UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2017/176

Ethnic fragmentation and school provision in India

Nishant Chadha[1] and Bharti Nandwani[2]

October 2017

Abstract: In this paper, we study the impact of ethnic fragmentation on the provision of private and public schools, separately. The distinction is made because the two types of schools have different objective functions, a factor which can influence the relationship between ethnic fragmentation and public goods provision. We find that ethnic fragmentation has a negative impact on the provision of schools overall, but this effect manifests differently for the two types of schools considered. To explain our findings we show that ethnic fragmentation lowers collective action, and because of the different objectives of provision of private and public schools, lack of collective action results in a differential impact. While private schools are shown to be lower in number, public schools are of lower quality in fragmented districts.

Keywords: Ethnic fragmentation, private schools, public good provision, public schools
JEL classification: H4, O2, I0

Acknowledgements: We would like to thank Rachel Gisselquist, Miguel Niño-Zarazúa, Arne Bigsten, Alok Kumar, seminar participants at UNU-WIDER, DSE winter school, and University of Gothenburg for helpful comments. We would also like to show our gratitude to UNU-WIDER, where a part of this research was conducted. The usual disclaimer applies.

An earlier draft of this paper was presented at the winter school, Delhi School of Economics 2016.

All tables appear at the end of the paper.

---

[1] India Development Foundation, Haryana, India; [2] Shiv Nadar University, Dadri, India; corresponding author: bn666@snu.edu.in.

This study has been prepared within the UNU-WIDER PhD Internship programme.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-402-5 https://doi.org/10.35188/UNU-WIDER/2017/402-5

Typescript prepared by Gary Smith.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Denmark, Finland, Sweden, and the United Kingdom.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

This paper looks at the impact of ethnic fragmentation on the provision of schools in Indian districts and examines how the impact differs across different mechanisms of provision of schools. There is extensive work in the economics and political science literature that has established negative economic and political consequences of ethnic diversity.[1] Ethnic diversity has been shown to lower economic growth rates and lead to bad public policy choices (Easterly and Levine 1997), lower public good provision (Alesina et al. 1999), higher economic inequality (Chadha and Nandwani 2016) and outbreaks of civil conflicts (Elbadawi and Sambanis 2002). Our paper is placed within the broad literature on the impact of ethnic fragmentation on public goods. A general consensus has been built in this area of research on the negative association between ethnic fragmentation and provision of public goods, both in the developed and developing country contexts. Alesina et al. (1999) document the negative relationship between ethnic diversity and productive public goods for US cities. Banerjee and Somanathan (2007) and Banerjee et al. (2005) show that a similar relationship holds for India.

However, this paper questions the nature of the established relationship and is, therefore, in line with the emerging body of work that has criticized conventional wisdom in the existing literature (Gisselquist et al. 2016; Singh 2011).[2] We build on the observation that the existing work has not focused on the mechanism of provision of the public good while considering the impact of ethnic fragmentation. For example, while some studies have looked at the impact of ethnic fragmentation on a single type of public good, others have not made a clear distinction between the different public goods that they study.[3] We believe that incorporating the distinction between mechanisms of public good provision and identifying the type of public good that is most affected can help us better understand the relationship between ethnic fragmentation and public good provision. For instance, since ethnic fragmentation is essentially a local demographic characteristic, public goods that are not provided by the local community/government might not be expected to be affected by local ethnic diversity, an idea also espoused by Gisselquist et al. (2016).

Our paper focuses on the impact of fragmentation on different mechanisms of providing a particular public good—schools. We decided to study provision of schools because, apart from providing opportunities to invest in human capital, in an earlier paper (Chadha and Nandwani 2016) we have shown that provision of schools and health centres can mitigate some of the negative impacts of ethnic fragmentation. Schools are provided both by the government (public schools) and by private players in India. Public schools are provided by either the central, state, or local government, whereas private schools are normally of two types: one which gets financial support from the government (usually referred to as aided schools) and the other that does not (referred to as unaided private schools).

[1] See Alesina and La Ferrara (2005) for a literature review.

[2] The results of both of these papers indicate that ethnic heterogeneity is not necessarily a deal breaker for development.

[3] It is probably one of the reasons why the mechanism suggested by the current studies for the impact of ethnic diversity on public goods is also not consistent between studies. Among the various mechanisms suggested, some evidence points towards heterogeneity in preferences across groups (Alesina et al. 1999; McClendon and Lieberman 2012) as the channel, others point towards weak social networks in ethnically heterogeneous societies (Habyarimana et al. 2007; Miguel and Gugerty 2005).

We test the relationship between ethnic fragmentation and the provision of these types of schools for a cross-section of districts using two large datasets, namely the Indian population census of 2011 and District Information System for Education (DISE) 2013–14 data. We focus on the 2011 census because it is the first census to categorize the schools into private and public, allowing us to study the differential impact. However, the census does not have school-specific information and so to exploit information on infrastructure, enrolment, and the number of teachers in a school, we turn to a much richer dataset provided through the DISE data.[4] We use the DISE data for the year 2013–14 as it was one of the first attempts to combine the data on elementary and secondary/higher secondary schools. Also, it is temporally closest to the 2011 census data. To measure our main independent variable, ethnic fragmentation, we construct the Ethnic Fractionalization Index (EFI), a popularly used measure of fragmentation in the literature, by making use of data from the 1931 census. For our purposes, we consider a caste group as a separate ethnic group. This is reasonable in the Indian setting because Indian society is divided along caste lines, with deep social and economic cleavages.

Our empirical results indicate that schools are negatively impacted in fragmented districts and that this impact is differential for private and public schools. While fragmented districts have a lower number of private unaided schools, these schools are not inferior in terms of quality compared to homogeneous districts. In contrast, ethnic fragmentation has no impact on the number of state or central public schools, but a weak negative impact is observed for schools provided by local government councils. Importantly, all public schools suffer in terms of quality in diverse places. This suggests that while private schools are affected on the extensive margin, public schools are affected on the intensive margin; hence the results suggest that the impact of ethnic fragmentation is sensitive to the mechanism of provision of public goods.

To explain our results, we hypothesize that the negative impact of fragmentation is driven by lack of collective action within the local population. We also suggest that the effect is differential because the two types of schools have very different objectives and therefore collective action is important at different stages for their provision. While private schools have profit maximization as their objective, public schools have equity considerations and aim to provide access to education for all. We therefore do not expect the public schools to vary much in their degree of fragmentation. However, for maintenance of infrastructure and to develop school plans, public schools need the participation of community members. To show that lack of collective action in the local population affects maintenance and quality of public schools, we provide evidence that fragmentation negatively affects the participation of community and school management teams in School Management Development Committee (SMDC) meetings. SMDC meetings, an important component of the Right to Education Act of 2009, represent a platform through which community members work together with the school administration to build plans for school development. Inactive participation by community members in these meetings is suggestive of inefficient coordination among them to maintain the education infrastructure.

Provision of private schools, on the other hand, is guided by a profit motive. Therefore, it is possible that in fragmented areas a lack of collective action might affect mobilization of resources, which is important for the provision of private schools. This is relevant for India as many private schools are run by charitable trusts or societies that require voluntary payments from the community members for their provision. To show that the mechanism of voluntary payments does not work well in diverse places due to weak social networks, we look at the impact of ethnic fragmentation on self-help groups and agricultural societies. These organizations represent a

[4] Detailed description of the DISE data is provided in the data section

platform through which community members voluntarily come together and make monetary contributions that are used for a collective goal. Presence and effective functioning of these groups is an indication that the mechanism of voluntary contributions works well in that area. We then show that the correlation between our indication of the mechanism of voluntary payments and provision of private schools is positive and significant, confirming that weak social networks make it difficult to mobilize resources.

