A teacher unlike me: Social distance, learning, and intergenerational mobility in developing countries

Karachiwalla, Naureen

International Food Policy Research Institute

19 September 2013

A teacher unlike me: Social distance, learning, and intergenerational mobility in developing countries*

Naureen Karachiwalla[†]

May 22, 2015

Abstract

Same-type teachers are extolled as a way to improve learning outcomes of socially disadvantaged students. This paper uses a relatively understudied social characteristic, caste, to study whether same-type teachers improve learning in a low-income country. Rich longitudinal data from Pakistan allows identification of causal effects using child fixed effects specifications. Low caste boys have significantly higher learning outcomes when taught by high caste teachers. Low caste boys have higher aspirations, and their parents spend significantly more time helping them with homework, when taught by these teachers. These results illustrate that, contrary to previous findings, in some settings different-type teachers may also promote educational attainment and aspirations, and thus intergenerational mobility.

JEL Codes: I24, I25, J15

Key words: social distance, learning outcomes, complementarities, caste

*I thank Harold Alderman, Kate Ambler, Jishnu Das, Stefan Dercon, Robert Garlick, Paul Glewwe, Berber Kramer, Sonya Krutikov, Eliana La Ferrara, Annamaria Milazzo, Albert Park, Simon Quinn, Abhijeet Singh, Kate Vyborny, and conference and seminar participants at IFPRI, NEUDC, the Lahore School of Economics and the University of Oxford for helpful comments. This paper is produced as part of the project Actors, Markets, and Institutions in Developing Countries: A micro-empirical approach (AMID), a Marie Curie Initial Training Network (ITN) funded by the European Commission under its Seventh Framework Programme Contract Number 214705 PITN-GA-2008-214705.

†International Food Policy Research Institute (IFPRI). 2033 K Street NW, Washington, DC, 20006, USA. Tel:+1(202)862-6476. Email: N.Karachiwalla@cgiar.org

1 Introduction

Social and economic inequality are substantial impediments to intergenerational mobility (Corak, 2013; Chetty et al., 2014). Social exclusion and marginalization can lead to low aspirations and to underinvestment in human capital (Black et al., 2015; Huillery and Guyon, 2014; Hoxby and Avery, 2012; Jacoby and Mansuri, 2011)[1], which can then lead to poverty traps (Becker and Tomes, 1979; Akerlof, 1997; Dalton et al., 2010; Kim and Loury, 2012). Both low aspirations and underinvestment in human capital contribute to the large observed gaps in learning for disadvantaged groups (Neal, 2006; UWEZO, 2013; Andrabi et al., 2007). Considerable evidence from high-income countries, and the limited evidence from low-income countries, shows that teachers who share the same characteristics (race, ethnicity, gender, caste) as their students can help marginalized students to learn better (Dee, 2004, 2005; Fryer and Levitt, 2004, 2006; Fairlie et al., 2014; Hanushek and Rivkin, 2006; Lindahl, 2007; Hanushek et al., 2009; Kingdon and Rawal, 2010; Muralidharan and Sheth, 2015). These findings highlight complementarities between teacher and student characteristics in the education production function, and they suggest a potential avenue by which children from marginalized groups can catch up to their relatively more advantaged peers.

While the evidence on same-type teachers is extensive in high-income countries, there is very little empirical evidence from low-income countries. This paper provides new evidence on complementarities between teacher and student characteristics in the education production function and their effect on learning and potential mobility using a rare lognitudinal data set from a low-income country that includes data on both teacher and student type (race/ethnicity/caste), as well as learning outcomes. I identify the effect of the interaction between a social characteristic of students and teachers, caste, on the learning outcomes of children in rural Pakistan. In South Asia, caste plays a similar role to that of race in high-income countries. I focus on whether child-teacher caste interactions affect the learning outcomes of low caste students, as they are the relatively

[1] Bowles et al. (2014) show theoretically that when returns to education depend on whether others in the same group invest, an equilibrium whereby nobody in the marginalized group invests and everyone in the high status group invests can persist.

disadvantaged group. Contrary to much of the previous evidence, I identify a positive effect of high caste teachers on low caste students. I also identify the effects of child and teacher caste separately in an attempt to understand the independent effects of these characteristics in the production function. I then investigate seven potential mechanisms that help explain the main findings.

I use the Learning and Educational Attainment in Punjab Schools (LEAPS) panel data set from Punjab, Pakistan. This unique and detailed longitudinal study tracks learning outcomes from independent tests of the same children over four years, and records the identity and characteristics of their teachers. Child fixed effects are used to identify the effect of the interaction between child and teacher caste through the same child switching between high and low caste teachers over time. Classroom (teacher-year) fixed effects are used to identify the effect of child caste by comparing high and low caste children in the same classroom. These specifications ameliorate many concerns faced by studies attempting to identify the causal effects of child and teacher characteristics on student outcomes. I also provide evidence on the independent effect of teacher caste through these specifications.

Identification concerns are addressed by showing that both the matching of students to teachers within schools, as well as the switching of children between teachers over time is quasi-random. Using Monte Carlo simulation methods, I show that the observed distribution of high and low caste matches of children to teachers is not significantly different from what would be observed under random assignment. Furthermore, a student's previous learning outcomes, as well as other observable characteristics, are orthogonal to whether the student switches caste-type of teachers. In this context, switching between teachers occurs naturally as children progress through grades as a vast majority of teachers tend to teach the same grade over several years (68% teach the same grade(s) for all years they are observed).

I report six main findings in the paper. First, estimating the child fixed effects specification, I find that contrary to previous literature, low caste children, and in particular low caste boys, have significantly higher test scores when they are taught by high caste teachers rather than by low caste teachers. This result is robust to restricting the sample to schools that pass that Monte Carlo simulation tests for random assignment, as well as to the addition of child, teacher, and school controls. Second, comparing children in the same classroom (i.e. taught by the same teacher in the same year), low caste boys have the highest learning outcomes relative to their peers. This finding, as well as results showing that teacher caste does not matter for high caste children but only for low caste children, provides evidence of the third finding: that high caste teachers are not just of higher quality in some unobserved way.

Next, I examine potential mechanisms. There are several theories why child-teacher caste interactions may matter for learning. The interaction between child and teacher caste could be important if either the same or different caste-type teachers can serve as role models to children (Dee, 2005; Hanushek and Rivkin, 2006; Kingdon and Rawal, 2010; Lindahl, 2007), or may discriminate against some children (Botelho et al., 2010; Hanna and Linden, 2012). Hanna and Linden (2012) found in a field experiment in India, that low caste children were graded between 0.03 and 0.09 standard deviations below high caste children with the same answers, and that it was actually low caste teachers who were discriminating. This finding is consistent with the notion that same-type teachers are not always beneficial. Child and teacher caste may also have independent effects on learning. Child caste could have negative effects if a school consists predominantly of either high or low caste children, and the other caste-type feels as if they are outsiders; outsiders tend to have lower learning outcomes (Akerlof and Kranton, 2010). Drawing attention to caste differences could also affect children's confidence and thus performance on a task through stereotype threat (Hoff and Pandey, 2006, 2011). Alternatively, there may be positive effects for a group if there are differing returns to education for high and low caste children; studies in India have found that low caste children have surprisingly high returns to human capital (Munshi and Rosenzweig, 2006; Luke and Munshi, 2007). Teacher caste could matter if high or low caste teachers are simply higher quality teachers.

The fourth and fifth findings help explain why low caste boys have higher learning outcomes when taught by high caste teachers. Low caste boys taught by high caste teachers have higher aspirations; they aspire to highly skilled and highly paid jobs, and they aspire to complete 1.5 additional years of schooling. In addition, low caste families invest more: the parents of low caste boys taught by high caste teachers spend an additional hour per week (16% more than the average number of hours spent per week) helping them with homework. Thus, mechanisms for improved learning outcomes for low caste children taught by high caste teachers exist on both the intensive (aspirations) and extensive (more time spent) margins. The sixth finding explains why low caste boys perform better relative to their peers in general. I use data on the children's parents and I find that returns to education are high for low caste men. This is consistent with the findings of Munshi and Rosenzweig (2006). In these data, I do not find support for other mechanisms such as children feeling like outsiders, role model effects, discrimination, or teacher quality.

The findings in this paper have important implications for both theory and for policy. Complementarities in social characteristics in the education production function are generally modelled with same-type teachers having positive effects on learning. The social dynamics between high and low caste groups is different from the ethnic and racial groups that have typically been studied, and it is important to study varying dimensions of social inequality. These results show that caste is different from previously studied aspects of social distance such as race or gender. Although I do not find evidence of discrimination by same-type teachers in these data with my measure of discrimination, this does not mean that there are not potentially other forms of discrimination at work, for example, in grading such as in Hanna and Linden (2012). Both this paper and Hanna and Linden (2012) show that caste complementarities may work differently from other social characteristics. Not all complementarities in social characteristics need be modeled in the same manner. Further, a policy focused on creating matches between teacher and student characteristics may be misplaced in some contexts; the implications of different types of matches on different groups must first be understood. The paper's findings illustrate that different-type teachers may also promote educational achievement and aspirations, and thus attainment and intergenerational mobility.

This paper makes several contributions. First, this paper contributes to the literature on teacher characteristics in low-income countries by explicitly examining complementarities between teacher and student characteristics. The assumption that teacher quality and teacher characteristics affect all students in the same way has recently been challenged. For example, Aucejo (2011) acknowledges and studies differences in teacher abilities, and differences in the way teachers allocate time to students of varying ability. In Pakistan, Rawal et al. (2013) finds that teacher attitudes towards teaching matter more for the learning of girls than for boys. Acknowledging and understanding complementarities in teacher and student characteristics can help us to improve learning by identifying potential gaps, as well as ways in which to close them.

The literature on social distance and learning has focused on the United States and other high-income countries; the evidence is sparse from low-income countries where inequality and lack of mobility are more of a hindrance to growth and development (Barro, 2000; Maoz and Moav, 1999). The South Asian equivalent of race in high-income countries is caste. The way in which caste affects children in India has been studied with regards to the issue of stereotype threat in performing a task (Hoff and Pandey, 2006, 2011), returns to education (Munshi and Rosenzweig, 2006; Luke and Munshi, 2007), and discrimination in grading in a field experiment (Hanna and Linden, 2012). The only paper studying learning outcomes and social distance in India (using caste) is Kingdon and Rawal (2010). There have been no such studies on Pakistan. One paper looks at education and caste in Pakistan (see Jacoby and Mansuri (2011)), but focuses on enrollment rather than on learning. There are large disparities between high and low caste groups in Pakistan (Gazdar, 2007; Mohmand and Gazdar, 2007; Jacoby and Mansuri, 2011; Vyborny and Chaudhury, 2012), particularly in Punjab where this social institution carried over during the conversion of Hindus to Islam. Although Islam is not meant to have caste, in Punjab, caste groups are well-defined; the distinction between high and low caste groups is clear and consistent across the province. It is also salient; both adults and children are aware of their caste group and status, as well as that of others. Thus, this paper fills an important gap by providing among the first estimates of the impact of social distance between teachers and students on learning in a low-income country setting.

