Educational Inequalities across Castes: The Supply Side Challenges

Neha Bailwal

Indian Institute of Technology Delhi, New Delhi

Email: neha.bailwal95@gmail.com

and

Sourabh Bikas Paul

Indian Institute of Technology Delhi, New Delhi

Email: sbpaul@hss.iitd.ac.in

Corresponding author: Neha Bailwal

Abstract

A range of policy interventions such as Sarva Siksha Abhiyan and Right to Education Act has been taken in India to bridge the social and gender gaps in educational outcome. In this paper, we show the extent of caste discrimination in the provision of public schools in rural India. Using two rounds of census data (2001 and 2011) we find that the villages inhabited by the higher share of socially backward groups have a lower chance of having public schools. The students in villages with high concentration of backward communities also travel farther to reach the nearest school. While the policy intervention helped to reduce the gaps in the provision of primary schools, the gaps in middle and secondary school provision are still very high. Moreover, the newly constructed middle and secondary schools are mostly concentrated in villages inhabited by higher caste. Finally, we show that the gap in educational outcome across social groups is largely explained by the gap in the provision of public schools.

Keywords: education; inequality; caste; rural India; public school provision

JEL Classification: I2; I20; I25

1 Introduction

In this paper, we look into the extent of caste discrimination in the provision of public schools in Indian villages and its effect on educational disparities between social groups. Despite several affirmative action programmes in place to safeguard the interest of socially and economically backward castes, provided in a constitutional schedule known as "Scheduled Castes" (SC) and "Scheduled Tribes"(ST), educational disparities still persist between SC, ST and others (Desai et al., 2009; Desai and Kulkarni, 2008). On an average SC and ST household are less educated than non-SCST households however their intergenerational mobility is improving over time (Hnatkovska et al., 2012). In rural India, the gap in educational attainment between SC and others has increased (Deshpande, 2011). Considering this persistent gap in educational outcome, our interest in discrimination in provision of public schools arises from the fact that though there is now a vast literature identifying various spheres of exclusions and discrimination (Desai and Thorat, 2012; Watch, 2014; Balagopalan and Subrahmanian, 2003; Nambissan, 2009), explanations for such unequal educational outcomes across caste groups are still limited. A major part of such explanations is linked to supply-side challenges. We question whether belonging to certain social groups acts as a hindrance to accessing the essential public services like public education in India. In this paper, we intend to explore this discrimination faced by socially and economically backward groups. The objectives of this paper are to find the inter-castes differences in the provision of public schools in rural India and whether these differences can be attributed to caste composition at a village level. If there is a systematic gap in the provision of schools, does it really affect

educational outcome?

India has a long history of public provisions in the forms of direct provision (food security, health care, and education), subsidised prices (energy, fertilisers), state monopoly (rail transport), and production and marketing by state enterprises (energy, telecommunication). Public expenditure on these services is increasing, yet they are not able to achieve favourable socio economic outcomes for all sections of the society (Filmer et al., 2000; Dongre et al., 2014). There are many reasons for the poor performance of public services like rampant corruption, high absenteeism of local service provider, low quality of service, and lack of accountability (Hammer et al., 2007). Yet another possible reason could be that these public services are benefiting only some sections of the society and neglecting others. A recurring theme in scholarly work is the political economy aspect of it and how it is linked to social diversity. Discrimination and exclusion of groups because of their ethnicity, religion, gender have been a serious problem in India (Thorat and Newman, 2012; Deshpande, 2011).

A fundamental question that social scientists have long tried to answer is why some community have better provision of public goods and services while others do not. Political fragmentation and party preference (Blakeslee, 2013), geographical location and land ownership (Banerjee and Somanathan, 2001), community mobilization (Björkman and Svensson, 2009) play important roles in the distribution of public goods. A Weak legal system, lack of accountability, insufficient resources, information asymmetry, elite capture, and bureaucratic political institutions are some of the reasons for under provision of public goods and services (Olken, 2010; Bardhan and Mookherjee, 2006). Evidence from both

developed and developing countries support that ethnic diversity is an important factor resulting in the differential provision of public goods and services in a community. There is a broad consensus on the negative relationship between ethnic fragmentation and provision of public goods, lower economic outcomes and higher overall inequality (Alesina et al., 1999; Easterly and Levine, 1997; Miguel and Gugerty, 2005; Chadha and Nandwani, 2018). Caste and religion play an important role in lowering the provision of educational and health services in India (Banerjee, 2004; Betancourt and Gleason, 2000; Chaudhary, 2009). Understanding why heterogeneity in terms of socially diverse groups undermines the provision of public goods and services needs identification of factors or channels through which this relationship works. Habyarimana et al. (2007) emphasise on differential preferences; Banerjee and Somanathan (2007) highlight political power; Miguel and Gugerty (2005) talk about the failure of collective action in explaining how ethnicity is related to underprovision of public goods and services. To a large extent lack of political will is also responsible for the uneven distribution of public goods. Communities which are politically connected may be in a strong position to demand better quality and quantity of public goods (in comparison to groups who are politically weak for examples SC and ST in India). Communities sharing the same ethnic or caste identity with those who are in power to make the decision of provision of public goods at a local level have a higher likelihood of receiving a better share of public goods (Kumar et al., 2017). Literature, therefore, suggests that ethnic diversity (caste, race, religion) provides a significant explanation of the difference in allocation and provision of public goods and services and lower economic outcome.

India had undertaken significant reforms in its education system through the introduction of centrally sponsored scheme, *"Sarva Shiksha Abhiyan"* in 2001 which was later reinforced with Right to Education Act in 2009. Under this scheme, the central, as well as state governments, increased budgetary allocation from Rs. 12,931 crore in 2005-06 to 59,835 crore in 2011-12 to construct new schools, to expand the capacities of existing schools, for a door to door campaign to increase enrolment rate and reduce drop out, to hire more human resources, etc (MHRD, Govt. of India). One of the important objectives of this reform is to bridge the social and gender gaps at the primary school level. Therefore, it becomes imperative to find whether such expansion programme of elementary schools, particularly in rural areas, between 2001 and 2011 has benefited the most deserving section i.e. socially and economically backward castes.

In this backdrop, our study contributes to the above literature in two important ways. First, we look into the supply side challenges in explaining the gaps in educational outcomes of caste groups. Extant literature relates other supply-side factors like teacher quality and absenteeism, physical infrastructure, quality of schools(private and public) to the lower educational outcome (Kremer et al., 2005; Dreze and Kingdon, 2001); however, there are a very limited number of studies on how the availability of public schools is systematically biased towards higher castes and its overall implication on educational gaps. Our paper complements this literature by stressing on the availability of public schools and by providing the extent of caste discrimination. Second, our study, motivated by the earlier results of Banerjee and Somanathan (2007), looks at systematic bias in the existence of public goods in recent times at a more disaggregated

level. Banerjee and Somanathan (2007) use parliamentary constituency level data in 1971 and 1991 to show that social heterogeneity undermines the provision of different types of public goods in rural India whereas we estimate the gaps using village level census data of 2001 and 2011, the period of rapid expansion of public spending through *"Sarva Shiksha Abhiyan"*. We believe that caste dynamics is so subtle that it is better captured at the village level. We also find the link between provision of public schools and educational outcome in rural India. Does public schools matter for literacy rate? If the educational outcome in a village is not significantly correlated with the existence of public school then the supply side discrimination that we try to capture in this paper do not ultimately matter in explaining the education gap between SC, ST, and nonSCST. We address the issue of endogeneity between public schools provision and literacy rate using system generalised method of moments(GMM) estimation. Source of endogeneity between public schools and literacy rate arises from reverse causation (simultaneity) as the existence of public school in a village is determined partly by the overall education level of the village. It is possible that villages with higher number of literates may be in a better position to raise their voice and demand greater provision of public schools.

Our findings show that there is systematic bias against socially and economically backward castes in the provision of public school. The villages dominated by SC and ST population have a lower chance of having public schools compared to the villages inhabited by non-SCST. In 2011, villages which have less than 25 percent SC population, the likelihood of having a primary school in those villages is 87 percent. However, this probability decreases to 72 percent if the

share of SC population in a village is more than 75 percent. Similarly, for STs, the probability of having a primary school is more than 90 percent in villages which has less than 20 percent ST population but this probability reduces to 87 percent as the share of ST population increases to 80 percent and above. Also, the students in villages with high concentration of backward communities travel farther to reach the nearest school. Chadha and Nandwani (2018) found that higher caste diversity(ethnic fragmentation) is systematically related to inequitable economic outcome. They demonstrated empirically that the provision of schools and health centers can actually reduce the adverse impact of ethnic fragmentation. Our paper also exhibits that systematic bias in the provision of public schools explains gaps in educational outcome across castes. Banerjee and Somanathan (2007) found that STs are the more disadvantageous group in terms of accessing public schools whereas SCs improved their representation in national politics since the 1980s that resulted in better political power to extract more public goods and services. The proportion of villages having middle and high public schools in a parliamentary constituency is negatively associated with the share ST population, but not with the share of SC population in the constituency. Contrary to their results, our findings suggest that STs and SCs are both discriminated against in terms of receiving public schools in rural India. We also find that public schools matters for literacy rate, particularly in those villages which are largely inhabited by SC and ST population. Constructing a new public school improves the overall literacy rate of a village by 4 percent. In addition, provision of a new public school in SC dominated village in a district reduces the literacy gap between non-SCST and SC by almost 3 percentage points.

