0.46) | (0.31) | (0.31) | (0.20) |
| Education × SC Men/Open SC | 11.56*** | 14.32*** | 11.91*** | 7.18*** |
| | (2.64) | (1.92) | (1.48) | (0.91) |
| Education × SC Women/SC Women | 17.76*** | 18.11*** | 14.84*** | 13.61*** |
| | (3.89) | (2.03) | (1.94) | (2.39) |
| Education × ST Men/Open ST | 7.45*** | 10.19*** | 6.67*** | 5.71*** |
| | (2.62) | (1.49) | (1.73) | (1.21) |
| Education × ST Women/ST Women | 5.20 | 16.26*** | 8.83*** | 5.20** |
| | (3.99) | (2.73) | (2.95) | (2.37) |
| Education × SC Pop. Share | 0.60 | 1.88** | 1.01 | 0.62* |
| | (1.14) | (0.78) | (0.86) | (0.37) |
| Education × ST Pop. Share | 6.19*** | 5.99*** | 5.74*** | 2.22*** |
| | (1.31) | (0.82) | (0.81) | (0.30) |
| Education × UC Women/Open Women × SC Pop. Share | 1.10 | 0.21 | 1.34 | -0.51 |
| | (1.89) | (1.32) | (1.43) | (0.87) |
| Education × SC Men/Open SC × SC Pop. Share | -18.59** | -27.40*** | -23.82*** | -15.84*** |
| | (7.99) | (5.69) | (4.25) | (2.70) |
| Education × SC Women/SC Women × SC Pop. Share | -23.83** | -30.02*** | -31.76*** | -30.73*** |
| | (11.91) | (5.66) | (5.33) | (7.23) |
| Education × ST Men/Open ST × SC Pop. Share | 10.37 | -4.10 | 5.35 | 5.96 |
| | (11.00) | (5.44) | (5.57) | (4.32) |
| Education × ST Women/ST Women × SC Pop. Share | 25.25** | 9.41 | 10.84 | 3.00 |
| | (12.56) | (8.25) | (9.22) | (7.77) |
| Education × UC Women/Open Women × ST Pop. Share | 0.38 | 0.12 | 2.11 | 1.13 |
| | (3.58) | (1.48) | (1.62) | (0.81) |
| Education × SC Men/Open SC × ST Pop. Share | -5.36 | -2.57 | -2.54 | 1.52 |
| | (9.62) | (6.30) | (3.89) | (2.55) |
| Education × SC Women/SC Women × ST Pop. Share | 23.03 | 1.57 | 14.90* | 4.68 |
| | (19.06) | (7.55) | (7.67) | (6.06) |
| Education × ST Men/Open ST × ST Pop. Share | -13.37*** | -16.40*** | -12.35*** | -7.93*** |
| | (3.21) | (1.97) | (2.19) | (1.44) |
| Education × ST Women/ST Women × ST Pop. Share | -9.75* | -23.85*** | -14.77*** | -5.20* |
| | (5.43) | (3.46) | (3.62) | (2.91) |
| # Observations | 981,228 | 2667506 | 2382874 | 5066468 |
| # GPs | 970 | 2,932 | 2,585 | 2,746 |
| Fixed Effects | GP | GP | GP | GP |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP clustered standard errors are reported in parentheses. The table reports the relationship between identity/eligibility for a GP's reservation status, education, GP SC and ST population shares, and selection as the Sarpanch following specification 2. The sample includes all elected politicians and the population in their GP from their identity group (identity models) or who were eligible for the reservation (reservation models). The dependent variable equals 10,000 for politicians and 0 for everyone else to aid interpretation. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects included as specified. We exclude people who were not 2022 at the time of data collection to ensure accurate education data.

position of chairperson. Profiles randomly vary in gender, caste, age, education, income, education, occupation, political experience, access to transportation, number of children, whether or not they are a member of a Self-Help Group, and expectations of family support if elected. To directly test the effect of reservation type on voter selection, the experiment randomly varied at the respondent level whether the electoral race was unreserved, reserved for a woman, or reserved for an SC or ST. Candidate gender and caste were constrained to match the reservation type.[21] Respondents were instructed to imagine these are two actual candidates in a local government election and asked to choose which of the two profiles they would vote for.

Figure 2 reports the average marginal conditional effects for education and political experience (model 1) and how these effects are moderated by reservation status (model 2). The results show that citizens prioritize education and experience in candidate selection, with somewhat greater emphasis on education over having previously been the elected chairperson (though not significantly so).[22] Candidates who had completed at least 10th-grade education were nine percentage points more likely to be selected than those with less than 10th-grade education. The preference for educated politicians, however, only exists in reserved-seat races. Model 2, which interacts candidate education and election type, reveals that there is no relationship between education and candidate selection in open-seat races but a positive and significant relationship for candidates in women and SC/ST reserved races.

An alternative explanation for this positive education penalty for reserved seat candidates is that, rather than voter discrimination, education serves as a stronger signal of quality for marginalized candidates than for dominant candidates. We think this is unlikely considering the effects of past political experience: Figure 2 shows that open-seat politicians are more likely to be selected if they have previously served as a Sarpanch. However, there is also a marginal benefit to political experience for SC/ST politicians, suggesting that the higher education standard applied to reserved seat politicians is not met by other higher qualifications standards, as would be expected if different qualifications serve as signals of quality for different types of candidates.

We validate these findings with an additional conjoint experiment on a different sample of 2,832 citizens who were asked to select between two female candidates vying for a women-reserved seat.[23] While the choice is constrained to women candidates, profiles vary in their caste. Appendix Figure D4 reports the average conditional marginal effect for education holding all other attributes

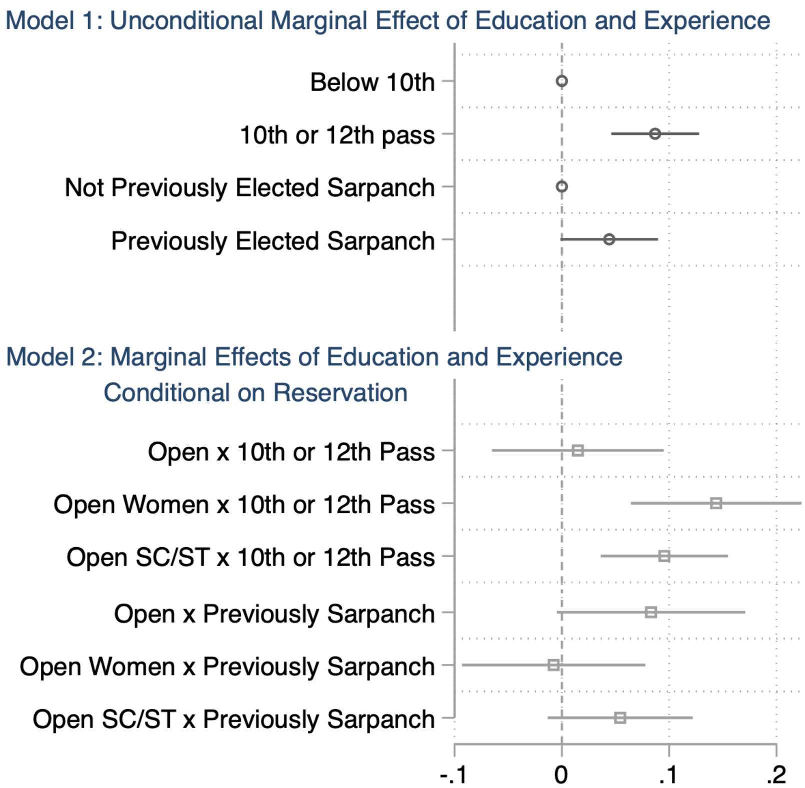
[21] In unreserved seats or SC/ST reserved seats, 80% of profiles were male to match observed reality. In unreserved or women-reserved seats, 45% of the profiles were upper caste, 45% were OBC, 5% were SC, and 5% were ST. As a result, reservation closely but imperfectly aligns with identity groups.

[22] The average marginal conditional effects for all attributes are reported in Appendix Figure D2.

[23] This survey was also conducted in Madhya Pradesh, and we sampled six citizens in each village using a random walk procedure, ensuring even balance across men and women and representation across education categories and age groups.

constant (model 1) and conditional on candidate caste (model 2). Citizens' preference for educated politicians only exists for minority candidates.

**Figure 2.** Citizens Prefer More Educated Candidates in Reserved Seats




*Note*: The figure reports the average marginal conditional effects of education and political experience by reservation status from a conjoint experiment conducted with a random sample of citizens in Madhya Pradesh. 95% confidence intervals from OLS regressions with standard errors clustered at the respondent level reported. The sample includes 1,251 citizens.

While our conjoint results present clear evidence that voters would discriminate against reserved seat politicians based on education in line with our observation data, one concern may be that our observational results on the higher standard imposed on reserved seat politicians occur at earlier stages in the political pipeline. First, the local elections that we consider are formally nonpartisan, removing concerns of discrimination in candidate nomination by parties. Second, we consider whether the differential selection of reserved seat politicians by education occurs at the state of candidate entry or voter selection in Appendix G using candidate data from the All India survey.[24] Appendix Table G2 correlates political selection with education, comparing elected representatives to candidates (column 3). It shows that elected politicians are positively

[24] Candidate data is not available for the Odisha sample.

selected from the pool of candidates on education, but there is no observable difference between the degree of positive selection across identity groups. It also shows that *candidates* in reserved seat races are more positively selected on education vis-a-vis their identity group in the population (column 1). Overall, this suggests that the positive selection we observe in our main results may happen at the stage of candidate emergence. However, this does not challenge our argument of voter discrimination: in equilibrium and in line with endogenous candidacy models, we would expect that candidates would predict voters' disproportionate preference for education among the marginalized and select into contestation accordingly.

## 5 The Role of Structural Discrimination

Next, we consider the role of structural discrimination in explaining the lower average levels of education between marginalized politicians and upper-caste male politicians as seen in Table 2. To do so, we leverage GP-level variation in the supply of educated marginalized group members. Our theory predicts that, under structural discrimination, GPs with a larger share of highly educated marginalized people should demonstrate less of a negative education gap between quota and non-quota-elected politicians. To examine this, we calculate the share of each GP's population that had completed 12th-grade education by identity group to indicate the supply of education and interact this with politician identity/GP reservation status in Specification 1. The full set of results for all samples can be found in Appendix Table D9.

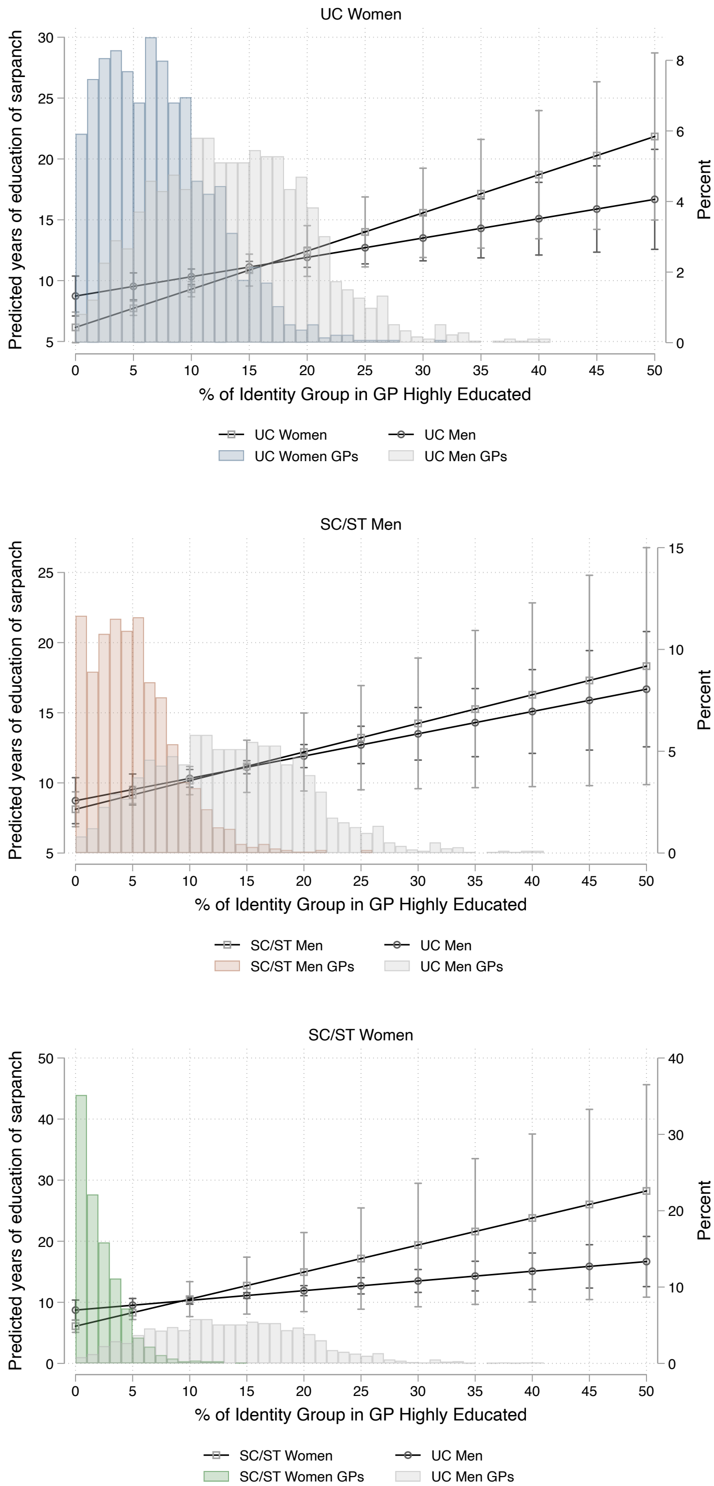
Using results from the Odisha panel of elections (our preferred specification), Figure 3 reports the predicted years of education for politicians from each identity group by the share of highly educated people from their identity group in their GP. For comparison, each panel also includes the predicted years of education for upper-caste men politicians by the share of highly educated upper-caste men in their GP. Each figure also shows the distribution of GPs by their share of highly educated people in the population (see the right-side y-axis). A comparison of these distributions reveals that, in all cases, the distribution of highly educated upper-caste men is to the right of that of women and SCs/STs, suggesting substantial structural discrimination. In the average GP, 13% of upper-caste men are highly educated, whereas only 7% of upper-caste women, 5% of SC/ST men, and 2% of SC/ST women are highly educated.