Furthermore, the paper improves upon previous studies in low-income countries by employing an identification strategy that has not previously been possible. The closest analogue to this paper also uses child fixed effects; however, the variation is across teachers who teach different subjects rather than over time (Kingdon and Rawal, 2010). Unobserved subject-related differences that are also related to caste and gender are quite easy to envisage; for example, it has been well-documented that female and minority students are judged by teachers to be less able at mathematics (Good et al., 2003; Tomasetto et al., 2011; Beilock et al., 2010). This study uses variation over time, and averages the test scores from three different subjects to reduce the potential for bias from such effects. Because high caste teachers are no more likely to teach higher grades (including for the sample of switchers) and because attrition does not differ either by caste or by previous switches and types of switches, child fixed effects is appropriate in this setting. This is in addition to simulation tests that show that caste matching of students and teachers is pseudo-random, and to tests that show that switching of children between teachers is random.

The breadth of information collected in the LEAPS panel also allows several mechanisms to be explored. Previous studies do not explore several possible mechanisms in one context, rather, they specifically test for the existence of one mechanism. In this paper, I am able to test for seven different mechanisms (the caste composition of schools, returns to education, discrimination, role model effects, aspirations effects, household investments, and teacher quality). The LEAPS project also developed and administered independent tests. Thus, the results in this paper reflect actual learning, rather than potentially being confounded by teacher grading.

Finally, studying the effects of an existing allocation system of teachers and students has the advantage of representing what actually occurs in the schooling system. It reflects local politics and dynamics, and it also reflects the true management practices of schools and teachers. Pakistani educational markets are also increasingly similar to many other parts of the world. Several public and private schools exist in a single village, and there is a large degree of school choice. This pattern is increasingly prevalent both in other countries in South Asia, as well as other low-income countries (Andrabi et al., 2009). Thus, in addition to being internally valid, this study is also representative of the external context.

The next section describes the history and meaning of caste in Pakistan, as well as the data used in this study. The empirical model to be tested is described in Section 3, followed by the results from estimating the empirical model in Section 4. Section 5 concludes.

2 Caste and Education in Pakistan, and Data

2.1 Caste

Most of the literature on caste focuses on India, where caste is codified in religion. The notion of caste in Pakistan is very different from that of India. Officially, caste does not exist in Pakistan, because caste is not recognised in Islam. However, people in rural Punjab (the province on which I focus) are very cognizant of their caste group and that of others, and are also aware of the social conventions surrounding caste that have endured despite the official stance of the government. Several authors have noted its importance in relation to land ownership, employment, political disempowerment, health and education, and access to services (Gazdar, 2007; Mohmand and Gazdar, 2007; Jacoby and Mansuri, 2011; Vyborny and Chaudhury, 2012). Caste is certainly a salient social division in Punjab.

In Pakistan, caste is a social, rather than a religious institution, and more appropriate terms include clan, kin, and tribe (*zaat/biraderi*). Castes are based on 'lineage endogamy', with patrilineal cousin marriage forming the basis for the kin networks. Intermarriage across caste groups is very rare. A child inherits the caste of his/her father, so caste is fixed over generations. People in a particular caste group tend to inherit the occupation of their predecessors. This is now becoming less prevalent as low caste groups are able to take up other professions, but the social implications have endured.

Castes in Punjab are classified according to historical land ownership and occupation. The hierarchy consists of land owners (*zamindars*), tenants/cultivators, service/artisan professions (such as weavers, iron smiths etc.), and finally outcast groups (menial labour professions) (Ibbetson, 1974). Land owners and tenants are considered high caste, and service and menial labour professions are considered low caste (Mohmand and Gazdar, 2007; Jacoby and Mansuri, 2011).

Within both the high and low caste categories, there is a range, with some castes considered to be of higher status than others. For example, the group '*Syed*' carries the highest status across Punjab, and other groups such as '*Jat*' and '*Rajput*' are considered of lower status than *Syeds*, but both are universally considered high caste. Here, I do not distinguish between the order of high and low caste groups, I distinguish only between those considered low caste and those not considered low caste. This is because a *Rajput* person will not treat a *Jat* person differently, but they may treat a low caste individual differently (Jacoby and Mansuri, 2011). Consequently, from now on any reference to caste can be thought of as caste-type (high or low) as this is the relevant social distinction.

In Punjab, caste is closely connected to social status and dignity, and there is a large degree of social stereotyping, which is derogatory to lower caste groups. There are historical rules of social interaction that govern relationships between groups. For example, high caste groups will not eat with lower caste groups (Ibbetson, 1974). Such practices are evident in schools as well. Jacoby and Mansuri (2011) note that low caste children are often made to sit either on the floor or outside the classroom, and are not permitted to use the same latrines as high caste children. This is despite the fact that many caste groups will live together in the same village, and schools will also contain a mix of high and low caste children (Andrabi et al., 2007).[2]

[2]The average school consists of six caste groups.

2.2 Data

This study uses the Learning and Education Attainment in Punjab Schools (LEAPS) survey, which is a longitudinal data set collected in each year from 2003-2006. There are very few longitudinal studies in low-income countries that collect test score data as well as rich household level data that allows for the inclusion of many controls.

112 rural villages in the three districts of Attock, Faisalabad, and Rahim Yar Khan were surveyed. The LEAPS project contained two survey components: a household survey and a school survey. A detailed household survey was conducted for a representative sample of 1,800 households. The households were oversampled to contain a child in grade 3. As a result, many of the children in the household survey who were in grade three in 2003 were also tested as part of the concurrent LEAPS school survey testing all children in grade 3 (an overlap of almost 1,200 children were administered both surveys and tests). Tests were administered in mathematics, English and Urdu (the official language in Pakistan). These tests were meant to be general, and a similar test was administered each year. The tests were scored using Item Response Theory (IRT) in order to make the tests comparable over time, and to measure underlying ability. IRT does this by estimating different weights to correctly answered questions depending on the difficulty of the question. The knowledge scores generated by this process represent a student's underlying knowledge in a particular subject. The knowledge scores are then converted into standard deviations from the mean for ease of interpretation. Teacher tests in the same three subjects, as well as head-teacher and teacher questionnaires, were also administered along with the school survey.

Both teachers and students are grouped into high and low caste groups, based on the caste reported by the household or teacher. I individually classify 22 caste group names into the high or low caste categories by consulting two historical sources. The first is H. A. Rose's *A Glossary of the Tribes and Castes of Punjab and North-West Frontier Province*, which is the most extensive work detailing all castes in Punjab (Ibbetson and Maclagan, 1911). The second source is the District Gazetteers for the three districts in the sample, which were compiled by the British during the colonial era and include detailed descriptions of the occupations, customs, and land ownership of castes and subcastes ((Pakistan), 1932, 1996; Dīn, 2001).[3] These two sources were consistent in their classifications of caste group names. This procedure ensures that the caste variables are pre-determined.

Panel A of Table 1 provides descriptive statistics of the 1,166 children in the sample, who are observed for between one and four years from 2003-2006 (2,582 child-year observations in total). High caste children comprise 75% of the sample, and this is consistent with the population of Punjab[4], and there are also more boys than girls (since these are all children enrolled by grade 3). High caste children tend to be enrolled in private school more than low caste children, and girls tend to have enrolled in school later. High caste households are much more likely to own land, and parents are more educated on average.

Although there are differences in characteristics between high and low caste groups, there is also a substantial degree of overlap. Figure 1 plots the densities of some household characteristics for high and low caste households. In both high and low caste groups, there are households who are wealthy and households who are not wealthy, households who have large families and those with small families, households with educated fathers and with uneducated fathers, and households who do own land and those who do not. As a result, differences between high and low caste groups are not solely attributable to background characteristics, and more and less wealthy people in the same caste group will have similar social standing in the village (Mohmand and Gazdar, 2007). Caste is a social concept, and it does not map perfectly onto wealth and income.

Table 1 also contains descriptive statistics on teachers in the sample by caste type, in Panel B. High caste teachers tend to be female, more experienced, and teaching in the same village from which they come. However, they tend to have lower levels of education than low caste teachers. Test scores of high and low caste teachers do not differ in

[3]High caste groups include: Aarain, Abbasi, Awaan, Baloch, Butt, Chachar, Gujjar, Jat, Laar, Mughal, Naich, Pathan, Qureshi, Rajput, Samija, and Syed. Low caste groups include: Ansari, Mohana, Muslim Sheikh, Rehmani, Sheikh, and Solangi. The category 'other' consists of very small groups and they are generally not considered high caste, so are classified as low caste. However, all the results are robust to dropping the 'other' category.

[4]This is also consistent with other studies, both in Pakistan and in India. In Pakistan, Jacoby and Mansuri (2011) also find that 75% of their sample of children is high caste. In India, Kingdon and Rawal (2010) find that 76% of students are 'high caste' (not Scheduled Caste or Scheduled Tribe).

any of the three subjects, nor are there large differences in comparing absenteeism for high and low caste teachers. Table 1 also provides information on the schools in the sample. Schools are generally small, with fewer than 200 students, and approximately 5-6 teachers. Approximately 5 different caste groups tend to be represented on average in a school, and the caste fractionalization index[5] is approximately 0.5, indicating that there is a moderate degree of fractionalization. High caste children tend to more frequently attend high caste dominant schools, and there are no differences between school facilities that high and low caste children attend.

2.3 Assignment of students to teachers and switching

There are two potential concerns with the data: non-random matches of teachers and students, and non-random switching of students between teachers. Table 2 provides summary data on matches and switches between caste types of children and teachers. Panel A gives the number of male/female children of high/low caste-type matched with male/female teachers of high/low caste-type. Panel B provides the number of children who switch from either a low caste teacher to a high caste teacher, or vice-versa. Since public schools are segregated by gender, gender matches are very common. The number of switches of students between teachers of different caste types is low; there are 104 switches over the four years. One reason that the number of switches is low is that a high proportion of students, and also of teachers, are high caste. However, given the proportions observed in the data, the proportions of matches are very close to those expected from random matching. Switching between caste-types of teachers in this context comes mainly through children moving through grades. Teachers tend to teach a particular grade for long periods of time, and so when a child is promoted to the next grade, that is when the switch between teachers occurs. In these data, 68% of teachers teach the same grade(s)[6] over all years they are observed in the study (I use only teachers who are observed more than once).

[5]Constructed analogously to ethno-linguistic fractionalization indices.

[6]The survey asked all teachers in the school which grades they taught, split into three categories: grade 3 and below, grades 4 and 5, and grade 6 and above.

In order to uncover the presence of non-random matching of teachers and students based on caste, I perform two tests based on Monte Carlo simulations. I first calculate the frequencies of each of the four types of caste matches in each school (teacher high caste, child high caste; teacher high caste, child low caste; teacher low caste, child high caste; teacher low caste, child low caste) that is observed in the data. Subsequently, within schools, I pool all four years and randomly assign teachers to classes within the school.[7] This procedure takes the structure of each class as given and the number of classes within a school as given, since we cannot have third and sixth grade children in the same class, for example. In addition, the number of children in a class is then held constant. I then compare the observed proportions of matches to matches from the randomly assigned teachers and students in two ways. I first use the Chi-squared goodness of fit test for each school (266 schools). Under the null hypothesis of random matching, the distribution of simulated caste-type matches approximates the actual distribution of matches. For all schools except one, the Chi-squared p-value is above 0.05[8], suggesting that these data do resemble random matches of teachers and students within schools. In the second test, I use the simulated matches to construct 95% confidence intervals for each of the four match types in a school, and check what proportion of schools fall within this interval. For all four categories, this proportion of schools is above 85%, and for the categories of low caste children matched to high and low caste teachers, the proportion of schools passing the test is above 90%.[9] 46 schools (17.5%) fail the test in at least one category. These tests suggest that the matching process of students and teachers is pseudo-random at least within schools. As a result, all specifications include fixed effects at the school level or smaller.[10]

I also check whether teachers match to schools based on caste. Not all teachers in

[7] I omit schools for which I have data on only one teacher. This leaves 266 schools within which children and teachers are randomly assigned to one another. The procedure amounts to asking, 'what if this teacher taught class x this year instead of last year'. This random matching is performed 1,000 times.