The gap between non-SCST and ST literacy rate is reduced by 5.7 percentage points with the provision of one new public school in ST dominated village in a district. We interpret the results with some caveats. We do not intend to identify the supply function of public schools. It is hard to distinguish whether the non-existence of public schools in some villages is due to low demand or for some other reasons (which we attribute it to caste discrimination in this study). We assume here that demand for public schools across geography is similar and there is no reason to believe that demand for schools is correlated with the share of SC and ST in a village. If at all it has to be correlated, it should be correlated positively because given that they are economically backward, affording private schools will be harder. If it is true then our results would be an underestimation.

The rest of the paper is organised as follows. In section 2, we describe the data and provide descriptives analysis. Section 3 illustrates the empirical strategy. In section 4 we present the results of our two main empirical hypotheses: how caste matters in under-provision of public schools in India and whether the existence of school matters for the overall educational outcome. Section 5 concludes.

2 The data and descriptive statistics

We use two rounds of Indian Census data (Office of the Registrar General and Census Commissioner, Government of India, 2001 and 2011). Indian census, held every 10 years, is the single largest source of information on different characteristics of the people of India. We use village level population composition and amenities data from 19 major states[1]. Census data give us information

[1]We dropped the North Eastern states (except Assam) and Jammu and Kashmir

about different caste composition of a village. It also has information on the number public schools in a village. In the absence of a public school in a village, the amenities data also contain information about the distance to the nearest public school in another village. We also have other village characteristics like distance to district head quarter and area of villages, etc. The census data do not give break up of education level by caste groups at a village level. However, it provides district level aggregates of education level and population attending educational institutions by age and social groups. We also use Indian Human Development Survey(IHDS) 2011-12 to calculate drop out rates across castes. IHDS is a nationally representative survey of 42,152 households across India. IHDS is organised jointly by National Council of Applied Economic Research, New Delhi (NCAER) and the University of Maryland. They collected data in 2004-05 for the first round and again reinterviewed the same households in 2011-12.

2.1 Population composition and public school provision

We begin with the spatial distribution of population composition and public schools provision in rural India. Panels a and b of Figure 1 display the share of SC and ST population respectively at a district level. Darker districts show the domination of SC and ST population. It is evident that STs are concentrated in a belt stretching from west to east passing through central India. SC dominated districts are mostly visible in northern states including Punjab, some parts of Uttar Pradesh, Bihar and Tamil Nadu in southern India. We contrast the concentration of SC and ST with concentration of public schools at district level in

panel c and d of Figure 1. Once again darker districts highlight a higher proportion of villages having primary, middle and secondary schools. Gujarat, Kerala, central Karnataka and Haryana have higher concentration of primary and middle schools. The north-south distinction is very clearly visible in this figure. Panel d showing the concentration of villages with secondary schools depicts a very grim picture. Only 2 out of 10 villages have secondary schools in districts of northern and central India. The proportion of villages having public schools is not uniform across states. In 2011, for instance, 95 % of villages in Gujarat have primary schools whereas only 62 % in Uttar Pradesh have primary schools. The variation in the provision of middle and secondary schools is also very wide across states. This district level school density maps lend some support to our conjecture that population composition and school provision are correlated. For instance, Karbi Anglong and Dima Hasao district of Assam are highly ST dominated areas and more than 70 percent of the villages in these districts have no primary, middle and secondary schools. Some parts of eastern Uttar Pradesh which are relatively SC dominated has less number public schools.

2.1.1 Likelihood of having schools in areas dominated by SC and ST population

We now find the probability of having a public school in villages and plot it against village level population composition. Our hypothesis here is that as the share of socially backward caste population in a village increases, the odds of having a public schools decreases. We estimate a locally weighted smooth curve of the probability of having a school against the share of population of SC

and ST in village. Figure 2 shows the probability of having primary schools in villages for SC(upper panel) and ST(lower panel) population respectively. The downward sloping curves imply that as the share of SC and ST population in a village increases, the probability of having primary schools falls. This shows that villages dominated by economically and socially backward population has less likelihood of having primary schools. However, if we compare over a decade, we see that curves are flatter in 2011 which hints towards some signs of improvement.

Figure 3 shows that the likelihood of having middle schools in villages dominated by SC and ST population is quite low. Villages where more than half of the population are SCs, the probability of having a middle school is as low as 28 percent in 2011, the figure in 2001 was 15 percent (Table 1). Figure 4 shows the likelihood of having secondary schools against the share of SC and ST population. These figures are quite unexpected because even with the introduction of flagship Sarva Siksha Abhiyan programme which was operational since 2001 and aims to construct new schools in underdeveloped areas, we find that there are many SC and ST dominated villages having no public schools at all. We also plot the chances of constructing a new school during 2001 to 2011 in a village against the proportion of SC and ST population in Figure 5. The newly constructed primary schools are in general in areas dominated by SCs and STs as depicted with an upward sloping curve in panel a and b of figure 5. However, for secondary schools, we see clear evidence of discrimination. The villages having higher concentration of backward castes are less likely to have a newly constructed secondary schools during this period (panel e and f of Figure 5). These results clearly show supply side discrimination in the provision of schools

in rural India.

Do the children of SC and ST dominated villages travel farther to reach the nearest public school? We look at the correlation between distance to the nearest public school and share of SC and ST population in villages. Figure 6 confirms our conjecture that living in areas dominated by backward caste is disadvantageous for both non-SCST, SC and ST population as the distance to the nearest school is higher for them. Studies find that parents are usually reluctant to send their children to far off schools to study especially girls as parents deem it unsafe. This results in irregular attendance and higher dropout rates particularly for students from socially backward castes. Comparing over a decade, we find that there is a marginal decrease in the distance to the nearest school, but the discrimination still exists(upward slope). Therefore, there is a dual burden of higher concentration of backward castes in a village - chances of having public school is lower and the students travel farther to reach the nearest public school. These two phenomena result in other adverse effects in human capital accumulation, particularly for the disadvantaged groups.

2.2 Gaps in educational outcome

This section of the paper reviews the differences in educational outcomes of different social groups using district level averages of the Census data. We use district level averages because published census data do not have village level disaggregated figures for social groups. We look at four variables to compare SC, ST and general categories: i) share of children attending schools ii) drop out rate iii) literacy rate iv) educational attainment. Drop out rates are calculated

using Indian Human Development Survey, 2011-12 because census data do not publish drop out figures. The inequality exists both in educational outcome of the adult population as well as for the children in school going age.

2.2.1 Share of population attending schools

Figure 7 shows the percentage of population attending schools by different age groups. The overall trend is the same across social groups. The gap between non-SCST and SC and ST widen as age increases. Comparison over a decade shows that the SC and ST population is catching up with non-SCST. In 2001, less than 70 % of the population in the age group of 7-12 years were in school. There is a significant improvement in 2011 with more than 80 % of the population attending schools.

2.2.2 Dropout rate

Figure 8 shows that while India has made a progress in terms of almost universal enrolment in education, however, it is less successful in terms of preventing students from dropping out even at the early stages of schooling. According to the Indian Human Development Survey(IHDS) 2011-12, drop out rate is highest amongst STs followed by SCs. Difference between ST, SC, and non-SCST widens after primary level. Once students complete primary schooling, the decision to continue to middle and secondary level is influenced by many factors like distance to school, augmenting family livelihood, disinterest in formal education and discrimination in school environment. Some of the factors are very significant for socially backward castes, thus explaining the wide gap in drop out

rates. This has some serious implications as they will not benefit much in terms of returns to education if they are only educated till primary level (Duraisamy, 2002). Hence, the focus should be on creating such an atmosphere for children especially from stigmatised section, so that they are motivated to complete their schooling. High enrolment figures are less likely to translate into regular attendance and better learning outcomes unless proper intervention is made to bring children back to schools.

2.2.3 Literacy rate

Figure 9 shows the average literacy rate of SC, ST and non-SCST population for 19 large states in India. In 2001 only 3.8 out of 10 STs were literate. The situation has marginally improved with 5 out of 10 STs being literate in 2011. There is an improvement in SCs average literacy rate from 4.5 out of 10 to around 5.5 out of 10 during this period. SC average literacy rate in 2011 is at the level where non-SCST was 10 years back. SC and ST intergenerational mobility in education is also low but improving over time(Hnatkovska et al., 2013, 2012). The difference in average literacy rate between the historically disadvantageous groups and upper castes still persist.