The results of the paper indicate that ethnic fragmentation negatively impacts collective action, which affects the provision and quality of schools. We confirm this result with various robustness checks and also rule out a number of alternative explanations that could be responsible for our observed results. The rest of the paper is structured as follows: Section 2 provides the literature review; Section 3 discusses the data used and the methodology followed to perform our tests. In Section 4 we discuss our results, and present the mechanism in Section 5. We conclude in Section 6.

2 Literature review

The fragmentation–public goods literature is extensive, with many cross-country and country-specific studies. In this review we limit ourselves to an emerging body of work questioning the conventional wisdom in this literature. Singh (2011) emphasizes the use of multidimensional subjective feeling of integrity/division to capture ethnic diversity, a concept she employs to show that in spite of high levels of ethnic heterogeneity, Kerala, an Indian state, has managed to make impressive advancements in the social sector. Gisselquist et al. (2016) show that in the Zambian context the negative relationship between provision of public goods and ethnic heterogeneity does not necessarily hold at the subnational level, specifically when the public good is funded centrally. The idea of this paper is similar to ours in that they hypothesize that the implication of ethnic fragmentation for the provision of public goods would be different if the good is funded centrally as opposed to being provided locally.

However, we have not come across many papers that specifically look at the differential effect of ethnic diversity on public good provision by the structure of ownership. Schündeln (2011) recognizes that the impact of fragmentation can differ by the technology of provision of public goods and shows that diversity in fact increases the willingness to privately contribute for provision of public goods. However, no explicit comparison is provided for the different technologies of provision of public goods. Also, the finding of this paper contradicts our own, probably because, unlike Schundeln, we do not look at the willingness to pay but the actual provision of private and public schools.

The papers that come closest to our findings are those by Miguel and Gugerty (2005) and Chaudhary (2009). Miguel and Gugerty (2005) document that ethnic diversity is associated with lower primary private school funding and school facilities in rural western Kenya. They show that it is difficult to impose social sanctions in ethnically diverse areas leading to collective action failures. However, the funding considered in that paper is private funding for schools and hence it is not clear whether the mechanism suggested would hold for schools that are provided publicly. Chaudhary (2009) finds a lower provision of private primary schools in ethnically diverse districts in colonial India, whereas no impact is observed for schools provided by the provincial government or the local board. The differential impact that she observes is very close to the results of this paper. However, even though she argues that the negative impact on private schools could be due to low demand for schooling by disadvantaged groups or due to difficulty in mobilizing resources in such areas, no formal results are presented in favour of either argument. It is not clear why public schools are not affected, whereas a clear negative effect is observed for private primary schools.

3 Data and methodology

We work with a cross-section of 479 Indian districts in 18 major states[5] to conduct our empirical analysis. Two distinct datasets, the 2011 census and data from the DISE 2013–14, have been used to construct school-level variables. We use the 2011 census because, unlike previous population censuses, village and town directories in this census provide a separate classification for private and public schools, which is what we need to study our research question. We add the number of schools in all the villages and towns in a district to arrive at the district-level aggregate. However, the census does not have school-level information about the quality of schools (in terms of infrastructure and teachers hired) and enrolment, and hence we turn to a much richer dataset provided under the DISE. To review the performance of schools in India and to monitor policies targeted towards schools, information on all the registered schools started to be maintained from 1995 under DISE. For the DISE data, the schools are asked to supply detailed information on a number of school characteristics like infrastructure, enrolment, results, etc. We use the data collected in 2013–14 as it was one of the first attempts to combine the data on elementary schools with secondary and higher secondary schools.[6] Table 1 contains the list of variables that have been constructed using the DISE and the census data. Table 14 provides a summary of the average difference in school characteristics in high and low fragmented districts, where high (low) corresponds to the above (below) median fragmentation levels. The summary statistics indicate that, on average, highly fragmented districts have fewer private schools whereas, quite contrary to findings of the previous literature, the number of public schools is no different to less fragmented districts. In addition, fragmented districts also have a lower proportion of schools with furniture and library facilities, implying lower-quality schools. Thus, the descriptive statistics indicate worse school provision in fragmented places.

The independent variable in all our tests is ethnic fragmentation. For the purpose of this paper, the different ethnic groups that we consider are those divided by caste. In India, the Hindu population (the major religious group) is divided into a number of castes with deep social cleavages that govern social and economic interaction between these caste groups.[7] We follow two criteria for considering a group as a separate ethnic group, as was done by Chadha and Nandwani (2016). First, following Chandra's definition, ethnicity implies nominal membership to an ascriptive category like race or caste in which membership is inherited, which is very well followed in the Indian caste system (Chandra 2007)). Second, ethnic group is salient in determining access to economic and social opportunities. There is ample evidence that disadvantaged castes in India are discriminated against in the labour market (Banerjee and Knight 1985); Madheswaran and Attewell 2007), have lower representation in the political sphere (Pande 2003), and have worse education and health outcomes (Maitra and Sharma 2009; Thorat and Neuman 2012). This indicates that

---

[5] This covers about 93.5 per cent of the population of the country. We do not work with the seven north-eastern states and the union territories.

[6] Before 2012–13, DISE data collected information only on elementary schools in India.

[7] Caste is also called *jati* and there are thousands of *jatis* in India. For our purposes, we include a caste in the sample if it constitutes more than 1 per cent of the state population. This leaves us with about 185 caste groups.

caste in India satisfies the two above-mentioned criteria—it is therefore reasonable to consider a caste as a separate ethnic group.

However, it is important to stress that here we are not considering the identity-based explanation that is common in the Indian context, given the hierarchical nature of caste. That is, our explanation is not driven by the size of disadvantaged castes residing in diverse districts. Instead, we are trying to capture the ability to work together in a heterogeneous setting, which is not necessarily the same as (or can even be over and above) the impact driven by discrimination against the disadvantaged caste. In other words, we are considering heterogeneity as opposed to hierarchy.

Besides caste fragmentation, one could argue for religious fragmentation since religion also follows the two criteria that we consider to classify a group as ethnic. However, there is a reason why we consider caste and not religious fragmentation. According to the 2011 census, religious groups other than the majority Hindu religion make up only 20 per cent of the population. Out of this 20 per cent, the majority (14 per cent) are Muslims, implying that we will essentially be looking at the impact of relative shares of Hindu–Muslim population if we base our fractionalization on religion. While this is an important question in itself, in this exercise we are interested in heterogeneity rather than identity. We also do not consider heterogeneity based on language because it does not satisfy the second criteria as there is no clear evidence of impact of language on attainment of economic outcomes. Also, it is common for states in India to have their own official language, therefore leaving little variation in the language spoken within a state.

Thus, we work with only caste fragmentation and to measure that we need a variable that captures the underlying caste diversity in Indian districts. Unfortunately, caste-level population data are not made available by any census conducted after Independence (in 1947). The data that we can make use of using the latest census is the population proportion of the three broad social groups, Scheduled Castes (SC), Scheduled Tribes (ST), and the rest of the population. Broadly, all the disadvantaged castes are clubbed together in the census within a constitutional category called SC; the middle and upper castes are clubbed into a category named 'others', and the disadvantaged tribes who are placed outside the Hindu caste system are put into a category named ST. These social categories are very broad, with many layers of hierarchy within them; therefore it is not appropriate to use them to capture the actual caste heterogeneity that operates at the ground level. Additionally, there would also be severe endogeneity concerns with these data. Both the mobility of caste groups and provision of public goods are likely to be driven by variables that are difficult to capture, like state institutions or policies. It is also possible that the availability of schools and other public goods itself could result in a particular distribution of caste groups. Therefore, we look for a proxy that captures caste diversity at the district level and is devoid of endogeneity concerns. We make use of the 1931 census to construct a proxy for the ethnic fragmentation measure.