[8] The smallest p-value is 0.019, and the average p-value is 0.964.

[9] The proportions of schools falling within the 95% confidence interval are: 85.2% for high child high caste, teacher high caste; 93.2% for teacher high caste, child low caste; 85.2% for teacher low caste, child high caste; and 96.2% for teacher low caste, child low caste.

[10] Although the main results use specifications with fixed effects at the classroom and child level, due to data limitations, some of the exploration of mechanisms will require school fixed effects.

a school were surveyed, but the sample of teachers surveyed in each school is random. A regression of a dummy variable for a high caste teacher on over fifty school controls produces only seven significant correlates (three at the 10% level, three at the 5% level, and one at the 1% level, see Appendix Table 6). This finding alleviates concerns of unobserved teacher factors influencing caste-based matching to schools.

Switching (or lack thereof) of students to different caste-types of teachers could be the result of several possible factors: grade promotion, switching schools, and parental lobbying. Grade promotion is the most common. It is common for schools in Punjab to have teachers teach the same grade for several years, so as a child progresses through to higher grades, this is the natural way to switch caste-types of teachers. Encouragingly, high caste teachers are no more likely to teach grade five and above than low caste teachers. Very few students and teacher switch schools. Only 2.6% of children in the sample switch schools between rounds, and only 2.7% of teachers switch schools between rounds. There are only nine cases in which the switch between high and low caste teachers is because of a child switching schools. I omit these observations from the analysis (leaving 104 switches).[11][12] Given that the matches of teachers and students to one another approximates what would occur if matched randomly, parental lobbying should not be a major concern in this context.

Finally, I run a regression of switching between high and low caste teachers on lagged test scores, as well as the many child, household, teacher, and school controls that are also used in the main specifications. The regressions are run both with school fixed effects, and with child fixed effects, and are run for the entire sample, as well as separately by child caste and gender. The results are contained in Table 3. Lagged test scores do not predict switching between high and low caste teachers. Observed child-level and household-level variables also do not predict whether a child will switch between caste types of teachers in a systematic and predictable way. The children who do switch are not different from those who do not switch. These tests suggest that although students

[11] The results are robust to omitting all children that switch schools (61), regardless of whether they switched caste-type of teachers.

[12] Teachers for the sample of 104 students who switch are also no more likely to teach grade 5 and above.

were not randomly allocated to teachers in the sense of an experiment, the matching of teachers and students, and the switching of children between caste-types of teachers, is pseudo-random.

3 Testing for Caste Effects

The econometric model is estimated from an education production function of the form:

$$T_{ijst} = f(\boldsymbol{X_i}, \boldsymbol{X_{it}}, \boldsymbol{X_j}, \boldsymbol{X_{jt}}, \boldsymbol{X_s}, \boldsymbol{X_{st}}, \boldsymbol{t}, \alpha_i, \alpha_j, \alpha_s, \epsilon_{ijts}) \tag{1}$$

where T_{ijst} is the test score of child i, taught by teacher j, in school s at time t. I use the average of math, English, and Urdu test scores.[13] $\boldsymbol{X_i}$ and $\boldsymbol{X_j}$ are fixed child and teacher observable characteristics, respectively. They include H_i and H_j, which are dummy variables for a high caste child and a high caste teacher, respectively. $\boldsymbol{X_s}$ are fixed school observables. $\boldsymbol{X_{it}}$, $\boldsymbol{X_{jt}}$, and $\boldsymbol{X_{st}}$ are time-variant child, teacher and school observables, respectively. $\boldsymbol{t}$ are time dummies, α_i, α_j and α_s are the child-specific, teacher-specific and school-specific time-invariant unobserved components, and ϵ_{ijts} is a random component.

The main threats to the validity of OLS estimation of equation (1) are non-random matching of children and teachers to schools, and non-random matching of children and teachers to one another within schools. To address these issues, I employ a number of checks and techniques to control for non-random sorting.

3.1 Child caste

I begin with identifying the effect of child caste on its own, in order to understand the independent effect of this characteristic. To estimate the effect of child caste on learning outcomes, a teacher-year (classroom, i.e. jt) fixed effects specification will be used:

$$T_{ijst} = \beta_1 H_i + \beta_2 \boldsymbol{X}_i + \beta_3 \boldsymbol{X}_{it} + \alpha_{jt} + \epsilon_{ijst} \tag{2}$$

[13]The results are consistent when separated by subject.

where both teacher and school observed and unobserved time variant and invariant factors have dropped out, as has the time trend, as they are collinear with the fixed effect. H_i allows us to compare high and low caste children taught by the same teacher in the same year (in the same classroom). This specification addresses many concerns of estimations of education production functions, namely, the possibility of non-random matches of teachers and students both across and within schools, as well as teacher unobservables such as innate ability.[14] However, in this specification, child unobservables do remain unaccounted for.

Another potential concern is that caste may be proxying for other characteristics. Consequently, I control for observed child and household-level characteristics that could be correlated with caste and with unobservables. $\boldsymbol{X}_t$ and $\boldsymbol{X}_{it}$ include the age and gender of the child, a household asset index, dummy variables for the grade level of the child[15], a dummy variable for whether the child is repeating the grade, dummy variables for each parent being uneducated, and household size.[16] In a separate specification, I also interact child caste with child gender, in which case three dummy variables are included instead of H_i: high caste female, high caste male, and low caste female. The omitted category is then low caste male children. These specifications will give us an idea of how low caste children perform relative to others in the presence of social inequalities.

3.2 Child-teacher caste interactions

Next, I will focus specifically on the disadvantaged group (low caste children)[17] and on the effects of social distance (the interaction between child and teacher caste). To estimate the effects of the interaction between child and teacher caste, a child fixed effects specification will be used:

[14]For example, if higher ability high caste children sort into better schools or classrooms, the effect of the high caste child dummy variable will be overestimated.

[15]Some schools have multi-grade teaching, particularly private schools.

[16]There is some missing data on parental education. For those cases, I include a missing data dummy variable.

[17]Results for both high and low caste children are contained in the appendix, and discussed in Section 4.

$$T_{ijst} = \beta_1 H_{jt} + \beta_2 \boldsymbol{X_{it}} + \beta_3 \boldsymbol{X_j} + \beta_4 \boldsymbol{X_{jt}} + \boldsymbol{t} + \alpha_i + \epsilon_{ijst} \tag{3}$$

where H_{jt} is a dummy variable for a high caste teacher in year t. Since this is a within-child specification[18], β_1 represents the effect of being taught by a high caste teacher relative to being taught by a low caste teacher. Identification is off children who switch between high and low caste teachers over time.

I include the same child and household controls in (3) as in (2). High and low caste teachers may also differ systematically in their observable and unobservable characteristics, and so it is important to control for observable characteristics. As teacher controls I include the teacher's age, number of years of teaching experience, dummy variables for the highest level of education, and dummy variables for whether the teacher is female and is originally from the same village in which the school is located. Teachers were also given the same tests as the children, and I include the average of the teacher's three scores. As school controls I include the number of children enrolled in the school, an index measuring caste fractionalisation at the school[19], a dummy variable for private school, and a school facilities index.[20] Finally, time dummies are included to control for the fact that children tend to perform better on the test over time.

Child fixed effects control for non-random sorting of children and teachers to schools, but do not control for children or teachers switching schools. The only concern with child fixed effects is non-random switches between high and low caste teachers. If low caste children who perform well switch to a better school, or to a school with more high caste teachers, then this will overestimate the effect of high caste teachers. I omit the nine switches between teacher caste types that are a result of switching schools.

A key assumption for child fixed effects to be valid is that past error terms are not

[18]Cassan (2012) has noted that caste groups in Pakistan have manipulated their identity in the past. However, in order for this type of manipulation to affect the results, it would need to be the case that households mis-report their caste in response to the switching of their child's teacher over the years of the survey. This is not the case in these data.

[19]The index is defined as: $1 - \sum_i s_i^2$ where s_i is the share of caste group i in the school.

[20]This index is constructed from a number of questions on the survey that ask about whether the school has the following facilities: library, computer, sports ground, hall, surrounding wall, fans, electricity, the type of toilet, whether potable water is available, and the seating arrangements for students (desks, mats, etc.). Principal Components Analysis was used to aggregate these measures.

correlated with the switching of teachers. For example, if children who perform better (or worse) last year switch to high caste teachers, this would bias coefficients. As shown above, the children who do switch are not different from those who do not switch, including in terms of their past test scores. Child fixed effects also requires that unobserved inputs such as motivation, preferences for an own caste-type (or opposite caste-type) teacher, and any others, be constant over time and not affect switching of teachers. Finally, the fixed effect (for example, child ability) also must be constant over time for child fixed effects to difference it out.

Finally, non-random attrition in the sample could bias coefficient estimates. If the lowest ability low caste children drop out of school, then the remaining ones will induce positive matching of teachers and students by caste. Most children are observed in at least three rounds of the survey, and there are no significant differences in attrition between high and low caste children. However, if the reasons for which children drop out of school or the sample are different for high and low caste children, or between children who do and do not switch cate-type of teachers, attrition could still bias estimates. To explore whether this is a potential threat to validity, I check the reasons for dropout of children both from the sample and from school. There are no significant differences between high and low caste children in dropping out of the sample due to switching schools and not being found, moving villages, being absent on the day of the test, or in the reason for not being found being unknown. To check what drives dropout from school, I regress having dropped out of school on the high caste dummy variable, lagged average test scores, high caste dummy interacted with lagged test scores, a dummy for whether the child previously switched caste-type of teacher, high caste dummy interacted with child previously switched caste-type of teacher, a dummy for whether the child currently switches caste-type of teacher or would have had she not dropped out of school[21], high caste child interacted with whether the child currently switches caste-type of teacher (or would have) as well as the child and household variables used in (2) and (3) (age, gender, grade, household

[21]This variable is constructed by assigning the caste-type of the teacher the child would have received assuming the child stayed in the same school and was promoted to the next grade, and then checking if a switch between caste-types would have occurred.

asset index, not-promoted dummy, father uneducated, mother uneducated).[22] None of these variables significantly predict dropout (see appendix Table 7, column (1)). I also regress having dropped out of school on dummy variables for whether the child previously switched to a high caste teacher and whether the child previously switched to a low caste teacher (separately for high and low caste children), dummy variables for whether the child currently switches or would have switched to a high caste teacher or to a low caste teacher (again separately for high and low caste children), as well as a dummy for high caste children, lagged average test scores, and high caste interacted with lagged average test scores, and the child and household controls. None of the switching variables, nor the test score or child and household variables predict dropout (see appendix Table 7, column (2)).[23] Thus, bias due to attrition should not be a concern in this context.

4 Results

This section will discuss the results of estimating the empirical models presented in the previous section. I will begin with identifying the effects of child caste, and the interaction of teacher caste with student caste for low caste children, and then will look into some potential mechanisms by which the findings could be occurring.