2.2.4 Educational attainment

A comparison of two rounds of census data reveals that the pattern in educational attainment among the literates is consistent with overall literacy gaps. The SC and ST continue to lag behind in the education ladder. Figure 10 shows the distribution of educational attainment of all age groups by castes in

2001 and 2011. In bottom category (primary education) SC and ST are slightly over-represented while in category 5 (secondary education or above) they are relatively under-represented. More than 80 percent of SC and ST still concentrated in secondary or below level of education in 2011. There is only marginal improvement in educational attainment of SC and ST over a decade; the proportion of SC and ST having more than secondary and above education improved from 25 % to 30 % and from 22 % to 24 % respectively.

The descriptive analysis above shows two important results: First, in general provision of public schools is quite low in Indian villages. Villages largely inhabited by socially and economically backward castes have a lower provision of public schools. Second, considering education as one of the important indicator of human development, quite expectedly SC and ST are not better performers. Now, it would be interesting to look whether the difference in educational outcomes among SC, ST and non-SCST can be attributed to lack of provision of public schools. The above analysis do not control for other important village level characteristics that might affect the probability of having public schools in a village and educational outcome. Therefore, in next section, we check the above hypotheses using a simple econometric model.

3 Empirical strategy

We undertake the empirical analysis in two steps. In the first section, we test our main hypothesis that whether the existence of public schools is less likely in villages predominantly inhabited by the socially and economically backward

castes. Let us define a binary random variable y_v as

$$y_v = \begin{cases} 1, & \text{if village } v \text{ has a public school} \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

The outcome variable y_v depends on host of village level characteristics including village level population composition (share of SC and ST). In order to test the effects of share of ST and SC on public school provisioning in a village, we estimate the following Probit model:

$$\begin{aligned} P(y_v = 1) &= \Phi(\beta_0 + \beta_1 share\_pop\_sc_v + \beta_2 share\_pop\_st_v \\ &+ \beta_3 tot\_pop + \boldsymbol{X}_v \boldsymbol{\delta} + \epsilon_v) \end{aligned}$$

where $share\_pop\_sc_v$, $share\_pop\_st_v$ and tot_pop are share of population of SC, ST and the total population in village v respectively (nonSCST share is reference). Usually, the concentration of social groups at village level is historically determined. We assume that the population measures in Census are error free. It is very unlikely that our outcome variable (having school or not) would influence the social group composition in a village. However, there are several other factors which may introduce the endogeneity problem. For example, a village may be well connected to the authorities of district or state education departments who are responsible for the decision of creating new schools. We control for this using a proxy by the distance to district headquarter. The geographical size of a village is also an important factor that may be correlated with our main explanatory variables. If both village and district is dominated by non-SCST,

it may have a different effect compared to a situation where both village and district is dominated by SC. We control for this by including a dummy of village and district dominance. There are many villages where the share of SC and ST population are zero(around a quarter of villages have no SC population at all, around 55 percent of villages have no ST population in 2001 and around 11 percent of villages are completely inhabited by non-SCST population). We control for these type of villages by including a dummy for such villages. The vector $\boldsymbol{X}_v$ includes control for distance to district headquarter from village, area of village, district village caste dominance, villages with zero share of SC and ST population and state dummies. Our main identification assumption is that population composition of a village is exogenous after controlling these factors. In multiple regression setting, to detect any discrimination we expect β_1 and β_2 to be negative.

In our next specification, we divide the villages based on the dominant social group by population share. For example, a village is dominated by SC if the share of SC is more than the share of ST and the share non-SCST in total village population. We include this village dominance dummies as an explanatory variable and specify the model as follows:

$$\begin{aligned} P(y_v = 1) &= \Phi(\beta_0 + \beta_1 village\_dominance\_sc_v + \beta_2 village\_dominance\_st_v \\ &+ \beta_3 tot\_pop_v + \boldsymbol{X}_v\boldsymbol{\delta} + \epsilon_v) \end{aligned}$$

where *village_dominance* are dummies for SC and ST dominated village (non-SCST dominance is reference). We control for total population, distance to district head quarter from village, area of village, district village caste domi-

nance, villages with zero shares of SC and ST population and state dummies as in the previous specification. If β_1 and β_2 turn out to be negative then there is statistical evidence of discrimination against SCs and STs.

In the third specification, we model the conditional expectation of y_v using a latent variable capturing the effect of the population size of different social groups as the main explanatory variable. In a resource-constrained economy, the decision to allot a public school to a village is an equilibrium outcome of several demand and supply side processes. Without explicitly modelling the structure, we assume that a village has latent power, y^*, to influence the probability of allocating a school in the village. This attraction power for allocation of a public school depends on the population composition of a village. We assume that the latent power is generated by a Cobb-Douglas function:

$$y^* = A(sc\_pop)^{\alpha}(st\_pop)^{\beta}(nonscst\_pop)^{\gamma} \tag{2}$$

where *sc_pop*, *st_pop* and *nonscst_pop* are population sizes of *sc*, *st* and *nonscst* respectively. All other factors have multiplicative effect on latent power. We assume constant but differentiated elasticities of group sizes. Our main hypothesis is that the elasticity of *nonscst* population is significantly higher than the other groups. Taking logarithmic transformation,

$$ln(y^*) = \alpha ln(sc\_pop) + \beta ln(st\_pop) + \gamma ln(nonscst\_pop) + lnA. \tag{3}$$

where lnA includes other controls, random error term ϵ, and a constant. There-

fore, the latent variable specification becomes

$$ln(y^*) = \alpha_0 + \alpha ln(sc\_pop) + \beta ln(st\_pop) + \gamma ln(nonscst\_pop) + \boldsymbol{X}\boldsymbol{\delta} + \epsilon \quad (4)$$

The error term has a normal distribution with mean zero and constant variance. There is a threshold level of latent power which determines the final binary outcome variable, y_v. As long as a constant is included, we can specify

$$y_v = \begin{cases} 1, & \text{if} \quad ln(y_v^*) > 0 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

The final estimable probit specification is

$$\begin{aligned} P(y_v = 1) &= \Phi(\beta_0 + \beta_1 ln(pop\_nonscst_v) + \beta_2 ln(pop\_sc_v) + \beta_3 ln(pop\_st_v) \\ &+ \boldsymbol{X}_v\boldsymbol{\delta} + \epsilon_v) \end{aligned}$$

We expect that if caste matters in the provision of public schools, the change in population size of non-SCST will affect the probability of having a public school in a village more than that of SC and ST.

With the above specifications if we arrive at a conclusion that there exist systematic gaps in the provision of public schools at a village level, we next test whether the provision of public schools matters for overall educational outcome. In other words, can gaps in the provision of public schools explain gaps in educational outcome across castes?

A simple OLS estimates of literacy rate on public school provision may suf-

fer from endogeneity issue. The source of correlation between public schools status and the error primarily arises from reverse causation (simultaneity) as the existence of public school in a village is partly determined by the overall education level of the village. It is possible that villages with higher number of literates may be in a better position to raise voice and demand greater provision of public schools. Therefore, in order to test the importance of public schools for educational outcome, we use system generalised method of moments(GMM) to estimate the following two equation simultaneously.

$$\begin{aligned} literacy\_rate_v &= \delta_0 + \delta_1 public\_school\_status_v + \delta_2 private\_school\_status_v \\ &+ state\_dummies + \epsilon_v \\ public\_school\_status_v &= \beta_0 + \beta_1 literacy\_rate_v + \boldsymbol{X}_v \boldsymbol{\delta_1} + \epsilon_v \end{aligned}$$

where $public\_school\_status_v$ and $private\_school\_status_v$ are dummy variables of having public and private school in village v respectively (reference: no school). In the second equation, $public\_school\_status_v$ is an outcome variable. The vector $\boldsymbol{X}_v$ includes population sizes of SC, ST and nonSCST, area of the village and state dummies. We assume that at least one control variable in the second equation is exogenous for identification of the effect of *public_school_status* on *literacy_rate*. We assume that the total area of the village is the exogenous variable which satisfies the exclusion restriction. Area of village influence the outcome of whether a village has a public school or not but it does not affect the literacy rate. This section of the paper is based on 2011 census data only as private school information is available for 2011.