Figure 3 provides evidence in support of the role of structural discrimination in shaping average differences in qualifications across reserved and open seats: the difference in education between upper-caste men politicians and marginalized politicians is declining in the supply of highly educated marginalized group members. The top panel shows that, in GPs where at least 15% of upper-caste women have completed 12th grade, the predicted education level of upper-caste women politicians is statistically identical to that of upper-caste male politicians (as evidenced by the crossing of the two predicted probability lines). The same expectation is obtained for SC/ST men politicians in GPs where at least 15% of SC/ST men have completed 12th grade education (middle panel) and for SC/ST women politicians in GPs where at least 10% of SC/ST women have completed education (see the bottom panel). In all panels, we see further evidence suggesting a stronger preference for highly educated minorities, as the slope of the line for marginalized politicians is greater than that for upper-caste male politicians. This means that as the supply of education in the group population increases, the average education of selected politicians increases at a greater rate for women and SC/ST politicians, though the limited existence of GPs with high rates of minority education suggests caution in interpreting these results. Table D9 shows that these results are robust and somewhat conservative across our other samples and less demanding specifications across Odisha. We also observe similar patterns of steeper educational selection gradients among women and minority politicians in the All India Sample, where there is greater variation in the supply of education to minorities. However, the results are noisy, given a small number of GPs and limited power to detect heterogeneous effects. Finally, Appendix Figure D1 shows that, for all levels of the supply of education, the Mincer Residual for marginalized politicians is the same or larger than that of upper-caste men politicians (and generally shows a similar greater positive gradient for marginalized politicians). Ultimately, these findings suggest that interpreting the overall average differences in education between open-seat and reserved-seat politicians as evidence of a negative impact of reservations on quality would be incorrect. Once the supply of education is accounted for, we observe either no difference in average qualifications or a positive difference.

A key challenge is interpreting these results as causal, given that the underlying distribution of education in the population may reflect other confounding factors that also shape selection decisions. We consider the robustness of our findings across minority men and women, as well as with the All India data, as suggestive of our interpretation of structural discrimination. Further work, however, is needed to confirm the size of the role of the supply of education in qualifications-based patterns of political selection.[25]

[25] To support the argument that the link between the supply of education and politicians' education is causal, we leverage data from the 1991 census of India to identify the GPs in Odisha that had secondary schools before the creation of village governments in 1992. We instrument the share of a GP population that has completed secondary school with an indicator for whether the GP had a secondary school in 1991 and then estimate the impact of population education levels on politician education levels. Appendix Table D10 reports the results from this instrumental variables analysis and shows that the supply of highly educated people, resulting from the early presence of secondary schools, increases the average educational attainment of politicians. These results should be interpreted with caution as it is possible that other unobserved factors determined the allocation of schools in 1991 and also shape political selection patterns today.

**Figure 3.** Greater Supply of Highly Educated Minorities Reduces the Education Gap




*Note*: The figure reports the predicted values of politician education from a model where politician identity is interacted with the share of the identity group in the GP that has completed 12th grade education following specification 1. 95% confidence intervals from based on standard errors clustered at the GP level. The results use the sample from the Odisha identity panel model with GP fixed effects (see column 1 in Appendix Table D9). Additional covariates include SC and ST population shares (and their squares) and politicians' group occupation shares in agriculture, business, service, manual work, and others.

## 6 Alternative Explanation: Costs to Entry

An alternative explanation commonly found in the literature is that the observed patterns of political selection are not the result of discrimination but because marginalized identity groups face higher costs to running for office (Fox and Lawless, 2004; Ashworth et al., 2024; Teele et al., 2018; Dal Bó et al., 2017). We control for these concerns by focusing on reservations, where identity is held constant and, therefore, so are identity-based costs to political entry. We further confirm that differential selection into contestation is unlikely to explain our results using our candidate data from the All India sample. Appendix Table G1 shows that, on average, 7.2 candidates contest across all races, with only marginally more candidates in open seat races (7.5, on average) than reserved seat races (6.6 in open women races and open SC/ST races and 6.1 in SC/ST women races), suggesting that reservations are not met with substantially lower contestation.

Under reservations, costs to entry would only explain our results if they differed along the intersection of identity and education. We consider this concern with respect to labor market opportunity costs in Appendix F, the domain in which we think these differential costs would be most likely given historic caste-based segmentation in labor markets.[26] Appendix Table F1 shows that the greater positive selection of women and minorities on education holds in both size and significance in GPs with more favorable labor environments. Appendix Figures F2 and F3 show that the expected level of politician education for all groups is unconditioned by the GP-level supply of manual work and service sector work. This suggests that the labor market opportunities available to different groups are unlikely to explain our results.

## 7 Conclusion and Implications

Do electoral quotas worsen politician quality? The evidence presented in this paper points to a resounding no. While it is true that women and lower-caste local politicians in rural India have, on average, fewer years of education than their upper-caste male counterparts, we show that voters hold these groups to a higher standard. Quota-elected politicians are more positively selected relative to their group's education distribution and are positively selected on education relative to the education distribution of the historically privileged upper castes. They also exhibit no less and potentially even more latent quality based on a measure derived using Mincer residuals. Further, once the supply of education in the population is accounted for, differences in average education levels among politicians disappear and, if anything, reverse. Taken together, this evidence suggests that once structural discrimination resulting from the uneven supply of qualifications in the population

[26] Our inclusion of either group-specific labor market opportunities at the GP-level or GP fixed effects as controls in our models accounts for local labor environments.

is accounted for, reserved-seat politicians are either of the same or higher quality as open-seat politicians.

Why are reserved seat politicians of the same or higher quality even amidst inequalities of opportunity? We argue and provide several pieces of evidence in support of the role of voter discrimination. In data from two candidate choice experiments, we show that voters hold reserved seat politicians and minorities to higher educational standards, even after accounting for political experience. We further show that the greater relative positive selection of reserved seat politicians is most present in communities dominated by non-coethnics, who are expected to rely more on stereotypes and exhibit greater taste-based preferences.

Since quotas are frequently used to right historical wrongs, quota recipients tend to suffer from precisely the kinds of discrimination that condition the effects of quotas in India's elections. We expect our results to travel to other countries and sectors (such as education and employment) with quotas where similar constraints on opportunity meet deep-seated stereotypes and stigmas (Bhavnani and Lee, 2021). Our theory would suggest that last-mile discrimination should improve candidate quality, even if qualifications are worse. Presuming quotas are unmeritocratic by observing their aggregate effects on qualifications can severely misrepresent their impact.

Our framework helps recast the long-running debate on affirmative action. By restricting competition, affirmative action does not mechanically worsen qualifications. Rather, affirmative action worsens qualifications if it emerges in contexts where structural discrimination restricts the entry of qualified candidates. In terms of policy, this underlines the importance of reforms such as mass education that could alleviate structural inequalities and, therefore, serve as a critical complement to affirmative action.

## References


Afridi, Farzana , Abhishek Arora, Diva Dhar, and Kanika Mahajan (2023). Women's work, social norms and the marriage market. Technical report, IZA Discussion Papers.

Anzia, Sarah F and Rachel Bernhard (2022). Gender stereotyping and the electoral success of women candidates: New evidence from local elections in the united states. *British Journal of Political Science 52*(4), 1544–1563.

Anzia, Sarah F and Christopher R Berry (2011). The Jackie (and Jill) Robinson effect: why do congresswomen outperform congressmen? *American Journal of Political Science 55*(3), 478–493.

Asher, Sam , Paul Novosad, and Charlie Rafkin (2018). Intergenerational mobility in India: Estimates from new methods and administrative data. *World Bank Working Paper*.

Ashworth, Scott , Christopher R. Berry, and Ethan Bueno de Mesquita (2024). Modeling Theories of Women's Underrepresentation in Elections. *American Journal of Political Science 68*(1), 289–303.

Auerbach, Adam Michael and Adam Ziegfeld (2020). How do electoral quotas influence political competition? evidence from municipal, state, and national elections in india. *The Journal of Politics 82*(1), 397–401.

Baltrunaite, Audinga , Piera Bello, Alessandra Casarico, and Paola Profeta (2014). Gender quotas and the quality of politicians. *Journal of Public Economics 118*, 62–74.

Bamezai, Apurva , Siddharth George, M. R. Sharan, and Borui Sun (2024). Who becomes a local politician? evidence from rural india. *Unpublished Manuscript*.

Ban, Radu and Vijayendra Rao (2008). Tokenism or agency? The impact of women's reservations on village democracies in South India. *Economic Development and Cultural Change 56*(3), 501–530.

Bauer, Nichole M (2020). *The qualifications gap: Why women must be better than men to win political office*. Cambridge University Press.

Beaman, Lori , Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova (2009). Powerful women: does exposure reduce bias? *The Quarterly journal of economics 124*(4), 1497–1540.

Besley, Timothy , Olle Folke, Torsten Persson, and Johanna Rickne (2017). Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden. *American Economic Review 107*(8), 2204–2242.

Besley, Timothy , Jose G. Montalvo, and Marta Reynal-Querol (2011). Do Educated Leaders Matter? *The Economic Journal 121*(554), F205–227.

Bhavnani, Rikhil R. (2009). Do Electoral Quotas Work after They Are Withdrawn? Evidence from a Natural Experiment in India. *American Political Science Review 103*(1), 23–35.

Bhavnani, Rikhil R (2017). Do the effects of temporary ethnic group quotas persist? Evidence from India. *American Economic Journal: Applied Economics 9*(3), 105–123.

Bhavnani, Rikhil R and Alexander Lee (2021). Does affirmative action worsen bureaucratic performance? evidence from the indian administrative service. *American Journal of Political Science 65*(1), 5–20.

Brulé, Rachel E (2020). Reform, representation, and resistance: The politics of property rights' enforcement. *The Journal of Politics 82*(4), 1390–1405.

Bush, Sarah Sunn (2011). International politics and the spread of quotas for women in legislatures. *International Organization 65*(1), 103–137.

Carnes, Nicholas and Noam Lupu (2016). What Good Is a College Degree? Education and Leader Quality Reconsidered. *The Journal of Politics 78*(1), 35–49.

Carreri, Maria and Julia Payson (2021). What makes a good local leader? evidence from us mayors and city managers. *Journal of Political Institutions and Political Economy 2*(2), 199–225.

Chattopadhyay, Raghabendra and Esther Duflo (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica 72*(5), 1409–1443.

Chauchard, Simon (2017). *Why representation matters: The meaning of ethnic quotas in rural India*. Cambridge University Press.

Chaudhuri, Ananish , Vegard Iversen, Francesca R Jensenius, and Pushkar Maitra (2024). Time in office and the changing gender gap in dishonesty: Evidence from local politics in india. *American Journal of Political Science 68*(1), 106–122.

Clayton, Amanda and Pär Zetterberg (2018). Quota shocks: Electoral gender quotas and government spending priorities worldwide. *The Journal of Politics 80*(3), 916–932.

Cruz, Cesi (2019). Social networks and the targeting of vote buying. *Comparative Political Studies 52*(3), 382–411.

Dal Bó, Ernesto and Frederico Finan (2018). Progress and perspectives in the study of political selection. *Annual Review of Economics 10*(1), 541–575.

Dal Bó, Ernesto , Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne (2017). Who becomes a politician? *The Quarterly Journal of Economics 132*(4), 1877–1914.

Das, Sabyasachi , Abhiroop Mukhopadhyay, and Rajas Saroy (2023). Does affirmative action in politics hinder performance? evidence from india. *Journal of Economic Behavior & Organization 214*, 370–405.

Desai, Zuheir , Varun Karekurve-Ramachandra, and Sergio Montero (2024). Are women better politicians? discrimination, gender quotas, and electoral accountability. Working Paper.

Dunning, Thad and Janhavi Nilekani (2013). Ethnic quotas and political mobilization: caste, parties, and distribution in Indian village councils. *American Political Science Review 107*(1), 35–56.

Fox, Richard L. and Jennifer L. Lawless (2004). Entering the Arena? Gender and the Decision to Run for Office. *American Journal of Political Science 48*(2), 264–280.

Fujiwara, Thomas , Hanno Hilbig, and Pia Raffler (2024). Biased party nominations as a source of women's electoral underperformance. OSF Working Paper.

Goyal, Tanushree (2024a). Local political representation as a pathway to power: A natural experiment in india. *American Journal of Political Science*, 1–15.

Goyal, Tanushree (2024b). Representation from below: How women's grassroots party activism promotes equal political participation. *American Political Science Review 118*(3), 1415–1430.

Gulzar, Saad (2021). Who Enters Politics and Why? *Annual Review of Political Science 24*(1), 253–275.

Gulzar, Saad , Nicholas Haas, and Benjamin Pasquale (2020). Does political affirmative action work, and for whom? Theory and evidence on India's scheduled areas. *American Political Science Review 114*(4), 1230–1246. Publisher: Cambridge University Press.

Habyarimana, James , Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein (2007). Why Does Ethnic Diversity Undermine Public Goods Provision? *American Political Science Review 101*(4), 709–725.

Hanna, Rema N and Leigh L Linden (2012). Discrimination in grading. *American Economic Journal: Economic Policy 4*(4), 146–168.

Jain, Chandan , Shagun Kashyap, Rahul Lahoti, and Soham Sahoo (2023). The impact of educated leaders on economic development: Evidence from India. *Journal of Comparative Economics*.

Karekurve-Ramachandra, Varun (2023). Gender quotas and upward political mobility in india. *Working paper*.

Karekurve-Ramachandra, Varun and Alexander Lee (2020). Do Gender Quotas Hurt Less Privileged Groups? Evidence from India. *American Journal of Political Science 64*(4), 757–772.

Karekurve-Ramachandra, Varun and Alexander Lee (2024). Can gender quotas improve public service provision? evidence from indian local government. *Comparative Political Studies*, 1–39.

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

Larson, Jennifer M and Janet I Lewis (2017). Ethnic networks. *American Journal of Political Science 61*(2), 350–364.

Mosse, David (2018). Caste and development: Contemporary perspectives on a structure of discrimination and advantage. *World development 110*, 422–436.

Murray, Rainbow (2010). Second Among Unequals? A Study of Whether France's "Quota Women" are Up to the Job–ERRATUM. *Politics & Gender 6*(4), 643–669.

O'Brien, Diana Z (2012). Quotas and qualifications in Uganda. In S. Franceschet, M. L. Krook, and J. M. Piscopo (Eds.), *The impact of gender quotas*, pp. 57–71. Section: 4.

Prillaman, Soledad Artiz (2023). *The Patriarchal Political Order: The Making and Unraveling of the Gendered Participation Gap in India*. Cambridge University Press.

