I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10595

Where's the Teacher? How Teacher Workplace Segregation Impedes Teacher Allocation in India

Sonja Fagernäs
Panu Pelkonen

FEBRUARY 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10595

Where's the Teacher? How Teacher Workplace Segregation Impedes Teacher Allocation in India

Sonja Fagernäs
University of Sussex

Panu Pelkonen
University of Sussex, CEE at LSE and IZA

FEBRUARY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Where's the Teacher? How Teacher Workplace Segregation Impedes Teacher Allocation in India*

Social or ethnic segregation can impede the equitable allocation of public resources in developing countries. We study an under-explored dimension; the allocation of public sector teachers in India. Using a register database for 2006-12, we construct indicators for the equality of teacher allocation and workplace segregation of teachers by gender and caste within districts. While pupil-teacher ratios have improved, the equality of teacher allocation has not. We show that allocation and segregation are connected; in districts with a higher degree of initial teacher segregation, a lower share of schools met pupil-teacher norms imposed by the Right to Education Act (2009).

JEL Classification: H75, I24, J45, M54

Keywords: teachers, public service delivery, segregation, caste, India, right to education

Corresponding author:
Panu Pelkonen
University of Sussex
Brighton, BN1 9RF
United Kingdom

E-mail: p.o.pelkonen@sussex.ac.uk

* We would like to thank the staff at NUEPA for providing us with the DISE database and for very helpful discussions; in particular Vimala Ramachandran, Arun Mehta, Naveen Bhatia and Pranati Panda. We would like to thank Suman Bhattacharyya at the ASER Centre for providing us with the ASER data. We are grateful to our colleagues at Sussex and seminar participants at ESOP, Oslo as well as conference participants at the ISI Development and Growth conference, Delhi, for useful comments. We acknowledge financial support from the British Academy. We are fully responsible for any remaining errors.

1 Introduction

The persistence of low levels of learning, despite increases in enrollment and school resources, has led to the need for a better understanding of the management of school systems and the motivation of teachers in many developing countries (e.g. Pritchett et al., 2013). Poor systems management can manifest itself for example as inefficiencies and distortions in the allocation of resources across the school system (Glewwe and Kremer, 2006). These inefficiencies and distortions on the other hand, can render new initiatives and programs ineffective.

This study highlights an overlooked dimension of school management: the spatial distribution of public sector teachers. Teachers are a key educational resource and well functioning school systems should be able to distribute teachers in proportion to the students across areas and schools. Large deviations from this aim could be considered as evidence of inefficient management of personnel, which is likely to influence the effectiveness of the delivery of education.

In this study we focus on the allocation of Indian primary school teachers using a large, annual register database of Indian public sector primary schools, the District Information System for Education (DISE), that covers entire India, for the 2006-2012 period. Regular Indian public primary school teachers are civil servants, who are hired and employed by Indian states, generally on permanent contracts. The allocation and transfers of teachers generally take place at the sub-state level, typically within a district, or sometimes within a block (see e.g. NUEPA, 2016).[1]

A major education initiative, the Right of Children to Free and Compulsory Education

1 A district is an administrative sub-unit of a state and a block an administrative sub-unit of a district.

Act (RTE) of 2009 came into force in India in April 2010.[2] This sets a number of requirements for the provision of schooling across India, including clearly stated norms on the numbers of teachers a school should have as a function of the number of pupils.

It is generally acknowledged by administrators that teachers have strong location preferences, which can create inertia, and unwillingness to transfer from a preferred area. Research shows that teachers and public sector workers are often averse to remote areas in developing countries (see e.g. Serneels et al., 2010, Fagernäs and Pelkonen, 2012 and Lemière et al., 2013).[3] As the pay of teachers is inflexible in India, the location of employment is considered to be an important non-pecuniary aspect of work.

A recent major report on teachers in India indicates that transfer rules are often not transparent and transfers are often implemented in an *ad hoc* fashion (NUEPA, 2016). Teachers may influence their own initial postings and subsequent transfers in both legitimate or illegitimate ways, depending on how the states and districts function in practice.[4] The preferences or constraints of teachers can therefore create a challenge to the administrators attempting to fill schools with a sufficient number of teachers, or allocate teachers equally across schools.

We begin this study by analyzing the distribution of teachers and whether the allocation has responded to the RTE Act. In addition to pupil-teacher ratios and the RTE allocation rules, we introduce a self-constructed, more 'objective' index to assess how equally teachers are allocated per pupils within Indian districts. Our analysis suggests that there is

2 See e.g. http://righttoeducation.in/. Last accessed 16th August 2016.

3 Earlier research has shown that location preferences of teachers can vary quite dramatically by gender or caste of teachers in India (Fagernäs and Pelkonen, 2012).

4 The role of politics in the postings and transfers has also been highlighted for instance by Béteille (2009) and Kingdon and Muzammil (2009) and Fagernäs and Pelkonen (2016). The DISE database does not include indicators on political connections of teachers.

substantial room for improvement in the allocation of teachers. Despite improving pupil-teacher ratios at the level of districts, the distribution of teachers across schools has not become more equitable over time. There is also a correlation between the equality of teacher allocation and primary school learning outcomes across Indian districts.

Secondly, we study the degree of segregation of teachers by gender and caste across schools within districts, using an adjusted index of dissimilarity. There is evidence of a varying degree of segregation by gender and caste, which correlates with socioeconomic characteristics. This observed degree of segregation will at least partly proxy for location and peer preferences of teachers, and resulting frictions in mobility.

Having established these patterns in the distribution and segregation of teachers, we proceed to study their association with each other. Does the need to segregate pose a potential constraint to the allocation of teachers within districts, and thus to policy initiatives such as the RTE? The answer is affirmative; we find that districts that had higher levels of workplace segregation of teachers prior to the RTE, especially by caste, made significantly smaller improvements in the allocation of teachers in the years after the implementation of the RTE.

Our study makes several contributions to the existing literature. While the distribution of health workers has received some attention (see e.g. Lemière et al., 2013 and Munga and Maestad, 2009), quantitative, large-scale studies on imbalances in the distribution of teachers are scarce in the context of developing countries.[5] Studies have highlighted inefficiencies driven by high teacher absence rates (e.g. Chaudhury et al., 2006, Kremer et al., 2005, Muralidharan et al., 2014) and the role of incentives in service delivery (e.g. Dal Bó et al., 2003, Muralidharan and Sundararaman, 2011).

5 See e.g. Zhang and Kanbur (2005) for a more general analysis on spatial inequalities in the provision of health and education in China. Jaramillo (2012) studies the determinants of the location of a teacher's first job in two regions of Peru.

The literature on workplace segregation in general has mostly focused on the measurement of segregation and its determinants (see e.g. Hellerstein and Neumark, 2008 for the US or Åslund and Skans, 2010 for Sweden). Teacher segregation has received rather little attention in this literature, especially in developing countries (for racial segregation of teachers in the US, see e.g. Frankenberg, 2009).

Our findings on segregation also have a connection with the literature on the role that a gender or caste match between pupils and teachers has on learning (see e.g. Dee, 2005, Dee, 2007, Rawal and Kingdon, 2010, Muralidharan and Sheth, 2015, Karachiwalla, 2015). For example, the geographic segregation of teachers by gender in India is significant enough for urban pupils to be much more likely to be taught by women, and remote rural pupils by men. The study also contributes to the literature on the role of social fractionalization in the provision of public goods (see e.g. Alesina et al, 1999, Banerjee et al., 2005).

We begin with a description of the data (Section 2), which is followed by a description of the indicators used to measure the equality of teacher allocation across districts (Section 3). Section 4 discusses the Right to Education Act and its general effects on teacher allocation. Section 5 focuses on measures of teacher workplace segregation. In Section 6 we explore the association between the equality of teacher allocation and teacher segregation during the implementation of the RTE, and Section 7 concludes and discusses the broader implications of our findings for policy and related research.

2 Data

The data used in this study come from a register database of all Indian public-sector primary schools, the District Information System for Education (DISE) by the National University of

Education and Planning (NUEPA). We utilize the data for the academic years of 2006/07 to 2012/13. This is an unbalanced panel database of schools and includes a separate file for teachers in the schools. For teachers, the database includes information on the name, age, caste, gender, date of birth, starting point of career as a teacher and indicators on educational qualifications. The database also includes variables on school resources and management, but it does not include reliable, or comparable data on learning. It covers all public sector schools as well as teachers, and is thus a promising source for studying the composition and distribution of teachers in the Indian primary education sector. A short discussion of the quality of the database can be found in the Appendix. The data are collected in the latter half of a calendar year, around September.