Next, we estimate a model that tests whether gaps in provision of public school explain gaps in educational outcome between groups. If it is true then the supply side discrimination is an important explanation for the gaps in educational outcome. The literacy rate is the only educational outcome variable available at the village level. However, the literacy rate mainly captures the educational outcome of adult population, which may not be a good educational outcome variable that is determined by the current public school status. Moreover, literacy rate is not available by caste groups at a village level. We cannot estimate village level educational gaps from the census data. The census data provide literacy rates and percentage of children attending schools (6 to 14 years of age group) at the district level by caste groups. Therefore, in addition to literacy rate gaps, we use gaps in share of children attending school at district level to find whether educational outcome gaps could be attributed to gaps in public school status. We define four outcome variables: i) gap between SC and non-SCST and ii) gap between ST and non-SCST in terms of literacy and school attendance. As these outcome variables are at the district level and our main hypothesised variable is the status of public school at the village level, we face a difficulty in aggregation of gap in provision of public school. We follow the following approach to test the hypothesis using the district level data. When the outcome variable is the SC and non-SCST gap, we categorise the villages into four groups based on the share of SC population: 0 - 25% (q1), 25 - 50% (q2), 50-75% (q3) and 75-100% (q4). If the probability of having a public school in $q4$ villages in a district is negatively associated with gaps in literacy rate and school attendance, we may conclude that the discrimination in the provision of public

school at village level may affect the gaps in educational outcome at district level. A similar approach is followed when the outcome variable is educational gap between ST and non-SCST. Formally, we specify the empirical model as

$$\begin{aligned} y_d &= \beta_0 + \beta_1 P(public\_school_v = 1 | v \in q1) \\ &+ \beta_2 P(public\_school_v = 1 | v \in q4) + \beta_2 tot\_pop_d + \boldsymbol{X}_d \boldsymbol{\delta} + \epsilon_d \end{aligned}$$

where y_d is literacy rate gaps and school attendance gaps, $P(public\_school_v = 1|v \in q4)$ is the probability of having a public school in $q4$ villages in district d. The vector $\boldsymbol{X}_v$ is a vector of district level covariates that are expected to affect district level gap in educational outcome across castes which includes the share of SC population, share of ST population and state dummies. If β_2 turns to be negative, there is a statistical evidence that provision of public schools in highly SC and ST dominated villages can reduce literacy and school attendance gaps at the district level.

4 Results

4.1 Does population composition matter for under-provision of public schools?

Results from estimating probit specifications (marginal effects) are reported in Table 2-4. Table 2 depicts results for all three types of public schools (primary, middle and secondary) and for two years (2001 and 2011). To interpret the results reference category is share of non-SCST. Our estimates highlight the dis-

advantaged position of SC and ST relative to non-SCST. As share of SC and ST population in a village increases, the probability of having public schools falls. These findings are in line with unconditional probability plot in the section 2.1.1 which suggests that even after controlling for important village level characteristics like total population, area of village, distance to district headquarter, etc the discrimination still exists. Comparison over a decade, show that after introduction of Sarva Siksha Abhiyan programme which aims to build public schools in those habitation which lacks schooling facilities, the extent of discrimination has reduced in 2011 because the coefficient is lower particularly for primary and middle schools but still negative. For secondary schools, we find that discrimination has in fact increased over a decade.

Next, we reestimate the model by including a dummy variable of village dominance. Negative coefficients on middle and secondary schools highlights that SC and ST dominated villages are discriminated. SC and ST dominated villages have lower probability of having middle and secondary schools in comparison to non-SCST dominated villages. However, for primary schools, we do not see such discrimination as SC and ST dominated villages has a relatively higher probability of having primary schools compared to non-SCST villages. Though we find that SC and ST dominated villages have a higher probability of having primary schools but still share of SC and share of ST has negative effects on having a primary schools[2].

[2]When we include village dominance as well as share of SC and ST as our regressors in a same specification. The village dominance variable gets dropped because of collinearity between district and village dominance dummy. We find that once again share of SC and ST has negative and significant effect on primary school provision. ST dominated villages have insignificant effect on primary school provision whereas SC has positive effect but at 10 % level of significance

Next, we try to capture the effects of change in population sizes on probability of public school provision in a village. Column 1-2 of Table 4 reports the marginal effects of having primary schools, while column 3-6 reports the same result for middle and secondary schools for 2001 and 2011 respectively. We see that coefficients of groups sizes are positive implying that population size affect provision of public schools positively, but effects are different for different groups. The marginal effect for non-SCST is higher than SC and ST for all kinds of public schools(primary, middle and secondary) and in both the years. The effects are robust even after controlling for important village level characteristics and state level aggregate factors. These results highlight caste discrimination against SC and ST which is being captured by differentiated elasticities.

4.2 Does public school matter for overall educational outcome?

Table 5 reports the result of system GMM estimation[3]. Results in Panel A of Table 5 capture the effects of change in population sizes on the probability of public school provision in a village. These results are in conformity with our previous results showing that the change in population size of non-SCST affect the probability of having a public school in a village more than that of SC and ST counterparts. Panel B of Table 5 depicts the results of our main hypothesis whether the provision of public schools matters for overall literacy rate or not. We find that constructing one new public school in a village increases the literacy rate of village by almost 14 percentage points. The statistically significant effect

[3]We have also tested this specification with share of SC and share of ST instead of population sizes, coefficient on public schools is still positive and significant

of public schools on literacy rate confirms the importance of public schools for educational outcome in rural areas. The positive coefficient on both public and private schools shows that schools in general matter for literacy rate but a higher coefficient on public school implies that in villages public school matter more for the educational outcome. The results are merely association between public schools and literacy rate and it should be not interpreted with causal inference.

4.3 Can gaps in the provision of public school explain the gap in the educational outcome?

Table 6 reports the result from a regression of the provision of public schools on gaps in educational outcome between non-SCST and SC. As mentioned above we take two types of educational outcome variable: i) literacy rate(panel A) ii) school attendance(panel B). Panel A and column 1 of Table 6 show that as the probability of having a public school in a q4 (highly SC dominated) village increases, the gap in literacy rate between non-SCST and SC declines. Column 2 and 3 control for other factors like total population, share of SC, share of ST population and state dummies to ensure that our results are not confounded because of all these factors. Even after controlling for important district level covariates, we see that literacy rate gaps are reduced significantly as the probability of having a public school in highly SC dominated villages increases. In Panel B of table 6 we present results for whether public schools matters for school going age children or not. Once again we find that even after controlling for state fixed effects and population composition, the gaps in the proportion of school going children between non-SCST and SC reduces as the

probability of having a public school in highly SC dominated villages increases.

Table 7 highlights the results of the impact of public schools on the gap between non-SCST and ST educational outcome. Once again, panel A depicts results for literacy rate gaps and Panel B presents results for school attendance gaps. We find that even after controlling for state fixed effects and other district level control, the gaps in literacy rate between nonSCST and ST at the district level is reduced by more than 5 percentage points as the probability of having a public school in q4 village increases by 1 percent point. We find that as the probability of having a public school in ST dominated village increases, the gap between nonSCST and ST in terms of share of population between 6-14 years age attending schools decreases. This is true even after controlling important district level controls and state aggregate factors.

5 Conclusion

This paper demonstrates empirically the extent of caste discrimination in provision of public schools in rural India and its effect on educational outcome. We find that as the share of SC and ST in a village increases, the probability of having a public school falls. In addition, the students in villages with the high concentration of lower caste communities travel farther to reach the nearest school. These results highlight the disadvantaged position of socially and economically backward castes in rural India. However, comparison over two rounds of census (2001 and 2011) shows that extent of caste discrimination in the provision of primary and middle schools has reduced which is also being reflected in better school attendance across social groups. For secondary schools, the extent of caste discrimination has increased. We find that villages dominated by socially and economically backward castes have a lower chance of having a middle and secondary school compared to villages inhabited by higher caste. Moreover, population size affects the provision of public schools differently for different social groups, the effects for non-SCST is higher than SC and ST. These results are robust after controlling for important village level characteristics and state level aggregate factors. The strong association between public schools and literacy rate highlights the importance of public schools. This paper empirically demonstrates that provision of public schools in SC and ST dominated villages helps in reducing gaps in educational outcome across castes.

In 2001, India had undertaken a significant education reform with introduction of centrally sponsored scheme "Sarva Shiksha Abhiyan" which was later reinforced with Right to Free and Compulsory Education Act 2009 (RTE). In

order to bridge the social and gender gaps, one of the main objective of this programme was to build new primary and middle schools in those habitations which do not have any schooling facilities. Given this backdrop, our findings show how far we have achieved this goal. Comparison over a decade shows that the number of public schools (particularly primary and middle) has increased in rural areas. We also see signs of improvement in educational outcome particularly school attendance rate where we find that catching up is taking place between SC,ST, and non-SCST.

Despite these improvements and policy intervention(SSA) in place which aims to bridge the social gaps, we find that educational disparities still exist between SC, ST, and non-SCST. The supply side discrimination (in terms of provision of public schools) captured by us in this paper explains one the reason for these educational disparities across castes. This suggests that this policy intervention has limited benefit to the most deserving section of the society because discrimination in middle and secondary schools is still very high. These results along with recent studies suggest that in order to ensure equitable economic outcome and reduce the impact of ethnic fragmentation, unbiased and equitable provision of public goods is important (Chadha and Nandwani, 2018).

Success of SSA and RTE has also been depicted by where the new schools are constructed between 2001 and 2011. We find that most of the new primary schools are constructed in those areas which have a higher concentration of SC and ST population. These results once again show that extent of caste discrimination has decreased over time particularly for primary schools. However, for middle schools, the probability of constructing new schools is lower in the village

where the share of SC population is high. This raise some serious concern as the policy intervention itself has some sort of discrimination. In secondary school provision, we find that over time the extent of discrimination has increased. This can have serious consequences as secondary school presence is an important determinant of primary school educational outcome (Mukhopadhyay and Sahoo, 2016). Since SSA focused particularly on elementary education (primary and middle schools), and it did not cover secondary schools, that could be one possible reason that discrimination in the provision of secondary schools has increased over time.