Purohit, Bhumi (2022). Gendered bureaucratic resistance against female politicians: Evidence from telangana, india. *Unpublished Manuscript*.

Schwarz, Susanne and Alexander Coppock (2022). What have we learned about gender from candidate choice experiments? a meta-analysis of sixty-seven factorial survey experiments. *The Journal of Politics 84*(2), 655–668.

Teele, Dawn Langan , Joshua Kalla, and Frances Rosenbluth (2018, August). The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics. *American Political Science Review 112*(3), 525–541.

Weeks, Ana Catalano (2022). *Making gender salient: From gender quota laws to policy*. Cambridge University Press.

Weeks, Ana Catalano and Lisa Baldez (2015). Quotas and qualifications: the impact of gender quota laws on the qualifications of legislators in the Italian parliament. *European Political Science Review 7*(1), 119–144.

# Appendix for "Does Affirmative Action Worsen Quality? Theory and Evidence to the Contrary from Elections"

Rikhil R. Bhavnani    Alba Huidobro    Soledad Artiz Prillaman

## Contents



*We thank Stuart Turnbull-Dugarte, Rajeshwari Majumdar, Tine Paulsen and participants at Leiden University, EPSA, APSA and MPSA for comments on previous versions of the paper, and Diego Tocre and Armelle Grondin for yeoman's work with the data analysis. Human subjects research in this article was reviewed and approved by the UW-Madison IRB (submission ID 2022-1281) and the Stanford IRB (submission IDs 67578 and 68998). Bhavnani acknowledges the support of the University of Wisconsin–Madison Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation. Prillaman acknowledges the support of the Stanford King Center on Development.

†Professor and Glenn B. and Cleone Orr Hawkins Chair, Department of Political Science, University of Wisconsin–Madison. Email: bhavnani@wisc.edu.

‡Postdoctoral Research Fellow, Chair of Political Behavior, Humboldt-Universität zu Berlin. Email: alba.huidobro@hu-berlin.de.

§Assistant Professor, Department of Political Science, Stanford University. Email: soledadp@stanford.edu.

# A Ethics and transparency

The authors confirm that their research is compliant with APSA's Principles and Guidance for Human Subjects Research and declare that they have no potential or perceived conflicts of interest arising from their research. Human subjects research was reviewed and approved by the XXX IRB (submission ID 2022-1281) and the XXX IRB (submission IDs 67578 and 68998).

Author XXX acknowledges the support of the University XXX Office of the Vice Chancellor for Research and Graduate Education with funding from the XXX Alumni Research Foundation. Author XXX acknowledges the support of the XXX Center on Development.

Data collection and analysis procedures are explained in the main text and Data appendix. The quantitative data and code necessary to produce the results will be made publicly available on the Dataverse, with the exception of data from the 2011-12 Socioeconomic and Caste Census (SECC) for Odisha.

SECC microdata cannot be released due to ethical considerations and under the terms that the data were obtained. Other studies that use census micro-data do not release them. See, for example, Ernesto Dal Bó, Frederico Finan, Olle Folke, Torsten Persson, Johanna Rickne. 2017. "Who Becomes A Politician?" *The Quarterly Journal of Economics* 132(4): 1877–1914. That said, the paper and replication code provide the guidance needed to replicate the analysis once the data have been obtained.

# B Data Appendix

To describe the nature of political selection in rural Indian villages, our study employs a systematic procedure to match two unique and comprehensive data sources. These sources include the 2011-12 Socioeconomic and Caste Census (the Odisha census), conducted nationwide to determine eligibility for government programs and provided for Odisha, and data on elected Gram Panchayat (GP) members from the two most recent elections in 2017 and 2022. The data integration process unfolded in four key stages. First, we scraped the names of the elected GP members from the state election commission's website. This dataset included only the names of the 6,770 elected chairpersons for 2017 and 6,932 for 2022, forming the foundation for subsequent steps.

Second, using data from India's Local Government Directory (LGD), we constructed a village-to-GP crosswalk to facilitate the merge of GP-level election data with village-level census data. Given the regularly changing village boundaries in India, this mapping was not perfect, though we achieved a very high match rate of close to 100%. We then dropped GPs associated with multiple blocks, those in which the identification variables are missing, and the 4% of villages that shared names within the same sub-district where it was impossible to match the villages across the two databases definitively. We then merge village-level datasets, including the 2001 and 2011 Censuses and the Socioeconomic High-resolution Rural-Urban Geographic Platform for India (SHRUG). The 2011 Census is merged using census codes, while SHRUG data is merged through a fuzzy match based on village names.

Third, using the final LGD village-GP crosswalk, We fuzzy-merged the 2017 and 2022 election data, including Sarpanch and PS Member data, by first merging at the block level, then at the GP level using block-GP names. We achieved a match rate of 98% for both elections. Next, we fuzzy-merged the Odisha census information using village names, achieving a 96% match rate for census observations in both years. We drop nameless people (around 8%).

Four, we executed a fuzzy matching process to merge politician data with the Odisha census data. We first excluded individuals from the census who were ineligible for candidacy, including those under 21 at the time of the elections and those with more than two children. We exclude GPs where more than one resident shared the same name as the politician as we had no way to adjudicate the match. Then we conducted the fuzzy matching using the Masala merge algorithm in Stata, a specialized algorithm designed to handle transliterated Indian names. The only available information for the merge was name and GP, given the sparseness of the electoral data. For the fuzzy merge, we used a threshold of 0.5, representing the maximum allowable Levenshtein distance for matches (adjusted for Hindu misspellings). Increasing the threshold from 0.5 to 2 boosts the match rate by 5%, but also introduces more inaccuracies, so this adjustment may not be beneficial.

First, we merged Sarpanch/PS Member names at the GP level from LGD with Odisha census names, saving both matched and unmatched names. Next, we removed middle names from the unmatched observations and performed another fuzzy merge. We then append the results from both merges. Fuzzy matches were visually validated for accuracy. Finally, we included all original individuals from the Odisha census, even those who were dropped due to eligibility criteria. This yielded a final match rate of 48% of politicians.

Tables B1 and B2 compare the characteristics of the GPs where the 2017 and 2022 politicians were matched with those where the politicians were not matched respectively. There are no significant differences between the matched and unmatched samples, suggesting that our matched sample is fairly representative of the population of Odisha.

**Table B1.** Balance table Comparing GPs with and without matched 2017 Politicians

| | (1) Unmatched GPs | | (2) Matched GPs | | Difference in Means | |
|---|---|---|---|---|---|---|
| Variable | N | Mean | N | Mean | N | (1)-(2) |
| Total Population | 2578 | 5071.462 | 2919 | 5087.052 | 5497 | -15.590 |
| Women (% of GP Population) | 2578 | 0.528 | 2919 | 0.529 | 5497 | -0.001 |
| SC (% of GP Population) | 2578 | 0.172 | 2919 | 0.170 | 5497 | 0.002 |
| ST (% of GP Population) | 2578 | 0.222 | 2919 | 0.230 | 5497 | -0.008 |
| Employed (% of GP Population) | 2578 | 0.346 | 2919 | 0.350 | 5497 | -0.003 |
| Unemployed (% of GP Population) | 2578 | 0.119 | 2919 | 0.120 | 5497 | -0.001 |
| Student (% of GP Population) | 2578 | 0.066 | 2919 | 0.065 | 5497 | 0.001 |
| Illiterate (% of GP Population) | 2578 | 0.331 | 2919 | 0.343 | 5497 | -0.012 |
| Average Years of Education in GP | 2578 | 4.869 | 2919 | 4.758 | 5497 | 0.112 |
| Average Age in GP | 2578 | 38.642 | 2919 | 38.622 | 5497 | 0.020 |
| Earns Less than 5k INR (% of GP Population) | 2578 | 0.855 | 2919 | 0.863 | 5497 | -0.008 |
| Earns 5k-10k INR (% of GP Population) | 2578 | 0.085 | 2919 | 0.081 | 5497 | 0.004 |
| Earns More than 5k INR (% of GP Population) | 2578 | 0.060 | 2919 | 0.056 | 5497 | 0.004 |
| Number of Primary Schools | 2526 | 1.781 | 2874 | 1.787 | 5400 | -0.006 |
| Number of Middle Schools | 2526 | 0.925 | 2874 | 0.926 | 5400 | -0.001 |
| Number of Secondary Schools | 2526 | 0.546 | 2874 | 0.527 | 5400 | 0.019 |
| GP Has Domestic Power Supply | 2526 | 0.912 | 2874 | 0.905 | 5400 | 0.006 |
| Share of Land used for Agriculture | 2522 | 0.579 | 2873 | 0.574 | 5395 | 0.004 |
| Mean Night Lights (Calibrated) in 2013 | 2556 | 5.240 | 2896 | 4.993 | 5452 | 0.247 |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are used when calculating the significance of the difference in means. Block fixed effects are included in the difference in means specification. The table compares the means of observable covariates across GPs where the 2017 electoral data was and was not matched to the SECC. Data on total population, population shares, and average socioeconomic characteristics come from the SECC data. Data on the number of schools, power supply, and land use come from the 2011 census village directory. Data on night lights come from the Earth Observation Group.

**Table B2.** Balance table Comparing GPs with and without matched 2022 Politicians

| | (1) Unmatched GPs | | (2) Matched GPs | | Difference in Means | |
|---|---|---|---|---|---|---|
| Variable | N | Mean | N | Mean | N | (1)-(2) |
| Total Population | 2924 | 5086.801 | 2573 | 5071.717 | 5497 | 15.084 |
| Women (% of GP Population) | 2924 | 0.529 | 2573 | 0.528 | 5497 | 0.001 |
| SC (% of GP Population) | 2924 | 0.170 | 2573 | 0.172 | 5497 | -0.002 |
| ST (% of GP Population) | 2924 | 0.229 | 2573 | 0.222 | 5497 | 0.007 |
| Employed (% of GP Population) | 2924 | 0.350 | 2573 | 0.346 | 5497 | 0.003 |
| Unemployed (% of GP Population) | 2924 | 0.120 | 2573 | 0.119 | 5497 | 0.001 |
| Student (% of GP Population) | 2924 | 0.065 | 2573 | 0.066 | 5497 | -0.001 |
| Illiterate (% of GP Population) | 2924 | 0.343 | 2573 | 0.331 | 5497 | 0.012 |
| Average Years of Education in GP | 2924 | 4.760 | 2573 | 4.867 | 5497 | -0.108 |
| Average Age in GP | 2924 | 38.621 | 2573 | 38.643 | 5497 | -0.022 |
| Earns Less than 5k INR (% of GP Population) | 2924 | 0.863 | 2573 | 0.855 | 5497 | 0.008 |
| Earns 5k-10k INR (% of GP Population) | 2924 | 0.081 | 2573 | 0.085 | 5497 | -0.004 |
| Earns More than 5k INR (% of GP Population) | 2924 | 0.056 | 2573 | 0.060 | 5497 | -0.004 |
| Number of Primary Schools | 2878 | 1.787 | 2522 | 1.782 | 5400 | 0.005 |
| Number of Middle Schools | 2878 | 0.927 | 2522 | 0.924 | 5400 | 0.002 |
| Number of Secondary Schools | 2878 | 0.527 | 2522 | 0.546 | 5400 | -0.018 |
| GP Has Domestic Power Supply | 2878 | 0.905 | 2522 | 0.912 | 5400 | -0.006 |
| Share of Land used for Agriculture | 2877 | 0.575 | 2518 | 0.578 | 5395 | -0.004 |
| Mean Night Lights (Calibrated) in 2013 | 2900 | 4.993 | 2552 | 5.241 | 5452 | -0.248 |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are used when calculating the significance of the difference in means. Block fixed effects are included in the difference in means specification. The table compares the means of observable covariates across GPs where the 2022 electoral data was and was not matched to the SECC. Data on total population, population shares, and average socioeconomic characteristics come from the SECC data. Data on the number of schools, power supply, and land use come from the 2011 census village directory. Data on night lights come from the Earth Observation Group.