Our main sample covers all regular or para-teachers who teach lower primary school students in all Indian states. These teachers teach grades 1 to 5, and in some states, grades 1 to 4.[6] Regular teachers are civil servants, generally with permanent contracts. In addition to regular teachers, many states hire temporary contract (para) teachers to staff schools that lack teachers. Para-teachers tend to be local to the areas, less likely to have teacher training and work on short-term contracts and lower pay (see e.g. Kingdon and Sipahimalani-Rao, 2010). The use of para-teachers itself can be indicative of challenges in the equal allocation of formal teachers. Our indicators on pupil-teacher ratios and the allocation of teachers include both types of teachers. Summary statistics for the main variables of interest for our sample of schools are shown in Table 1.

6 Teachers may be in schools that include classes above lower primary ones, but only teachers who report teaching "primary" or "mostly" lower primary grades are included in the sample. In schools that include both lower and upper primary levels, the following categories have been reported for the classes taught by the teachers: "Primary", "Mostly primary" "Upper primary" or "Mostly upper primary". In such schools, we include only pupils in grades 1-5 (1-4 in some states), and teachers who report that they teach "Primary" or "Mostly primary", and exclude teachers who teach "Upper primary" or "Mostly upper primary"

TABLE 1

3 Measuring the equality of teacher allocation

Teachers are employed by states, while their allocation and transfers typically take place at the sub-state level. We study the allocation at the level of districts, which is arguably the most natural level of focus (see e.g. NUEPA, 2016). The district is often also the level at which the training of teachers takes place in the so called DIETs (District Institute for Education and Training). Transfers of primary school teachers can take place due to administrative reasons, such as the rationalization of teacher-pupil ratios, upon the request of teachers, in the form of mutually agreed transfers, or on disciplinary grounds, but many states lack clear transfer rules (NUEPA, 2016, Ramachandran et al., 2005 and 2008, and Sharma and Ramachandran, 2009).[7]

The recommended pupil-teacher ratio (PTR) under earlier government guidelines (Sarva Shiksha Abhiyan) was set at 40:1. Since the Right to Education Act (or RTE), these norms changed and a legal requirement was established. For smaller schools, the new norm is 30:1 pupils per teacher, and the minimum number of teachers in any school is two. However, for larger schools, the required PTR is 40:1 (NUEPA, 2016, RTE Forum, 2015). The precise RTE norms on the number of teachers per pupils can be found in the Appendix, Table A2. They have been applicable from April 2010 onwards.

Our first indicator for the allocation of teachers is based on the RTE norms. These requirements set clearly defined norms on pupil-teacher ratios for which administrators such as the District Education Officers could be held accountable. We measure the compliance of

7 There is no unique teacher identifier in the DISE database. It can be created for a specific school based on teacher characteristics and it is then possible to determine when a specific teacher exist, or enters a schools, but the destination of exiting teachers cannot be convincingly determined.

districts with these pupil-teacher ratio norms by computing the proportion of schools in a district that have at least the required number of teachers, given the number of pupils.

The RTE norms do not minimize the variance of pupil-teacher ratios across schools, as they do not uniformly require the pupil-teacher ratio to be at 30:1, and can lead to very low pupil teacher ratios in the smallest schools, which are supposed to have two teachers.

A more objective manner of allocating teachers would be to aim to equalize PTRs across schools.[8] We construct an indicator to study the inequality of teacher allocation within districts and label it the 'Share of misallocated teachers'. It describes the share of teachers that could be moved to minimize the variability of PTRs across schools. To our knowledge, this measure has not been used in the education literature before.

The indicator captures the proportion of teachers that could be transferred *within a district*, so as to minimize the variability in pupil-teacher ratios across schools *within the district.* The deficiency (or surplus) of teachers in each school (s) in a given year (t) is defined as the number of teachers the school would need to gain (or lose if negative), Δ_{st}, in order to have the same pupil-teacher ratio as the district (d) on average:

$$(1) \quad \frac{P_{st}}{T_{st}+\Delta_{st}} = \frac{P_{dt}}{T_{dt}} \Rightarrow \Delta_{st} = \frac{P_{st}T_{dt}}{P_{dt}} - T_{st} \quad ,$$

where P refers to pupils and T to teachers. The sum of deficits across schools equals the sum of surpluses across schools. As such, the potential within-district transfers, or the number of misallocated teachers (M_{dt}) is the sum of the deficits

8 If one assumes that lowering the pupil-teacher ratio (PTR) brings diminishing marginal benefits in terms of learning, the equality of PTRs across schools also implies efficiency in the sense that it maximises total learning. In practice, a more important argument for the equality for PTRs across schools is the aim for equality in the quality of public service provision.

(2) $M_{dt} = \sum^{n} \Delta_{st} | \Delta_{st} > 0$,

where the sum is taken over all *n* schools within district *d* in year *t*. This measure is further divided by the total number of teachers in the district to give our measure of the rate, or the proportion of teachers that could be moved

(3) $MR_{dt} = \frac{M_{dt}}{T_{dt}}$.

This indicator is informative of how teachers are distributed in practice at the district level. It is important to note that this is a measure of imbalance across schools, not of the degree of teacher shortages as such. The assumption is that in a well functioning state, the hiring and distribution of teachers should be needs based: the state should aim to maintain similar pupil-teacher ratios across districts, and districts (or lower levels) aim to distribute teachers across schools in similar proportion to enrollment, so that the pupil-teacher ratios would be roughly equal across schools.[9]

Figure 1 shows the distribution of this measure for cross sections of districts in 2006 and 2012. For comparison, Figure 1 also shows the distribution of corresponding numbers calculated for primary schools of 150 Local Education Authorities (LEA) in England for 2015.

Since pupil-teachers ratios are never precisely similar across schools, districts will, perhaps unfairly, always be labeled by this index as having a certain proportion of

9 In this study, we overlook the possibility that the desirable pupil-teacher ratio across schools may vary due to issues such as special educational needs.

misallocated teachers. We therefore carry out a simple simulation exercise with hypothetical districts, where the distribution of pupil numbers across schools is drawn from a Poisson distribution with the means varying between 100-200. We then apply a varying maximum class size cut-off of 30 to 50 pupils per teacher in different simulations.[10] For these 'perfectly managed' hypothetical school systems, the share of 'misallocated' teachers would be between .03-.12, which is in the same range as for England in Figure 1.

FIGURE 1

Figure 1 shows that the overall quality of teacher allocation in India has not changed much from 2006 to 2012. We will return to this observation. Table 2 summarizes the average pupil-teacher ratios and our indicators for the equality of teacher allocation, the proportion of schools that meet the requirements of the RTE Act on the numbers of teachers per pupils and the share of misallocated teachers. Table 2 shows the averages for pre- and post-RTE years, and indicates a decline in the PTR from 38 to 31.5 pupils per teacher, and an increase in the proportion of schools that satisfy the PTR norms of the RTE, from 52 to 66 percent. Teacher allocation by state for the panel of districts is the Appendix Table A3.

TABLE 2

Our indicators for teacher allocation include both regular and para-teachers.[11]

10 If the number of pupils is below the cut-off, say 30, the school has one teacher. The next teacher is brought in when the pupil numbers exceed the next threshold in multiples of 30, so that a school with 31 pupils will have two teachers.

11 Para-teachers are hired locally on temporary contracts and lower pay.

However, the use of para-teachers can be a reactionary measure to the difficulties of filling specific posts with regular teachers. Therefore, the imbalances in the allocation of teachers would be somewhat larger if para-teachers were to be excluded from the figures.

To our knowledge, no systematic descriptions of the quality of teacher allocation are available for India or for other major developing countries. Given that the DISE database does not include reliable data on learning, we cannot credibly explore the connection between imbalances and learning in detail. However, to provide some indication of the association, we describe the correlation between our indicator for the share of misallocated teachers and the levels of learning.