Therefore, if the aim is to bridge the social gaps, increasing number of public schools in areas which are easily accessible to students from marginalised caste should be the main push of educational policy. We agree that just having schools is not a sufficient condition for improving educational outcomes but it certainly a necessary condition. The first and foremost step to ensure equal educational outcome is to make sure that public schools are available to every student irrespective of their social identity.

References

Alesina, A., Baqir, R., and Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114(4):1243–1284.

Balagopalan, S. and Subrahmanian, R. (2003). Dalit and adivasi children in schools. *IDS Bulletin*, 34(1):43–54.

Banerjee, A. (2004). Who is getting the public goods in india? some evidence and some speculation. *India?s Emerging Economy*, pages 183–213.

Banerjee, A. and Somanathan, R. (2001). Caste, community and collective action: The political economy of public good provision in india. Technical report, mimeo, Massachusetts Institute of Technology.

Banerjee, A. and Somanathan, R. (2007). The political economy of public goods: Some evidence from india. *Journal of Development Economics*, 82(2):287–314.

Bardhan, P. and Mookherjee, D. (2006). Decentralisation and accountability in infrastructure delivery in developing countries. *The Economic Journal*, 116(508):101–127.

Betancourt, R. and Gleason, S. (2000). The allocation of publicly-provided goods to rural households in india: on some consequences of caste, religion and democracy. *World Development*, 28(12):2169–2182.

Björkman, M. and Svensson, J. (2009). Power to the people: evidence from a randomized field experiment on community-based monitoring in uganda. *The Quarterly Journal of Economics*, 124(2):735–769.

Blakeslee, D. S. (2013). Politics and public goods in developing countries: Evidence from india. *Unpublished Working Paper*.

Chadha, N. and Nandwani, B. (2018). Ethnic fragmentation, public good provision and inequality in india, 1988–2012. *Oxford Development Studies*, pages 1–15.

Chaudhary, L. (2009). Determinants of primary schooling in british india. *The Journal of Economic History*, 69(1):269–302.

Desai, S., Adams, C. D., and Dubey, A. (2009). Segmented schooling: inequalities in primary education. *Blocked by Caste: Discrimination and Social Exclusion in Modern India*, pages 230–52.

Desai, S. and Kulkarni, V. (2008). Changing educational inequalities in india in the context of affirmative action. *Demography*, 45(2):245–270.

Desai, S. and Thorat, A. (2012). Social inequalities in education. *India Infrastructure Report*, pages 44–52.

Deshpande, A. (2011). *The grammar of caste: economic discrimination in contemporary India*. Oxford University Press.

Dongre, A., Kapur, A., and Tewary, V. (2014). How much does india spend per student on elementary education?

Dreze, J. and Kingdon, G. G. (2001). School participation in rural india. *Review of Development Economics*, 5(1):1–24.

Duraisamy, P. (2002). Changes in returns to education in india, 1983–94: by gender, age-cohort and location. *Economics of Education Review*, 21(6):609–622.

Easterly, W. and Levine, R. (1997). Africa's growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics*, 112(4):1203–1250.

Filmer, D., Hammer, J. S., and Pritchett, L. H. (2000). Weak links in the chain: a diagnosis of health policy in poor countries. *The World Bank Research Observer*, 15(2):199–224.

Habyarimana, J., Humphreys, M., Posner, D. N., and Weinstein, J. M. (2007). Why does ethnic diversity undermine public goods provision? *American Political Science Review*, 101(4):709–725.

Hammer, J., Aiyar, Y., and Samji, S. (2007). Understanding government failure in public health services. *Economic and Political Weekly*, pages 4049–4057.

Hnatkovska, V., Lahiri, A., and Paul, S. (2012). Castes and labor mobility. *American Economic Journal: Applied Economics*, 4(2):274–307.

Hnatkovska, V., Lahiri, A., and Paul, S. B. (2013). Breaking the caste barrier intergenerational mobility in india. *Journal of Human Resources*, 48(2):435–473.

Kremer, M., Chaudhury, N., Rogers, F. H., Muralidharan, K., and Hammer, J. (2005). Teacher absence in india: A snapshot. *Journal of the European Economic Association*, 3(2-3):658–667.

Kumar, H., Somanathan, R., Chakravarthy, S., Baul, T., Rosenblat, T., Wang, Q., Banerjee, R., Gupta, N. D., and Villeval, M. C. (2017). Caste connections and government transfers: The mahadalits of bihar. Technical report.

Miguel, E. and Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in kenya. *Journal of public Economics*, 89(11):2325–2368.

Mukhopadhyay, A. and Sahoo, S. (2016). Does access to secondary education affect primary schooling? evidence from india. *Economics of Education Review*, 54:124 – 142.

Nambissan, G. B. (2009). Exclusion and discrimination in schools: Experiences of dalit children.

Olken, B. A. (2010). Direct democracy and local public goods: Evidence from a field experiment in indonesia. *American political science review*, 104(2):243–267.

Thorat, S. and Newman, K. S. (2012). *Blocked by caste: economic discrimination in modern India*. Oxford University Press.

Watch, H. R. (2014). They say we're dirty: Denying an education to india's marginalized. Technical report.

Table 1: Probability of having public school in village

Population composition in village	Primary 2001	Primary 2011	Middle 2001	Middle 2011	Secondary 2001	Secondary 2011
More than 50 % SC	0.68	0.73	0.15	0.28	0.05	0.08
More than 50 % ST	0.75	0.83	0.14	0.29	0.05	0.08
More than 50 % NonSCST	0.80	0.83	0.31	0.45	0.13	0.17

Table 2: Marginal effects: Probit model of having primary, middle and secondary school in a village

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Primary: margins 2001	Primary: margins 2011	Middle: margins 2001	Middle: margins 2011	Secondary: margins 2001	Secondary: margins 2011
Share SC population	-0.103***	-0.058***	-0.334***	-0.148***	-0.146***	-0.212***
	(0.000)	(0.005)	(0.000)	(0.000)	(0.001)	(0.000)
Share ST population	-0.091***	-0.007	-0.467***	-0.315***	-0.256***	-0.323***
	(0.000)	(0.718)	(0.000)	(0.000)	(0.000)	(0.000)
Square share SC population	0.022*	0.028**	0.047***	-0.040**	-0.016	0.013
	(0.059)	(0.014)	(0.010)	(0.026)	(0.613)	(0.685)
Square share ST population	0.044***	0.002	0.175***	0.122***	0.106***	0.111***
	(0.000)	(0.834)	(0.000)	(0.000)	(0.000)	(0.000)
Square share Non-SCST population	-0.081***	-0.048***	-0.223***	-0.143***	-0.112***	-0.157***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(Total population)	0.137***	0.121***	0.198***	0.231***	0.110***	0.129***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
All Controls	Yes	Yes	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.398	0.403	0.313	0.303	0.311	0.270
Observations	557,939	539,954	540,092	522,255	457,939	439,954
pval in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

[a] All controls include distance to district head quarter from village, area of village, district village caste dominance, villages with zero share of SC and ST population and state dummies.

Table 3: Marginal effects: Probit model of having primary, middle and secondary school in a village with village dominance

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Primary: margins 2001	Primary: margins 2011	Middle: margins 2001	Middle: margins 2011	Secondary: margins 2001	Secondary: margins 2011
Village_dominated = SC	0.005***	0.005***	-0.033***	-0.034***	-0.021***	-0.028***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Village_dominated = ST	0.019***	0.021***	-0.042***	-0.026***	-0.028***	-0.023***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(Total population)	0.138***	0.121***	0.200***	0.232***	0.111***	0.132***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
All Controls	Yes	Yes	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.398	0.403	0.312	0.302	0.310	0.271
Observations	557,939	539,954	540,092	522,255	457,939	439,954
pval in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

[a] All controls include distance to district head quarter from village, area of village, district village caste dominance, villages with zero share of SC and ST population and state dummies.

Table 4: Marginal effects: Probit model of having primary, middle and secondary school in a village

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Primary: margins 2001	Primary: margins 2011	Middle: margins 2001	Middle: margins 2011	Secondary: margins 2001	Secondary: margins 2011
Log(Non-SCST population)	0.060***	0.057***	0.091***	0.094***	0.063***	0.064***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(SC population)	0.042***	0.031***	0.051***	0.056***	0.027***	0.033***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(ST population)	0.045***	0.035***	0.022***	0.028***	0.011***	0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
All Controls[a]	Yes	Yes	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.356	0.361	0.282	0.268	0.295	0.255
Observations	557,939	539,954	540,092	522,255	557,939	539,954

pval in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[a] All controls include distance to district head quarter from village, area of village, district village caste dominance, villages with zero share of SC and ST population and state dummies.