# C Balance Tests

**Table C1.** Balance table Comparing GPs by Identity of Elected Chairperson in 2022

| | (1) Upper-caste Men | | (2) Upper-caste Women | | (3) SC/ST Men | | (4) SC/ST Women | | Difference in Means | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | N | Mean | N | Mean | N | Mean | N | Mean | N | (1)-(2) | N | (1)-(3) | N | (1)-(4) |
| Total Population | 892 | 5295.638 | 672 | 5166.994 | 744 | 4996.070 | 611 | 4805.399 | 1564 | 128.644 | 1636 | 299.568* | 1503 | 490.239*** |
| Women (% of GP Population) | 892 | 0.522 | 672 | 0.522 | 744 | 0.538 | 611 | 0.537 | 1564 | 0.000 | 1636 | -0.016 | 1503 | -0.015* |
| SC (% of GP Population) | 892 | 0.190 | 672 | 0.183 | 744 | 0.146 | 611 | 0.156 | 1564 | 0.006 | 1636 | 0.043** | 1503 | 0.034 |
| ST (% of GP Population) | 892 | 0.099 | 672 | 0.100 | 744 | 0.387 | 611 | 0.373 | 1564 | -0.001 | 1636 | -0.288*** | 1503 | -0.274*** |
| Employed (% of GP Population) | 892 | 0.338 | 672 | 0.342 | 744 | 0.363 | 611 | 0.360 | 1564 | -0.003 | 1636 | -0.024** | 1503 | -0.021 |
| Unemployed (% of GP Population) | 892 | 0.118 | 672 | 0.113 | 744 | 0.125 | 611 | 0.124 | 1564 | 0.005 | 1636 | -0.007 | 1503 | -0.006 |
| Student (% of GP Population) | 892 | 0.069 | 672 | 0.069 | 744 | 0.060 | 611 | 0.061 | 1564 | -0.000 | 1636 | 0.009* | 1503 | 0.008 |
| Illiterate (% of GP Population) | 892 | 0.278 | 672 | 0.281 | 744 | 0.429 | 611 | 0.405 | 1564 | -0.003 | 1636 | -0.151 | 1503 | -0.128 |
| Average Years of Education in GP | 892 | 5.270 | 672 | 5.279 | 744 | 4.110 | 611 | 4.226 | 1564 | -0.009 | 1636 | 1.160** | 1503 | 1.044 |
| Average Age in GP | 892 | 38.890 | 672 | 38.878 | 744 | 38.307 | 611 | 38.334 | 1564 | 0.013 | 1636 | 0.584 | 1503 | 0.557 |
| Earns Less than 5k INR (% of GP Population) | 892 | 0.838 | 672 | 0.836 | 744 | 0.894 | 611 | 0.891 | 1564 | 0.003 | 1636 | -0.056 | 1503 | -0.052 |
| Earns 5k-10k INR (% of GP Population) | 892 | 0.096 | 672 | 0.098 | 744 | 0.062 | 611 | 0.064 | 1564 | -0.002 | 1636 | 0.034 | 1503 | 0.032 |
| Earns More than 5k INR (% of GP Population) | 892 | 0.066 | 672 | 0.067 | 744 | 0.044 | 611 | 0.046 | 1564 | -0.001 | 1636 | 0.022 | 1503 | 0.020 |
| Number of Primary Schools | 874 | 1.931 | 665 | 1.817 | 737 | 1.657 | 598 | 1.704 | 1539 | 0.115* | 1611 | 0.275** | 1472 | 0.227 |
| Number of Middle Schools | 874 | 0.982 | 665 | 0.992 | 737 | 0.849 | 598 | 0.866 | 1539 | -0.011 | 1611 | 0.132 | 1472 | 0.115 |
| Number of Secondary Schools | 874 | 0.588 | 665 | 0.588 | 737 | 0.449 | 598 | 0.465 | 1539 | 0.000 | 1611 | 0.139 | 1472 | 0.123 |
| GP Has Domestic Power Supply | 874 | 0.946 | 665 | 0.934 | 737 | 0.851 | 598 | 0.881 | 1539 | 0.012 | 1611 | 0.095 | 1472 | 0.065 |
| Share of Land used for Agriculture | 874 | 0.618 | 665 | 0.619 | 736 | 0.523 | 598 | 0.524 | 1539 | -0.001 | 1610 | 0.096 | 1472 | 0.094 |
| Mean Night Lights (Calibrated) in 2013 | 884 | 5.864 | 669 | 5.491 | 739 | 4.100 | 604 | 4.259 | 1553 | 0.373 | 1623 | 1.764 | 1488 | 1.605 |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are used when calculating the significance of the difference in means. Block fixed effects are included in the difference in means specification. The table compares the means of observable covariates across GPs based on the identity of the elected chairperson in 2017. Data on total population, population shares, and average socioeconomic characteristics come from the SECC data. Data on the number of schools, power supply, and land use come from the 2011 census village directory. Data on night lights come from the Earth Observation Group.

**Table C2.** Balance table Comparing GPs by Identity of Elected Chairperson in 2022

| | (1) Upper-caste Men | | (2) Upper-caste Women | | (3) SC/ST Men | | (4) SC/ST Women | | Difference in Means | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | N | Mean | N | Mean | N | Mean | N | Mean | N | (1)-(2) | N | (1)-(3) | N | (1)-(4) |
| Total Population | 820 | 4887.849 | 674 | 5050.764 | 599 | 5241.701 | 480 | 5203.119 | 1494 | -162.915 | 1419 | -353.852*** | 1300 | -315.270** |
| Women (% of GP Population) | 820 | 0.523 | 674 | 0.522 | 599 | 0.535 | 480 | 0.536 | 1494 | 0.000 | 1419 | -0.013 | 1300 | -0.014 |
| SC (% of GP Population) | 820 | 0.183 | 674 | 0.178 | 599 | 0.161 | 480 | 0.157 | 1494 | 0.005 | 1419 | 0.022*** | 1300 | 0.027*** |
| ST (% of GP Population) | 820 | 0.107 | 674 | 0.111 | 599 | 0.372 | 480 | 0.389 | 1494 | -0.004 | 1419 | -0.266** | 1300 | -0.283** |
| Employed (% of GP Population) | 820 | 0.340 | 674 | 0.342 | 599 | 0.353 | 480 | 0.354 | 1494 | -0.002* | 1419 | -0.013 | 1300 | -0.014 |
| Unemployed (% of GP Population) | 820 | 0.117 | 674 | 0.112 | 599 | 0.126 | 480 | 0.124 | 1494 | 0.005 | 1419 | -0.010 | 1300 | -0.008 |
| Student (% of GP Population) | 820 | 0.068 | 674 | 0.069 | 599 | 0.063 | 480 | 0.064 | 1494 | -0.001 | 1419 | 0.005 | 1300 | 0.004 |
| Illiterate (% of GP Population) | 820 | 0.287 | 674 | 0.275 | 599 | 0.398 | 480 | 0.404 | 1494 | 0.012 | 1419 | -0.111 | 1300 | -0.117 |
| Average Years of Education in GP | 820 | 5.210 | 674 | 5.262 | 599 | 4.365 | 480 | 4.355 | 1494 | -0.051 | 1419 | 0.846 | 1300 | 0.855 |
| Average Age in GP | 820 | 38.840 | 674 | 38.863 | 599 | 38.331 | 480 | 38.387 | 1494 | -0.022 | 1419 | 0.509 | 1300 | 0.453 |
| Earns Less than 5k INR (% of GP Population) | 820 | 0.837 | 674 | 0.841 | 599 | 0.876 | 480 | 0.881 | 1494 | -0.004 | 1419 | -0.039 | 1300 | -0.044 |
| Earns 5k-10k INR (% of GP Population) | 820 | 0.096 | 674 | 0.093 | 599 | 0.073 | 480 | 0.068 | 1494 | 0.003 | 1419 | 0.023 | 1300 | 0.028 |
| Earns More than 5k INR (% of GP Population) | 820 | 0.067 | 674 | 0.066 | 599 | 0.051 | 480 | 0.051 | 1494 | 0.001 | 1419 | 0.016 | 1300 | 0.016 |
| Number of Primary Schools | 807 | 1.830 | 660 | 1.818 | 587 | 1.705 | 468 | 1.741 | 1467 | 0.012 | 1394 | 0.125 | 1275 | 0.089 |
| Number of Middle Schools | 807 | 0.957 | 660 | 0.983 | 587 | 0.864 | 468 | 0.861 | 1467 | -0.027 | 1394 | 0.093 | 1275 | 0.096 |
| Number of Secondary Schools | 807 | 0.564 | 660 | 0.621 | 587 | 0.470 | 468 | 0.502 | 1467 | -0.057* | 1394 | 0.094 | 1275 | 0.062 |
| GP Has Domestic Power Supply | 807 | 0.944 | 660 | 0.947 | 587 | 0.871 | 468 | 0.857 | 1467 | -0.003 | 1394 | 0.074 | 1275 | 0.087 |
| Share of Land used for Agriculture | 805 | 0.623 | 659 | 0.615 | 586 | 0.521 | 468 | 0.522 | 1464 | 0.008 | 1391 | 0.102 | 1273 | 0.101 |
| Mean Night Lights (Calibrated) in 2013 | 814 | 5.766 | 670 | 5.748 | 592 | 4.445 | 476 | 4.620 | 1484 | 0.018 | 1406 | 1.321 | 1290 | 1.145 |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are used when calculating the significance of the difference in means. Block fixed effects are included in the difference in means specification. The table compares the means of observable covariates across GPs based on the identity of the elected chairperson in 2022. Data on total population, population shares, and average socioeconomic characteristics come from the SECC data. Data on the number of schools, power supply, and land use come from the 2011 census village directory. Data on night lights come from the Earth Observation Group.

**Table C3.** Balance table Comparing GPs by Reservation Status in 2022

| | (1) Open Seats | | (2) Open Women | | (3) SC/ST Open | | (4) SC/ST Women | | Difference in Means | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | N | Mean | N | Mean | N | Mean | N | Mean | N | (1)-(2) | N | (1)-(3) | N | (1)-(4) |
| Total Population | 887 | 4924.362 | 612 | 5040.750 | 591 | 5290.580 | 413 | 5294.562 | 1499 | -116.388 | 1478 | -366.218*** | 1300 | -370.200*** |
| Women (% of GP Population) | 887 | 0.521 | 612 | 0.521 | 591 | 0.537 | 413 | 0.537 | 1499 | -0.000 | 1478 | -0.016 | 1300 | -0.016 |
| SC (% of GP Population) | 887 | 0.190 | 612 | 0.187 | 591 | 0.151 | 413 | 0.146 | 1499 | 0.003 | 1478 | 0.038*** | 1300 | 0.043** |
| ST (% of GP Population) | 887 | 0.113 | 612 | 0.113 | 591 | 0.366 | 413 | 0.386 | 1499 | -0.000 | 1478 | -0.254 | 1300 | -0.273 |
| Employed (% of GP Population) | 887 | 0.336 | 612 | 0.342 | 591 | 0.355 | 413 | 0.354 | 1499 | -0.007*** | 1478 | -0.020 | 1300 | -0.018 |
| Unemployed (% of GP Population) | 887 | 0.120 | 612 | 0.110 | 591 | 0.127 | 413 | 0.124 | 1499 | 0.010*** | 1478 | -0.007 | 1300 | -0.004 |
| Student (% of GP Population) | 887 | 0.069 | 612 | 0.069 | 591 | 0.063 | 413 | 0.063 | 1499 | -0.000 | 1478 | 0.005 | 1300 | 0.005 |
| Illiterate (% of GP Population) | 887 | 0.274 | 612 | 0.269 | 591 | 0.402 | 413 | 0.406 | 1499 | 0.006 | 1478 | -0.127 | 1300 | -0.131 |
| Average Years of Education in GP | 887 | 5.287 | 612 | 5.305 | 591 | 4.360 | 413 | 4.370 | 1499 | -0.018 | 1478 | 0.927 | 1300 | 0.917 |
| Average Age in GP | 887 | 38.913 | 612 | 38.878 | 591 | 38.192 | 413 | 38.280 | 1499 | 0.035 | 1478 | 0.721 | 1300 | 0.633* |
| Earns Less than 5k INR (% of GP Population) | 887 | 0.836 | 612 | 0.838 | 591 | 0.873 | 413 | 0.881 | 1499 | -0.002 | 1478 | -0.037 | 1300 | -0.044 |
| Earnes 5k-10k INR (% of GP Population) | 887 | 0.096 | 612 | 0.094 | 591 | 0.075 | 413 | 0.069 | 1499 | 0.002 | 1478 | 0.021 | 1300 | 0.027 |
| Earns More than 5k INR (% of GP Population) | 887 | 0.068 | 612 | 0.067 | 591 | 0.052 | 413 | 0.050 | 1499 | 0.000 | 1478 | 0.015 | 1300 | 0.017 |
| Number of Primary Schools | 872 | 1.839 | 598 | 1.833 | 580 | 1.700 | 402 | 1.704 | 1470 | 0.007 | 1452 | 0.139 | 1274 | 0.135 |
| Number of Middle Schools | 872 | 0.976 | 598 | 1.010 | 580 | 0.840 | 402 | 0.813 | 1470 | -0.034 | 1452 | 0.136 | 1274 | 0.162 |
| Number of Secondary Schools | 872 | 0.596 | 598 | 0.647 | 580 | 0.445 | 402 | 0.465 | 1470 | -0.051** | 1452 | 0.152 | 1274 | 0.131 |
| GP Has Domestic Power Supply | 872 | 0.959 | 598 | 0.953 | 580 | 0.860 | 402 | 0.828 | 1470 | 0.006 | 1452 | 0.098 | 1274 | 0.130 |
| Share of Land used for Agriculture | 872 | 0.626 | 597 | 0.617 | 577 | 0.514 | 402 | 0.513 | 1469 | 0.009 | 1449 | 0.112 | 1274 | 0.112 |
| Mean Night Lights (Calibrated) in 2013 | 881 | 5.828 | 606 | 5.912 | 585 | 4.479 | 410 | 4.480 | 1487 | -0.084 | 1466 | 1.348 | 1291 | 1.348** |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are used when calculating the significance of the difference in means. Block fixed effects are included in the difference in means specification. The table compares the means of observable covariates across GPs based on reservation status in 2022. Data on total population, population shares, and average socioeconomic characteristics come from the SECC data. Data on the number of schools, power supply, and land use come from the 2011 census village directory. Data on night lights come from the Earth Observation Group.