4.1 Child caste, and the interaction between child and teacher caste

Table 4 Panel A presents the independent effect of child caste on learning outcomes. The outcome is the average of math, English and Urdu test scores for each child. Column (1) includes the coefficient on the dummy variable for a high caste child, relative to a low caste child in the same class, and provides a basic correlation. Column (2) adds child, teacher and school controls to the specification of column (1). There are no differences in learning between high and low caste children in the same class. Columns (3) and (4) interact child caste with child gender. The three dummy variables represent the difference

[22] Village fixed effects, time dummies, and clustered standard errors at the village level are also included.
[23] Village fixed effects, time dummies, and clustered standard errors at the village level are also included.

between low caste boys (the omitted category) and high caste girls, high caste boys, and low caste girls in the same class, respectively. High caste girls perform no worse than low caste boys. Both high caste boys and low caste girls, however, perform less well than low caste boys. This is consistent with the raw data. These results show that, similar to the findings of Munshi and Rosenzweig (2006) and Kingdon (1998), conditional on enrolment in school, low caste children perform well.

Panel B of Table 4 presents results on the effect of the interaction between child and teacher caste (social distance) for low caste children. Column (1) presents a basic correlation, and column (2) adds child, teacher and school controls. Columns (1) and (2) show that social distance does matter. Low caste children have significantly higher learning outcomes when taught by high caste teachers. Switching from a low caste teacher to a high caste teacher increases the test scores of low caste children by 0.21 standard deviations on average, and this effect is significant at the five percent level.[24] There is disassortative complementarity between teacher and student caste. This finding is contrary to studies in the United States and in India where same-type teachers benefit students. In Pakistan, the relatively more advantaged group of teachers has a positive impact on the group of relatively disadvantaged students.[25]

As a robustness check, I re-run specification (3) in two ways. I restrict the sample first to schools that passed the Chi-squared distribution test. I then restrict the sample to schools for which the observed frequencies of all four possible caste matches fell within the 95% confidence interval for the caste match groups, which was constructed using the simulated data. Columns (3) and (4) contain the results. When the sample is restricted in these ways, the results are consistent. For low caste children, switching to a high caste teacher increases test scores by 0.48 standard deviations (significant at the 10% level) when the sample is restricted to schools that pass the Chi-square distribution test, and

[24] The regressions were also run separately for mathematics, English, and Urdu test scores (results not reported). The coefficients on the dummy variable for low caste children taught by high caste teachers was positive for all three subjects, and was significant at the ten percent level for mathematics and at the five percent level for English.

[25] There is a literature on teacher ability in the education production function, so it is interesting to look at in this setting. Surprisingly, the coefficient on the teachers' average test scores is quite small in comparison to the effect of teacher caste. The coefficient is approximately 0.04 and is not significant.

increases test scores by 0.22 standard deviations (significant at the 5% level) when the sample is restricted to schools that pass the confidence interval test.

To ensure that the relevant social distance measure is high and low caste, I consider whether it matters if the child and teacher belong to the same caste group (Panel C). It is possible that having a teacher of the same caste group matters, rather than the caste type. I construct a dummy variable equal to one if the child and teacher share the same caste (*biraderi*, endogamous kinship group) and include this variable in equation (3). For low caste children, test scores are 0.21 standard deviations lower when they belong to the same caste group as their teacher. This the same magnitude as the coefficient on caste-type of teachers instead of the exact caste group.[26]

Finally, Panel D of Table 4 investigates the heterogeneity of the child-teacher caste interaction effect further, and shows how the results differ by gender of the child. For each gender, the omitted category is being taught by a low caste teacher. Column (1) includes only child fixed effects and no controls, column (2) adds child controls, column (3) adds teacher controls, and column (4) adds school controls, to make the full set of controls. Although these results are based on small sample sizes, the results are relatively stable over specifications. The effect of a low caste child performing better when taught by a high caste teacher appears to be stronger for boys than for girls. For low caste boys, switching from a low caste to a high caste teacher increases test scores by 0.34 standard deviations, and this is statistically significant at the one per cent level.[27] It is also economically significant; in a recent review of several education interventions in low-income countries, McEwan (2013) finds that some of the best interventions produce effect sizes of 0.15 standard deviations.

What is Caste?

[26] I also test whether it matters that the child and teacher share the same gender. I run the same specification as (3) but with a dummy variable taking on the value of one if the child and teacher share the same gender. I do this separately for boys and for girls. The coefficients are not statistically significant.

[27] I also separate the sample into public and private schools (results not reported). The effect of a high caste teacher is significant in both types of schools. However, the p-value in a test for the equality of the coefficient on 'child low caste male, teacher high caste' between the public and private school specifications is 0.28, so they are not significantly different from one another.

Since there are important effects of caste on learning outcomes, it is important to understand what these social differences between castes mean. It is possible that this measure of social distance is capturing something else that is related to caste. Although there is much overlap in characteristics such as education, wealth, and land ownership between high and low caste households, are these characteristics proxying for caste inequalities? I focus on children and look at two alternatives: household wealth and parental education. Instead of splitting children into high and low caste, they are split into not poor and poor, and into educated and uneducated parents.[28] I use the household asset index and classify households as poor if they fall into the bottom quintile of this index. The average test scores of the children are then regressed on dummies analogous to the teacher-student caste interaction dummies, but with wealth and parental education for children, and high and low caste as before for teachers. Table 8 in the Appendix contains the results for wealth in column (1) and for parental education in column (2). The effect of a high caste teacher is not significantly different for boys who are 'poor' versus 'not poor' (p-value of test of equality between coefficients 0.164), nor between boys with educated versus uneducated parents (p-value 0.618). This shows that children's caste does not map completely onto wealth, nor onto household education. It is important to understand what caste means for teachers, as well.[29] This will be discussed in the following section.

4.2 Mechanisms

In this section, I discuss some patterns in the data that can help uncover the mechanisms behind the results in the previous section. Why do low caste boys perform well relative to their peers? Is it because they predominantly attend low caste dominant schools? Or perhaps they have high returns to education. Why are high caste teachers so productive for the learning outcomes of low caste boys? I look at a measure of teacher discrimination,

[28] A child is classified as having uneducated parents if both parents have never attended any school, and is classified as having educated parents if at least one of the two parents has attended school.

[29] This analysis is also conducted by, in addition to students, grouping teachers into educated and uneducated parents, and into poor or not poor by wealth quintile, and the results are consistent. Caste is not equivalent to education or wealth for teachers or students.

at whether teachers may be serving as role models to children, at whether they inspire high aspirations, and at whether children and parents invest relatively more. Finally, I consider whether teacher quality matters.

First, I look at mechanisms to explain why low caste boys have such high learning outcomes. If the schools that low caste boys attend are predominantly comprised of low caste children, this could be why we see that low caste boys outperform high caste children in the same class. The caste composition of a school may affect whether children feel like insiders or outsiders, and outsiders may have poor learning outcomes. Column (1) of Panel A in Table 5 displays results from a school fixed effects regression of average test scores for low caste children on a dummy variable for whether a school is high caste dominant. A school is considered high caste dominant if the proportion of children enrolled in grades one to five who are high caste is greater than the proportion of children enrolled in grades one to five who are low caste. The coefficient reflects the difference in test scores between low caste children in high caste dominant schools relative to low caste dominant schools. Identification comes from the changing caste composition of the school over time. The caste composition of the school does not affect learning outcomes; the coefficient on the dummy variable for high caste dominant schools is very small and insignificant.

I also test whether low caste children have high returns to education. If they do, this can help explain why we see that low caste children (especially boys) have higher test scores than high caste children. I look at the parents of the children in the sample and compare returns to education for high caste and low caste men in the same village.[30] Figure 2 plots the results from regressing the log of monthly earnings (including earnings from self-employment, farming, and wages) separately for high and low caste men on dummy variables for having completed primary school, middle school, high school, or more education than high school. The omitted category is either never having attended school, or not having completed primary school. I include many individual and household

[30] Only 105 women report positive earnings. In addition, it is boys whose behaviour is most affected. As a result, I estimate the earnings functions only of fathers.

level controls.[31] Returns to completing primary school are identical (and indistinguishable from zero) for both high and low caste men. Returns to completing middle school, high school, and more than high school appear higher for low caste men than for high caste men (see also Appendix Table 9).[32] However, the coefficients for high and low caste men are not individually or jointly significantly different from one another. What this does show is the importance of education for economic mobility for the low caste male children in the sample. Returns to education are as high for low caste boys as for high caste boys, so perhaps this is why they have high test scores.

Next, I look into mechanisms to explain why low caste boys perform so well when taught by high caste teachers. It is possible that the reason is that low caste teachers are discriminating against them. This would be consistent with the findings of Hanna and Linden (2012). I cannot test whether teachers are biased when they are grading children. However, the way in which teachers perceive and judge the performance of children could potentially be important for learning outcomes. To test this, I construct a measure that can be thought of as measuring biased perceptions. Teachers were asked to rate children on their academic performance on a scale of one to ten. I rank children in the same class according to these ratings. I also rank children in the same class according to an average of their three test scores. I then subtract the teacher's ranking from the test ranking, and create a dummy variable for whether this difference is negative (meaning that the teacher overrates the child's performance). For low caste children, this outcome is regressed on two dummy variables, as in Panel D of Table 4: low caste girls taught by a high caste teacher, and low caste boys taught by a high caste teacher. The same controls are included as in equation (3), along with child fixed effects and time dummies. The results are contained in column (2) of Table 5 Panel A. The coefficient on the dummy variable for low caste boys taught by high caste teachers is not significantly different from

[31]Controls include: age, unmarried dummy, subjective overall health (scale of 1-16), can read, can write, can count, can add, number of years to achieve highest grade, household size, household has audio, print, visual, internet media, number of deaths in household in past year, whether a member left the household in past year, household has member abroad, remittances, whether household owns land, whether household owns the house, total expenditure, asset index, whether harvest worse than previous year, whether earnings lower than previous year.

[32]This is the rate of return, not absolute earnings, that appear higher.

zero. It does not look as if high caste teachers are over- or under-rating the academic performance of low caste boys.[33]

I also test for the presence of role model effects and aspiration effects. One interpretation of role model effects is that teachers who are more similar to the child could elicit better performance. However, this does not seem to be the case in these data, as it is high caste teachers who produce high test score outcomes for low caste boys. Rather, opposite caste-type teachers may be serving as role models. Another way to interpret role model effects is that a child may like to emulate his/her teacher. I use data on what the children report wanting to do when they grow up. I construct a dummy variable for whether the child would like to be a teacher (27% of the sample overall). To look at aspiration effects, I also construct a dummy variable for whether the child reports wanting to have a skilled job apart from teaching (doctors, government jobs, politicians, engineers, or private sector jobs). High caste teachers may be able to inspire children to have high aspirations. Their family and friends may have well paid and highly respected jobs, and may be able to pass on information and experiences to children. I regress these outcomes on child and teacher caste-type and gender dummies within schools.[34] The omitted category is a low caste male child matched with a low caste teacher. I include the same controls as in equation (3) and also include the average test score, and measures for the parents' assessment of how intelligent and hardworking the child is, as these could also influence their aspirations. The results are reported in columns (4) and (5) of Table 5 Panel A.