Table 5: Impact of public schools on literacy rate

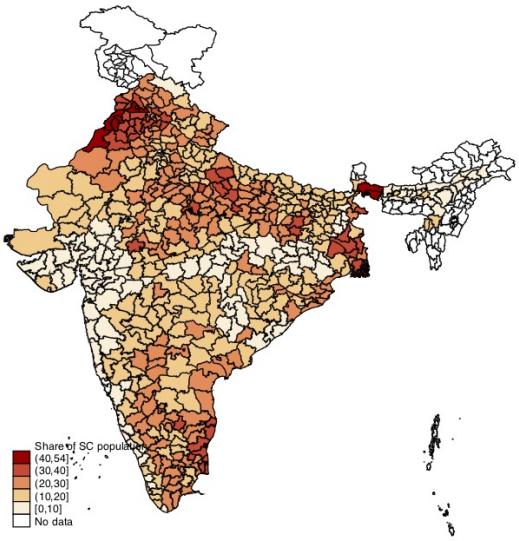
A. Dependent variable: Public primary school status		
Variables	Probit estimates	GMM estimates
Log(Non-SCST population)	0.298***	0.042***
	(0.002)	(0.000)
Log(SC population)	0.254***	0.027***
	(0.000)	(0.000)
Log(ST population)	0.229***	0.031***
	(0.000)	(0.000)
Literacy rate	0.002***	0.007**
	(0.000)	(0.000)
Log(area)	0.264***	0.047***
	(0.003)	(0.000)
Caste Fractionalization Index(Herfindahl)	-2.062***	-0.203***
	(0.000)	(0.005)
Private school status	-0.007	-
	(0.008)	
State Dummies	Yes	Yes
Observations	572,652	572,652
B. Dependent variable: Literacy Rate	OLS estimates	GMM estimates
Public school status	0.620***	13.790**
	(0.056)	(0.365)
Private primary school status	1.683***	0.258**
	(0.051)	(0.087)
State Dummies	Yes	Yes
Observations	572,652	572,652
standard error in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Table 6: Impact of public schools on gap in educational outcome between Nonscst and SC

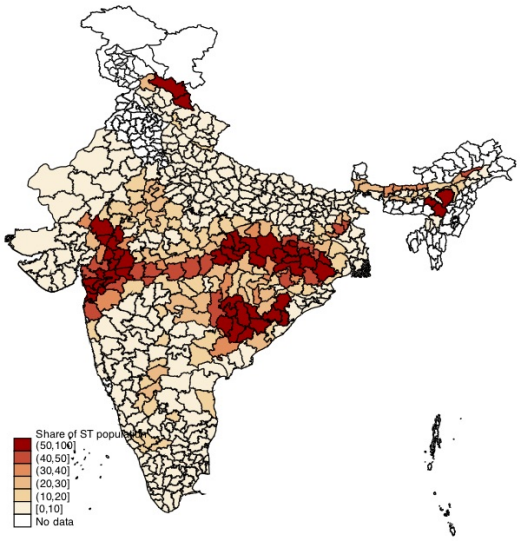
	(1)	(2)	(3)
Panel A. Dependent variable: Literacy rate gaps	model 1	model 2	model 3
$Probability(public\_school_v = 1 \mid v \in q1)$	-0.733	4.699**	2.367
	(2.435)	(2.323)	(2.207)
$Probability(public\_school_v = 1 \mid v \in q4)$	-4.051***	-6.312***	-3.223***
	(1.279)	(1.218)	(1.018)
Share of SC		35.165***	28.233***
		(4.082)	(4.071)
Share of ST		7.756***	5.392***
		(1.766)	(1.643)
Total Population		0.000**	0.000
		(0.000)	(0.000)
State Dummies	No	No	Yes
Adj R-squared	0.030	0.195	0.516
Observations	463	425	425
Panel B. Dependent variable: School attendance gaps	model 1	model 2	model 3
$Probability(public\_school_v = 1 \mid v \in q1)$	-0.015	0.013	0.047***
	(0.017)	(0.017)	(0.016)
$Probability(public\_school_v = 1 \mid v \in q4)$	-0.028***	-0.036***	-0.013*
	(0.009)	(0.009)	(0.007)
Total population		0.000**	0.000
		(0.000)	(0.000)
Share of SC		0.185***	0.154***
		(0.027)	(0.028)
Share of ST		0.038***	0.019
		(0.013)	(0.012)
State Dummies	No	No	Yes
Adj R-squared	0.036	0.132	0.498
Observations	463	463	463
standard error in parentheses			

Table 7: Impact of public schools on gap in literacy rate between Nonscst and ST

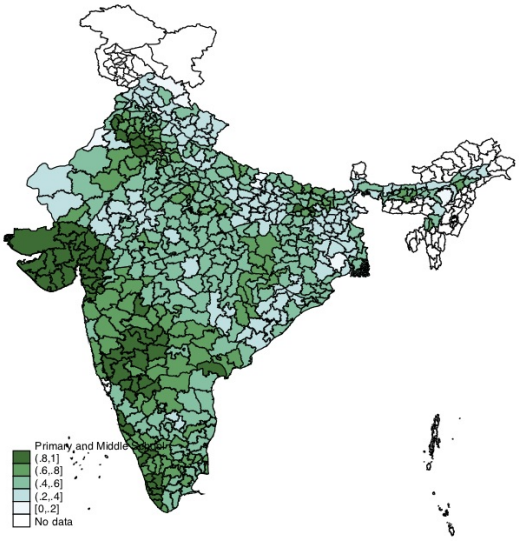
	(1)	(2)	(3)
Panel A. Dependent variable: Literacy rate gaps	model 1	model 2	model 3
$Probability(public\_school_v = 1 \mid v \in q1)$	13.563***	19.894***	0.992
	(4.402)	(4.358)	(5.338)
$Probability(public\_school_v = 1 \mid v \in q4)$	-6.201***	-9.861***	-5.834***
	(2.118)	(2.165)	(2.063)
Share of SC		37.836***	12.163
		(7.404)	(8.412)
Share of ST		19.707***	9.247***
		(2.893)	(2.864)
Total population		0.000*	0.000
		(0.000)	(0.000)
State Dummies	No	No	Yes
Adj R-squared	0.029	0.146	0.408
Observations	342	342	342
Panel B. Dependent variable: School attendance gaps	model 1	model 2	model 3
$Probability(public\_school_v = 1 \mid v \in q1)$	0.076*	0.113***	0.013
	(0.040)	(0.041)	(0.053)
$Probability(public\_school_v = 1 \mid v \in q4)$	-0.076***	-0.097***	-0.066***
	(0.018)	(0.020)	(0.020)
Total population		0.000	0.000
		(0.000)	(0.000)
Share of SC		0.221***	0.106
		(0.066)	(0.081)
Share of ST		0.110***	0.030
		(0.026)	(0.028)
State Dummies	No	No	Yes
Adj R-squared	0.043	0.087	0.262
Observations	342	342	342
standard error in parentheses			