Figure 2 shows a scatter plot of districts in 2010 with the proportion of misallocated teachers, and the share of rural government school pupils between the age of 9-11, who are able to read a story.[12] The data source for the learning outcomes is the ASER 2010 household survey, which tests children at home.[13] The ASER survey is restricted to rural areas, and therefore the comparison with the misallocation measure is not entirely valid, since it is computed for both rural and urban areas. An unequal distribution of teachers can also potentially benefit many pupils, but overall, Figure 2 suggests a negative average correlation ($r = -0.20$). While the correlation is not very strong, across Indian districts, a more unequal allocation of teachers is associated with lower levels of learning on average.

FIGURE 2

12 Ages 9-11 correspond to the typical ages at the end of the primary school.

13 http://www.asercentre.org/

4 Teacher allocation and the Right to Education Act

Recent assessments of the RTE (for example RTE Forum, 2015) uniformly suggest that pupil-teacher ratios have declined significantly. Little is known about the extent to which the distribution of teachers has improved.

For an understanding of developments over time, we estimate an OLS model, where the three indicators summarized in Table 2 are explained with year dummies (θ_t) and district fixed effects (λ_d) using a panel of districts. The model takes the following form

$$(4)\quad Outcome_{dt} = \lambda_d + \theta_t + e_{dt} \qquad t_{pre-RTE} \in [2006,2009], t_{RTE} \in [2010,2012] \ ,$$

and its purpose is to simply assess the significance of aggregate developments before and after the RTE. The results are shown in Table 3. In all of the models, the years are divided into pre- and post-RTE years. The RTE came into force in April 2010. The first year in our database for which the RTE norms of teachers per pupils are applicable is the year 2010, given that the DISE data are recorded in late 2010. The year 2009 is the reference year and the years 2006-2008 are viewed as pre-RTE years.

TABLE 3

Column 1 of Table 3 shows that there is a declining trend in pupil-teacher ratios, which is evident in both the pre- and post-RTE period. The dependent variable in column 2 is the proportion of schools that have at least the number of teachers required by the national RTE norms. Again, an improving trend is visible both before and after the RTE. The largest

single improvement takes place between 2010 and 2011, suggesting that the states have may started to react to the norms.

The results in Columns 1 and 2 together indicate that the RTE may not have done much to increase the recruitment of teachers beyond the pre-RTE trend: the trend improvements in columns 1 and 2 are roughly equal in magnitude three years before and after the implementation of the RTE. It may thus be that improvements in the numbers of schools satisfying the RTE norms have mechanically followed from the larger numbers of teachers, without any additional effort to rationalize the allocation of teachers.

Column 3 shows the development in the more 'objective' teacher allocation measure, based on the variability of the school-specific PTRs around the district mean PTRs. Significant improvements in this measure should be possible for most Indian districts even without improvements in the PTRs. However, there is no substantial improvement; a pattern that can already be visually detected in Figure 1.

Overall, this evidence suggests that the RTE has not led to objective improvements in the allocation of teachers. The larger share of schools satisfying the RTE norms seems to derive mostly from general improvements in pupil-teacher ratios.

There are potential reasons for why the objective distribution of teachers may not have improved, despite the sensible requirements of the RTE norms. Firstly, the RTE norms do not aim to minimize the variability in the PTRs; large and small schools aim for different PTRs (either 40 or 30), and very small schools are still always required to have two teachers.[14] Secondly, it is possible that some administrators aim to maximise the share of complying schools by maintaining some schools with high PTRs.

14 The precise teacher norms are in Appendix table A2.

5 Teacher segregation and geographic imbalances of teachers by characteristics

A study of the full array of reasons behind the misallocation of teachers is beyond the scope of this article. In the Sections that follow, we focus on one aspect; the workplace segregation of teachers by caste and gender. Section 5.1 provides some stylized facts on the geographic distribution of teachers and Section 5.2 describes indices of segregation of teachers by caste and gender. In Section 6, we show that the ability of states to meet the RTE norms and the equality of teacher allocation are constrained by the level of workplace segregation.

5.1 Segregation by remoteness

A special feature of teacher labor markets in India is that female teachers are disproportionately concentrated in urban and semi-urban areas. This can be driven by both cultural norms and the preferences of the teachers themselves (Fagernäs and Pelkonen, 2012). The data also suggests that there is similar sorting by caste along the remoteness dimension of the schools.

In the DISE data there are two variables to capture the geographic remoteness of a school: whether the school is in an urban area, and the school's distance to the block headquarters.[15] As Table 1 shows, 6.8% of the schools in the sample are urban schools, and the average distance to the block headquarters is 16 km.

Figure 3 shows the percentage of female teachers by the distance of the school from the block headquarters. It is evident that the presence of women declines steadily with the remoteness of the school. In addition to gender, there is also a geographic pattern in the caste

15 Block is an administrative sub-unit of a district. While the DISE database has GIS codes for some schools, the coverage varies greatly by state and may not be reliable.

distribution of teachers. We simplify the analysis on caste by focusing on two broader categories; general caste and lower caste (scheduled caste, scheduled tribe and other backward class - SC/ST/OBC[16]) teachers. Lower caste teachers are more likely to be found in remote locations than general caste teachers (Figure 4). This gradient by distance is nearly as steep as that for gender; moving from 0 to 40 km from the block headquarters increases the share of lower caste teachers by about 20 percentage points, and a similar change in Figure 3 would increase the proportion of male teachers by about 30 percentage points.[17] These facts will be known to educational practitioners in India, but their precise scope is not well documented in academic literature.

FIGURES 3 AND 4

Segregation along the remoteness dimension is obviously not the only form of workplace segregation in India, but its ubiquity in the data cannot be overlooked. Figures 3 and 4 together show that any segregation of teachers by gender and caste across schools will be at least partly driven by geographic remoteness. Next, we turn to more formal measures of segregation.

5.2 Measure for workplace segregation

To arrive at more general measures of workplace segregation, we construct district level indexes for the segregation of teachers across schools by gender and caste, using the DISE teacher level database, annually between 2006-2012.

16 Other backward class (OBC) also includes 'other reserved groups'.
17 The graphs refer to formal teachers, but the distribution of para-teachers by gender and caste and remoteness is similar as for formal teachers.

We being the analysis of segregation on the commonly used Index of Dissimilarity across schools. This index measures the degree of similarity of teachers by gender and caste in schools in a district. The higher the degree of similarity, the more segregated are the teachers in the district. The properties of this index, and a number of alternatives in measuring school segregation (of pupils) are discussed for example by Allen and Vignoles (2007). For gender segregation, the index is calculated as

$$(5) \quad D=\frac{1}{2}\sum_s \left|\frac{f}{F}-\frac{m}{M}\right| \quad ,$$

where s and d refer to school and district respectively, f and m refer to the number of female and male teachers in a school respectively, and F and M refer to the number of female and male teachers in a district respectively. The sum is computed over all schools in the district. The index ranges between [0,1], where zero indicates that all schools have the same proportion of female teachers, and one means that teachers are perfectly segregated by schools.

Since the number of teachers in schools can be small, even a random allocation of teachers by gender or caste will lead to a positive index of segregation. For example, if female teachers account for 50% of teachers and are randomly allocated to two-teacher schools, 50% of the schools will have teachers of both genders {F,M}, and another 50% will be 'perfectly segregated' with combinations {M,M} or {F,F}, leading to an Index of Dissimilarity of 0.5. We therefore adjust the index of dissimilarity as in Carrington and Troske (1997), and compute an adjusted index of dissimilarity (D_a) as follows

$$(6) \qquad D_a = (D - D_r) / (1 - D_r) \text{ if } D > D_r$$
$$= (D - D_r) / D_r \text{ if } D < D_r,$$

where D is the observed Index of Dissimilarity and D_r is the randomly occurring dissimilarity. D_r is computed by Monte Carlo simulation taking into account the size distribution of schools.[18] The values for the adjusted index lie between the range of [-1,1], and the positive figures indicate the share of the possible segregation that takes place over and beyond the randomly occurring segregation. Negative values on the other hand suggest that there is less segregation than expected.

Table 4 shows that in the case of gender segregation, the average simulated value for the index of dissimilarity in the panel of districts is .502.[19] The average value for the actual dissimilarity index is .612. The value for the adjusted index of .223 suggests that 22.3% of possible teacher gender segregation takes place in a typical Indian school district, but this figure ranges from negative segregation (possible attempts to match men with women) to high levels of gender segregation.