[33]Teachers may actually be able to rate children's performance quite accurately if they have more information on the child's ability than one year of test scores can provide (for example, they may be aware of previous performance or family circumstances). The test score measure is also noisy. If this is the case, there will be systematic measurement error in the outcome variable that is correlated with caste. If teachers of the same caste-type as a child have better information than teachers of the opposite caste-type as a child, then the coefficient for low caste children taught by high caste teachers will be biased downwards (upwards) if true ability is higher (lower) than ability demonstrated by the test score. As a check, I construct a test score measure that is the average of not just the three current test scores, but also all test scores that occurred in previous years for which I have data (in the first year, only the three current-year tests are used). I then rank the students in the same class again according to this new test measure and compare the difference between the ranking of the teachers rating and the ranking of this cumulative test score. If teachers are using information on past test scores in their assessment, then including this information in the discrimination measure should reduce the measurement error in their ranking (possibly by less for low caste children, however). The results are contained in column (3), and are almost identical to those of column (2), so this type of systematic measurement error should not be a problem.

[34]There are not enough observations to include child fixed effects since the question was only asked in the third and fourth rounds of the survey, and so I include school fixed effects instead.

It appears that low caste boys report wanting to become teachers significantly less when they are taught by high caste teachers than by low caste teachers. Low caste boys taught by high caste teachers are marginally more likely to report wanting to have a high skilled job than low caste boys taught by low caste teachers. It does not appear as if low caste boys would like to emulate their teachers by also becoming teachers. Instead, they have higher aspirations and would like to go into other high skilled professions, many of which pay better than teaching. I also check whether students report aspiring to a higher grade when taught by high caste teachers. Using the same specification as in columns (4) and (5) I regress the highest grade the child reports aspiring to attain on the teacher-student caste match dummy variables (results reported in column (6)). Low caste male children aspire to complete 1.5 additional grades when taught by a high caste teacher rather than by a low caste teacher, although the coefficient is imprecisely measured. This is also consistent with high caste teachers inspiring higher aspirations among low caste boys.

I also look at child and household investments to see if additional investments are made when low caste boys are taught by high caste teachers. It could be that parents invest more in their children's education. Parents can do this by increasing educational expenditures, by helping the child with their schoolwork, or by meeting with the child's teacher regarding their performance. It could also be that children work harder, and this could be of their own accord or with encouragement from parents.

The school survey asked teachers whether they had met the child's parents in the last month. The household survey asked if anyone in the household had helped the child with his/her homework, and if so, how many hours of help was given per week. It also asked parents about monthly education expenditures for each child, and collected extensive data on child time use, including the number of minutes per day spent on housework and paid work, homework, leisure (including sleep/rest, play, and media), and learning (time at school and on private tutoring). Panel B of Table 5 displays regressions of these household investment and child time use outcomes on the same specification as (3) with the full set of child, teacher and school controls, as well as time dummies and child fixed effects. I also include the average test scores of the child as an additional control, as these could affect child and parental investments. Column (7) shows that low caste boys received significantly more help with homework when taught by high caste teachers, on average 0.97 hours more per week, and this effect is significant at the one percent level.[35] The mean number of hours per week spent helping with homework (for parents of low caste boys who do help) is 6.18, so an additional hour is a 16% increase. None of the other outcomes have significant effects for low caste boys. Low caste households are investing more; either low caste children are asking for more help, or parents are offering it when their children are taught by high caste teachers. This result could also relate to the result on aspirations; perhaps low caste boys ask for more help because they have higher aspirations when taught by high caste teachers. Thus, mechanisms for improved learning outcomes for low caste children taught by high caste teachers exist on both the intensive (aspirations) and extensive (more time spent) margins.

Finally, it could be the case that high caste teachers are simply higher quality teachers; teacher caste may matter independently of student caste. There are three pieces of evidence regarding teacher quality. The first comes from Panel A of Table 4. Here, teacher unobservables such as innate ability are swept away by the teacher-year fixed effect. Low caste children outperform high caste children with the same teacher. Secondly, high caste children perform equally well when taught by high and low caste teachers (see Appendix Table 10). In an analogous regression to equation (3) that includes both high and low caste children, the dummy variable for high caste children taught by high caste teachers is insignificant. Panel D of Table 4 also shows that including teacher observables in the regression does not change the coefficient on the caste dummy variables very much at all (comparing columns (2) and (3)). This indicates that teacher observable characteristics that could be related to quality also do not vary systematically by caste. So it is not the case that high caste teachers are just better teachers; some children just perform significantly better when taught by them. These results are not consistent with the notion that teacher quality affects all students approximately equally.

To summarize, low caste boys have high returns to education, and this could be

[35]Using the Bonferroni correction (Bonferroni, 1936) to correct for the eight outcome variables, the coefficient is significant at the 10% level of confidence.

why they outperform their peers. Low caste boys have higher aspirations, and their parents spend more time helping them with homework when they are taught by high caste teachers. The results are consistent with low caste households leveraging education as a means to move out of poverty, and with social distance between teachers and students potentially improving the scope for educational achievement and attainment, and thus intergenerational mobility.

5 Conclusion

This paper examines the impact of social distance between teachers and students in the form of caste on learning outcomes in Punjab, Pakistan. It contributes to the literature on the role of social characteristics in learning, as well as to the literature on the complementarities between child and teacher characteristics in the education production function with both previously unavailable data, and with strong identification. It identifies the causal effects of child caste, teacher caste, and the interaction between child and teacher caste on learning outcomes. Many measures are employed in order to reduce bias from non-random switching and sorting of children and teachers both across schools and within schools. Some possible mechanisms of how caste differences between teachers and students may affect learning outcomes are then explored: social dynamics in schools, returns to education, discrimination, aspirations, role model effects, household investments, and teacher quality.

I find that low caste boys have the highest levels of learning compared to other types of children in the same class. This could be because they have high returns to education; returns are just as high for low caste boys as they are for high caste boys. Learning outcomes of low caste children do not vary with the caste composition of the school. Teacher caste on its own does not matter for learning outcomes. It is not the case that high caste teachers are simply higher quality teachers. The interaction between teacher and child caste does matter for learning outcomes. I find that low caste male children have significantly higher learning outcomes when taught by high caste teachers compared to low caste teachers, and that they also receive significantly more help with their homework when taught by high caste teachers. I also find that low caste boys have higher schooling and career aspirations when taught by high caste teachers. These data show no evidence of discrimination by teachers in perceptions of child ability, or of role model effects. Finally, caste does not map perfectly either onto inequalities in wealth or in parental education levels.

The results of this study suggest that there are complementarities in teachers' and students' social characteristics in the education production function, and that these complementarities may occur in surprising ways. The literature in the United States and in India has found that teachers of the same race/caste as the student produce higher learning outcomes. This paper finds the opposite, which suggests that caste may be different from previously studied aspects of social distance such as race and minority status, and even from caste in India. The findings also suggest that there can exist aspects of learning that positively affect disadvantaged groups. If this relative advantage in learning, enhanced by teacher-student complementarities, translates into progression through the schooling system and better jobs for these children, they can then serve as role models and a source of networks for younger low caste children. Such effects could result in the narrowing of the gap between these social groups over generations.

References

Akerlof, G. A. (1997, September). Social Distance and Social Decisions. *Econometrica 65*(5), 1005–1028.

Akerlof, G. A. and R. E. Kranton (2010). *Identity economics: How our identities shape our work, wages, and well-being.* Princeton University Press.

Andrabi, T., J. Das, A. Khwaja, T. Vishwanath, and T. Zajonc (2007). The Learning and Educational Attainment in Punjab Schools (LEAPS) Report. *World Bank, Washington, DC*.

Andrabi, T., J. Das, and A. I. Khwaja (2009). Report cards: The impact of providing school and child test scores on educational markets. *Working Paper*.

Aucejo, E. M. (2011). Assessing the role of teacher-student interactions. *Manuscript, Duke University*.

Barro, R. J. (2000). Inequality and growth in a panel of countries. *Journal of economic growth 5*(1), 5–32.

Becker, G. S. and N. Tomes (1979, December). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy 87*(6), 1153–89.

Beilock, S. L., E. A. Gunderson, G. Ramirez, and S. C. Levine (2010). Female teachers math anxiety affects girls math achievement. *Proceedings of the National Academy of Sciences 107*(5), 1860–1863.

Black, S., K. Cortes, and J. Lincove (2015). Academic Undermatching of High-Achieving Minority Students: Evidence from Race-Neutral and Holistic Admissions Policies. *American Economic Review: Papers and Proceedings, forthcoming*.

Bonferroni, C. E. (1936). *Teoria statistica delle classi e calcolo delle probabilita*. Libreria internazionale Seeber.

Botelho, F., R. Madiera, and M. Rangel (2010). Discrimination goes to school? Racial differences in performance assessments by teachers. *Working Paper, Department of Economics, University of Sao Paolo*.

Bowles, S., G. C. Loury, and R. Sethi (2014, 02). Group Inequality. *Journal of the European Economic Association 12*(1), 129–152.

Cassan, G. (2012). Identity based policies and identity manipulation: evidence from Colonial Punjab. Technical report.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014, January). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. NBER Working Papers 19843, National Bureau of Economic Research, Inc.

Corak, M. (2013, Summer). Income Inequality, Equality of Opportunity, and Intergenerational Mobility. *Journal of Economic Perspectives 27*(3), 79–102.

Dalton, P. S., S. Ghosal, and A. Mani (2010). Poverty and aspirations failure.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics 86*(1), 195–210.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review 95*(2), 158–165.

Dīn, M. M. (2001). *Bahawalpur State Gazetteer with Map 1904*. Sang-e-Meel Publications.

Fairlie, R., F. Hoffman, and P. Oreopoulos (2014). A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom. *American Economic Review 104*(8), 2567–2591.

Fryer, R. G. and S. D. Levitt (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics 86*(2), 447–464.

Fryer, R. G. and S. D. Levitt (2006). The black-white test score gap through third grade. *American Law and Economics Review 8*(2), 249–281.

Gazdar, H. (2007). Class, caste or race: veils over social oppression in Pakistan. *Economic and Political Weekly*, 86–88.

Good, C., J. Aronson, and M. Inzlicht (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology 24*(6), 645–662.

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

Hanushek, E. A., J. F. Kain, and S. G. Rivkin (2009). New evidence about brown v. Board of Education: The complex effects of school racial composition on achievement. *Journal of Labor Economics 27*(3), 349–383.

Hanushek, E. A. and S. G. Rivkin (2006). School quality and the black-white achievement gap. *National Bureau of Economic Research WPS*.

Hoff, K. and P. Pandey (2006). Discrimination, social identity, and durable inequalities. *American Economic Review 96*(2), 206–211.

Hoff, K. and P. Pandey (2011). Making up people: Experimental evidence on identity and development from caste India. *World Bank mimeo*.

Hoxby, C. M. and C. Avery (2012, December). The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students. NBER Working Papers 18586, National Bureau of Economic Research, Inc.

Huillery, E. and N. Guyon (2014). The Inequality in Aspirations: Why do French Teenagers from Low Social Background Limit their Ambition?

Ibbetson, D. B. and E. Maclagan (1911). Glossary of the tribes and castes of the Punjab and North-West Frontier Province,(ed.) Rose, H. A.

Ibbetson, S. D. (1974). *Punjab castes*. Sh. Mubarak Ali.

Jacoby, H. and G. Mansuri (2011). Crossing boundaries: gender, caste and schooling in rural Pakistan. *World Bank Policy Research Working Paper Series*.