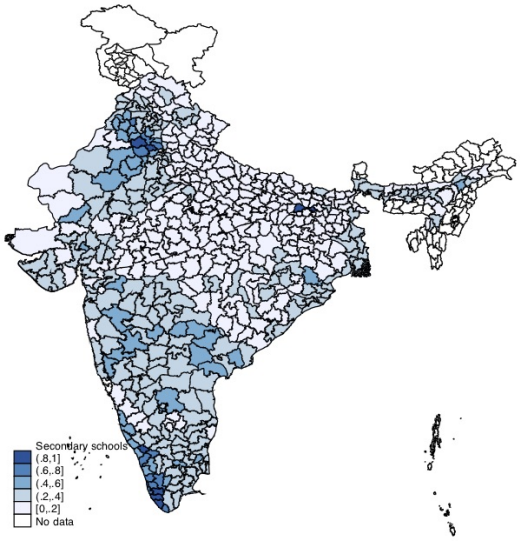
(a) Share of SC population



(b) Share of ST population



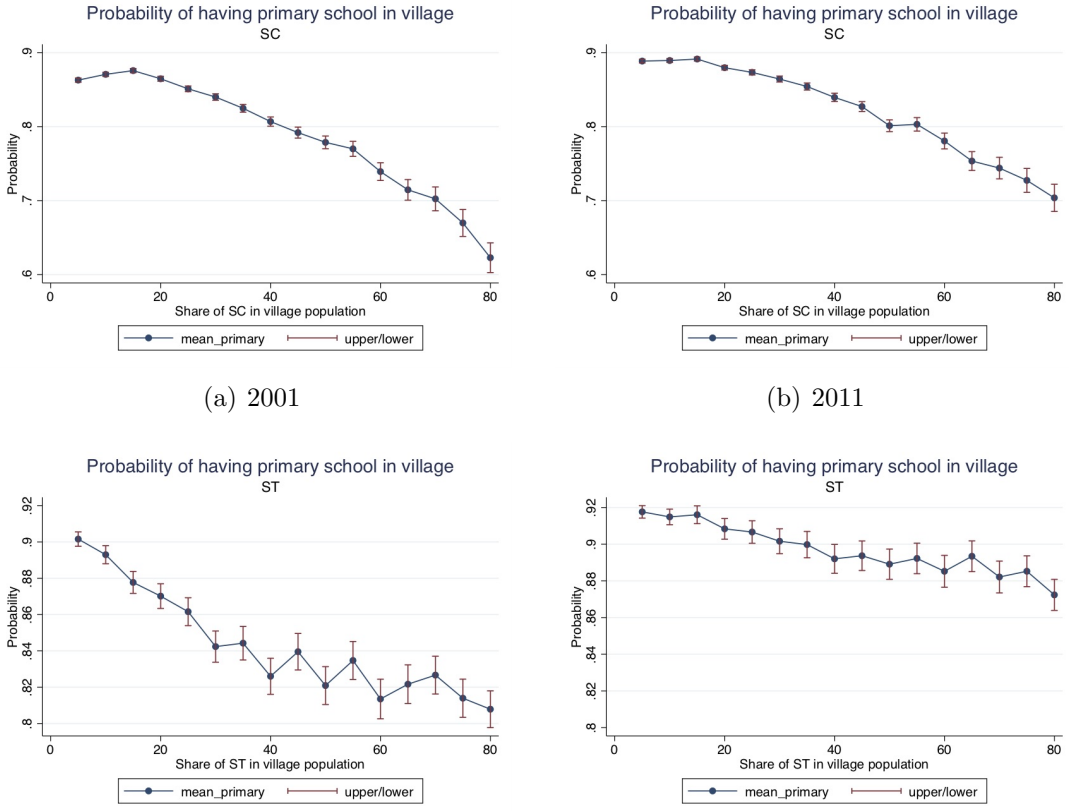
(c) Proportion of villages in a district with primary and middle schools



(d) Proportion of villages in a district with secondary schools

Source: Authors's calculation from Census data

Figure 1: Population composition and public school provision

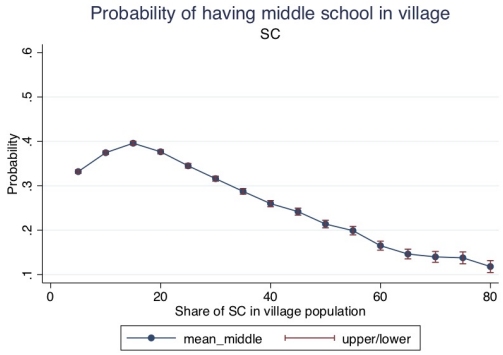


(a) 2001 (b) 2011

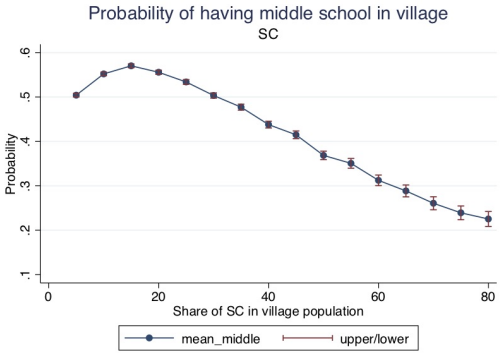
(c) 2001 (d) 2011

Source: Authors's calculation from Census data

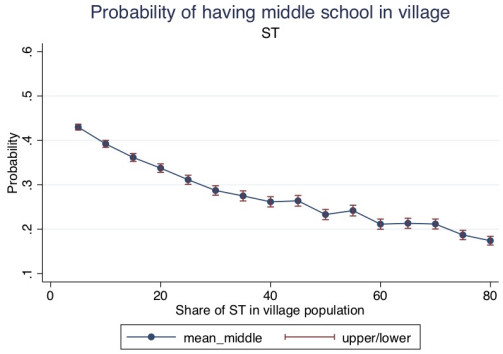
Figure 2: Probability of having a Public Primary School in Village



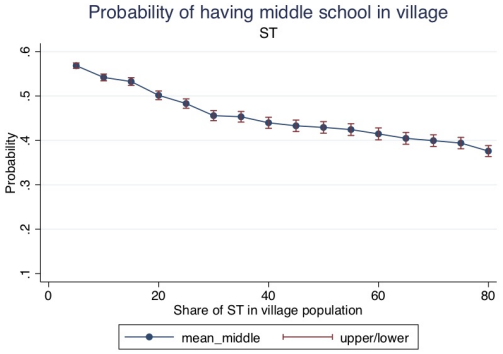
(a) 2001



(b) 2011



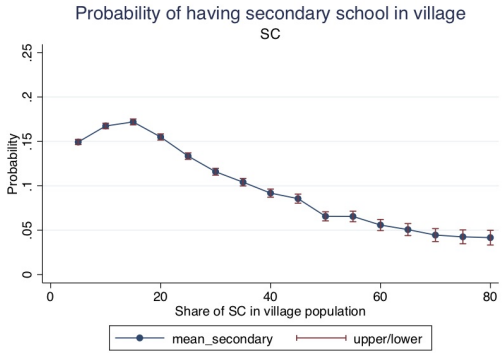
(c) 2001



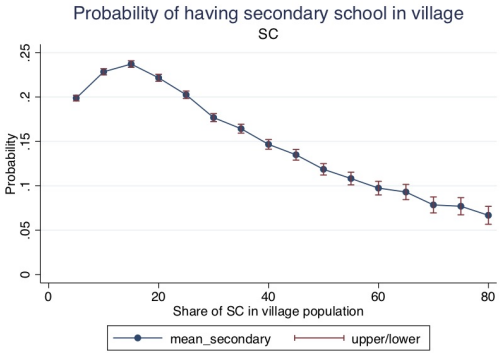
(d) 2011

Source: Authors's calculation from Census data

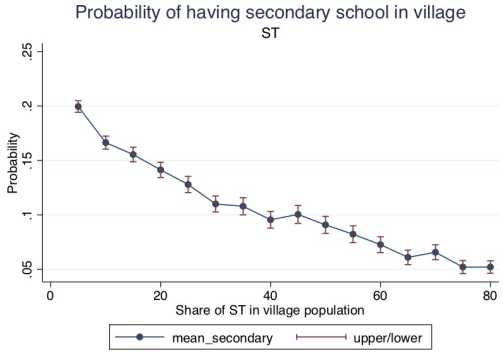
Figure 3: Probability of having a Public Middle School in Village