The degree of caste segregation, as we have defined it, is roughly similar to that of gender segregation. We recognize that we simplify the variability in castes by focusing on a binary indicator, but it provides a convenient comparison with segregation by gender.[20] Interestingly, gender and caste segregation are not correlated in the panel of districts. For

18 We take the number of teachers in schools as given. We then draw the gender of each teacher randomly from a binomial distribution, where the likelihood of being female is the same as in the district as a whole in that year. We then compute the segregation index for the district-year cell. This is repeated 100 times and the average index over the randomisations is used as the final D_r.

19 The value for the randomly occurring index can be above 0.5 when the proportion of women deviates from 50%.

20 The information on caste in DISE is in four categories. In principle, teacher names could be used to infer caste at more disaggregated level, but this would be potentially a large undertaking with uncertain benefits given the number of names involved. We conduct a robustness check to assess whether our results hold when districts with substantial tribal populations is excluded.

example, for the pooled pre-RTE panel of 2006-2009, the correlation is -.05.

TABLE 4

To provide some socioeconomic correlates for the segregation of teachers, we estimate simple district level OLS regressions with district level explanatory variables and state fixed effects and state trends. Table A4 in the Appendix shows the results. The results show that a lower level of female literacy, which also proxies for the level of development, is associated with a higher degree of segregation by gender and caste. Segregation of teachers by caste is associated further with pupil caste segregation, while gender segregation is associated with a higher rate of urbanization and a larger reliance on para-teachers. The larger reliance on para-teachers is likely to be connected with the presence of remote areas in the district, that formal female teachers are particularly averse to.

6 Does teacher segregation constrain the equality of teacher allocation?

Having established that the RTE has not improved the quality of teacher allocation, and the presence of significant and varying degree of workforce segregation, we test the hypothesis that such segregation can pose a constraint for the allocation of teachers. Teachers are known to have strong preferences with respect to location, and those preferences can vary by gender and caste.

In the estimations that follow, we utilize the onset of the RTE, together with the degree of pre-RTE variation in the gender and caste segregation of teachers to estimate whether policies on allocation of teachers is constrained by workplace segregation of

teachers. We rely on a district level panel data set, constructed from the school-level data. The dependent variables of interest are the district-level shares of schools that satisfy the minimum requirements of the RTE norms on PTRs, and our self-constructed measure of teacher allocation.

In the estimation framework, we allow the effect of the RTE policy to vary with respect to the pre-policy levels of teacher segregation. A similar identification strategy has been employed by for instance Bleakley (2007), Acemoglu and Johnson (2007) and Bhalotra and Venkataramani (2015). Effectively, we study whether districts with a higher degree of segregation have a lower rate of improvement in the equalization of PTRs than those with a lower degree of segregation, as the RTE is implemented throughout the country. We estimate the following model

$$(7)\quad Outcome_{dt} = \beta_1 RTE_t \times PreCSeg_d + \beta_2 RTE_t \times PreGSeg_d + \boldsymbol{X_{dt}}\Gamma + \lambda_d + \theta_t + e_{dt} \quad ,$$

where λ_d refers to district fixed effects, and θ_t to year effects. $PreCSeg_d$ is the adjusted caste segregation index and $PreGSeg_d$ is the adjusted pre-policy gender segregation index for the districts. The values for these indexes are averages for the pre-policy years 2006-2009. RTE_t is a dummy that takes the value of one for the years 2010-12, and zero for 2006-09. Vector $\mathbf{X_{dt}}$ stands for additional district-specific controls.

The main coefficients of interest are β_1 and β_2 - the interaction terms between the RTE policy and the pre-policy levels of gender and caste segregation. This reflects the degree to which the effect of the RTE policy varies depending on the pre-RTE level of segregation. The district-level fixed effects control for all district specific differences in fixed characteristics,

such as remoteness, demographic characteristics and the level of development. The year effects control for all annual shocks, as well as for all general effects of the RTE policy.

We begin by estimating equation 7 for the district level pupil-teacher ratios, and the share of schools that comply with the RTE norms. These results are presented in columns 1 and 2 of Table 5. The first column shows that the impact of the RTE policy on pupil-teacher ratios is not affected by the initial degree of segregation. This implies that for a given level of pupils, *the level of teacher resources* that the districts have at their disposal doesn't vary significantly with the initial level of segregation. On the other hand, the second column shows that the *way in which the teaching resources are allocated*, does vary by the initial level of segregation: districts with a higher degree of segregation prior to the RTE, especially by caste, comply less with the RTE requirements on teachers per pupils in the post-RTE period.

TABLE 5

To assess the magnitude of the interaction effect in column 2, let us suppose that the adjusted levels of segregation would be zero instead of the average of 0.221 for caste and 0.223 for gender. In such a case, the share of schools satisfying the RTE norms in the post-RTE period, would be up to 10 percentage points higher (-.333*.221 - .151*.223 = -.107). This is a large effect, which, if true, would make a significant difference for the perceived success of the RTE.

In columns 3 and 4 we assess the robustness of the result. Firstly, in column 3, we interact the RTE policy variable with a pre-RTE (2006-2009) linear trend in the dependent

variable. Here we aim to account for the fact that districts may have made progress towards RTE compliance already prior to the RTE due to improving PTRs, and ensure that our findings in column 2 are not simply explained by pre-RTE trends that may continue and be correlated with levels of segregation. Secondly, we interact the RTE policy variable with the average pre-RTE PTR for the district to account for different levels of initial teacher resources for districts. The main results are robust to these additions, although the negative effect of caste segregation becomes smaller.

In the most rigorous specification in the final column, we allow the effect of the RTE policy to vary depending on the variables that were found to be associated with segregation within states (Appendix Table A4). The results are shown in column 4 of Table 5. The negative effect of gender segregation is significantly mitigated, but this doesn't alter the negative effect of caste segregation much.

Figure 1 and Table 2 together suggested that the more objective indicators for the equality of teacher allocation have not changed significantly with the RTE. In Table 6, we study whether pre-RTE teacher segregation influences the degree to which the equality of teacher allocation changes with the RTE. We estimate equation 7 using the measure of the share of misallocated teachers as the dependent variable. The results are shown in Table 6 with similar robustness checks as in Table 5.

TABLE 6

The results in Table 6 show that initial caste segregation is associated with a significantly higher level of misallocated teachers in the post-RTE period. For gender

segregation no such effects are present.

Since we have divided caste into general caste versus all others, there is a warranted concern that in areas with substantial tribal populations (scheduled tribe), the potential segregation between scheduled caste and scheduled tribe teachers is being overlooked, and this might bias our current results.[21] In the Appendix Table A6 we address this concern by showing that our results with respect to caste are to a large extent robust to the exclusion of districts with substantial tribal populations.

Overall, we find robust evidence that the rationalization of teacher resources in the post-RTE period has been constrained by the initial degree of segregation of teachers, especially by caste. For gender segregation, the evidence points to the same direction, but the effects are weaker.

More generally, the results in Tables 3, 5 and 6 indicate that the way in which policy targets are formulated, matters. The proportion of schools that satisfy the RTE allocation requirements increases with the RTE, while the objective measure of the equality of teacher allocation does not improve. Despite this, the objective measures of allocation have improved significantly more in districts with a lower level of caste segregation.

7 Discussion

This study begins by providing a number of stylized facts about the distribution of primary school teachers in India. We find that at the local level, the district, the allocation of teachers according to the pupil numbers is fairly unequal is most places, and that there would be significant scope for moving teachers to improve the distribution. Pupil-teacher ratios have

21 About one out of six of lower caste primary teachers are Scheduled Tribe (ST).

improved over the period 2006-2012, and this has led to a mechanical increase in the share of schools that meet the current legal norms, set by the Right to Education Act of 2009 (RTE). However, our self-constructed measure of allocative efficiency of teachers does not reveal improvement over time.

The distribution of teachers is also characterized by significant levels of workplace segregation by gender and caste, based also on a binary categorization of teachers into general caste teachers versus all others. Gender segregation especially, is a function of remoteness of schools, and is likely to arise from cultural norms and preferences. On average across Indian districts, segregation by caste is not correlated with segregation by gender. The segregation of teachers and its implications for service delivery have to date received little attention in the literature.