The 1931 census was the last census to record detailed district-level information on population shares of different castes. Construction of ethnic fragmentation using historical data has been done before (Banerjee and Somanathan 2007); Banerjee et al. 2005); Anderson 2011; Suryanarayan 2017) as it has its own advantages. It greatly reduces the endogeneity concern as it is highly unlikely that the caste composition of districts in 1931 would be affected by factors influencing the provision of schools provided much later in time. However, an important assumption we make when working with the 1931 census is that caste proportions in 1931 have not changed much over time and therefore they are representative of present-day caste diversity. To justify this assumption, we rely on the existing evidence and claim the proportions seen in 1931 seem to be a good proxy for present-day fragmentation.

One way in which our assumption would be violated and present-day diversity would be different from historical diversity is if there had been caste-based migration. However, there is evidence that caste-based migration in India is quite low (Munshi and Rosenzweig 2009), probably due to reliance on sub-caste networks of mutual insurance which rarely go beyond village boundaries (Anderson 2011). Although women migrate to different villages for the purpose of marriage, it almost always happens within the same caste. In general, migration in India is low and whatever migration happens, about 62 per cent of it is intra-district, while 24 per cent is inter-district, and 13 per cent is inter-state. A very high proportion of intra-district migration would mean that migration should not substantially affect caste proportion within districts over time. In order to show that caste proportions have not substantially changed over time, Anderson (2011) matched the caste proportion obtained from the 1921 census with the caste proportion obtained from a recent dataset that she collected in 1997–98 with the World Bank.[8] The district caste proportions using the two datasets were very similar, providing additional confidence about the representativeness of the 1931 proportions. Along with the existing evidence, we also checked the correlation between the fragmentation measure based on present-day proportions of social groups and the measure based on the 1931 proportions, and found a positive and highly significant correlation between them, further providing evidence that the 1931 proportions would serve as a good proxy for present-day caste diversity.

As is quite standard in the literature, we measure ethnic diversity using the EFI, given by $1-\sum \beta_i^2$ where β_i is the population share of the ith ethnic group. Since 1931, a lot of new districts have been created, so for the districts that were formed after Independence, we weight the caste figures from the original district according to the area of the new district that was taken from them, following Banerjee and Somanathan (2007).

Using the above-mentioned datasets, we perform two sets of tests to conduct our empirical analysis. The first set of tests corresponds to exploiting the variation in the number and the quality of public and private schools across districts. We estimate the following regression equation:

$$School_{ds} = \alpha_s + \beta_1 cfrag_{ds} + X'_{ds}\delta + \epsilon_{ds} \tag{1}$$

where d indexes the district and s indexes the state. The dependant variable in Equation 1 is a measure of the number and quality of public and private schools. The main parameter of interest is β_1, the coefficient of $cfrag_{ds}$, which measures ethnic fragmentation. X_{ds} contains all the control variables. In Equation 1, the dependent variables have been obtained by aggregating the school-level information up to the district level. But since we have school-level data, we can get more power in our regressions by estimating the relationship between caste fragmentation and school provision at the school level using Equation 2:

$$School_{ds} = \gamma_s + \beta_2 cfrag_{ds} + \phi private_{ids} + \beta_3(cfrag \times private) + X'_{ids}\delta_1 + X'_{ds}\delta_2 + \epsilon_{ids} \tag{2}$$

where i indexes the school, d indexes the district, and s indexes the state. The variable *Private* is a dummy which takes a value of 1 if the school is private and 0 otherwise. The coefficient of the interaction term, β_3, gives us the differential impact of fragmentation on private and public schools.

[8] Note that this exercise was done for two states, Bihar and Uttar Pradesh, and the caste data for the year 1997–98 was collected for the seven castes. Essentially, caste proportions were matched for seven castes in two states over the two datasets.

In this test, we allow the possibility that shocks to school provision and quality could be correlated within a district by clustering the standard errors at the district level.

In both sets of tests we include state dummies, as policies and guidelines for the provision of schools could differ across states. X_{ds} indicates district-level controls, namely urbanization rate, work participation rate, and number of colleges in a district. Economic intuition for adding urbanization rate is that the pattern and the rate at which a district gets urbanized might have some correlation both with the degree of fractionalization (for example, more rural districts might be more fragmented) and the number and quality of schools (urban areas might see more private schools while rural more public). Hence, we do not want fragmentation to pick up the effect of urbanization.[9] We add number of colleges and work participation rate in the regression because if a district, on average, has a higher work participation rate and number of colleges, this might indicate better labour outcomes and in general higher preference for education in that district. Thus, these variables, to some extent, capture the demand for education in a district. Since our motive is to study the impact of ethnic fragmentation on the supply of schools, we control these variables in the regression to partially isolate the impact of ethnic fragmentation on the supply of schools from their demand. X_{ids} indicates school-level controls, namely: age of the school, dummy; *Urban*, indicating whether the school is located in urban areas; and a dummy, *Roadaccess*, indicating whether the school is approachable by all-weather roads. Data on all the district-level controls come from the 2011 census and data on school-level controls have been taken from the DISE dataset.

3.1 Data concerns

Since we rely on the DISE data for most of our empirical analysis, in this section we address some of the concerns related with this dataset to make sure our results are not picking up any bias in the data. The DISE data, which started to be collected from 1995–96, was envisioned to be a census of all existing registered schools. It relies on *self-reported* information on a range of school characteristics, such as the medium of instruction, year of establishment, whether the school is approachable by all-weather roads, whether funds available to the schools are granted under government schemes, information on instruction hours, teachers, enrolment, and school results, among others. Thus, it offers detailed data on school particulars which are compiled at the district level. One of the concerns with the DISE data that has started to be recognized is regarding the coverage of schools. If the schools that report information in the DISE data are systematically different from the ones that do not report information and if this misreporting has a correlation with the fractionalization index, then our results would be biased.

To see whether the sample of schools covered under the DISE data is representative and there is no correlation of the under-reporting of data with the fractionalization index, we construct the difference between the number of schools covered by the census and the DISE. A positive difference would imply that the census covers more schools than the DISE and that divergence in reporting of schools between the census and the DISE is greater. We regress this divergence on the fractionalization index to check the correlation between them. Results are reported in Table 13. We find the coefficient of the fractionalization index to be negative and significant for two of the three dependent variables. This indicates that, if anything, divergence in the reporting of data is lower in diverse districts, and therefore the concern of under-reporting of data is reduced, which

[9] Apart from the urbanization rate, we also performed our analysis with a dummy that indicated whether the district has more than 50 per cent of its population living in urban areas. Our results remain the same with this variable as well.

seems to be good news for us. Also, the negative coefficient seems reasonable as the gap between the two datasets is about three years. In a rapidly growing environment there would be more schools in the DISE since it has data from later years.

The other concern with the DISE data is that it is self-reported. It is therefore likely that schools misreport the information provided in order to project themselves as 'good-quality' schools. This is plausible given that grants and other government benefits to schools in India generally increase with enrolment and teachers hired. An implication of this would be that the DISE captures quality that is better than what is true in practice. However, this would mean that our results are an underestimate of the true effect of fragmentation, at least for public schools, since the results indicate that schools are of poor quality in fragmented places. Therefore, schools projecting themselves as better-quality schools should not be a big concern for us.

4 Results

We begin by discussing district-level regressions. Table 2 reports the regression results from estimating Equation 1 using the census data. The table is divided into two sections, with the first section reporting results without the control variables while second has all the controls. The main reason for presenting a specification without the control variables is the concern that our control variables are likely to be endogenous. As indicated by the table, ethnic fragmentation seems to have a weak negative impact on *only private* elementary and secondary schools,[10] with very similar coefficients across two specifications. The coefficient of elementary private school indicates that if we were to move from a completely homogeneous district to a perfectly fragmented district, we would see a fall in elementary schools (per 1000 population) by 0.19. This amounts to a 68 per cent decline in the number of schools when compared with the average number of elementary schools in our sample (0.276). Clearly, ethnic fragmentation has a detrimental impact on the provision of private schools. However, there is no evidence of negative impact on public schools. Thus, there seems to be a weak differential impact of ethnic fragmentation on private and public schools.