# D Robustness

## D.1 Tables with Control Coefficients Reported

**Table D1.** Replication of Table 2 reporting Control Covariates

| **Dependent Variable:** | Years of Education | | | | | | Mincer Residual | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Sample:** | Odisha Census | | | | All India Census | | Odisha Census | | | |
| **Group Measure:** | Identity Panel | Identity 2017 | Identity 2022 | Reservation 2022 | Identity | Reservation | Identity Panel | Identity 2017 | Identity 2022 | Reservation 2022 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| UC Women/Open Women | -2.43*** | -2.72*** | -3.92*** | -1.08*** | -1.81 | -1.55* | 0.04* | -0.01 | 0.01 | -0.01 |
| | (0.34) | (0.72) | (0.82) | (0.39) | (1.21) | (0.92) | (0.02) | (0.06) | (0.06) | (0.03) |
| SC/ST Men/Open SC/ST | -1.60*** | -1.21*** | -1.02*** | -1.01*** | -1.97 | -2.28** | -0.01 | 0.00 | 0.04* | 0.03 |
| | (0.47) | (0.32) | (0.31) | (0.31) | (1.22) | (1.07) | (0.03) | (0.02) | (0.02) | (0.02) |
| SC/ST Women/SC/ST Women | -3.73*** | -4.13*** | -5.27*** | -3.30*** | -3.53*** | -2.86* | 0.05 | -0.01 | 0.04 | 0.02 |
| | (0.44) | (0.66) | (0.78) | (0.43) | (1.19) | (1.64) | (0.03) | (0.05) | (0.05) | (0.03) |
| SC Population Share | | 2.18 | 2.40 | 2.72 | -1.91 | -4.36 | | -0.02 | -0.08 | -0.12 |
| | | (2.64) | (2.87) | (2.96) | (5.48) | (5.51) | | (0.19) | (0.20) | (0.20) |
| SC Population Share Squared | | -0.98 | -9.13* | -10.03* | 8.88 | 10.02 | | -0.22 | 0.09 | 0.16 |
| | | (4.61) | (5.08) | (5.22) | (6.02) | (6.36) | | (0.32) | (0.34) | (0.35) |
| ST Population Share | | -0.85 | -0.85 | -1.23 | 15.93** | 12.94** | | 0.21* | 0.26* | 0.27** |
| | | (1.74) | (1.91) | (1.96) | (7.23) | (5.80) | | (0.12) | (0.13) | (0.14) |
| ST Population Share Squared | | -2.07 | 0.03 | 0.26 | -14.64* | -11.34* | | -0.26** | -0.41*** | -0.42*** |
| | | (1.90) | (2.14) | (2.19) | (8.17) | (6.10) | | (0.13) | (0.13) | (0.14) |
| Total GP Population | | 0.00* | 0.00 | 0.00 | | | | 0.00 | -0.00 | -0.00 |
| | | (0.00) | (0.00) | (0.00) | | | | (0.00) | (0.00) | (0.00) |
| Share of Population Employed in Agriculture | | -1.15 | -3.24** | 0.71 | | | | -0.11 | -0.05 | -0.06 |
| | | (1.17) | (1.41) | (0.89) | | | | (0.09) | (0.10) | (0.06) |
| Share of Population Employed in Manual Work | | -2.06** | -2.98*** | 0.37 | | | | -0.06 | -0.11 | -0.11** |
| | | (0.99) | (1.12) | (0.62) | | | | (0.08) | (0.08) | (0.04) |
| Share of Population Employed in Service Work | | 1.53 | 10.40** | 15.80*** | | | | 0.30 | 0.02 | -0.08 |
| | | (3.73) | (4.29) | (4.44) | | | | (0.35) | (0.32) | (0.30) |
| Share of Population Employed in Business | | 1.10 | -9.92** | -2.91 | | | | -0.08 | 0.10 | 0.02 |
| | | (3.61) | (3.87) | (3.56) | | | | (0.27) | (0.27) | (0.25) |
| Share of Population Not-Employed | | 0.45 | -4.84* | -1.88 | | | | -0.08 | 0.13 | 0.17 |
| | | (2.10) | (2.64) | (2.50) | | | | (0.15) | (0.18) | (0.17) |
| E(Y \| UC Men/Open) | 10.87 | 10.61 | 10.72 | 10.71 | 11.02 | 10.48 | 0.07 | 0.08 | 0.07 | 0.08 |
| # Observations | 2,522 | 2,932 | 2,583 | 2,498 | 126 | 157 | 2,522 | 2,932 | 2,583 | 2,498 |
| # GPs | 1,261 | 2,932 | 2,583 | 2,498 | 126 | 157 | 1,261 | 2,932 | 2,583 | 2,498 |
| Fixed Effects | GP | Block | Block | Block | State | State | GP | Block | Block | Block |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are reported in parentheses. The table reports the relationship between politician identity/GP reservation status and education. The sample for each column includes only politicians. Models where the group measure is identity report the relationship between politician identity and education. Models where the group measure is reservation report the relationship between GP reservation status and politician education. For the models using the Odisha Census data, the sample includes the set of GPs where politicians were able to be matched. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects and controls included as specified.

## D.2 Identity Rotation in the Odisha Panel Sample

**Table D2.** Politician Identity Rotation in the Odisha Panel Sample

| | | **2022 Politician Identity** | | | | **Total** |
|---|---|---|---|---|---|---|
| | | UC Men | UC Women | SC/ST Men | SC/ST Women | |
| **2017 Politician Identity** | UC Men | 31 | 192 | 24 | 86 | 333 |
| | UC Women | 198 | 59 | 80 | 21 | 358 |
| | SC/ST Men | 15 | 107 | 12 | 129 | 263 |
| | SC/ST Women | 114 | 32 | 116 | 45 | 307 |
| **Total** | | 358 | 390 | 232 | 281 | 1,261 |

*Notes*: The table shows the number of GPs by the identity of their elected chairperson in 2017 and 2022. These data comprise the sample of the Odisha panel models.

## D.3 Robustness to Excluding Population Squares

**Table D3.** Replication of Table 2 without Controlling for SC and ST Population Shares Squared

| **Dependent Variable:** | Years of Education | | | | | Mincer Residual | | |
|---|---|---|---|---|---|---|---|---|
| **Sample:** | Odisha Census | | | All India Census | | Odisha Census | | |
| **Group Measure:** | Identity 2017 | Identity 2022 | Reservation 2022 | Identity | Reservation | Identity 2017 | Identity 2022 | Reservation 2022 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| UC Women/Open Women | -2.74*** | -3.91*** | -1.07*** | -1.63 | -1.64* | -0.01 | 0.01 | -0.01 |
| | (0.72) | (0.82) | (0.39) | (1.23) | (0.89) | (0.06) | (0.06) | (0.03) |
| SC/ST Men/Open SC/ST | -1.22*** | -1.00*** | -0.99*** | -2.32* | -2.26** | 0.00 | 0.04* | 0.02 |
| | (0.32) | (0.31) | (0.31) | (1.21) | (1.06) | (0.02) | (0.02) | (0.02) |
| SC/ST Women/SC/ST Women | -4.15*** | -5.25*** | -3.27*** | -3.47*** | -3.20** | -0.01 | 0.04 | 0.02 |
| | (0.66) | (0.79) | (0.43) | (1.19) | (1.53) | (0.05) | (0.06) | (0.03) |
| E(Y \| UC Men/Open) | 10.61 | 10.72 | 10.71 | 11.02 | 10.48 | 0.08 | 0.07 | 0.08 |
| # Observations | 2,932 | 2,583 | 2,498 | 126 | 157 | 2,932 | 2,583 | 2,498 |
| # GPs | 2,932 | 2,583 | 2,498 | 126 | 157 | 2,932 | 2,583 | 2,498 |
| Fixed Effects | Block | Block | Block | State | State | Block | Block | Block |
| Population Controls | X | X | X | X | X | X | X | X |
| Occupation Controls | X | X | X | | | X | X | X |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are reported in parentheses. The table reports the relationship between politician identity/GP reservation status and education. The sample for each column includes only politicians. Models where the group measure is identity report the relationship between politician identity and education. Models where the group measure is reservation report the relationship between GP reservation status and politician education. For the models using the Odisha Census data, the sample includes the set of GPs where politicians were able to be matched. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects and controls included as specified. Population controls include SC and ST population shares and the total population size. Occupation controls include the share of the population from the politicians' identity group in agriculture, business, service, manual work, and others.

## D.4 Robustness to Interacting Block Fixed Effects and SC and ST Population Shares

**Table D4.** Replication of Table 2 Interacting Block Fixed Effects and SC and ST Population Shares

| **Dependent Variable:** | Years of Education | | | | | Mincer Residual | | |
|---|---|---|---|---|---|---|---|---|
| **Sample:** | Odisha Census | | | All India Census | | Odisha Census | | |
| **Group Measure:** | Identity 2017 | Identity 2022 | Reservation 2022 | Identity | Reservation | Identity 2017 | Identity 2022 | Reservation 2022 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| UC Women/Open Women | -2.04** | -3.66*** | -1.24** | -1.74 | -1.89* | 0.00 | 0.07 | 0.01 |
| | (0.97) | (1.15) | (0.55) | (1.48) | (1.08) | (0.07) | (0.08) | (0.04) |
| SC/ST Men/Open SC/ST | -1.01** | -0.79* | -1.20*** | -3.36** | -3.63*** | 0.00 | 0.04 | 0.04 |
| | (0.40) | (0.43) | (0.44) | (1.36) | (1.33) | (0.03) | (0.03) | (0.03) |
| SC/ST Women/SC/ST Women | -3.61*** | -5.13*** | -3.64*** | -3.78*** | -3.73** | -0.02 | 0.09 | 0.03 |
| | (0.87) | (1.11) | (0.61) | (1.43) | (1.79) | (0.06) | (0.08) | (0.04) |
| E(Y \| UC Men/Open) | 10.61 | 10.72 | 10.71 | 11.02 | 10.48 | 0.08 | 0.07 | 0.08 |
| # Observations | 2,932 | 2,583 | 2,498 | 126 | 157 | 2,932 | 2,583 | 2,498 |
| # GPs | 2,932 | 2,583 | 2,498 | 126 | 157 | 2,932 | 2,583 | 2,498 |
| Fixed Effects | Block × SC/ST Pop. | Block × SC/ST Pop. | Block × SC/ST Pop. | State × SC/ST Pop. | State × SC/ST Pop. | Block × SC/ST Pop. | Block × SC/ST Pop. | Block × SC/ST Pop. |
| Population Controls | X | X | X | X | X | X | X | X |
| Occupation Controls | X | X | X | | | X | X | X |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are reported in parentheses. The table reports the relationship between politician identity/GP reservation status and education. The sample for each column includes only politicians. Models where the group measure is identity report the relationship between politician identity and education. Models where the group measure is reservation report the relationship between GP reservation status and politician education. For the models using the Odisha Census data, the sample includes the set of GPs where politicians were able to be matched. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects and controls included as specified. Population controls include Block × SC population share and Block × ST population share and the total population size. Occupation controls include the share of the population from the politicians' identity group in agriculture, business, service, manual work, and others.

## D.5 Robustness to the Inclusion of Politicians under 21 at the time of SECC data collection

**Table D5.** Replication of Table 2 without Subsetting to Politicians who were 22 at the time of data collection

| **Dependent Variable:** | Years of Education | | | | Mincer Residual | | | |
|---|---|---|---|---|---|---|---|---|
| **Sample:** | Odisha Census | | | | Odisha Census | | | |
| **Group Measure:** | Identity Panel (1) | Identity 2017 (2) | Identity 2022 (3) | Reservation 2022 (4) | Identity Panel (5) | Identity 2017 (6) | Identity 2022 (7) | Reservation 2022 (8) |
| UC Women/Open Women | -2.32*** | -2.57*** | -3.64*** | -1.07*** | 0.03 | 0.00 | 0.04 | 0.01 |
| | (0.30) | (0.66) | (0.68) | (0.34) | (0.02) | (0.05) | (0.05) | (0.03) |
| SC/ST Men/Open SC/ST | -1.44*** | -1.10*** | -0.89*** | -0.96*** | 0.01 | 0.01 | 0.05** | 0.02 |
| | (0.41) | (0.30) | (0.28) | (0.28) | (0.03) | (0.02) | (0.02) | (0.02) |
| SC/ST Women/SC/ST Women | -3.51*** | -3.89*** | -4.79*** | -2.87*** | 0.04 | -0.01 | 0.07 | 0.03 |
| | (0.39) | (0.60) | (0.65) | (0.36) | (0.03) | (0.04) | (0.05) | (0.03) |
| E(Y \| UC Men/Open) | 10.73 | 10.56 | 10.52 | 10.58 | 0.07 | 0.08 | 0.07 | 0.08 |
| # Observations | 3,202 | 3,213 | 3,067 | 2,970 | 3,202 | 3,213 | 3,067 | 2,970 |
| # GPs | 1,603 | 3,213 | 3,067 | 2,970 | 1,603 | 3,213 | 3,067 | 2,970 |
| Fixed Effects | GP | Block | Block | Block | GP | Block | Block | Block |
| Population Controls | | X | X | X | | X | X | X |
| Occupation Controls | | X | X | X | | X | X | X |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are reported in parentheses. The table reports the relationship between politician identity/GP reservation status and education. The sample for each column includes only politicians. The sample includes all matched politicians, not just those aged 22 at the point of education data collection. Models where the group measure is identity report the relationship between politician identity and education. Models where the group measure is reservation report the relationship between GP reservation status and politician education. For the models using the Odisha Census data, the sample includes the set of GPs where politicians were able to be matched. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects and controls included as specified. Population controls include Block × SC population share and Block × ST population share and the total population size. Occupation controls include the share of the population from the politicians' identity group in agriculture, business, service, manual work, and others.

**Table D6.** Replication of Table 3 without Subsetting to Politicians who were 22 at the time of data collection

| **Dependent Variable:** | Sarpanch ($\times$ 10,000) | | | |
|---|---|---|---|---|
| **Sample:** | Odisha Census | | | |
| **Group Measure:** | Identity Panel | Identity 2017 | Identity 2022 | Reservation 2022 |
| | (1) | (2) | (3) | (4) |
| Education | 1.38*** | 1.64*** | 1.71*** | 0.95*** |
| | (0.10) | (0.06) | (0.06) | (0.03) |
| Education $\times$ UC Women/Open Women | 0.24 | 0.10 | -0.11 | 0.37*** |
| | (0.17) | (0.09) | (0.09) | (0.06) |
| Education $\times$ SC/ST Men/Open SC/ST | 1.47*** | 1.56*** | 1.08*** | 0.97*** |
| | (0.23) | (0.14) | (0.12) | (0.07) |
| Education $\times$ SC/ST Women/SC/ST Women | 2.00*** | 2.61*** | 1.02*** | 1.68*** |
| | (0.31) | (0.18) | (0.14) | (0.14) |
| E(Y \| UC Men/Open) | 1.01 | 1.56 | 1.38 | 0.79 |
| # Observations | 1156358 | 3640890 | 3969716 | 7526382 |
| # GPs | 835 | 3,214 | 3,071 | 2,976 |
| Fixed Effects | GP | GP | GP | GP |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP clustered standard errors are reported in parentheses. The table reports the relationship between identity/eligibility for a GP's reservation status, education, and selection as the Sarpanch. The sample includes all elected politicians and the population in their GP from their identity group (identity models) or who were eligible for the reservation (reservation models). The sample includes all matched politicians, not just those aged 22 at the point of education data collection. The dependent variable equals 10,000 for politicians and 0 for everyone else to aid interpretation. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects included as specified.