We then proceeded to test the hypothesis that the need to segregate constrains the allocation of teachers; in particular whether districts with a higher level of workplace segregation of teachers have more difficulties in distributing teachers according to needs. The underlying assumption is that the observed degree of segregation proxies for the location preferences of teachers and possibly also their ability to resist or influence transfers. In India, a recent major report into school teachers' work and transfer policies in nine states concluded that there are no transparent or uniform rules on how teachers are distributed, and that teachers transfers are often implemented in an *ad hoc* fashion (NUEPA, 2016).

Our estimates show that in the post-RTE period, when the new rules on pupil-teacher ratios became applicable, the pre-RTE degree of segregation, especially by caste, is negatively associated with the equity of teacher allocation. This result is robust to the inclusion of all district-specific fixed factors, underlying district-specific pre-policy trends

and a number of other district level socioeconomic controls.

The broad motivation for studying the distribution of teachers arises from the persistently poor learning outcomes, and prevalent teacher absences across the developing world (Chaudhury et al, 2006). Our results suggest that the social mosaic typical of many developing countries, including India, can hinder personnel policies and service delivery and draws attention to the importance of specifically taking it into account in the design of service delivery. These problems are likely to be larger in countries with large regional differences in development, deeper segregation by ethnicity or social status, or a weaker institutional framework to rigorously implement the existing rules.

References

Acemoglu, Daron, and Simon Johnson (2007). Disease and Development: The Effect of Life Expectancy on Economic Growth. *Journal of Political Economy,* 115 (6): 925-65.

Alesina, Alberto, Reza Baqir, and William Easterly (1999). Public Goods and Ethnic Divisions. *Quarterly Journal of Economics*, 114: 1243-1284.

Allen, Rebecca and Anna Vignoles. (2007). What Should an Index of School Segregation Measure? *Oxford Review of Education*, 33(5): 643-668.

Åslund, Olof and Oskar Nordström Skans (2010).Will I See You at Work? Ethnic Workplace Segregation in Sweden, 1985–2002. *Industrial & Labor Relations Review*, 63(3): 471-493.

Banerjee, Abhijit, Lakshmi Iyer and Rohini Somanathan (2005). History, Social Divisions, and Public Goods in Rural India. *Journal of the European Economic Association*, 3(2): 639-647.

Béteille, Tara. Absenteeism, Transfers and Patronage: The Political Economy of Teacher

Labor Markets in India. PhD diss., Stanford University, 2009.

Bhalotra, Sonia R. and Venkataramani, Atheendar (2015). Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions. Working Paper, available at SSRN: https://ssrn.com/abstract=1940725.

Bleakley, Hoyt (2007). Disease and Development: Evidence from Hookworm Eradication in the American South. *Quarterly Journal of Economics,* 122 (1): 73-117.

Carrington, William J. and Kenneth R. Troske (1997). On Measuring Segregation in Samples With Small Units. *Journal of Business & Economic Statistics* , 15(4): 402-409.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers (2006). Missing in Action: Teacher and Health Worker Absence in Developing Countries. *Journal of Economic Perspectives*, 20(1): 91-116.

Dal Bó, Ernesto, Frederico Finan and Martin A. Rossi (2003). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *The Quarterly Journal of Economics*, 128(3): 1169-1218.

Dee, Thomas (2005). A Teacher Like Me: Does Race, Ethnicity, or Gender Matter? *American Economic Review*, 95: 158-165.

Dee, Thomas (2007). Teachers and the Gender Gaps in Student Achievement. *Journal of Human Resources*, 42: 525-554.

Fagernäs, Sonja and Panu Pelkonen (2012). Preferences and Skills of Indian Public Sector Teachers. *IZA Journal of Labor & Development*, 1:3.

Fagernäs, Sonja and Panu Pelkonen (2016). Teachers, Electoral Cycles and Learning in India. *Mimeo.*

Frankenberg, E. (2009). The Racial Segregation of American Teachers, *Education Policy*

Analysis Archives.

Glewwe, Paul and Michael Kremer (2006). Schools, Teachers and Education Outcomes in Developing Countries. *Handbook of the Economics of Education,* Volume 2, pp. 945-1017.

Hellerstein, Judith K. and David Neumark (2008). Workplace Segregation in the United States: Race, Ethnicity, and Skill. *Review of Economics and Statistics*. Vol. 90(3), pp. 459-477.

Jaramillo, Miguel (2012). The Spatial Geography of Teacher Labor Markets: Evidence from a Developing Country. *Economics of Education Review*, 31(6): 984-995.

Karachiwalla, Naureen (2015). A Teacher Unlike Me: Social Distance, Learning, and Intergenerational Mobility in Developing Countries. Mimeo.

Kingdon, Geeta, and Mohammed Muzammil (2009). The Political Economy of Education in India: The Case of Uttar Pradesh. *Oxford Development Studies,* 37(2): 123-144.

Kingdon, Geeta Gandhi and Vandana Sipahimalani-Rao (2010).Para-teachers in India: Status and Impact, Economic & Political Weekly, March 20, Vol. XLV, No 12.

Kremer, Michael, Karthik Muralidharan, Nazmul Chaudhury, Jeffrey Hammer and F. Halsey Rogers (2005). Teacher Absence in India: A Snapshot. *Journal of the European Economic Association*, 3(2-3): 658-67.

Lemière, Christophe, Christopher H. Herbst, Carmen Dolea, Pascal Zurn, and Agnes Soucat (2013). Rural-Urban Imbalance of Health Workers in Sub-Saharan Africa. The Labor Market for Health Workers in Africa: A New Look at the Crisis. The World Bank.

Munga, M.A. and O. Maestad. (2009). Measuring Inequalities in the Distribution of Health Workers: The Case of Tanzania. *Human Resources for Health*, 7:4.

Muralidharan, Karthik and Venkatesh Sundararaman (2011). Teacher Performance Pay:

Experimental Evidence from India. *Journal of Political Economy*, 119(1), 39-77.

Muralidharan, Karthik and Ketki Sheth (2015). Bridging Education Gender Gaps in Developing Countries: The Role of Female Teachers. Forthcoming, *Journal of Human Resources.*

Muralidharan, Karthik, Jishnu Das, Alaka Holla and Aakash Mohpal (2014). The Fiscal Cost of Weak Governance: Evidence from Teacher Absence in India. NBER Working Paper 2029.

NUEPA (2016). Teachers in the Indian Education System. NUEPA Research Reports Publications Series 001/2016. National University of Educational Planning and Administration, New Delhi.

Pritchett, Lant, Rukmini Banerji and Charles Kenny (2013). Schooling is Not Education! Using Assessment to Change the Politics of Non-learning. A Report of the Center for Global Development Study Group on Measuring Learning Outcomes.

Ramachandran, Vimala, Madhumita Pal, Sharada Jain, Sunil Shekar and Jitendra Sharma (2005). Teacher Motivation in India. Educational Resource Unit, New Delhi.

Ramachandran, Vimala, Suman Bhattacharjea and K.M. Sheshagiri (2008). Primary School Teachers. The Twists and Turns of Everyday Practice. Educational Resource Unit, New Delhi.

Rawal, Shenila and Geeta Kingdon (2010). Akin to My Teacher: Does Caste, Religious or Gender Distance Between Student and Teacher Matter? Some Evidence from India. Department of Quantitative Social Science Working Paper 1018, Institute of Education, University of London.

RTE Forum (2015). Status of Implementation of the Right of Children to Free and Compulsory Education Act, 2009: Year Five (2014-15), March 2015.

Serneels Pieter, Jose Garcia-Montalvo, Gunilla Pettersson, Tomas Lievens, Jean Damascene Butera, Aklilu Kidanu (2010). Who Wants to Work in a Rural Health Post? The Role of Intrinsic Motivation, Rural Background and Faith-Based Institutions in Ethiopia and Rwanda. *Bulletin of the World Health Organisation*, 88(5): 342–349.

Sharma, Rashmi and Vimala Ramachandran (2009). *The Elementary Education System in India*. Routledge, New Delhi.

Volmink, John, Michel Gardiner, Siyabonga Msimang, Paul Nel, Amelia Moleta, Gerhard Scholtz and Tommy Prins (2016). Report of the Ministerial Task Team Appointed by Minister Angie Motshekga to Investigate Allegations into the Selling of Posts of Educators by Members of Teachers Unions and Departmental Officials in Provincial Education Departments. Department of Basic Education, South Africa, May 2016.