Since the school information in the census data is limited, we now turn to the DISE data to perform the rest of our empirical analysis. We begin by re-estimating Equation 1 to see if a similar differential effect exists with the DISE data. Our finding with the census data is substantiated (both with and without control variables), as reported in Table 3. Ethnic fragmentation lowers the level of private elementary and private secondary schools without having any impact on public schools. The magnitude of the coefficients is also very similar using both datasets. Note that the observed result of no impact on public schools is inconsistent with the established result in the literature, which has shown negative linkages between provision of education and ethnic diversity. However, we think this result is reasonable in the Indian setting as the provision of most public schools in India is the responsibility of either the state or the central government.[11] It does not depend on the ability of communities to act together to raise funds for schools (a common practice for the provision of schools by private players or by the local government), a mechanism commonly suggested for the negative association between public provision of schools and fragmentation. In addition, the governments are bound by equity considerations since many

[10] Elementary schools have classes up to standard 8, so they include both primary and middle schools. Standard 8 means eight years of schooling—students begin elementary school around the age of six and graduate at around 14.

[11] Barring the ones provided by the local government.

policies and programmes in India (Sarva Shiksha Abhiyan, Right to Education Act) aim at universalization of elementary education in the country, which probably leaves little variation in the number of public schools across districts.

Public schools, using the DISE data, can further be classified into those that are managed by local bodies, state government, and central government. Similarly, private schools can be divided into those that receive aid for their operation from the government and those that are unaided private schools. We test the impact of ethnic fragmentation on this finer classification of schools at the elementary level.[12] Results, reported in Table 4, indicate that while local government and aided schools are negatively affected by ethnic fragmentation, there is no effect for either aided schools or schools that are run by state/central government. The observed lack of effect for aided private and government schools suggests that the schools that receive funding from the state or central government for their operation and management seem to be not affected by the local caste composition. On the other hand, private schools that require voluntary payments from community members for their provision, and local government schools that are funded and managed by local government bodies, are affected by the local caste fragmentation. This suggests that schools that depend on the local community for their provision are the ones that are fewer in number in fragmented districts.

Apart from looking at the provision of schools, we also want to understand how *existing* schools in fragmented places perform in terms of quality, as captured by a number of variables. Note that in this exercise we look at the truncated distribution of private schools in fragmented districts. We test the impact of fragmentation on hiring of teachers and infrastructure in Table 5. As indicated in the table, public schools perform poorly in hiring regular teachers, maintaining furniture, and library facilities in fragmented locations. On the other hand, infrastructure in private schools is not inferior. These findings, along with the results of Tables 2 and 3, indicate that while the number of public schools does not fall in fragmented districts, existing schools are of inferior quality and infrastructure. The result for private schools is exactly the opposite: there are fewer overall in fragmented places, but once they are provided they do not seem to be inferior (in comparison to homogeneous districts) in terms of quality. Thus, a differential impact is observed at the intensive margin as well. Explanations for these findings are provided in the mechanism section.

The results presented in Table 5 correspond to district-level regressions in which we have not been able to factor the variation in school-level characteristics within a district. In order to get more power in the regressions and to test the differential impact observed with the district-level data, we present the results for school-level regressions. Table 6 tests the impact of ethnic fragmentation on the number of teachers. Results indicate that public schools are of somewhat lower quality in fragmented places as captured by lower teachers per number of classes, a lower proportion of teachers that are regular, and a higher number of single-teacher public schools.[13] Additionally, we

---

[12] We run the test only at the elementary level because at secondary or higher secondary levels most of the public schools are run by the state government and most of the private schools are unaided, leaving little variation to exploit.

[13] Note that there are many fewer observations in columns 1 and 2 compared to columns 3 and 4. This is because information on enrolment in the DISE data is missing for 15 per cent of the schools in our sample. This could raise concerns of selection bias due to misreporting. To check how different the two groups of schools are, we conducted a comparison of the means test and found that schools with missing enrolment data have a higher number of teachers per classrooms, furniture, are more urban, a higher proportion of them are private schools, and they are accessible by roads. We partially address this issue by controlling for some of these factors like *roadaccess*, *age*, *urban*, and *private* dummy in all our school-level regressions. However, we recognize that this does not completely solve the issue as we cannot control for infrastructure variables since they serve as our dependent variables. To get around this issue further, in the following columns/tables, we also look at variables that do not depend on the enrolment data for their construction.

observe a positive differential impact for private schools, implying that the negative impact of fragmentation for public schools goes down when one looks at private schools. Thus, the school-level results are consistent with what was observed with district-level regressions.

5 Mechanism

We now present the plausible mechanism through which, we believe, ethnic fragmentation lowers the number of private schools and negatively impacts the quality of public schools. We hypothesize that the negative impact on schools is driven by a low tendency for collective action in fragmented places. The effect is differential because the two types of schools have very different objectives due to which collective action is important at different stages in their provision. Public schools in India are provided by the state and the provision is guided by equity considerations, and therefore local demographic structure might not play a role in their provision. However, public schools require the participation of local community members to maintain infrastructure and to build school development plans. This is even more true after the enactment of the Right to Education Act of 2009, which mandated the constitution of a school development body mainly consisting of the local population to build school development plans. On the other hand, private schools need to raise money to be able to provide schools. Many private schools in India are provided by charitable trusts/societies, which require voluntary payments from its members. Therefore, if private players find it difficult to raise money in fragmented places, due to weak social networks, then their provision is likely to be affected.

We provide evidence for the lack of collective action in the case of public schools by looking at the impact of ethnic fragmentation on SMDC meetings. All public and aided schools are required to constitute SMDCs as per the Right to Education Act of 2009. An SMDC is an institutional mechanism to allow community members, members of the local bodies, parents, SC/ST members, along with the school management team, to build and implement school development plans. Effective implementation of SMDCs and an active participation of community members in SMDC meetings is an indication of teamwork and coordination among the local community to build school plans. If our hypothesis is true, then fragmented districts should see less participation by community members in SMDC meetings, and this is exactly what we find in Table 7. The coefficient of SMDC meetings is negative and significant, indicating low frequency of these meetings in fragmented districts. Additionally, participation by local body members, parents, and members of SC/ST groups in these meetings is low. We also perform this test at the school level; the results are reported in Table 8. Table 8 shows that, consistent with district-level regressions, there are fewer SMDC meetings in fragmented districts, and these meetings see lower participation of its members, specifically members of local bodies. These findings confirm that fragmented locations lack participation by community members for the development and management of school plans, lending support to our hypothesis.

For private schools, our conjecture is that collective action in fragmented districts is not strong enough, due to which the mechanism of voluntary payments by community members to raise private schools does not work well. We test the first part of this assertion by looking at the impact of caste fragmentation on the number of self-help groups and agricultural credit societies.[14] Self-help groups and agricultural credit societies represent one of the many ways by which community members voluntarily come together to save a particular amount, which is then used in times of

[14] We obtain the data on these two variables from the village directories of the 2011 Census.

distress for mutual help. Even though this is not the best way to test the mechanism of voluntary payments to raise schools in fragmented places, it gives us an indication of the ability to work collectively as these are voluntary groups in which members are connected to each other and can use this network to raise money. Table 9 shows that fragmentation lowers the number of self-help groups and agricultural societies, lending evidence to the fact that the collective action for voluntary payment mechanism is weak in fragmented districts.

In Table 10 we look at the correlation between our measure of collective action to raise funds and the number of private and public schools. The results indicate that self-help groups and agricultural societies are positively correlated with private schools, but there does not seem to be a clear impact on public schools. This indicates that homogeneous districts with well-functioning self-help groups and agricultural societies—for our purpose a proxy for collective action to raise funds—see higher numbers of private schools. This would then imply that fragmented districts are not that closely knit and do not have strong social networks to facilitate the mobilization of funds for private schools. We also try to explain our finding that existing private schools are not of inferior quality in fragmented districts. An explanation for this additional result could be that existing private schools in fragmented places are a result of successful collective action, leading to mobilization of funds. This successful collective action, which is responsible for the provision of schools, could also result in provision of better infrastructure and teachers in fragmented places.