## D.6 Robustness to only including Non-Scheduled Areas

**Table D7.** Replication of Table 2 only for Non-Scheduled Areas

| **Dependent Variable:** | Years of Education | | | | Mincer Residual | | | |
|---|---|---|---|---|---|---|---|---|
| **Sample:** | Odisha Census | | | | | | | |
| **Group Measure:** | Identity Panel (1) | Identity 2017 (2) | Identity 2022 (3) | Reservation 2022 (4) | Identity Panel (5) | Identity 2017 (6) | Identity 2022 (7) | Reservation 2022 (8) |
| UC Women/Open Women | -2.34*** | -2.36*** | -3.42*** | -0.81** | 0.04 | -0.00 | 0.00 | -0.01 |
| | (0.35) | (0.82) | (0.94) | (0.40) | (0.02) | (0.06) | (0.06) | (0.03) |
| SC/ST Men/Open SC/ST | -1.43*** | -1.39*** | -0.85** | -0.93*** | -0.01 | -0.01 | 0.04 | 0.04* |
| | (0.53) | (0.34) | (0.34) | (0.34) | (0.03) | (0.02) | (0.03) | (0.02) |
| SC/ST Women/SC/ST Women | -3.91*** | -4.17*** | -4.79*** | -2.98*** | 0.03 | -0.00 | 0.04 | 0.01 |
| | (0.47) | (0.76) | (0.90) | (0.45) | (0.03) | (0.06) | (0.06) | (0.04) |
| E(Y \| UC Men/Open) | 10.95 | 10.75 | 10.83 | 10.71 | 0.07 | 0.08 | 0.07 | 0.08 |
| # Observations | 2,092 | 2,327 | 2,072 | 1,994 | 2,092 | 2,327 | 2,072 | 1,994 |
| # GPs | 1,046 | 2,327 | 2,072 | 1,994 | 1,046 | 2,327 | 2,072 | 1,994 |
| Fixed Effects | GP | Block | Block | Block | GP | Block | Block | Block |
| Population Controls | | X | X | X | | X | X | X |
| Occupation Controls | | X | X | X | | X | X | X |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP Clustered standard errors are reported in parentheses. The table reports the relationship between politician identity/GP reservation status and education. The sample for each column includes only politicians. Models where the group measure is identity report the relationship between politician identity and education. Models where the group measure is reservation report the relationship between GP reservation status and politician education. For the models using the Odisha Census data, the sample includes the set of GPs where politicians were able to be matched. The identity panel models in columns (4) and (10) include the 1,261 GPs where we match politicians in both 2017 and 2022, allowing for a within-GP comparison. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects and controls included as specified. Population controls include SC and ST population shares and their squared values and the total population size. Occupation controls include the share of the population from the politicians' identity group in agriculture, business, service, manual work, and others.

**Table D8.** Replication of Table 3 only for Non-Scheduled Areas

| **Dependent Variable:** | Sarpanch (× 10,000) | | | |
|---|---|---|---|---|
| **Sample:** | Odisha Census | | | |
| **Group Measure:** | Identity Panel (1) | Identity 2017 (2) | Identity 2022 (3) | Reservation 2022 (4) |
| Education | 1.78*** | 1.94*** | 2.23*** | 1.29*** |
| | (0.11) | (0.07) | (0.08) | (0.04) |
| Education × UC Women/Open Women | 0.57*** | 0.32*** | 0.21* | 0.68*** |
| | (0.20) | (0.12) | (0.13) | (0.09) |
| Education × SC/ST Men/Open SC/ST | 3.32*** | 3.43*** | 3.07*** | 2.24*** |
| | (0.38) | (0.27) | (0.26) | (0.18) |
| Education × SC/ST Women/SC/ST Women | 6.62*** | 5.82*** | 4.81*** | 4.13*** |
| | (0.87) | (0.41) | (0.47) | (0.43) |
| E(Y \| UC Men/Open) | 1.48 | 1.41 | 0.92 | 0.83 |
| # Observations | 833,302 | 2319206 | 2048351 | 4447753 |
| # GPs | 766 | 2,331 | 2,079 | 2,174 |
| Fixed Effects | GP | GP | GP | GP |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP clustered standard errors are reported in parentheses. The table reports the relationship between identity/eligibility for a GP's reservation status, education, and selection as the Sarpanch. The sample includes all elected politicians and the population in their GP from their identity group (identity models) or who were eligible for the reservation (reservation models) from GPs in non-Scheduled Areas. The dependent variable equals 10,000 for politicians and 0 for everyone else to aid interpretation. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects included as specified.

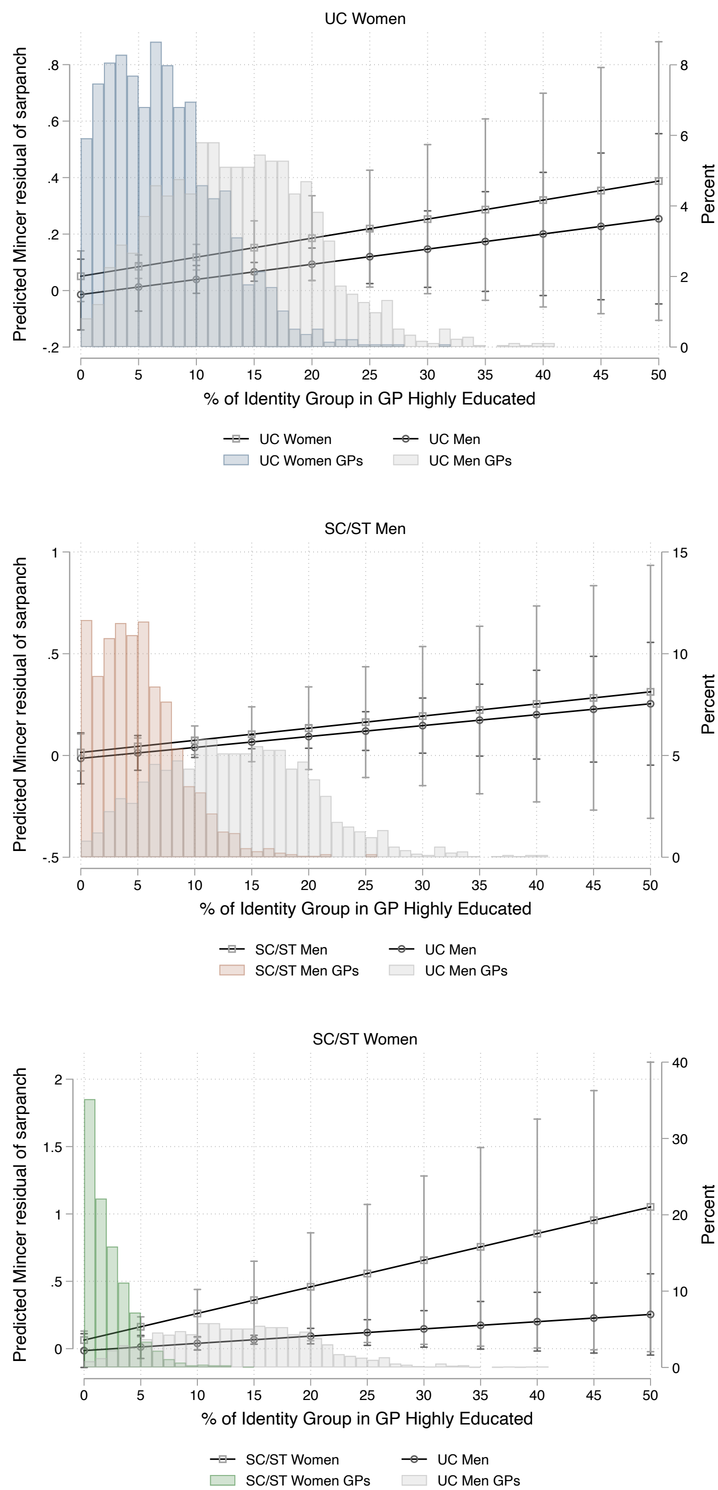
## D.7 Additional Structural Discrimination Analysis

**Table D9.** Table Corresponding to Figure 3

| **Dependent Variable:** | Years of Education | | | | | | Mincer Residual | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Sample:** | Odisha Census | | | | All India Census | | Odisha Census | | | |
| **Group Measure:** | Identity Panel | Identity 2017 | Identity 2022 | Reservation 2022 | Identity | Reservation | Identity Panel | Identity 2017 | Identity 2022 | Reservation 2022 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| | (0.06) | (0.03) | (0.03) | (0.03) | (0.04) | (0.04) | (0.00) | (0.00) | (0.00) | (0.00) |
| UC Women/Open Women | -2.57*** | -1.50* | -3.16*** | -1.71*** | -2.11 | -0.22 | 0.06 | 0.01 | 0.02 | -0.01 |
| | (0.69) | (0.81) | (0.90) | (0.52) | (2.17) | (2.33) | (0.05) | (0.06) | (0.07) | (0.03) |
| SC/ST Men/Open SC/ST | -0.61 | -1.16** | -1.80*** | -1.14** | -1.06 | -1.14 | 0.03 | -0.00 | 0.07** | 0.03 |
| | (0.99) | (0.46) | (0.57) | (0.53) | (1.98) | (1.99) | (0.08) | (0.03) | (0.04) | (0.04) |
| SC/ST Women/SC/ST Women | -2.64*** | -2.82*** | -4.28*** | -3.14*** | -0.89 | 1.17 | 0.08 | 0.01 | 0.05 | 0.00 |
| | (0.91) | (0.75) | (0.86) | (0.53) | (1.92) | (2.34) | (0.07) | (0.06) | (0.06) | (0.04) |
| UC Women/Open Women × % Highly Educated | 0.15*** | -0.01 | 0.08* | 0.11*** | 0.17 | 0.11 | 0.00 | -0.00 | -0.00 | 0.00 |
| | (0.06) | (0.05) | (0.04) | (0.04) | (0.13) | (0.16) | (0.00) | (0.00) | (0.00) | (0.00) |
| SC/ST Men/Open SC/ST × % Highly Educated | 0.05 | 0.13*** | 0.22*** | 0.09* | 0.02 | 0.04 | 0.00 | 0.00 | -0.00 | -0.00 |
| | (0.10) | (0.05) | (0.06) | (0.05) | (0.06) | (0.09) | (0.01) | (0.00) | (0.00) | (0.00) |
| SC/ST Women/SC/ST Women × % Highly Educated | 0.28 | 0.24** | 0.28*** | 0.13* | -0.09 | -0.18 | 0.01 | 0.01 | 0.01 | 0.01 |
| | (0.18) | (0.11) | (0.10) | (0.08) | (0.20) | (0.27) | (0.01) | (0.01) | (0.01) | (0.00) |
| SC Population Share | | 0.90 | 1.38 | 1.62 | -6.66 | -7.83 | | -0.04 | -0.09 | -0.12 |
| | | (2.61) | (2.87) | (2.91) | (5.31) | (6.71) | | (0.19) | (0.19) | (0.20) |
| SC Population Share Squared | | 0.15 | -7.56 | -8.29 | 14.20** | 15.28** | | -0.20 | 0.09 | 0.17 |
| | | (4.55) | (5.14) | (5.19) | (5.91) | (7.21) | | (0.32) | (0.34) | (0.35) |
| ST Population Share | | -0.35 | -0.49 | -0.90 | 16.24** | 16.53** | | 0.22* | 0.26* | 0.27** |
| | | (1.72) | (1.87) | (1.92) | (7.36) | (7.42) | | (0.12) | (0.13) | (0.14) |
| ST Population Share Squared | | -2.09 | 0.14 | 0.34 | -16.81** | -17.14** | | -0.26** | -0.40*** | -0.41*** |
| | | (1.89) | (2.09) | (2.14) | (8.06) | (8.20) | | (0.12) | (0.13) | (0.14) |
| Total GP Population | | 0.00 | 0.00 | 0.00 | | | | 0.00 | -0.00 | -0.00 |
| | | (0.00) | (0.00) | (0.00) | | | | (0.00) | (0.00) | (0.00) |
| Share of Population Employed in Agriculture | | 0.35 | -1.30 | 0.52 | | | | -0.07 | -0.04 | -0.06 |
| | | (1.18) | (1.41) | (0.87) | | | | (0.09) | (0.10) | (0.06) |
| Share of Population Employed in Manual Work | | -0.60 | -1.31 | 0.69 | | | | -0.03 | -0.09 | -0.11** |
| | | (0.99) | (1.12) | (0.61) | | | | (0.08) | (0.09) | (0.04) |
| Share of Population Employed in Service Work | | -7.16* | 2.64 | 4.38 | | | | 0.12 | -0.05 | -0.13 |
| | | (3.94) | (4.28) | (4.25) | | | | (0.38) | (0.33) | (0.33) |
| Share of Population Employed in Business | | -4.89 | -12.09*** | -8.52** | | | | -0.19 | 0.05 | -0.00 |
| | | (3.74) | (3.90) | (3.68) | | | | (0.27) | (0.28) | (0.26) |
| Share of Population Not-Employed | | 0.07 | -4.27* | -2.36 | | | | -0.09 | 0.15 | 0.17 |
| | | (2.08) | (2.57) | (2.46) | | | | (0.15) | (0.18) | (0.17) |
| # Observations | 2,522 | 2,932 | 2,583 | 2,498 | 126 | 121 | 2,522 | 2,932 | 2,583 | 2,498 |
| # GPs | 1,261 | 2,932 | 2,583 | 2,498 | 126 | 121 | 1,261 | 2,932 | 2,583 | 2,498 |
| Fixed Effects | GP | Block | Block | Block | State | State | GP | Block | Block | Block |

*Notes*: $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01. Standard errors clustered at the GP level in parentheses. The sample includes only politicians. Covariates include SC and ST population shares and politicians' group occupation shares in agriculture, business, service, manual work, and others.