Zhang, Xiaobo and Ravi Kanbur (2005). Spatial Inequality in Education and Health Care in China. *China Economic Review*, 16(1): 189-204.

Table 1 Summary statistics on public sector primary schools, panel data for 2006-2012

	Obs.	Mean	SD	Min	Max
Number of teachers	5903148	2.87	1.83	1	239
Number of formal teachers	5903148	2.35	1.88	0	239
Number of parateachers	5903148	.446	.827	0	26
Required RTE teacher norm satisfied	5903148	.466	.499	0	1
Share of SC/ST/OBC teachers	5900916	.659	.377	0	1
Share of female teachers	5903090	.375	.363	0	1
Urban school	5903032	.068	.253	0	1
Distance to Block HQ (km)	5872827	16.0	13.3	0	90
Enrolment	5902981	104.3	95.3	0	819
Share of pupils SC/ST/OBC	5755188	.770	.308	0	1

Notes: Source DISE 2006-2012. SC/ST/OBC = Scheduled caste, Scheduled tribe, Other backward class (the last category also includes other reserved groups). A block is an administrative sub-unit of a district. HQ refers to Headquarters. Outliers have been removed (see Appendix), as well as districts with less than 30 primary schools.

Table 2 Summary Statistics for District level data set, measures of teacher allocation 2006-2012

N = 3441	Pre-RTE (2006-09)				Post-RTE (2010-12)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Measures of teacher allocation								
Pupil / Teacher ratio (PTR)	38.0	14.8	4.3	99.6	31.5	14.2	2.9	133.9
Share of schools that satisfy RTE norms	.523	.248	0	1	.657	.237	0	1
Share of misallocated teachers	.216	.045	.073	.484	.212	.045	.093	.386

Notes: The measures of imbalances are computed using school level data for each district and year. The share of misallocated teachers refers to equation 3. Sample has been restricted to districts which have not split during the period.

Table 3 Development of teacher allocation in a panel of districts over 2006-2012

	[1] P/T ratio	[2] % Satisfy RTE norms	[3] Share mis-allocated
Year 2006	6.57**	-.118**	.00568**
	[.345]	[.00572]	[.00162]
Year 2007	4.23**	-.0889**	-.00075
	[.329]	[.00563]	[.00113]
Year 2008	2.1**	-.0418**	.00157
	[.29]	[.00521]	[.0011]
Year 2010 (RTE)	-.290	.0134**	-.00124
	[.341]	[.00504]	[.0011]
Year 2011 (RTE)	-4.33**	.0843**	-.0058**
	[.309]	[.00587]	[.0011]
Year 2012 (RTE)	-5.49**	.124**	-.00121
	[.403]	[.00607]	[.00135]
Observations	3441	3441	3441
R-squared	.870	.873	.805

Notes: The reference year is 2009. All models include district fixed effects. 'RTE' refers to years when the Right to Education act is in force. **, *, + refer to p = .01, .05 and .10 statistical significance. Robust standard errors are in parentheses.

Table 4 Summary Statistics for measures of segregation, District level data set, 2006-2012

	Obs.	Mean	SD	Min	Max
Measures of segregation					
Teacher gender segregation	3441	.612	.096	.329	.875
Teacher gender segregation, random	3441	.502	.088	.276	.832
Teacher gender segregation, adjusted	3441	.223	.114	-.251	.661
Teacher caste segregation	3441	.663	.135	.303	1
Teacher caste segregation, random	3441	.568	.159	.301	.997
Teacher caste segregation, adjusted	3441	.221	.139	-.208	1
Pupil caste segregation	3441	.153	.121	.021	.938
Pupil caste segregation, random	3441	.074	.023	.038	.266
Pupil caste segregation, adjusted	3441	.063	.166	-.698	.915

Notes: The segregation indexes are based on the Index of Dissimilarity, equation 6. Sample has been restricted to districts which have not split during the period. Pupil caste segregation is measured similarly to teacher caste segregation, using school-level data.

Table 5 The Effect of the teacher segregation on the allocation of teachers, as required by the RTE Act

	[1] P/T ratio	[2] % Satisfy RTE norms	[3] % Satisfy RTE norms	[4] % Satisfy RTE norms
Interactions				
RTE x pre-policy caste segregation	-4.78	-.333**	-.204**	-.191**
	[6.8]	[.0237]	[.0251]	[.0249]
RTE x pre-policy gender segregation	5.9	-.151**	-.146**	-.0496+
	[7.38]	[.0285]	[.0289]	[.0301]
Further controls				
RTE x pre-policy trend in dependent			.524**	.55**
			[.0318]	[.0321]
RTE x pre-policy P/T ratio			-.000266**	.000818**
			[.0000888]	[.000157]
RTE x Female literacy rate				.051
				[.0387]
RTE x Male-female literacy gap				-.106
				[.0732]
RTE x Urbanisation rate				.0703**
				[.0208]
RTE x pre-policy pupil caste segregation				.004
				[.0198]
RTE x Share of teachers parateachers				-.222**
				[.0261]
District fixed effects	Y	Y	Y	Y
Year effects	Y	Y	Y	Y
Observations	3441	3441	3433	3433
R-squared	.652	.881	.900	.904

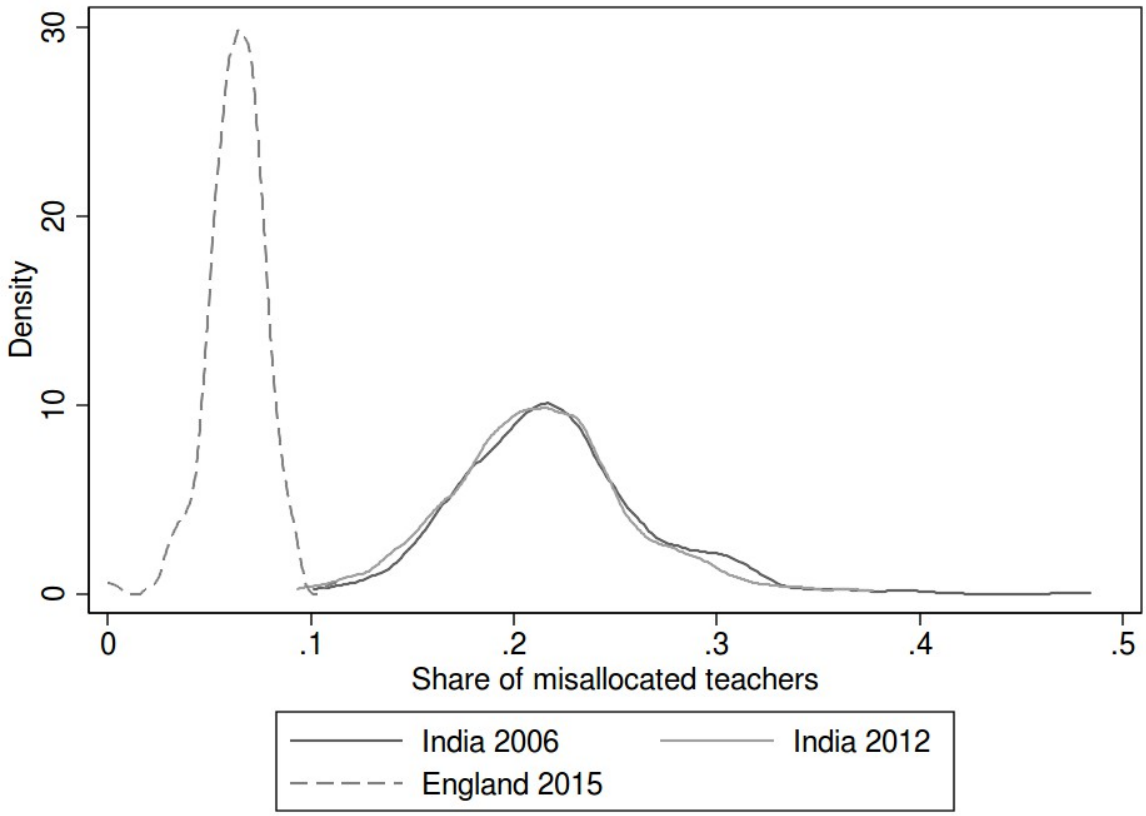
Notes: **, *, + refer to p = .01, .05 and .10 statistical significance. Robust standard errors are in parentheses. The variables for the Female literacy rate, male-female literacy gap and the urbanization rate come from the Indian Census of 2011, and are only available for this year. Summary statistics are shown in the notes to Table A5. The pre-policy trend in the depending variable is calculated for the 2006-09 period and the pre-policy PTR is an average for 2006-09. RTE is a dummy variable indicating years 2010-2012.