Kim, Y.-C. and G. Loury (2012, July). Social Externalities, Overlap and the Poverty Trap. MPRA Paper 54953, University Library of Munich, Germany.

Kingdon, G. G. (1998). Does the labour market explain lower female schooling in India? *The Journal of Development Studies 35*(1), 39–65.

Kingdon, G. G. and S. Rawal (2010). Akin to my teacher: Does caste, religious and gender distance between student and teacher matter? Some evidence from India. *Department of Quantitative Social Science Working Papers*, 10–18.

Lindahl, E. (2007). Gender and ethnic interactions among teachers and students–evidence from Sweden. Technical report.

Luke, N. and K. Munshi (2007). Social affiliation and the demand for health services: Caste and child health in south India. *Journal of development economics 83*(2), 256–279.

Maoz, Y. D. and O. Moav (1999). Intergenerational mobility and the process of development. *The Economic Journal 109*(458), 677–697.

McEwan, P. J. (2013). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Wellesley, United States: Wellesley College. Mimeographed document*.

Mohmand, S. and H. Gazdar (2007). Social structures in rural Pakistan. *Thematic paper prepared under TA4319, Determinants and Drivers of Poverty Reduction and ADBs Contribution in Rural Pakistan. ADB, Islamabad*.

Munshi, K. and M. Rosenzweig (2006). Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy. *American Economic Review 96*(4), 1225–1252.

Muralidharan, K. and K. Sheth (2015). Bridging education gender gaps in developing countries: The role of female teachers. *Journal of Human Resources, forthcoming*.

Neal, D. (2006, June). *Why Has Black-White Skill Convergence Stopped?*, Volume 1 of *Handbook of the Economics of Education*, Chapter 9, pp. 511–576. Elsevier.

(Pakistan), P. (1932). *Gazetteer of the Attock District, 1930*. Sang-e-Meel Publications.

(Pakistan), P. (1996). *Gazetteer of the Chenab Colony*. Sang-e-Meel Publications.

Rawal, S., M. Aslam, and B. Jamil (2013). Teacher characteristics, actions and perceptions: What matters for student achievement in Pakistan? *CSAE Working Paper Series, 2013-19*.

Tomasetto, C., F. R. Alparone, and M. Cadinu (2011). Girls' math performance under stereotype threat: the moderating role of mothers' gender stereotypes. *Developmental psychology 47*(4), 943.

UWEZO (2013). Are our children learning? Literacy and Numeracy Across East Africa 2013.

Vyborny, K. and A. Chaudhury (2012). Patronage and the poor: Evidence from rural Punjab, Pakistan. *mimeo*.

Figure 1: Characteristics of High and Low Caste Households

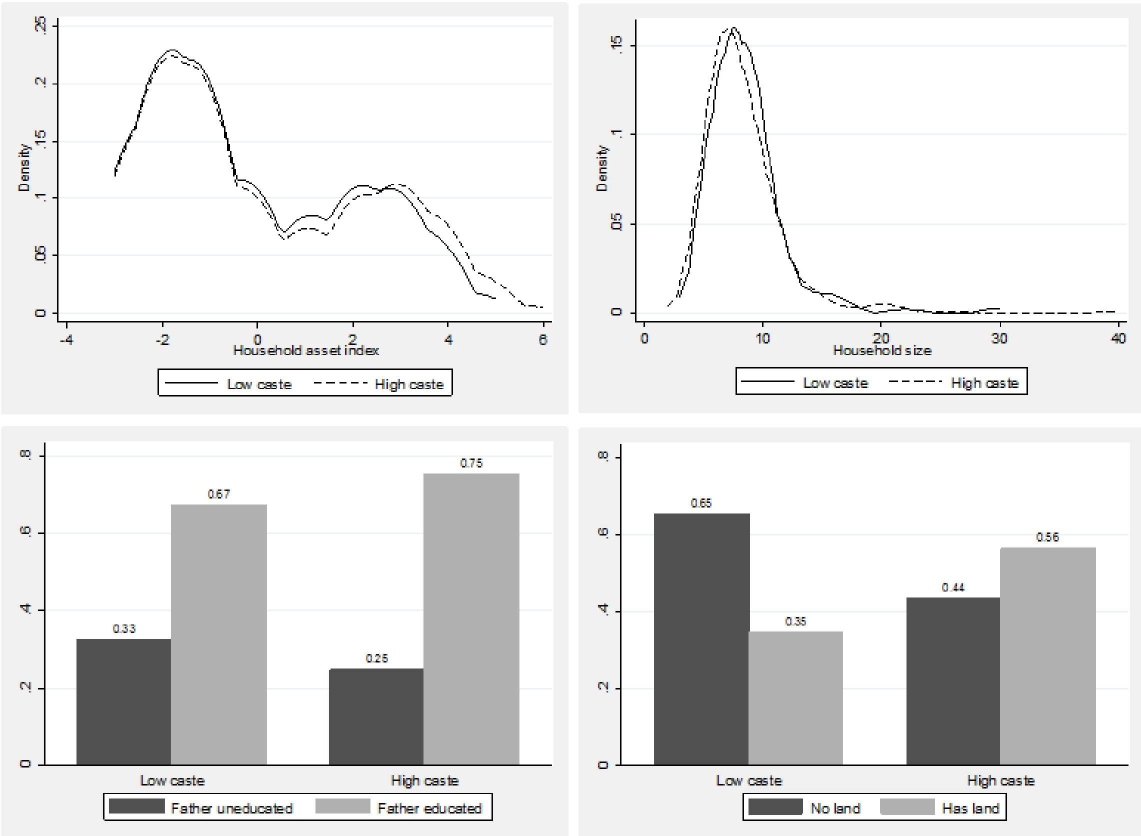

Figure 2: Returns to Education

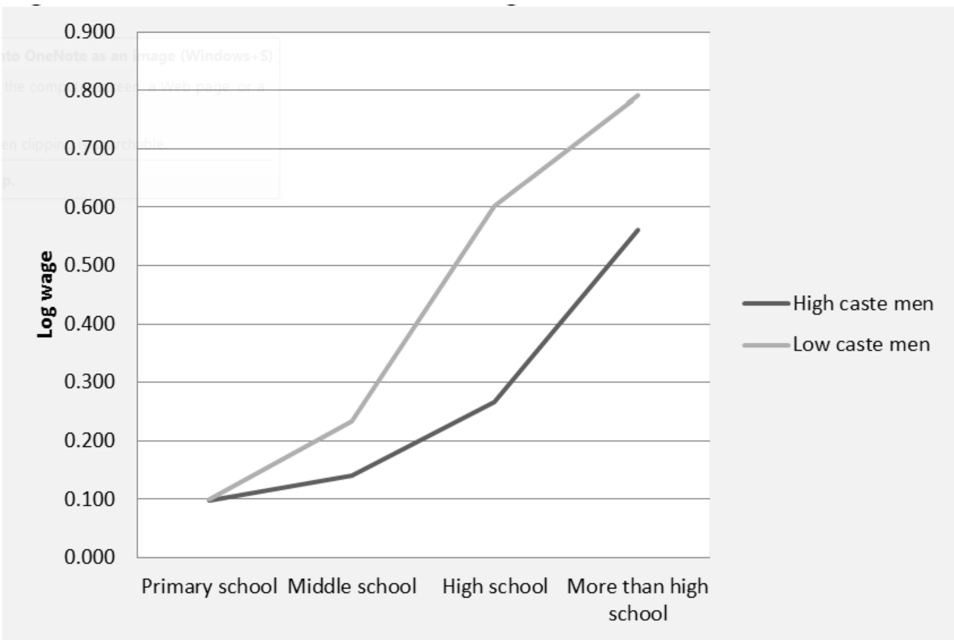

Table 1: Descriptive Statistics

	Panel A: Children							
	High caste				**Low caste**			
Observations	**Boys**		**Girls**		**Boys**		**Girls**	
Number of children*time	1,063		937		346		236	
Number of children	477		425		157		107	
Characteristics	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Average age	10.16	1.68	10.30**	1.65	10.06	1.62	10.51**	1.66
Grade	3.91	0.82	3.89	0.82	3.92	0.79	3.83	0.81
Enrolled in private school	0.26	0.44	0.28	0.45	0.22	0.42	0.20	0.40
Held back	0.03	0.17	0.03	0.17	0.02	0.13	0.03	0.18
Households	Mean		Sd		Mean		Sd	
Household size	7.87		3.01		8.22**		2.98	
Owns land	0.56		0.50		0.36***		0.48	
Asset index	-0.09		2.29		-0.24		2.21	
Mother uneducated	0.41		0.49		0.47***		0.50	
Father uneducated	0.25		0.43		0.34***		0.47	
	Panel B: Teachers							
	High Caste				**Low Caste**			
	Mean		Sd		Mean		Sd	
Number of teachers*time	2,313				269			
Number of teachers	733				94			
Age	34.537		9.487		35.361		7.851	
Female	0.555		0.497		0.387***		0.488	
From village	0.389		0.488		0.353		0.479	
Teaching experience	6.724		18.252		3.695***		9.735	
Education: matric	0.402		0.490		0.238***		0.427	
Education: FA	0.279		0.449		0.271		0.445	
Education: BA	0.251		0.434		0.431***		0.496	
Education: MA	0.068		0.252		0.059		0.237	
Teacher training	0.793		0.405		0.877***		0.329	
Private school	0.257		0.437		0.253		0.435	
Test scores	2.557		0.836		2.602		0.734	
Days absent (past month)	1.975		2.540		2.086		3.223	
	Panel C: Schools							
	High Caste children				**Low Caste children**			
	Mean		Sd		Mean		Sd	
Number of schools	342				91			
Number of male teachers	2.76		4.24		3.40		4.60	
Number of female teachers	3.97		4.07		3.30		3.93	
Number of students	195.15		149.27		222.95		177.39	
Number of caste groups	4.84		2.11		5.10		2.31	
High caste dominant	0.854		0.354		0.736***		0.443	
Caste fractionalization index	0.51		0.232		0.484		0.264	
School facilities index	0.434		1.57		0.449		1.45	

Notes: Asterisks denote significant differences between boys and girls, between high and low caste households, and between high and low caste teachers. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Matches and Switches

Panel A: Matches						
			Teacher			
			High caste		Low caste	
			Male	Female	Male	Female
	High Caste	Male	765	216	79	8
Child		Female	27	823	21	61
	Low Caste	Male	227	54	62	4
		Female	11	190	3	31
Panel B: Switches						
			Number of switches			
	High Caste	Male	43			
Child		Female	26			
	Low Caste	Male	27			
		Female	8			
	Total		104			