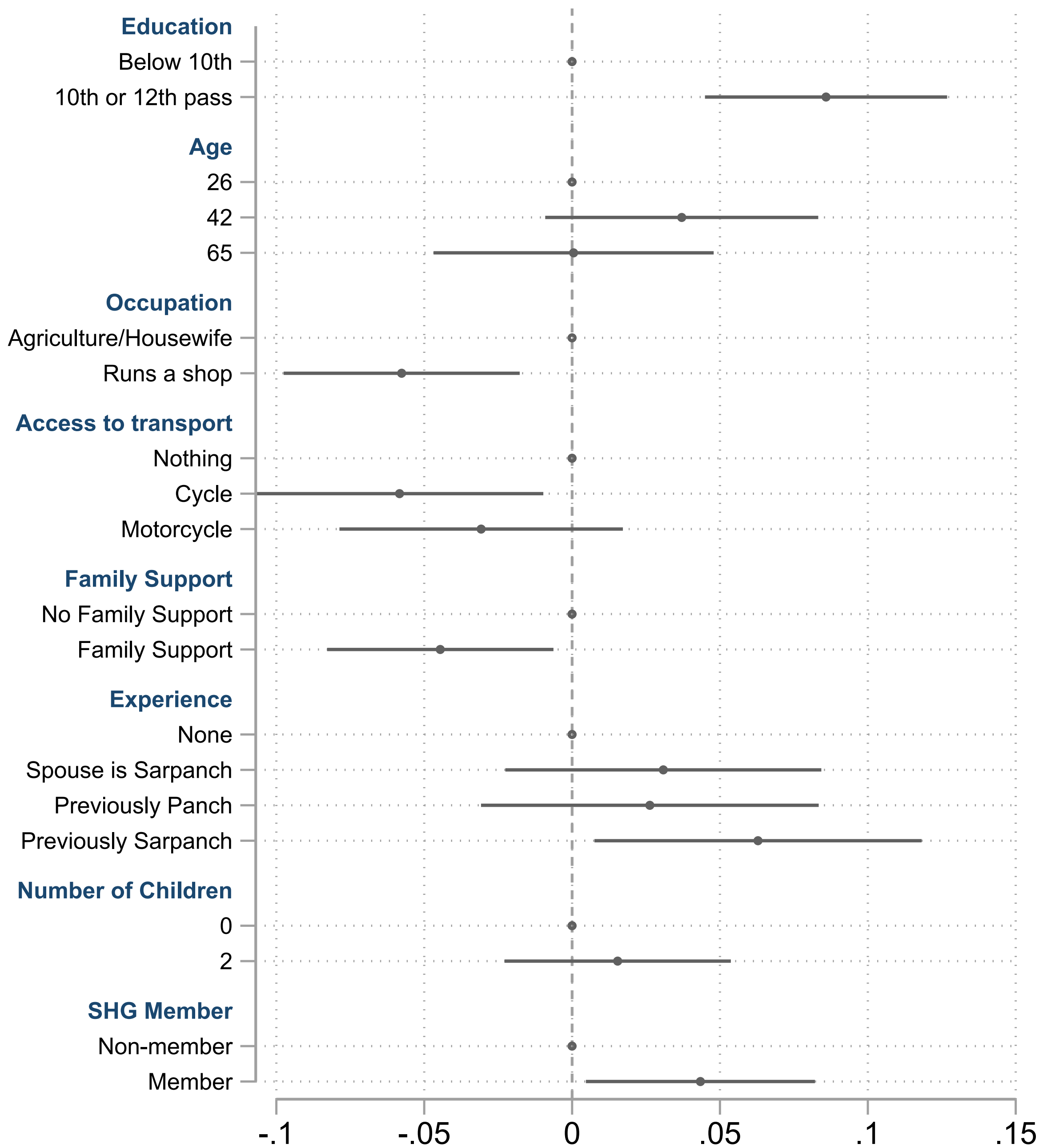
**Figure D1.** Replication of Figure 3 using the Mincer Residual




*Notes:* The figure reports the predicted values of politician Mincer residuals from a model where politician identity is interacted with the share of the identity group in the GP that has completed 12th grade education following specification 1. 95% confidence intervals from based on standard errors clustered at the GP level. The results use the sample from the Odisha identity panel model with GP fixed effects (see column 1 in Appendix Table D9). Additional covariates include SC and ST population shares (and their squares) and politicians' group occupation shares in agriculture, business, service, manual work, and others.

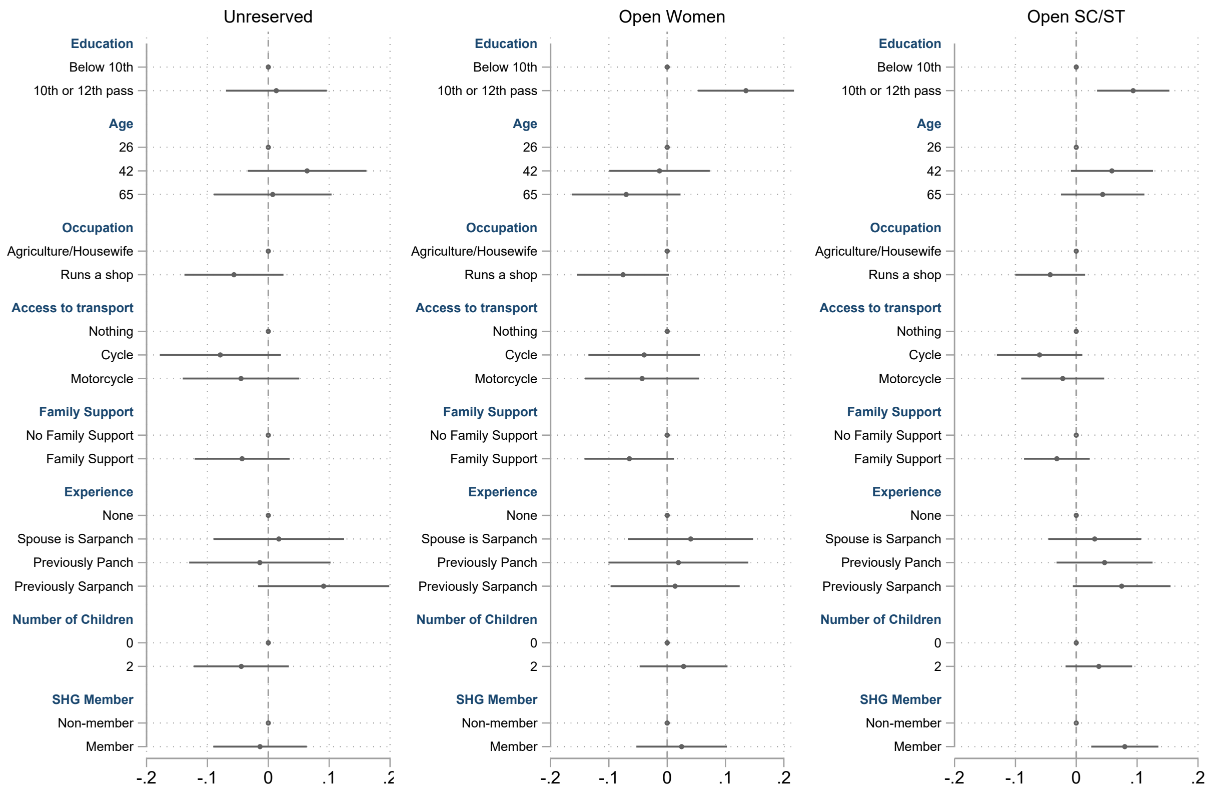
## D.8 Robustness of Conjoint Analysis

**Figure D2.** AMCEs for All Attributes in Citizen Conjoint Experiment



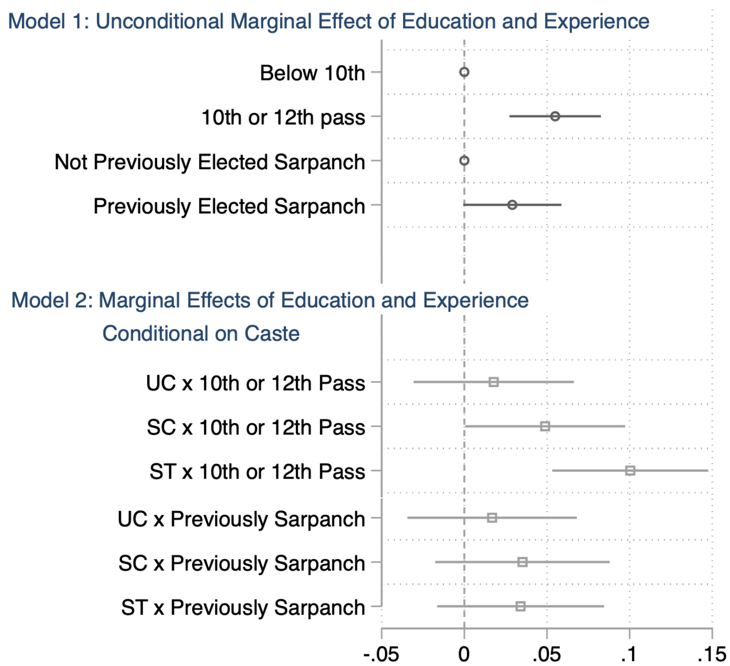

*Notes:* The figure reports the average marginal conditional effects of all attributes from a conjoint experiment conducted with a random sample of citizens in Madhya Pradesh. 95% confidence intervals from OLS regressions with standard errors clustered at the respondent level reported. The sample includes 1,251 citizens.

**Figure D3.** AMCEs for All Attributes in Citizen Conjoint Experiment by Election Type




*Notes:* The figure reports the average marginal conditional effects of all attributes by randomized election type from a conjoint experiment conducted with a random sample of citizens in Madhya Pradesh. 95% confidence intervals from OLS regressions with standard errors clustered at the respondent level reported. The sample includes 1,251 citizens.

**Figure D4.** Impact of Politicians' Education on Citizen Preferences, Mediated by Caste from a Second Citizen Conjoint Experiment




*Note*: The figure reports the average marginal conditional effects of education and political experience by candidate caste from a conjoint experiment conducted with a random sample of citizens in Madhya Pradesh. 95% confidence intervals from OLS regressions with standard errors clustered at the respondent level reported. The sample includes 2,832 citizens.

## D.9 Instrumental Variables Analysis on Impact of the Supply of Education

**Table D10.** The Effects of Secondary Schools on Politician Education

| | First-stage | Reduced Form | IV |
|---|---|---|---|
| | GP Average Education | Politician Education | Politician Education |
| | (1) | (2) | (3) |
| Secondary School in 1991 | 0.178*** | 0.720* | |
| | (0.060) | (0.382) | |
| GP Average Education | | | 4.123* |
| | | | (2.316) |
| Observations | 2936 | 2886 | 2886 |
| $R^2$ | 0.856 | 0.274 | 0.081 |
| Number of Blocks | 427 | 427 | 427 |
| First Stage F | | | 8.054 |

*Notes*: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the GP level in parentheses. The models include block fixed effects and controls for SC and ST population shares. This table leverages data from the 1991 census of India to identify the GPs in Odisha that had secondary schools prior to the creation of village governments in 1992. It instruments the share of a GP's population that has completed secondary school with an indicator for whether the GP had a secondary school in 1991 and then estimates the impact of population education levels on politician education levels.

## E Politician Performance

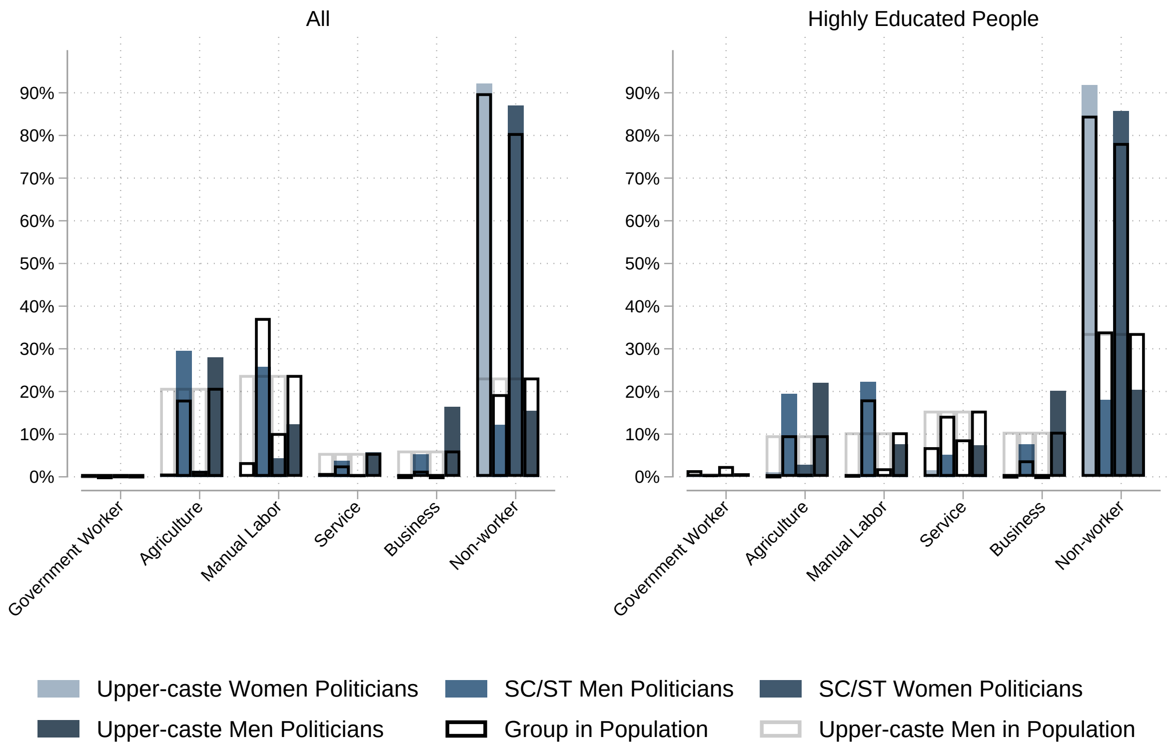
**Table E1.** The Relationship between Education, Identity, and Politician Performance

| **Dependent Variable:** | Ratio of Expenditures (Payments/Receipts) (Payments/Receipts) | | No. of People at Gram Sabha | |
|---|---|---|---|---|
| **Sample**: | 2017 | | 2022 | |
| | (1) | (2) | (3) | (4) |
| Years of Education | 0.001* | 0.001 | 0.360 | -0.853 |
| | (0.001) | (0.001) | (0.621) | (1.371) |
| **Identity** | | | | |
| UC Women | | -0.001 | | -20.453 |
| | | (0.016) | | (19.227) |
| SC/ST Men | | -0.018 | | -18.026 |
| | | (0.016) | | (22.017) |
| SC/ST Women | | -0.016 | | -49.488** |
| | | (0.015) | | (21.969) |
| Observations | 2687 | 2687 | 2353 | 2353 |
| $R^2$ | 0.495 | 0.498 | 0.814 | 0.815 |
| Number of Blocks | 425 | 425 | 417 | 417 |
| Fixed Effects | Block | Block | Block | Block |

*Notes*: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the GP level in parentheses. The table shows the relationship between politician education, politician identity, and performance outcomes. The sample includes only politicians in Odisha who were elected in the year shown in the table. Upper-caste men/open seats are the omitted identity group in all models. All models include block fixed effects and control for SC and ST population shares (and their squares), the total population size, and the share of the population from the politicians' identity group in agriculture, business, service, manual work, and others.