Table 6 The effect of teacher segregation on the misallocation of teachers since the RTE Act

	[1] Share mis-allocated	[2] Share mis-allocated	[3] Share mis-allocated
Interactions			
RTE x pre-policy caste segregation	.0138*	.0213**	.0188**
	[.00641]	[.00668]	[.00662]
RTE x pre-policy gender segregation	0.0109	0.00952	0.0135
	[.00625]	[.00614]	[.0072]
Further controls			
RTE x pre-policy trend in dependent		.256**	.253**
		[.0352]	[.0361]
RTE x pre-policy P/T ratio		.0000371*	.000
		[.000018]	[.0000321]
RTE x Female literacy rate			.0248*
			[.0101]
RTE x Male-female literacy gap			-.026
			[.0201]
RTE x Urbanisation rate			-.006
			[.00521]
RTE x pre-policy pupil caste segregation			-.002
			[.00644]
RTE x Share of teachers parateachers			.0146*
			[.00593]
District fixed effects	Y	Y	Y
Year effects	Y	Y	Y
Observations	3441	3433	3433
R-squared	.805	.816	.818

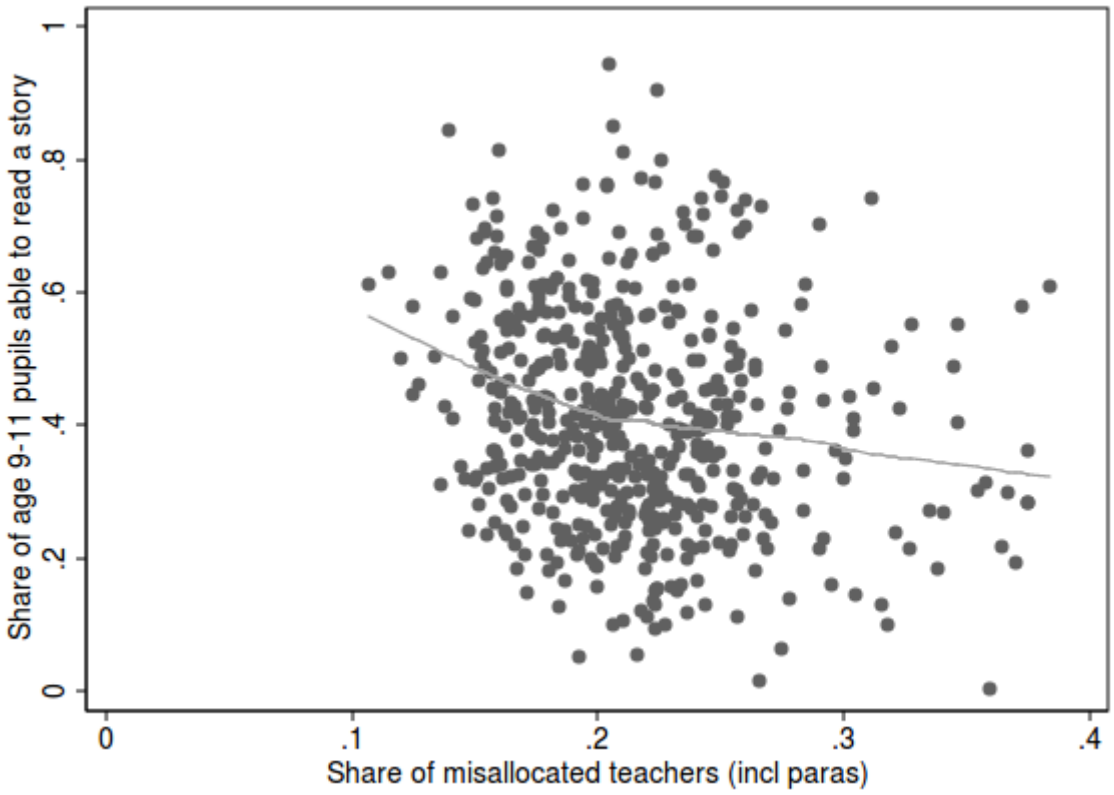
Notes: **, *, + refer to p = .01, .05 and .10 statistical significance. Robust standard errors are in parentheses. he variables for the Female literacy rate, male-female literacy gap and the urbanization rate come from the Indian Census of 2011, and are only available for this year. Summary statistics are shown in the notes to Table A5. The pre-policy trend in the depending variable is calculated for the 2006-09 period and the pre-policy PTR is an average for 2006-09. RTE is a dummy variable indicating years 2010-2012.

Figure 1 The proportion of teachers that could be transferred within Indian districts in 2006 and 2012 and English Local Education Authorities in 2015.



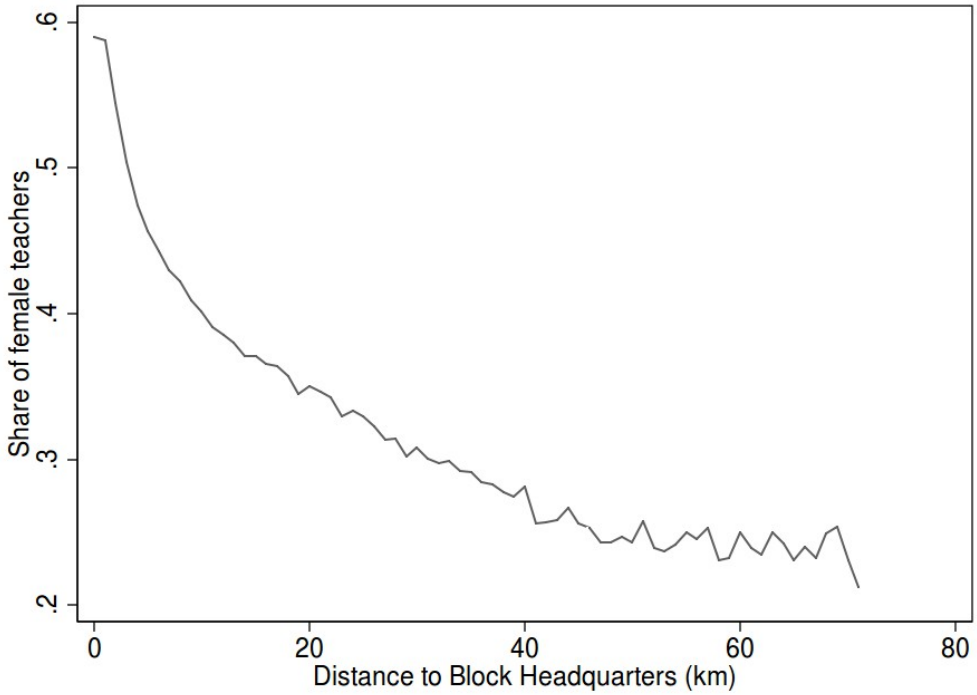
Notes: Figures for England have been calculated for primary schools in 150 Local Education Authorities in 2015, excluding teaching assistants. (Source: Department for Education, UK). Figures for India are based on Indian districts in 2006 and 2012.