Table 3: Switching between high and low caste teachers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All children		**High caste male**		**High caste female**		**Low caste male**		**Low caste female**	
Lagged test scores	-0.010	-0.009	-0.038	-0.018	0.003	-0.000	-0.026	-0.021	-0.029*	0.003
	(0.020)	(0.009)	(0.039)	(0.016)	(0.028)	(0.014)	(0.047)	(0.024)	(0.017)	(0.010)
Age	-0.012	-0.000	-0.030	-0.012	0.016	0.014*	-0.063**	-0.001	0.049***	0.002
	(0.012)	(0.005)	(0.023)	(0.008)	(0.017)	(0.008)	(0.028)	(0.013)	(0.017)	(0.004)
Female	.	0.024	.	.	.	.	.	.	.	.
	.	(0.019)	.	.	.	.	.	.	.	.
Child high caste	.	-0.003	.	.	.	.	.	.	.	.
	.	(0.012)	.	.	.	.	.	.	.	.
Grade 3	-0.138	.	-0.413	.	.	.	.	.	.	.
	(0.180)	.	(0.325)	.	.	.	.	.	.	.
Grade 4	-0.097	0.046	-0.062	0.136*	0.160	-0.022	16.085*	-0.038	-0.051	0.019
	(0.117)	(0.041)	(0.171)	(0.070)	(0.343)	(0.075)	(9.613)	(0.225)	(0.069)	(0.059)
Grade 5	-0.077	0.062	0.027	0.240**	0.172	-0.055	.	.	-0.035	.
	(0.083)	(0.057)	(0.150)	(0.112)	(0.475)	(0.078)	.	.	(0.072)	.
Grade 6	-0.036	0.083	0.007	0.296**	0.285	0.010	.	.	1.553***	.
	(0.089)	(0.070)	(0.157)	(0.129)	(0.470)	(0.077)	.	.	(0.290)	.
HH Asset Index	-0.005	-0.006	-0.001	-0.003	0.014	0.013	-0.098***	-0.047*	-0.016	0.001
	(0.010)	(0.007)	(0.011)	(0.009)	(0.016)	(0.010)	(0.030)	(0.028)	(0.011)	(0.010)
Held back	0.043	0.048	0.185**	0.108*	0.085	-0.009	-16.126*	-0.022	.	0.007
	(0.061)	(0.034)	(0.081)	(0.065)	(0.209)	(0.052)	(9.553)	(0.318)	.	(0.115)
HH size	0.001	0.002	-0.000	0.003	-0.001	-0.001	-0.037	0.002	-0.009	-0.003
	(0.005)	(0.001)	(0.008)	(0.002)	(0.004)	(0.001)	(0.032)	(0.004)	(0.012)	(0.003)
Father undeducated	-0.032	-0.014	-0.005	-0.011	0.039	-0.009	-0.225*	-0.138**	0.015	-0.020
	(0.042)	(0.012)	(0.050)	(0.015)	(0.041)	(0.015)	(0.124)	(0.068)	(0.050)	(0.015)
Mother uneducated	0.017	0.003	0.042	0.011	0.040	0.004	-0.034	0.074	-0.020	0.028
	(0.035)	(0.013)	(0.036)	(0.018)	(0.060)	(0.018)	(0.136)	(0.069)	(0.036)	(0.019)
Dad educ missing	-0.030	-0.004	0.013	0.006	0.067	0.008	-0.397**	-0.009	-0.025	-0.014
	(0.050)	(0.012)	(0.068)	(0.022)	(0.069)	(0.014)	(0.156)	(0.056)	(0.059)	(0.020)
Teacher age	0.010***	0.010***	0.013***	0.010**	0.014***	0.012**	-0.030***	-0.002	0.011***	0.021***
	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.009)	(0.008)	(0.004)	(0.005)
Teacher female	-0.071	-0.144	-0.004	-0.125	-0.082	-0.110	-0.564	-1.059***	.	.
	(0.079)	(0.115)	(0.103)	(0.130)	(0.094)	(0.117)	(0.478)	(0.287)	.	.

Continued on next page

Table 3 – *Continued from previous page*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All children		**High caste male**		**High caste female**		**Low caste male**		**Low caste female**	
Teacher from village	0.016	0.003	0.010	0.030	0.145**	0.020	0.104**	-0.052	0.095*	0.093*
	(0.031)	(0.032)	(0.046)	(0.067)	(0.065)	(0.039)	(0.052)	(0.060)	(0.048)	(0.048)
Teacher experience	-0.000	-0.001	-0.000	0.001	-0.001	-0.000	0.180*	-0.004	0.158***	0.000
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.107)	(0.003)	(0.044)	(0.001)
Teacher educ FA	-0.035	-0.013	-0.050	-0.055	0.044	0.028	-0.254***	-0.105	0.425***	0.202**
	(0.043)	(0.047)	(0.056)	(0.066)	(0.069)	(0.066)	(0.091)	(0.111)	(0.070)	(0.083)
Teacher educ BA	0.053	0.003	0.061	-0.041	0.171*	0.151	-0.205*	-0.225	0.296***	0.155
	(0.063)	(0.088)	(0.102)	(0.104)	(0.101)	(0.126)	(0.117)	(0.176)	(0.055)	(0.094)
Teacher educ MA	0.021	0.135	0.072	0.180	0.216	0.183	-0.413**	-0.014	0.051	0.042
	(0.072)	(0.094)	(0.136)	(0.138)	(0.137)	(0.154)	(0.162)	(0.163)	(0.065)	(0.092)
Teacher test scores	-0.018	-0.016	0.015	0.029	-0.085**	-0.075*	0.113	0.049	-0.049**	-0.055**
	(0.020)	(0.026)	(0.033)	(0.036)	(0.038)	(0.041)	(0.068)	(0.066)	(0.022)	(0.027)
Private school	0.093*	-0.027	0.045	-0.033	0.080	-0.058	0.228	-0.101	-0.079*	-0.027
	(0.048)	(0.021)	(0.054)	(0.063)	(0.095)	(0.043)	(0.930)	(0.170)	(0.044)	(0.065)
School size	-0.001**	-0.000	0.000	-0.000	-0.001***	-0.000	-0.001	0.002	0.002***	0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
School facilities index	0.023*	0.023	0.025	0.028	0.045**	0.049*	-0.038	-0.038	0.001	0.016
	(0.012)	(0.019)	(0.016)	(0.022)	(0.022)	(0.027)	(0.034)	(0.052)	(0.014)	(0.023)
ELF index	0.181**	0.159	0.132	0.151	0.058	0.071	0.679**	0.362	0.044	0.010
	(0.091)	(0.148)	(0.143)	(0.221)	(0.145)	(0.168)	(0.341)	(0.462)	(0.114)	(0.128)
ELF missing	0.043	0.160	0.059	0.276	0.018	0.008	0.344	0.189	.	.
	(0.068)	(0.120)	(0.117)	(0.191)	(0.140)	(0.154)	(0.251)	(0.252)	.	.
Teacher tests missing	-0.048	-0.019	0.069	0.102	-0.235	-0.207	0.267	0.257	-1.195***	-1.158***
	(0.066)	(0.087)	(0.091)	(0.120)	(0.143)	(0.137)	(0.257)	(0.269)	(0.143)	(0.220)
Observations	1457	1457	612	612	508	508	205	205	132	132
R^2	0.144	0.118	0.221	0.210	0.287	0.214	0.461	0.356	0.929	0.782
Fixed effect	Child	School	Child	School	Child	School	Child	School	Child	School

Notes: $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome is switching between a high and low caste teacher.
All specifications include time dummies, and standard errors clustered at the level of the fixed effect.

1.0

Table 4: Main Results

Outcome: average of math, English and Urdu test scores	(1)	(2)	(3)	(4)
Panel A: Child Caste				
	Basic	Controls	Basic	Controls
Child high caste	0.020	-0.037		
	(0.058)	(0.054)		
Child high caste, female			-0.038	-0.077
			(0.108)	(0.104)
Child high caste, male			-0.079	-0.140*
			(0.081)	(0.074)
Child low caste, female			-0.202*	-0.192*
			(0.117)	(0.110)
Observations	2,582	2,582	2,582	2,582
Number of children	1,166	1,166	1,166	1,166
R^2	0.000	0.072	0.004	0.076
Panel B: Teacher Caste (low caste children)				
	Basic	Controls	Robust 1	Robust 2
Teacher high caste	0.213**	0.214**	0.476*	0.219**
	(0.091)	(0.086)	(0.281)	(0.087)
Observations	582	582	290	573
Number of children	264	264	136	259
R^2	0.442	0.488	0.511	0.487
Panel C: Dyadic Caste Match (low caste children)				
	Basic	Controls		
Teacher and child belong to same caste group	-0.260**	-0.211**		
	(0.102)	(0.103)		
Observations	582	582		
Number of children	264	264		
R^2	0.442	0.486		
Panel D: Child Caste and Gender and Teacher caste Interaction (low caste children)				
	Basic	+ Child	+ Teacher	+ School
Child: female. Teacher: high caste	-0.100	-0.106	-0.198	-0.183
	(0.173)	(0.167)	(0.169)	(0.168)
Child: male. Teacher: high caste	0.314***	0.342***	0.344***	0.341***
	(0.096)	(0.100)	(0.085)	(0.089)
Observations	582	582	582	582
Number of children	264	264	264	264
R^2	0.447	0.471	0.492	0.495

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Panel A: Regressions include teacher*time (classroom) fixed effects. Standard errors clustered at the teacher*time (classroom) level. Panel B: Regressions include child fixed effects. Standard errors clustered at the child level. Robust 1 refers to a restriction on the sample to schools that passed the Monte Carlo χ^2 test, and Robust 2 refers to a a restriction on the sample to schools that passed the Monte Carlo confidence interval test. Panels C and D: Regressions include child fixed effects. Standard errors clustered at the child level. Child controls: age, female (only in Panel A column (2)), grade dummies, household asset index, dummy for non-promotion, mother's years of education, father's years of education, dummy variables for mother's and father's education missing. Teacher controls (except Panel A): age, female, from the village, years of teaching experience, education level dummies (FA, BA, MA), average of math, English, and Urdu test scores, dummy variable for test scores missing. School controls: private school dummy, school size, school facilities index, caste fractionalization index. Panels A, B, and C: child, teacher, and school controls included in all regressions with controls.

Table 5: Mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Outcome	Average of math, English, Urdu test scores	Diff: actual test score rank and teacher rank One year	Cumulative	Teaching Career	Skilled Job	Grade Aspiration		
High caste dominant school	0.068							
	(0.091)							
C: low, male. T: high		-0.168	-0.030	-0.459*	0.415*	1.507		
		(0.233)	(0.183)	(0.260)	(0.238)	(1.648)		
C: low, female. T: high		0.276	0.370	-0.198	0.226	-0.808		
		(0.940)	(0.691)	(0.282)	(0.263)	(1.709)		
C: high, male. T: high				-0.426	0.367	1.713		
				(0.263)	(0.236)	(1.661)		
C: high, male. T: low				-0.323	0.474**	1.128		
				(0.221)	(0.198)	(1.722)		
C: high, female. T: high				0.051	0.151	0.103		
				(0.276)	(0.256)	(1.727)		
C: high, female. T: low				0.086	0.058	-0.058		
				(0.305)	(0.290)	(1.847)		
C: low, female. T: low				-0.118	0.098	0.536		
				(0.406)	(0.371)	(3.049)		
Observations	582	260	260	726	726	726		
Number of children	264	135	135	323	323	323		
R^2	0.483	0.138	0.194	0.193	0.136	0.133		
Fixed effect	School	Child	Child	School	School	School		
Panel B: Outcome	House/paid work	Homework	Leisure	Learning	Parents met teacher	Parents helped w/ hwork	Hrs helping w/ hwork	Education Expenditure
C: low, female. T: high	-27.479	36.846	-17.785	-19.923	-0.145	0.350**	0.246	-27.359
	(32.983)	(25.226)	(51.604)	(26.606)	(0.231)	(0.161)	(0.697)	(38.142)
C: low, male. T: high	-21.988	3.995	30.404	-4.517	-0.013	-0.119	0.971***	-32.283
	(14.453)	(15.018)	(35.587)	(18.189)	(0.145)	(0.109)	(0.355)	(34.832)
Observations	567	567	567	567	348	575	575	574
Number of children	263	263	263	263	175	261	261	261
R^2	0.218	0.140	0.117	0.150	0.195	0.110	0.714	0.308

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. C denotes child, T denotes teacher. High denotes high caste, low denotes low caste. Panel A: All regressions include child, teacher and school controls, and time dummies. Child controls: age, grade dummies, household asset index, dummy for non-promotion, mother's years of education, father's years of education, dummy variables for mother's and father's education missing. Teacher controls: age, female, from the village, years of teaching experience, education level dummies (FA, BA, MA), average of math, English, and Urdu test scores, dummy variable for test scores missing. School controls: private school dummy, school size, school facilities index, caste fractionalization index. Columns (4) and (5) also include the child's average test score. Standard errors clustered at the level of the fixed effect. High caste dominant school is a school in which more than 50% of the students belong to high caste groups. Skilled jobs include: doctor, government job, politician, engineer, or private sector job. Panel B: All regressions include child, teacher and school controls, time dummies, and child fixed effects. Standard errors clustered at the child level. Outcomes in columns (1)-(4) are measured in minutes per day, column (7) is measured in hours per week.