5.1 Alternative explanation

To make sure that the observed effect of ethnic fragmentation on private and public schools is not driven by channels other than those suggested by us, we consider a number of alternative explanations. The first explanation that we consider is that of low economic activity in fragmented places. There is evidence that ethnic fragmentation negatively impacts economic growth (Easterly and Levine 1997). Low economic growth by suppressing the demand for schools and/or by lowering the capacity to provide schools, both by private players and the government, could lower the provision of schools in fragmented areas. To see if our results are robust to economic activity, we control district gross domestic product (GDP) in our main regression specification. Data for district GDP have been taken from the Open Government Database (OGD) of the Indian government. As reported in the first section of Table 11, the coefficient of private schools continues to remain negative and significant and there is no impact on the coefficient of public schools, indicating that the observed effect of ethnic fragmentation is not driven by low economic activity in fragmented places. The reason we do not include district GDP in our main test and present this result as a robustness check is that the data on district GDP for 2010–11 are available only for 12 out of the 18 states that we consider. For the remaining six states, we use the latest available district GDP data, which leads to some inconsistency in the timing of district GDP data. For this reason, we consider other indicators of economic activity like urbanization and work participation rate, both of which are highly correlated with district GDP, in our main tests.

Apart from the possibility of average district domestic product being affected by the local ethnic diversity, there is also evidence that ethnic diversity widens the distribution of consumption/income (Chadha and Nandwani 2016). An unequal distribution of income could result in different preferences over the distribution of schools, which could lead to low provision of schools in diverse areas. Therefore, we control the Gini coefficient of real per capita consumption expenditure obtained from the National Sample Survey round conducted in 2011–12. This allows us to capture the effect of diversity on schools over and above the effect that could be driven by an unequal distribution. The second part of Table 11 shows that districts with higher inequality have low provision of public schools and no impact on the provision of private schools. Importantly, results also show that inequality does not alter the relationship between fragmentation and the provision of schools, as we still observe negative and significant impacts for private schools and no impact for public schools.

As mentioned before, there has been social and economic discrimination in India along caste lines. Therefore, it is possible for fragmentation to mimic the economic disparity among the different caste groups, leading to a negative impact on the provision of schools. If this is the case, then the coefficient of the fractionalization index should become insignificant once we control for economic inequality between social groups. We do so in the third part of Table 11 by adding the group Gini coefficient constructed again using the NSS data from 2011–12, and find that the coefficients for private schools are still negative and significant, with only slight change in their magnitude. This confirms that economic disparity among the diverse population is not driving our result.

In the past 35 years or so, politics in India has become quite fragmented and it is believed to be even more so in socially diverse regions (due to the emergence of caste-based parties). Along with this apparent correlation with caste diversity, the nature of political competition (dominance of multiple or two parties) has also been shown to matter for the provision of public goods (Chhibber and Nooruddin 2004). Therefore, to make sure our results are robust to political make-up in a district, we add political fragmentation in the election year 2009, constructed using vote shares of different political parties in a district.[15] The results, reported in the fourth section of Table 11, show that ethnic fragmentation continues to exert negative influence on the provision of private schools even when political competition is added. We also employ other measures of the political environment, namely vote share of the winning party, vote share of national and state parties, and the proportion of voters per population in a district. However, our result remains robust to all the measures of political environment (results, not reported, can be obtained from the authors on request).

As mentioned earlier in the paper, we are not relying on the hierarchy within the caste structure to explain our results. However, the standard caste identity theory could also very well explain our results if heterogeneous districts have a high proportion of disadvantaged castes who either do not have the resources to provide schools or have low preference for education. To see if heterogeneity explains our results over and above hierarchy, we add the population proportion of disadvantaged groups, SC, and ST in our regression specification, reported in the fifth section of Table 11. The coefficients of private schools continue to remain negative and significant even after the addition of disadvantaged groups, confirming that results of the paper are driven due to failure of collective action in heterogeneous societies rather than discrimination.

Additionally, we test whether members of different caste groups do not want to study together in the same school; if this is true, this can cause segregation of schools by caste. In that situation, it might not be viable to provide schools (specifically by private players) for students belonging to a minority caste, a reason driving low provision of schools rather than the mechanisms that we consider. This mechanism has also been suggested by Alesina et al. (1999), who point out that neighbourhood segregation by ethnicity can negatively impact the provision of public goods. To test whether this is the case, we create enrolment fragmentation (that is, how heterogeneous is the enrolment by social groups in a given school in a class?) for each class by using information on enrolment by social groups in the DISE. The DISE data report enrolment from SC, ST, Other

[15] The unit at which parliamentary elections happen, parliamentary constituencies, are different from districts, which is the unit of our analysis. We match constituencies to districts using delimitation commission reports of the Election Commission of India. The authors can be contacted for more details on the matching procedure.

Backwards Classes (OBCs), and the general category. Since we consider very broad categories to construct our fragmentation measure of the school, our ethnic fragmentation measure for this test is also created using the broad social groups (SC, ST, others) reported in the 2011 census. The results reported in Table 12 show that districts that are fragmented (by social groups) seem to have a diverse class composition (with no negative differential impact for private schools). The results are only reported for classes 1, 5, 10, and 12 for brevity, even though the result holds for other classes also. This suggests that fragmentation does not lead to segregation in enrolment (at least for existing schools) and therefore the impact of ethnic fragmentation does not seem to be driven by segregation in enrolment.

Apart from these formal tests, we also make sure that any particular state in our sample is not driving the results. We sequentially drop one state at a time from our sample, but the results do not indicate sensitivity to any particular state. Thus, overall, we rule out all the possible alternate explanations that we can think of that could be responsible for a negative impact on the provision of schools.

6 Conclusion

This paper provides evidence that public schools are of lower quality in diverse districts in India and that private schools also have not been able to take up this space created by lower-quality public schools. This clearly shows that among a range of social and economic problems that diverse districts might face, one of them is fewer private and more lower-quality public schools. Given that schools represent an important investment for human capital accumulation, this finding has implications for performance of these districts in terms of achieving desirable future education outcomes. The results of this paper suggest that the reason schools are affected in diverse areas is that the ability to act collectively is low in fragmented districts. Since the two types of schools are provided with very different objectives, low collective action results in a differential effect for the two schools. Private schools are lower in number but are not impacted in terms of their quality in fragmented places. On the other hand, state and central schools that do not rely on local caste composition for their provision seem to be not affected by caste diversity, at least on the extensive margin. The results of the paper suggest that the relationship between ethnic fragmentation and public goods provision is not independent of the type of public good considered. The results also indicate that schools are not provided efficiently in districts in India that are highly fragmented and thus further research needs to be done to come up with the most efficient way to provide schools in diverse places.

References

Alesina, A., and E. La Ferrara (2005). 'Ethnic Diversity Economic Performance'. *Journal of Economic Literature*, 43(3): 762–800.

Alesina, A., R. Baqir, and W. Easterly (1999). 'Public Goods and Ethnic Divisions'. *The Quarterly Journal of Economics*, 114(4): 1243–84.

Anderson, S. (2011). 'Caste as an Impediment to Trade'. *American Economic Journal: Applied Economics*, 3(1), 239–63.

Banerjee, B., and J.B. Knight (1985). 'Caste Discrimination in the Indian Urban Labour Market'. *Journal of Development Economics*, 17(3): 277–307.

Banerjee, A., and R. Somanathan (2007). 'The Political Economy of Public Goods: Some Evidence from India'. *Journal of Development Economics*, 82(2): 287–314.