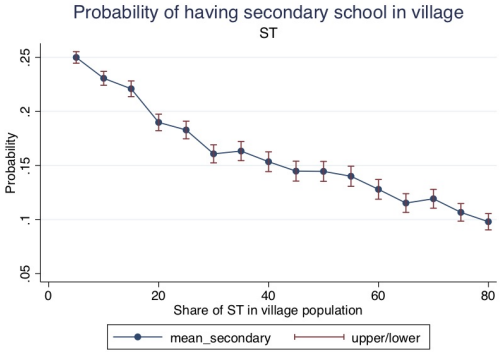
(a) 2001



(b) 2011



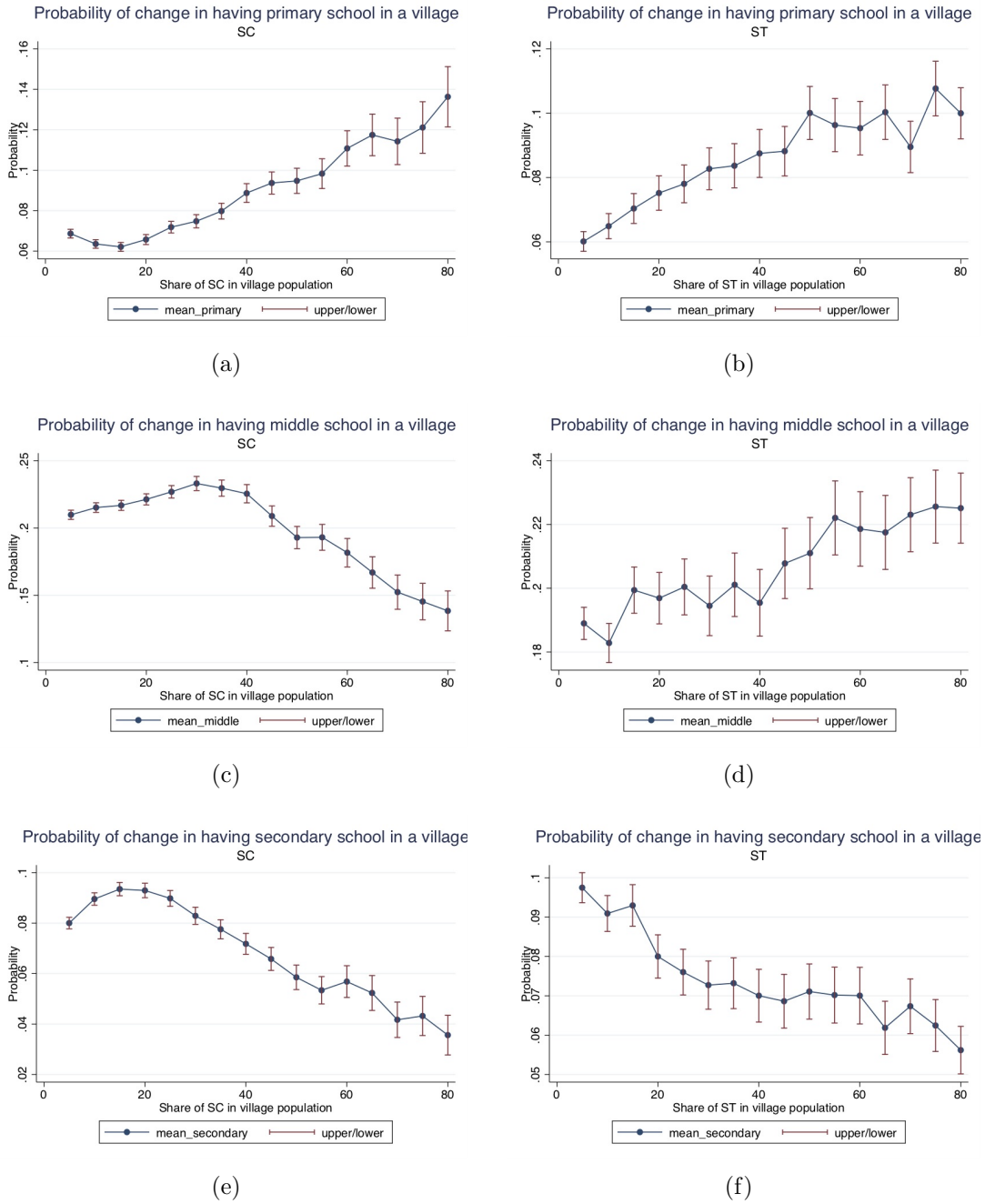
(c) 2001



(d) 2011

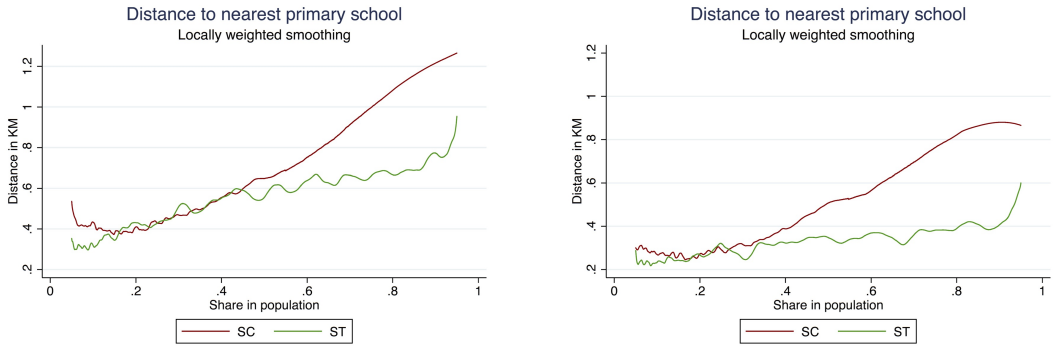
Source: Authors's calculation from Census data

Figure 4: Probability of having a Public Secondary School in Village



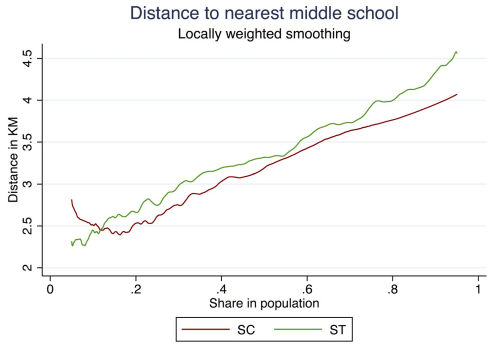
Source: Authors's calculation from Census data

Figure 5: Probability of construction of new school and share of SC, ST in village population

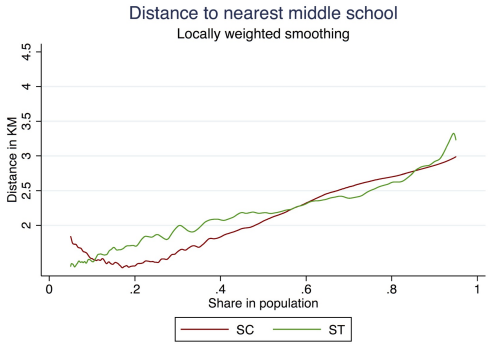


(a) 2001

(b) 2011



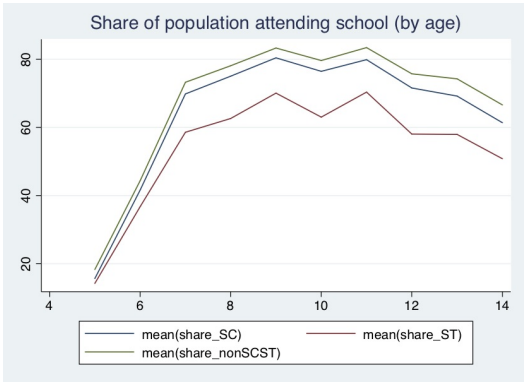
(c) 2001



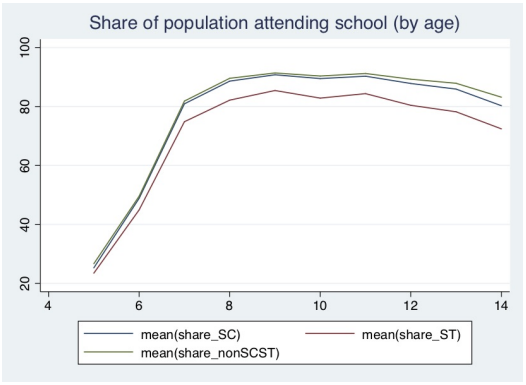
(d) 2011

Source: Authors's calculation from Census data

Figure 6: Average distance to public schools and share of SC, ST population



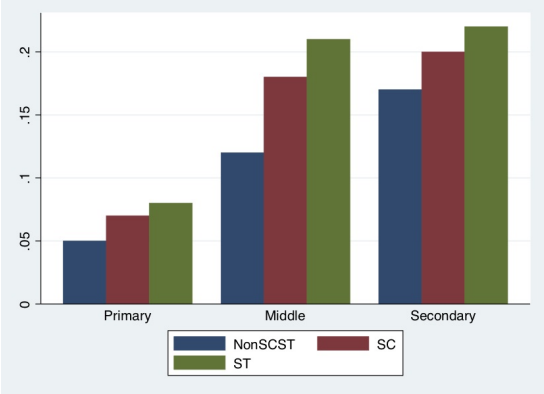
(a) 2001



(b) 2011

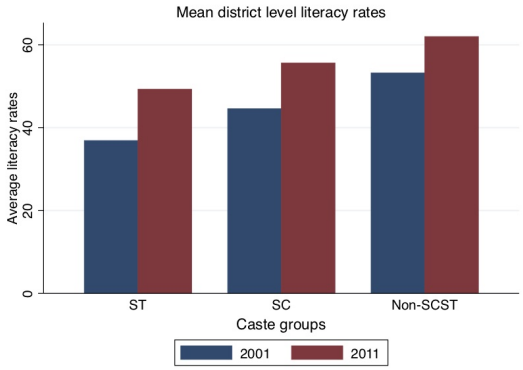
Source: Authors's calculation from Census data

Figure 7: Share of Population Attending School



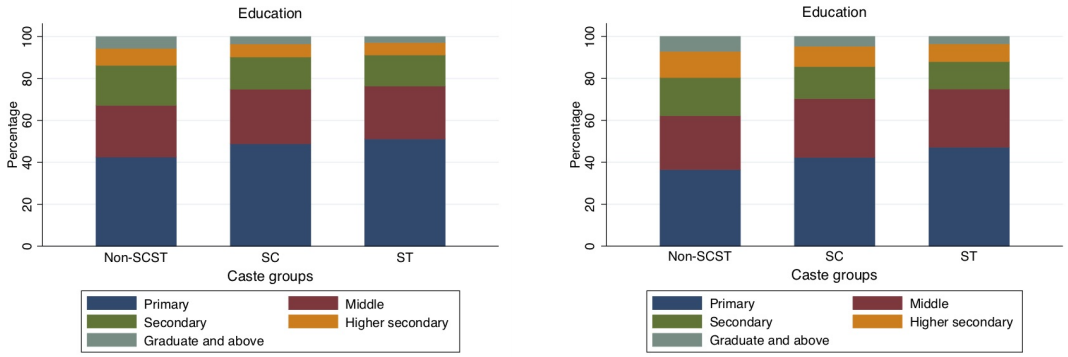
Source: Authors's calculation from India Human Development Survey-II (IHDS-II), 2011-12

Figure 8: Dropout rates



Source: Authors's calculation from Census data

Figure 9: Literacy Rate



(a) 2001 (b) 2011

Source: Authors's calculation from Census data

Figure 10: Distribution of Educational Attainment across caste groups