A Appendix: Additional Tables

Table 6: Teachers matching to schools

Outcome: High caste teacher				
	(1)		(2)	
	OLS		Probit	
	coef	se	coef	se
Private school	-0.091	(0.06)	-0.830	(0.79)
Facilities index	0.008	(0.05)	-0.475	(0.49)
Number of male teachers	0.004	(0.01)	0.087	(0.07)
Number of female teachers	-0.002	(0.01)	0.048	(0.07)
School size	-0.000	(0.00)	-0.000	(0.00)
Number of public schools within 5 min	-0.038*	(0.02)	-0.505**	(0.21)
Number of Islamic schools within 5 min	-0.011	(0.03)	-0.037	(0.22)
Number of private schools within 5 min	0.000	(0.02)	0.017	(0.23)
Number of public schools within 5-15 min	0.006	(0.01)	-0.118	(0.21)
Number of Islamic schools within 5-15 min	-0.013	(0.01)	-0.521***	(0.18)
Number of private schools within 5-15 min	0.002	(0.00)	0.245	(0.18)
Number of public schools greater than 15 min	-0.001	(0.00)	-0.093	(0.07)
Number of Islamic schools greater than 15 min	0.001	(0.00)	0.084	(0.08)
Number of private schools greater than 15 min	0.001	(0.00)	0.001	(0.02)
Time from school to nearest phone	0.058	(0.04)	0.655*	(0.37)
Time from school to nearest bank	-0.008	(0.02)	-0.114	(0.29)
Time from school to nearest health facility	-0.013	(0.02)	-0.289	(0.24)
Time from school to nearest transport facility	-0.024	(0.02)	-0.360	(0.28)
Time from school to nearest government office	-0.001	(0.00)	-0.081	(0.20)
School has a library	0.002	(0.05)	-0.559	(0.52)
School has a computer	-0.082	(0.09)	-1.186	(0.86)
School has a place for sports	-0.011	(0.05)	-0.822	(0.60)
School has a hall	-0.008	(0.05)	-1.123*	(0.62)
School has a wall surrounding it	0.002	(0.06)	-0.528	(0.59)
School has fans	0.001	(0.06)	-0.654	(0.74)
School has electricity	-0.018	(0.07)	-1.246	(0.80)
High caste dominant school	0.098*	(0.06)	0.994**	(0.48)
Contracts for teachers	-0.050	(0.05)	-0.683	(0.59)
School awards teachers bonuses	0.083**	(0.04)	0.781*	(0.43)
School has an SMC	0.057	(0.05)	0.761	(0.57)
Language of instruction - Urdu	0.133	(0.12)	0.466	(1.24)
Language of instruction - Punjabi	0.124	(0.11)	1.067	(0.89)
Language of instruction - Pashto	0.041	(0.12)	0.071	(0.96)
Language of instruction - Sindhi	0.102	(0.11)	0.707	(0.97)
Language of instruction - Seraiki	0.090	(0.12)	0.309	(1.08)
Building personally owned	0.210*	(0.12)	-0.831	(1.25)
Building rented	0.238*	(0.14)	-0.081	(1.32)

Continued on next page

Table 6 – *Continued from previous page*

Outcome: High caste teacher				
	(1)		(2)	
	OLS		Probit	
	coef	se	coef	se
Building owned by government	0.029	(0.11)	-2.697**	(1.13)
Building donated	0.149	(0.11)	.	.
Toilet - none	0.003	(0.05)	-0.602	(0.52)
Toilet - latrine	0.021	(0.05)	-0.205	(0.49)
Toilet - flush	-0.095	(0.11)	-1.017	(0.73)
Toilet - Tank	-0.015	(0.06)	-0.832	(0.62)
Water - well	0.021	(0.05)	-0.063	(0.56)
Water - pump	0.067	(0.04)	0.549	(0.56)
Water - official source	0.037	(0.04)	-0.065	(0.42)
Sitting arrangement - floor	-0.126	(0.12)	-0.748	(0.70)
Sitting arrangement - mats	-0.097	(0.11)	-0.613	(0.57)
Sitting arrangement - desks	-0.108	(0.09)	-0.564	(0.36)
Sitting arrangement - mix	-0.062	(0.10)	.	.
Observations	1624		610	
R^2	0.110			

Notes: $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Regressions include time dummies and village fixed effects.

Standard errors clustered at the village level.

1.0

Table 7: Dropout

	(1)		(2)	
	coef	se	coef	se
Child high caste	0.000	(0.000)	0.001	(0.001)
Lagged average test scores	-0.000	(0.000)	-0.000	(0.000)
High caste * lagged average test scores	0.001	(0.001)	0.001	(0.001)
Switched caste-type of teacher last year	-0.001	(0.001)		
High caste * switched caste-type of teacher last year	0.000	(0.001)		
Switched caste-type of teacher this year (or would have)	0.001	(0.001)		
High caste * switched (or would have) caste-type teacher	-0.001	(0.001)		
High caste: switched from low to high caste teacher this year (or would have)			0.001	(0.001)
Low caste: switched from low to high caste teacher this year (or would have)			0.001	(0.001)
High caste: switched from high to low caste teacher this year (or would have)			-0.000	(0.001)
Low caste: switched from high to low caste teacher this year (or would have)			0.002	(0.002)
High caste: switched from low to high caste teacher last year			-0.002	(0.002)
Low caste: switched from low to high caste teacher last year			-0.001	(0.001)
High caste: switched from high to low caste teacher last year			-0.002	(0.002)
Low caste: switched from high to low caste teacher last year			-0.007	(0.007)
Age	0.000	(0.000)	0.000	(0.000)
Female	-0.001	(0.001)	-0.001	(0.001)
Grade 2	-0.003	(0.003)		0.000
Grade 3	0.004	(0.004)	0.008	(0.008)
Grade 4	-0.000	(0.000)	-0.001	(0.001)
Grade 5	-0.000	(0.000)	0.001	(0.002)
Grade 6	-0.000	(0.000)	-0.002	(0.002)
Grade 7	-0.000	(0.000)	-0.004	(0.004)
Household asset index	0.001	(0.001)	0.001	(0.001)
Not promoted	-0.004	(0.004)	-0.008	(0.008)
Household size	-0.000	(0.000)	-0.000	(0.000)
Father uneducated	0.000	(0.000)	0.001	(0.001)
Mother uneducated	-0.000	(0.000)	-0.000	(0.000)
Father education missing	0.002	(0.002)	0.003	(0.003)
Constant	-0.000	(0.000)	-0.001	(0.002)
Number of observations	2,237		1,350	
R^2	0.050		0.073	

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome: child dropped out of school. Both regressions include village fixed effects, year dummy variables, and standard errors clustered at the village level.

Table 8: Wealth and parental education

Outcome: average of math, English and Urdu test scores			
(1)		(2)	
Wealth	coef/se	Education	coef/se
(1) Child: not poor, female.	-0.242	Child: educated parent, female.	-0.133
Teacher: high caste	(0.206)	Teacher: high caste	(0.199)
(2) Child: not poor, male.	0.263**	Child: educated parent, male.	0.310***
Teacher: high caste	(0.102)	Teacher: high caste	(0.095)
(3) Child: poor, female.	-0.105	Child: uneducated parents, female.	-0.182
Teacher: high caste	(0.190)	Teacher: high caste	(0.181)
(4) Child: poor, male.	0.397***	Child: uneducated parents, male.	0.364***
Teacher: high caste	(0.097)	Teacher: high caste	(0.112)
Number of observations	582	Number of observations	582
Number of children	264	Number of children	264
R^2	0.486	R^2	0.479
P-value difference coefficient (2) and (4)	0.164	P-value difference coefficient (2) and (4)	0.618

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include child, teacher and school controls, time dummies, child fixed effects. Child controls: age, grade dummies, household asset index, dummy for non-promotion, mother's years of education, father's years of education, dummy variables for mother's and father's education missing. Teacher controls: age, female, from the village, years of teaching experience, education level dummies (FA, BA, MA), average of math, English, and Urdu test scores, dummy variable for test scores missing. School controls: private school dummy, school size, school facilities index, caste fractionalization index. Standard errors clustered at the child level.

Table 9: Returns to Education

	Outcome: log of the monthly wage			
	(1)		(2)	
	High caste men		Low caste men	
	coef	se	coef	se
Primary school	0.099	0.072	0.099	0.136
Middle school	0.141*	0.084	0.234	0.195
High school	0.266**	0.115	0.602**	0.302
More than high school	0.562***	0.151	0.791**	0.353
Constant	7.627***	0.291	7.448***	0.47
Number of observations	1,020		314	
R^2	0.27		0.34	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Omitted category: less than primary school. Regressions include individual and household controls, as well as time dummies and village fixed effects. Controls: age, unmarried dummy, subjective overall health (scale of 1-16), can read, can write, can count, can add, number of years to achieve highest grade, household size, household has audio, print, visual, internet media, number of deaths in household in past year, whether a member left the household in past year, household has member abroad, remittances, whether household owns land, whether household owns the house, total expenditure, asset index, whether harvest worse than previous year, whether earnings lower than previous year. Standard errors clustered at the village level.

Table 10: Child and teacher caste, all children

	(1) Basic	(2) Controls
C: high, female. T: high	0.013	-0.007
	(0.098)	(0.098)
C: high, male. T: high	-0.089	-0.124
	(0.148)	(0.151)
C: low, female. T: high	-0.125	-0.155
	(0.168)	(0.164)
C: low, male. T: high	0.274***	0.253***
	(0.089)	(0.081)
Observations	2,582	2,582
Number of children	1,166	1,166
R^2	0.409	0.425

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All regressions include child fixed effects and standard errors clustered at the child level. Child controls: age, grade dummies, household asset index, dummy for non-promotion, mother's years of education, father's years of education, dummy variables for mother's and father's education missing. Teacher controls: age, female, from the village, years of teaching experience, education level dummies (FA, BA, MA), average of math, English, and Urdu test scores, dummy variable for test scores missing. School controls: private school dummy, school size, school facilities index, caste fractionalization index. C denotes child, T denotes teacher.