Banerjee, A., L. Iyer, and R. Somanathan (2005). 'History, Social Divisions, and Public Goods in Rural India'. *Journal of the European Economic Association*, 3(2–3): 639–47.

Chadha, N., and B. Nandwani (2016). 'Ethnic Fragmentation, Public Good Provision, and Inequality in India, 1988–2012'. Working Paper 2016/105. Helsinki: UNU-WIDER.

Chandra, K. (2007). *Why Ethnic Parties Succeed: Patronage and Ethnic Head Counts in India.* Cambridge: Cambridge University Press.

Chaudhary, L. (2009). 'Determinants of Primary Schooling in British India'. *The Journal of Economic History*, 69(1): 269.

Chhibber, P., and I. Nooruddin (2004). 'Do Party Systems Count? The Number of Parties and Government Performance in the Indian States'. *Comparative Political Studies*, 37(2): 152–87.

Easterly, W., and R. Levine (1997). 'Africa's Growth Tragedy: Policies and Ethnic Divisions'. *The Quarterly Journal of Economics*, 112(4): 1203–50.

Elbadawi, I., and N. Sambanis (2002). 'How Much War Will We See? Explaining the Prevalence of Civil War'. *Journal of Conflict Resolution*, 46(3): 307–34.

Gisselquist, R.M., S. Leiderer, and M. Niño-Zarazúa (2016). 'Ethnic Heterogeneity and Public Goods Provision in Zambia: Evidence of a Subnational Diversity Dividend'. *World Development*, 78: 308–23.

Habyarimana, J., M. Humphreys, D.N. Posner, and J.M. Weinstein (2007). 'Why Does Ethnic Diversity Undermine Public Goods Provision?'. *APSR*, 101(4), 709–25.

Madheswaran, S., and P. Attewell (2007). 'Caste Discrimination in the Indian Urban Labour Market: Evidence from the National Sample Survey'. *Economic and Political Weekly*, 42: 4146–53.

Maitra, P., and A. Sharma (2009). 'Parents and Children: Education Across Generations in India'. Paper presented at the 5th Annual Conference on Economic Growth and Development, ISI Delhi, Delhi.

McClendon, S., and G.H. Lieberman (2012). 'The Ethnicity–Policy Preference Link in Sub-Saharan Africa'. *Comparative Political Studies*, 46(5): 574–602.

Miguel, E., and M.K. Gugerty (2005). 'Ethnic Diversity, Social Sanctions, and Public Goods in Kenya'. *Journal of Public Economics*, 89(11–12): 2325–68.

Munshi, K., and M. Rosenzweig (2009). 'Why is Mobility in India So Low? Social Insurance, Inequality, and Growth'. Working Paper 14850. Cambridge, MA: National Bureau of Economic Research.

Pande, R. (2003). 'Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India'. *The American Economic Review*, 93(4): 1132–51.

Singh, P. (2011). 'We-ness and Welfare: A Longitudinal Analysis of Social Development in Kerala, India'. *World Development*, 39(2): 282–93.

Schündeln, M. (2013). 'Ethnic Heterogeneity and the Private Provision of Public Goods'. *The Journal of Development Studies*, 49(1): 36–55.

Suryanarayan, P. (2017). 'When Do the Poor Vote for the Right-Wing and Why: Status Inequality and Vote Choice in the Indian States'. Working Paper 2017/20. Helsinki: UNU-WIDER.

Thorat, S., and K.S. Neuman (2012). *Blocked by Caste: Economic Discrimination in Modern India*. Oxford: Oxford University Press.

Table 1: Variable definition and data source

Variable	Definition	Source
G (P) element	Public (private) elementary schools (per thousand population) in a district	Census and DISE
G (P) secondary	Public (private) secondary schools (per thousand population) in a district	Census and DISE
G (P) highsecondary	Public (private) higher secondary schools (per thousand population) in a district	Census and DISE
Local elementary	Elementary schools (per thousand population) in a district provided by local bodies	DISE
Aided elementary	Aided private elementary schools (per thousand population) in a district	DISE
Unaided elementary	Unaided private elementary schools (per thousand population) in a district	DISE
G (P) Reg teachers	Average number of regular teachers in a public (private) school in a district	DISE
G (P) Contract teachers	Average number of contract teachers in a public (private) school in a district	DISE
G (P) furniture	Proportion of public (private) schools with furniture in a district	DISE
G (P) library	Proportion of public (private) schools with library in a district	DISE
SMDC meetings	Average number of SMDC meetings held in a school in a district	DISE
Males (females)	Average number of male (female) members in SMDC	DISE
Local gvt males (females)	Average number of male (female) members from local bodies in SMDC	DISE
SCandST males (females)	Average number of SC/ST male (female) members in SMDC	DISE
Private	Dummy, equals 1 if the schools is private; 0 otherwise	DISE
(Reg) teachers enrol	Total number of (regular) teachers divided by total enrolment in a school	DISE
(Reg) teachers/classes	Total number of (regular) teachers divided by the highest class in a school	DISE
Reg/total teachers	Proportion of regular teachers in a school	DISE
Single (reg) teach schools	Dummy indicating if a school has only one (regular) teacher	DISE
Classrooms enrol (classes)	Total number of classrooms per total enrolment (classrooms) in a school	DISE

Good condn/total classrooms	Proportion of classrooms in good condition	DISE
Library	Dummy indicating if the school has a library	DISE
SHgroup	Proportion of villages having self-help groups in a district	Census
Agrisociety	Proportion of villages having agricultural credit societies in a district	Census
Urban	Dummy, equals 1 if the school is located in an urban area; 0 otherwise	DISE
Road access	Dummy, equals 1 if the school is accessible by all-weather roads; 0 otherwise	DISE
Estdyear	Age of the school	DISE
Toilets b (g)	Total number of toilets for boys (girls)	DISE
Enrol c_i	Total number of students enrolled in class i in a school	DISE
$frag_i$	Ethnic fragmentation in class i in a school	DISE

Source: authors.

Table 2: Effect of ethnic fragmentation on school provision

Panel A

	(1) G_element	(2) P_element	(3) G_second	(4) P_second	(5) G_hrsecond	(6) P_hrsecond
Ethnic frag	0.0948	−0.1951*	−0.1347⁺	−0.0533⁺	−0.0337	−0.0221
	(0.875)	(0.083)	(0.140)	(0.102)	(0.539)	(0.283)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No

Panel B

	(1) G_element	(2) P_element	(3) G_second	(4) P_second	(5) G_hrsecond	(6) P_hrsecond
Ethnic frag	0.3330	−0.1922*	−0.1224	−0.0548*	−0.0274	−0.0254
	(0.537)	(0.086)	(0.169)	(0.090)	(0.613)	(0.195)
Urbanization	−1.0617***	0.0570⁺	−0.0952***	0.0273**	−0.0325*	0.0262***
	(0.000)	(0.141)	(0.002)	(0.015)	(0.085)	(0.000)
Workrate	3.2785***	−0.1090	0.2183**	−0.0275	0.1746***	−0.0599**
	(0.000)	(0.425)	(0.045)	(0.486)	(0.009)	(0.013)
College_dis	−0.0000	0.0000	−0.0000	0.0000⁺	−0.0000	0.0000
	(0.792)	(0.511)	(0.212)	(0.140)	(0.505)	(0.286)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	405	405	405	405	405	405

Notes: *p*-values in parentheses.

This is a district-level regression. Fragmentation numbers are based on the 1931 census. Results presented in the first section of the table do not use any control variables, whereas the second part has all the controls.

$^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors based on data sources listed in Table 1.