# F Outside Options

**Figure F1.** Distribution of Occupations in the Population and among Politicians



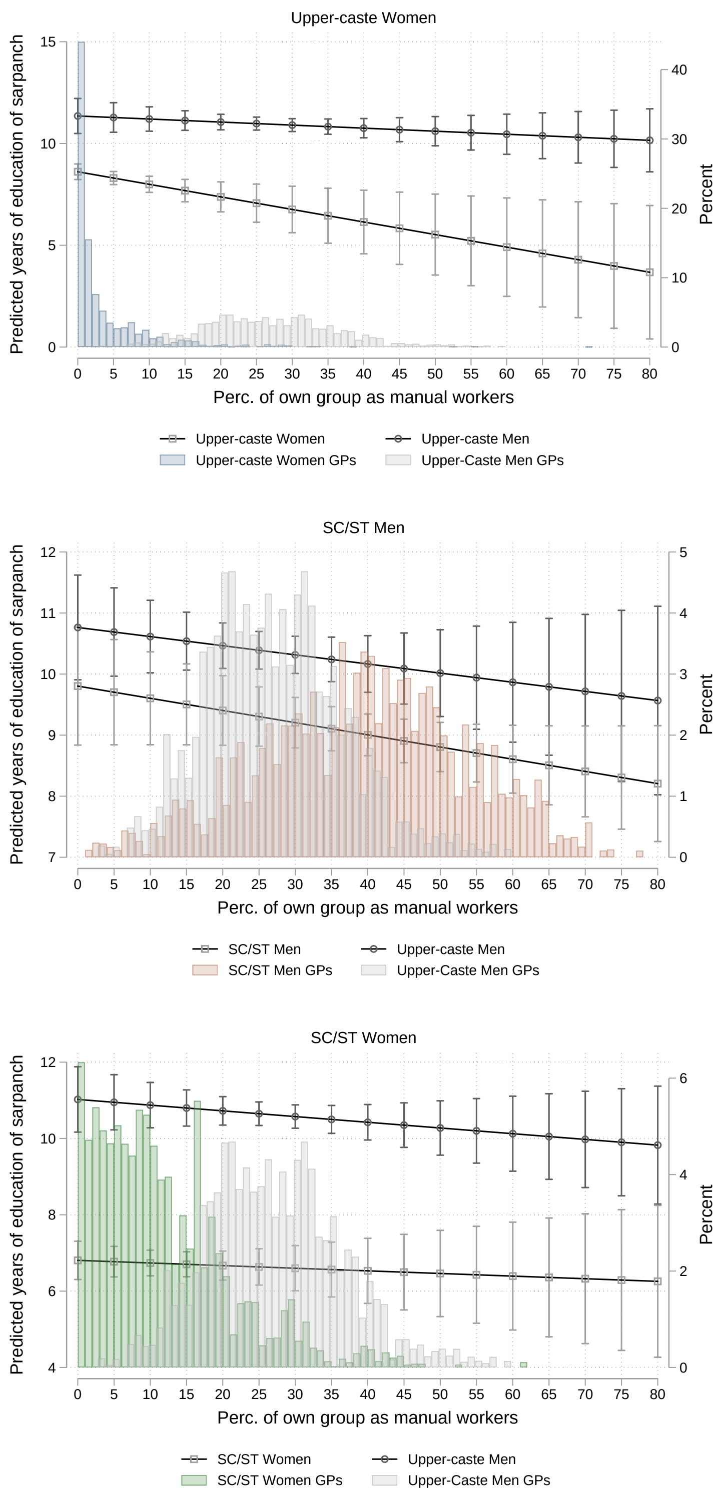

*Note*: This figure plots the occupation distribution for politicians and the population by their caste and gender. The blue bars depict the distribution of occupation for each politician identity group, the transparent black bars depict the population distribution for the corresponding identity group, and the transparent grey bars depict the population distribution for upper caste men. Data come from the Odisha census.

**Table F1.** Selection of Politicians by Identity, Education, and Job Supply

| **Dependent Variable:** | Selection as Politician | | | | |
|---|---|---|---|---|---|
| | All | >5% Business | >10% Manual | >10% Agriculture | >10% Non-workers |
| | (1) | (2) | (3) | (4) | (5) |
| Education | 0.009*** | 0.008*** | 0.010*** | 0.011*** | 0.009*** |
| | (0.000) | (0.001) | (0.000) | (0.001) | (0.000) |
| Education × UC Women | 0.001** | | 0.014*** | 0.027 | 0.001** |
| | (0.001) | | (0.003) | (0.028) | (0.001) |
| Education × SC/ST Men | 0.021*** | 0.040*** | 0.020*** | 0.018*** | 0.020*** |
| | (0.001) | (0.014) | (0.001) | (0.002) | (0.001) |
| Education × SC/ST Women | 0.033*** | | 0.036*** | 0.097*** | 0.033*** |
| | (0.002) | | (0.003) | (0.032) | (0.002) |
| Observations | 2909075 | 525249 | 1891873 | 1116802 | 2862083 |
| $R^2$ | 0.003 | 0.001 | 0.003 | 0.002 | 0.003 |
| Number of GPs | 3212 | 385 | 2159 | 1179 | 3121 |
| Fixed Effects | GP | GP | GP | GP | GP |

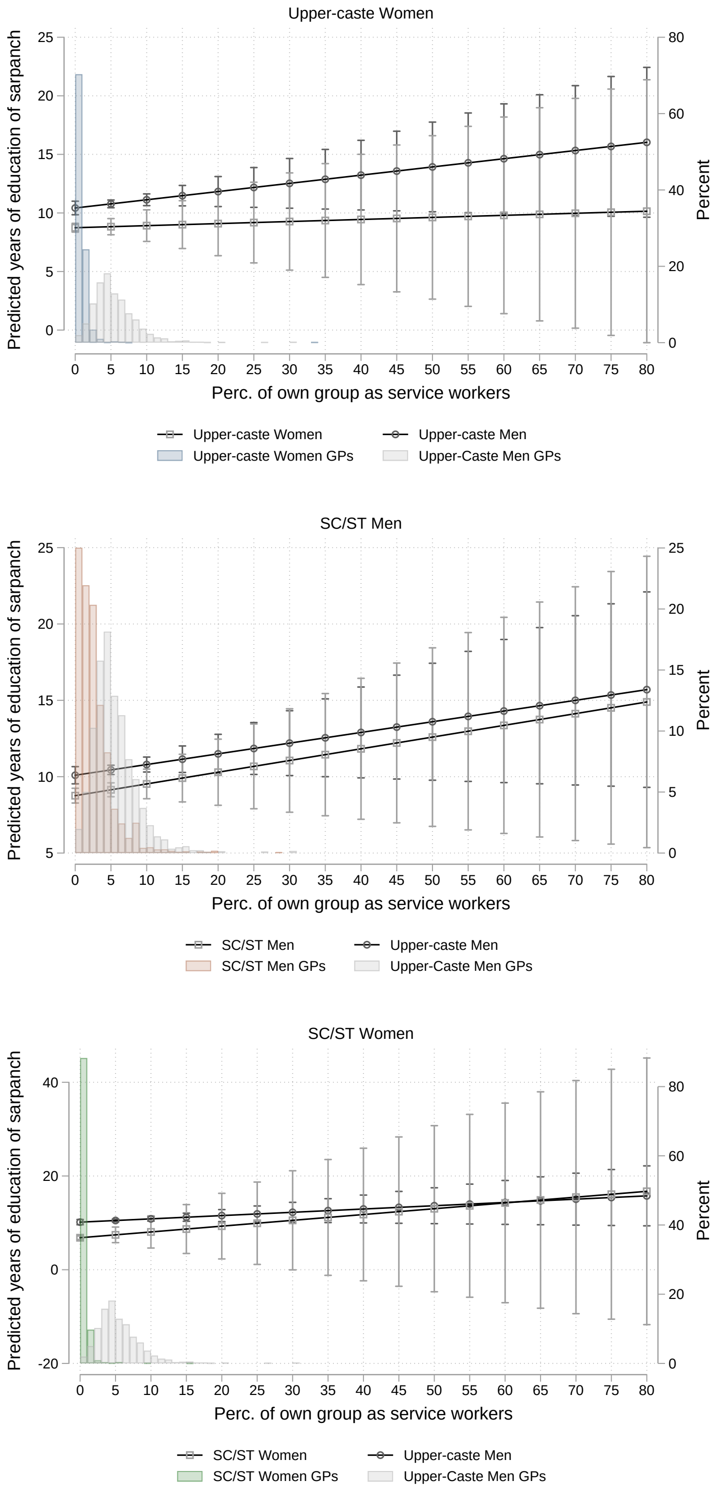
*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $* * * < 0.01$. GP clustered standard errors are reported in parentheses. The table reports the relationship between identity, education, and selection as the Sarpanch for GPs with different labor market environments. The sample includes all elected politicians and the population in their GP from their identity group. The dependent variable is standardized to aid interpretation. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects included as specified. The data are from the Odisha 2017 election.

**Figure F2.** Politician Education Condition on the Supply of Manual Work




*Note*: The figure reports the predicted values of politician education from a model where politician identity is interacted with the share of the identity group in the GP that is manually employed following specification 1. 95% confidence intervals from based on standard errors clustered at the GP level. The results use the sample from the Odisha identity panel model with GP fixed effects (see column 1 in Appendix Table D9). Additional covariates include SC and ST population shares (and their squares) and politicians' group occupation shares in agriculture, business, service, manual work, and others.

**Figure F3.** Politician Education Condition on the Supply of Service Work




*Note*: The figure reports the predicted values of politician education from a model where politician identity is interacted with the share of the identity group in the GP that is employed in the service sector following specification 1. 95% confidence intervals from based on standard errors clustered at the GP level. The results use the sample from the Odisha identity panel model with GP fixed effects (see column 1 in Appendix Table D9). Additional covariates include SC and ST population shares (and their squares) and politicians' group occupation shares in agriculture, business, service, manual work, and others.

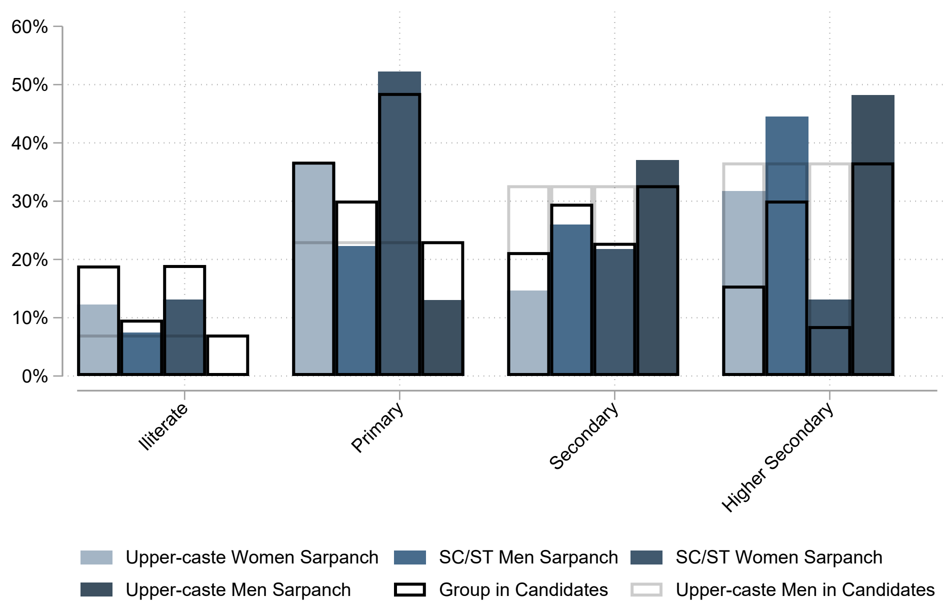
## G Candidates Analysis

**Table G1.** Descriptive Statistics on Candidates by GP Reservation Status

| | All Seats | Open Seats | Open Women Seats | Open SC/ST Seats | SC/ST Women Seats |
|---|---|---|---|---|---|
| Number of Candidates | 7.160 | 7.460 | 6.571 | 6.603 | 6.103 |
| | (3.460) | (3.519) | (2.927) | (3.714) | (3.922) |
| Total Votes | 2083.326 | 2292.469 | 1728.133 | 2281.634 | 1586.448 |
| | (1800.152) | (1850.397) | (1659.252) | (2186.026) | (978.497) |
| Number of Votes 1st | 851.431 | 924.643 | 680.517 | 1015.593 | 624.517 |
| | (850.067) | (801.699) | (636.208) | (1289.857) | (645.294) |
| Number of Votes 2nd | 511.529 | 498.044 | 542.356 | 605.777 | 437.552 |
| | (595.790) | (522.588) | (562.300) | (846.930) | (341.530) |
| Margin of Victory 1st vs. 2nd | 346.122 | 426.599 | 147.337 | 424.240 | 186.966 |
| | (503.394) | (548.782) | (354.201) | (583.339) | (350.129) |
| Observations | 748 | 309 | 233 | 131 | 29 |
| Number of GPs | 144 | 57 | 46 | 29 | 7 |

*Notes*: The table reports the mean and standard deviation (in parentheses) for each variable (row) for each GP reservation status (column). Data include all candidates from the All India survey.

**Figure G1.** Politicians are positively Selected on Education Relative to Candidates




*Note*: This figure plots the education distribution for politicians and candidates by their caste and gender. The blue bars depict the distribution of education for each politician identity group, the transparent black bars depict the candidate distribution for the corresponding identity group, and the transparent grey bars depict the candidate distribution for upper caste men. Illiterate corresponds to 0 years of education, literate to less than five years of education, primary to having completed at least five years of education, secondary to having completed at least ten years of education, and higher secondary to having completed at least 12 years of education. Data come from the All India census.

**Table G2.** Selection of Politicians from the Population and Candidates

| **Dependent Variable:** | Candidate (× 10,000) | Sarpanch (× 10,000) | Sarpanch (× 1) |
|---|---|---|---|
| **Sample:** | Population | Population | Candidates |
| | (1) | (2) | (3) |
| Education | 5.13*** | 1.77*** | 0.02*** |
| | (0.93) | (0.23) | (0.01) |
| Education × Open Women | 1.66 | 1.60** | 0.02 |
| | (1.91) | (0.67) | (0.02) |
| Education × Open SC/ST | 6.68* | 2.91** | -0.02 |
| | (3.61) | (1.18) | (0.01) |
| Education × SC/ST Women | 52.96*** | 31.51*** | -0.03 |
| | (19.18) | (11.73) | (0.09) |
| E(Y \| Open Seat) | 4.86 | 0.93 | 0.16 |
| # Observations | 141,382 | 141,345 | 590 |
| # GPs | 176 | 176 | 150 |
| Fixed Effects | GP | GP | GP |

*Notes*: *Levels of significance*: $* < 0.1$, $** < 0.05$, $*** < 0.01$. GP clustered standard errors are reported in parentheses. The table reports the relationship between eligibility for a GP's reservation status, education, and selection as the Sarpanch or emergence as a candidate. The dependent variable in column 1 is emergence as a candidate, and in columns 2 and 3 is selection as Sarpanch. The sample in columns 1 and 2 are the population eligible for the position (according to the reservation status), showing political selection vis-a-vis eligible constituents. The sample in column 3 is the set of all candidates, showing political selection vis-a-vis candidates. Upper Caste (UC) men/open seats are the omitted identity group in all models. Fixed effects included as specified.