Figure 2 Literacy of rural government school pupils between the age of 9-11 and the share of misallocated teachers across districts, 2010



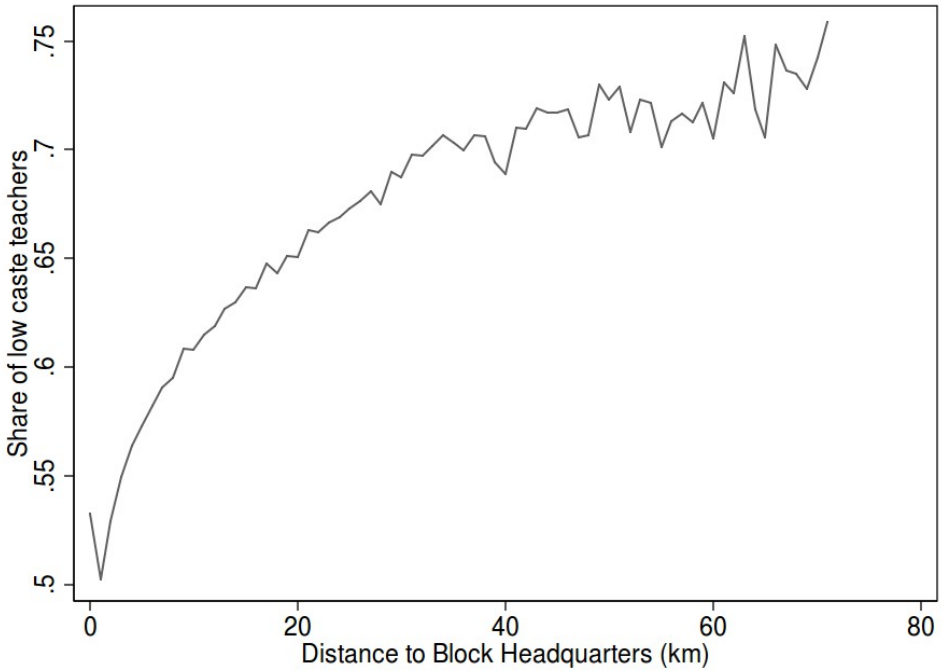
Notes: The data source for literacy is the ASER 2010 survey, restricted to government school pupils aged 9-11. Data includes 527 districts, with on average 200 pupils per district (min:21, max:472). Lowess smoothing line is computed with the defaults of Stata 14.

Figure 3 The share of female teachers by distance to block headquarters, 2011



Notes: Graphed using one kilometer bins. Sample: Formal teachers aged 18-55.

Figure 4 The share of lower caste teachers by distance to block headquarters, 2011



Notes: Graphed using one kilometer bins. Sample: Formal teachers aged 18-55.

Appendix

Cleaning of the DISE data

The sample of districts includes only districts that have not split over the time period, to guarantee comparability over time.

The DISE database was originally introduced for the purpose of planning and monitoring of national education programs in India, as such information systems were not available. The responsibility of reporting lies with the schools. The consistency of the DISE data is checked annually at the state level with 5% re-sampling, and should involve independent monitors. There are no other comparable large, or accurate data sources on Indian schools.[22]

For this study we constrain the data to non-private lower primary schools (grades 1-4 or 1-5 depending on the state). Some schools also have higher grades, but in such cases we have included in our sample only teachers who report that they teach 'Primary' or 'Mostly Primary' grades, and primary level enrollment numbers. With such restrictions, our raw data contains a panel of 7198460 teacher observations over 2006-12 (Table A1).

Some of the school characteristics in DISE include outliers. We have categorized observations as outliers for each year as follows: largest 0.5% of enrollment numbers for each elementary school, largest 0.5% of values for the distance to the block headquarters, and largest 0.5% of the values for the pupil-teacher ratio. If a school has outlier values for any of these variables, it is excluded from the analysis and summary statistics. We also set as outliers the schools that are in districts with less than 30 schools, or for which the number of pupils, or formal teachers is missing. Overall, outliers lead to the exclusion of about 1% of the observations, which are more likely to be urban schools (Table A1).

22 http://www.dise.in/

Table A1 Sample selection due to data cleaning, teacher level data 2006-2012

	Raw data	Outliers	Final sample
Observations	7198460	71808	5903148
Year	2009.1	2009.4	2009.0
Share women	.376	.472	.375
Share SC/ST/OBC	.66	.72	.66
Urban school	.072	.166	.068

Notes: The 'Raw data' is restricted to formal lower primary schools and teachers who report that they teach in such schools.

Table A2 Pupil-teacher ratio norms under the Right to Education Act.

Admitted children	Number of required teachers
Up to 60	2
61-90	3
91-120	4
121-200	5
Above 200	One per forty children

Notes: Source: RTE Forum (2015)

Table A3 Teacher allocation by state, district averages, 2006-2012.

State	P/T	RTE norms	MR
Andaman and Nicobar	13	.985	.208
Andhra Pradesh	33	.648	.198
Arunachal Pradesh	21	.818	.291
Assam	31	.665	.299
Bihar	63	.189	.224
Chandigarh	42	.491	.124
Chattisgarh	29	.679	.224
Dadra Nagar	35	.499	.163
Daman & Diu	17	.968	.186
Gujarat	29	.697	.176
Haryana	34	.559	.168
Himachal Pradeh	16	.958	.190
Jammu and Kashmir	12	.961	.260
Jharkand	47	.335	.215
Karnataka	42	.484	.244
Kerala	20	.909	.214
Madhya Pradesh	41	.438	.208
Maharashtra	29	.667	.200
Manipur	18	.874	.275
Meghalaya	18	.882	.232
Mizoram	16	.869	.237
Orissa	31	.646	.213
Pondicherry	23	.847	.223
Punjab	33	.604	.156
Rajasthan	32	.610	.235
Sikkim	13	.968	.226
Tamil Nadu	41	.461	.188
Tripura	25	.726	.254
Uttar Pradesh	25	.762	.246
Uttarakhand	45	.437	.221
West Bengal	35	.522	.176

Notes: All figures are means of district-level values over 2006-2012. MR refers to the share of misallocated teachers. All numbers include both formal and para-teachers. 'RTE Norms' refers to the share of schools that satisfy the minimum number of teachers as required by RTE norms. The sample has been restricted to districts that have not split over 2006-2012.

Table A4 Associations between the misallocation of teachers and gender and caste segregation, panel data set 2006-2012, OLS

	[1]	[2]
	Adjusted caste segregation	Adjusted gender segregation
Female literacy rate	-.116*	-.137*
	[.0491]	[.0581]
Male-female literacy gap	-.1130	.17+
	[.13]	[.10]
Urbanisation rate	.0459	.107**
	[.029]	[.0234]
Adjusted pupil caste segregation	.11*	.0405+
	[.0444]	[.0226]
Share of teachers parateachers	.0091	.141**
	[.0299]	[.0307]
Year effects	Y	Y
State effects	Y	Y
Observations	3441	3441
R-squared	.526	.554

Notes: **, *, + refer to p = .01, .05 and .10 statistical significance. Standard errors (in parentheses) are clustered by district. The summary statistics for the variables can be found in Table A5.

Table A5 Summary statistics of additional fixed control variables, district panel data, 2006-2012

	Obs.	Mean	SD	Min	Max
Female literacy	3441	.634	.122	.303	.977
Male-female literacy gap	3441	.172	.062	-.069	.340
Rate of urbanisation	3441	.246	.183	0	1
Share of para-teachers	3441	.140	.189	0	.892

Notes: The source for Literacy rates and the Rate of urbanization is the Indian Census of 2011 and the values are fixed over time. Sample has been restricted to districts which have not split during the period. The share of para-teachers is calculated at the teacher-level using the DISE database.

Table A6 Main results, excluding districts with substantial tribal (ST) populations

	[1] % Satisfy RTE norms	[2] Share mis-allocated
Interactions		
RTE x pre-policy caste segregation	-.28**	.0199*
	[.0317]	[.00824]
RTE x pre-policy gender segregation	0.011	-0.00226
	[.0365]	[.00767]
Further controls		
RTE x pre-policy trend in dependent	.593**	.25**
	[.0369]	[.0372]
RTE x pre-policy P/T ratio	.000819**	.000
	[.000191]	[.0000323]
RTE x Female literacy rate	-.022	.0625**
	[.0538]	[.0116]
RTE x Male-female literacy gap	-.281**	.0454*
	[.0975]	[.0228]
RTE x Urbanisation rate	.0698**	-.008
	[.0232]	[.00544]
RTE x pre-policy pupil caste segregatio	.000	.005
	[.0251]	[.00767]
RTE x Share of teachers parateachers	-.241**	.007
	[.0313]	[.00631]
District fixed effects	Y	Y
Year effects	Y	Y
Observations	2579	2579
R-squared	.900	.842

Notes: This Table replicates column 4 of Table 5 and column 3 of Table 6. The sample is altered by excluding 25% of districts with largest population share of Scheduled Tribes based on the 2011 census. In practice, we exclude districts where more than 18.98% of the population belongs to Scheduled Tribes. **, *, + refer to p = .01, .05 and .10 statistical significance. Standard errors (in parentheses) are clustered by district.