Table 3: Effect of ethnic fragmentation on school provision (DISE data)

Panel A

	(1) G_element	(2) P_element	(3) G_second	(4) P_second	(5) G_highsecond	(6) P_highsecond
Ethnic frag	−0.0695	−0.1265**	−0.0190	−0.0485**	−0.0011	−0.0011
	(0.794)	(0.011)	(0.596)	(0.034)	(0.964)	(0.936)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No

Panel B

	(1) G_element	(2) P_element	(3) G_second	(4) P_second	(5) G_highsecond	(6) P_highsecond
Ethnic frag	0.0122 (0.938)	−0.1586*** (0.001)	−0.0075 (0.777)	−0.0572** (0.011)	0.0054 (0.765)	−0.0072 (0.596)
Urbanization	−0.4299*** (0.000)	0.0712*** (0.000)	−0.0202** (0.031)	0.0254*** (0.001)	−0.0025 (0.698)	0.0204*** (0.000)
College_dis	0.0675*** (0.000)	−0.0007 (0.532)	0.0104*** (0.000)	−0.0001 (0.821)	0.0071*** (0.000)	0.0010*** (0.009)
Workrate	1.1305*** (0.000)	−0.1202** (0.041)	0.0629* (0.056)	−0.0805*** (0.004)	0.0368+ (0.121)	−0.0530*** (0.003)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	396	394	396	394	384	380

Notes: *p*-values in parentheses.

This is a district-level regression. Fragmentation numbers are based on the 1931 census. Results presented in the first section of the table do not use any control variables, whereas the second part has all the controls.

+ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.

Table 4: Effect of ethnic fragmentation on school provision

	(1) Local elementary	(2) G_elementary	(3) Aided_elementary	(4) Unaided_elementary
Ethnic frag	−0.1121+	−0.2039	−0.0004	−0.1583***
	(0.117)	(0.251)	(0.986)	(0.000)
Urbanization	−0.1739***	−0.1933***	−0.0097	0.0810***
	(0.000)	(0.002)	(0.181)	(0.000)
College_dis	0.0028*	0.0661***	−0.0003	−0.0004
	(0.094)	(0.000)	(0.558)	(0.679)
Workrate	0.0459	0.1911	−0.0713***	−0.0484
	(0.604)	(0.385)	(0.006)	(0.358)
State FE	Yes	Yes	Yes	Yes
Observations	396	396	396	396

Notes: *p*-values in parentheses.

This is a district-level regression. Fragmentation numbers are based on the 1931 census.

+ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.

Table 5: Effect of ethnic fragmentation on hiring of teachers and school infrastructure

	(1) G_furniture	(2) P_furniture	(3) G_single teach sch	(4) P_single teach sch	(5) G_electricity	(6) P_electricity
Ethnic frag	−0.2982**	−0.1706	0.1216**	−0.0126	−0.1160	−0.0698
	(0.028)	(0.236)	(0.036)	(0.855)	(0.304)	(0.540)
Urbanization	0.0996**	0.1412***	−0.0856***	−0.0182	0.1969***	0.3038***
	(0.037)	(0.005)	(0.000)	(0.454)	(0.000)	(0.000)
College_dis	0.0097***	0.0029	0.0064***	−0.0021	−0.0058**	0.0021
	(0.002)	(0.381)	(0.000)	(0.180)	(0.023)	(0.410)
Workrate	−0.2772*	−0.4488**	0.1624**	0.0260	−0.3635***	0.0238
	(0.100)	(0.013)	(0.024)	(0.763)	(0.010)	(0.867)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	407	405	407	405	407	405

Notes: *p*-values in parentheses.

This is a district-level regression. Fragmentation numbers are based on the 1931 census.

$^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.

Table 6: Impact of fragmentation on teachers

	(1) tch_enrol_all	(2) reg_tch_enrol_all	(3) Teachers/classes	(4) Reg teachers/classes	(5) Regular/total teachers
Ethnic frag	−0.0266	−0.0299	−0.1365	−0.1987*	−0.1995***
	(0.251)	(0.151)	(0.209)	(0.069)	(0.003)
Private	−0.0925***	−0.0801***	−0.1639	0.0719	0.1853
	(0.002)	(0.001)	(0.549)	(0.758)	(0.177)
Cfrag_private	0.1013***	0.0928***	0.6020**	0.3789+	−0.0760
	(0.001)	(0.000)	(0.040)	(0.132)	(0.608)
Urban	0.0038***	0.0040***	0.2044***	0.1965***	0.0085+
	(0.002)	(0.001)	(0.000)	(0.000)	(0.148)
Road access	−0.0028***	−0.0011	0.1016***	0.0972***	0.0177***
	(0.006)	(0.209)	(0.000)	(0.000)	(0.006)
Age	−0.0002***	−0.0002***	0.0039***	0.0040***	0.0003**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.038)
College_dis	0.0033***	0.0027***	−0.0101***	−0.0087***	−0.0010
	(0.000)	(0.000)	(0.000)	(0.000)	(0.486)
Workrate	−0.0101	−0.0129	−0.3027**	−0.3302***	−0.0506
	(0.545)	(0.417)	(0.011)	(0.006)	(0.417)
tch_enrol_all					−0.0234***
					(0.009)
State FE	Yes	Yes	Yes	Yes	Yes
Observations	412,948	412,948	505,881	505,881	408,726

Notes: *p*-values in parentheses.

This is a school-level regression. Fragmentation numbers are based on the 1931 census.

Standard errors are clustered at the district level.

$^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.

Table 7: Effect of ethnic fragmentation on SMDC

	(1) SMDC meetings	(2) Total members	(3) Local gvt	(4) SC/ST	(5) Parents
Ethnic frag	−3.2387**	−5.2882**	−0.7603*	−1.0904***	−2.0100***
	(0.014)	(0.039)	(0.050)	(0.009)	(0.001)
Urbanization	−0.8645*	−4.1807***	−0.6276***	−0.4560***	−0.7647***
	(0.061)	(0.000)	(0.000)	(0.002)	(0.000)
College_dis	−0.0158	−0.0078	0.0294***	0.0117	0.0469***
	(0.595)	(0.893)	(0.001)	(0.209)	(0.001)
Workrate	0.8689	3.5108	0.5427	2.7785***	1.9184**
	(0.593)	(0.269)	(0.259)	(0.000)	(0.012)
State FE	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407

Notes: *p*-values in parentheses.

This is a district-level regression. Fragmentation numbers are based on the 1931 census.

$^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.

Table 8: Impact of fragmentation on SMDC meetings and participation

	(1) SMDC meetings	(2) Local gvt males	(3) Local gvt females	(4) Males	(5) Females
Ethnic frag	−3.7650**	−0.4779+	−0.3459*	−3.4775+	−1.8948
	(0.027)	(0.115)	(0.075)	(0.135)	(0.284)
Urban	−0.7634***	−0.3260***	−0.1705***	−1.8975***	−0.8821***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Road access	0.0311	−0.0171	−0.0077	−0.0015	−0.0410
	(0.675)	(0.518)	(0.638)	(0.992)	(0.764)
Age	0.0185***	0.0069***	0.0037***	0.0392***	0.0209***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
College_dis	−0.0226	0.0151	0.0143	−0.0759+	−0.0146
	(0.286)	(0.383)	(0.206)	(0.124)	(0.642)
Workrate	0.9643	0.7848***	0.4149**	6.9122***	2.5685+
	(0.578)	(0.007)	(0.046)	(0.002)	(0.108)
State FE	Yes	Yes	Yes	Yes	Yes
Observations	79,308	79,308	79,307	79,308	79,308

Notes: *p*-values in parentheses.

This is a school level-regression. Fragmentation numbers are based on the 1931 census.

Standard errors are clustered at the district level.

$^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.

Table 9: Impact of fragmentation on self-help groups and agricultural credit societies

	(1) SHgroup	(2) Agrisociety
Ethnic frag	−0.5411***	−0.4178**
	(0.001)	(0.022)
Urbanization	0.0068	0.0980*
	(0.879)	(0.052)
Workrate	−0.1509	−0.2948⁺
	(0.349)	(0.104)
Litrate_11	0.0340	−0.1155
	(0.708)	(0.258)
State FE	Yes	Yes
Observations	274	274

Notes: *p*-values in parentheses.

This is a district-level regression. Fragmentation numbers are based on the 1931 census.

⁺ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.

Table 10: Impact of self-help groups and agricultural credit societies on school provision

	(1) P_element	(2) G_element	(3) P_second	(4) G_second	(5) P_highsecond	(6) G_highsecond
SHgroup	0.0711***	−0.3506***	0.0222***	0.0075	0.0069*	−0.0037
	(0.000)	(0.000)	(0.000)	(0.157)	(0.071)	(0.238)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Agrisociety	0.0491**	−0.2776***	−0.0014	−0.0012	0.0033	0.0097**
	(0.042)	(0.000)	(0.873)	(0.858)	(0.495)	(0.018)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	331	331	330	331	311	315

Notes: *p*-values in parentheses.

This is a district-level regression. Controls include urbanization, work participation rate, lagged net state domestic product, proportion of development expenditure in a state. Fragmentation numbers are based on the 1931 census.

Standard errors are clustered at the district level.

$^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.

Table 11: Alternative explanations

	(1) G_element	(2) P_element	(3) G_second	(4) P_second	(5) G_highsecond	(6) P_highsecond
Ethnic frag	−0.1214	−0.1550***	−0.0253	−0.0495**	−0.0057	−0.0012
	(0.634)	(0.003)	(0.508)	(0.039)	(0.828)	(0.937)
Dist GDP (in 100 cr.)	−0.0006***	0.0000*	−0.0000**	0.0000+	−0.0000	0.0000
	(0.000)	(0.053)	(0.028)	(0.106)	(0.308)	(0.208)
Ethnic frag	0.0361	−0.1496***	−0.0020	−0.0544**	0.0078	−0.0064
	(0.814)	(0.002)	(0.939)	(0.016)	(0.665)	(0.642)
Gini	−0.4795***	−0.0134	−0.0626***	0.0054	−0.0381***	0.0119
	(0.000)	(0.726)	(0.003)	(0.765)	(0.009)	(0.277)
Ethnic frag	0.1384	−0.1893***	0.0041	−0.0820***	0.0133	−0.0189
	(0.397)	(0.000)	(0.883)	(0.000)	(0.490)	(0.188)
Gini group	−0.3442**	−0.0391	−0.0383	−0.0278	−0.0186	0.0034
	(0.050)	(0.457)	(0.203)	(0.259)	(0.369)	(0.824)
Ethnic frag	0.0086	−0.1566***	−0.0079	−0.0569**	0.0052	−0.0074
	(0.956)	(0.001)	(0.766)	(0.011)	(0.775)	(0.591)
Political frag	0.0430	−0.0337	−0.0017	−0.0151	0.0040	−0.0099
	(0.700)	(0.323)	(0.930)	(0.347)	(0.758)	(0.310)
Ethnic frag	−0.0433	−0.1307***	−0.0181	−0.0461**	−0.0024	−0.0016
	(0.780)	(0.007)	(0.492)	(0.036)	(0.894)	(0.904)
SC_prop	−0.1666	0.0842*	−0.0205	0.0220	−0.0288+	0.0261*
	(0.293)	(0.089)	(0.446)	(0.327)	(0.123)	(0.061)
ST_prop	0.3776***	−0.0160	0.0232**	−0.0233**	0.0093	−0.0053
	(0.000)	(0.439)	(0.040)	(0.014)	(0.235)	(0.375)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	378	375	378	376	366	360

Notes: *p*-values in parentheses.

This is a district-level regression. Fragmentation numbers are based on the 1931 census. The first part of the table does not have any control variables as dist GDP and all the other district-level controls are highly correlated. The second part of the table has Gini coefficient of real per capita consumption expenditure computed using the NSS round conducted in 2011–12. The third part has group Gini coefficient capturing inequality between caste groups. The fourth part of the table has political frag, which is the fractionalization index created using the vote share of political parties in a district in the parliamentary elections held in 2009. The fifth part controls the population proportion of SC and ST groups.

+ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on sources noted above and in Table 1.

Table 12: Impact of fragmentation on distribution of students by caste

	(1) frag_1	(2) frag_5	(3) frag_10	(4) frag_12
Frag_cen	0.2251***	0.4032***	0.3060***	0.3470***
	(0.003)	(0.000)	(0.000)	(0.000)
Private	−0.0227	0.1132***	−0.0054	−0.0285
	(0.400)	(0.001)	(0.789)	(0.205)
Frag_private	0.1925***	−0.1339^{+}	−0.0049	−0.0570
	(0.005)	(0.121)	(0.919)	(0.315)
Enrol_c1	−0.0001^{+}			
	(0.102)			
Urban	0.0015	−0.0041	−0.0033	−0.0017
	(0.834)	(0.546)	(0.492)	(0.772)
Road access	0.0194***	0.0246***	0.0156***	0.0119*
	(0.009)	(0.000)	(0.002)	(0.094)
Age	0.0005***	0.0011***	0.0011***	0.0009***
	(0.001)	(0.000)	(0.000)	(0.000)
College .dis	−0.0005	−0.0015	−0.0031**	−0.0020
	(0.840)	(0.619)	(0.046)	(0.264)
Workrate	−0.4091***	−0.2303^{+}	−0.3232***	−0.3032***
	(0.001)	(0.113)	(0.001)	(0.004)
Enrol_c5		0.0000		
		(0.889)		
Enrol_c10			0.0001***	
			(0.000)	
Enrol-c12				0.0001***
				(0.000)
State FE	Yes	Yes	Yes	Yes
Observations	23,309	33,355	81,706	29,864

Notes: *p*-values in parentheses

This is a school-level regression. Fragmentation numbers are based on the 2011 census.

Standard errors are clustered at the district level.

$^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: authors, based on data sources listed in Table 1.

Table 13: Impact of fragmentation on data divergence

	(1) diff_cd_element	(2) diff_cd_second	(3) diff_cd_highsecond
Ethnic frag	−1,196.3765	−345.0466**	-131.5926^{+}
	(0.156)	(0.036)	(0.116)
State FE	Yes	Yes	Yes
Observations	393	393	372

Notes: *p*-values in parentheses.

Dependent variables in all three columns is the difference in the number of schools at the elementary, secondary and the higher secondary level, respectively. Fragmentation numbers are based on the 1931 census.

$^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.

Table 14: Summary statistics.

Variable	Mean	Mean (highly fragmented)	Mean (less fragmented)	Difference	t value
G_element	0.4907	0.5113	0.4525	0.059	(−1.77)
G_second	0.0451	0.0473	0.0418	0.006	(−1.05)
G_highsecond	0.0191	0.0187	0.0181	0.001	(−0.17)
P_element	0.1210	0.1015	0.1435	−0.0420***	(−5.93)
P_second	0.0469	0.0365	0.0591	−0.0226***	(−7.58)
P_highsecond	0.0236	0.0175	0.0296	−0.0121***	(−7.21)
Reg_teachers_gvt	4.0889	4.1213	4.0056	0.116	(−0.53)
Reg_teachers_pvt	8.4295	9.0967	8.0387	1.058**	(−3.00)
G_furniture	0.5620	0.4980	0.6079	−0.110***	(−4.87)
P_furniture	0.8228	0.7826	0.8645	−0.0820***	(−4.78)
G_library	0.6082	0.5253	0.6851	−0.160***	(−6.43)
P_library	0.5924	0.5031	0.6695	−0.166***	(−6.25)
Urbanization	0.2477	0.2402	0.2578	−0.018	(−1.03)
College_dis	3.3754	3.3575	3.3369	0.021	(−0.06)
Workrate	0.4117	0.4196	0.4025	0.0172*	(−2.42)

$^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors, based on data sources listed in Table 1.