ISHA GUPTA

University of Padova

THE ROLE OF SOCIAL IDENTITY AND PERCEIVED DISCRIMINATION IN HUMAN CAPITAL FORMATION: EVIDENCE FROM INDIA

October 2022

Marco Fanno Working Papers – 290

Department of Economics and Management "Marco Fanno"
University of Padova
Via del Santo 33, 35123 Padova

The Role of Social Identity and Perceived Discrimination in Human Capital Formation: Evidence from India*

ISHA GUPTA†

Abstract

This study examines the role of historically defined social identity in human capital development over time by focusing on a region from India where social identities are defined along the lines of castes and religious groups. It investigates the evolution of gaps across social groups in children's cognitive outcomes and parental investment in children's education from ages 5 to 15. Significant gaps in test scores and parental investment are found between children from lower and upper Hindu castes. These gaps, which originate early in childhood and persist throughout the 10 years of the study period, cannot be completely explained by the differences in socioeconomic status across social groups. Moreover, the perception of social discrimination is also found to be a significant contributor to the gaps in cognitive outcomes and parental investment across social groups. While parents' perceived social discrimination is associated with lower parental investment throughout childhood, it is negatively associated with children's cognitive outcomes only at later ages.

Keywords: *child development, parental investment, human capital formation, social identity, caste, perceived social discrimination, India.*

JEL classification: *I24, J15, J24*

* I would like to thank my advisors, Guglielmo Weber, Marco Bertoni and Daniele Paserman, for their priceless guidance and support. My gratitude also goes to Kevin Lang, Giorgio Brunello, Lorenzo Rocco, Roberto Bonfatti, Catherine Porter, Alessandro Tarozzi, Alessandro Bucciol, Danilo Cavapozzi for their helpful comments and suggestions. I am grateful to participants of seminars at the University of Padova; Development Reading Group at Boston University; Society of Household Economics (SEHO) Boston 2021; Warsaw International Economic Meeting 2021; PhD summer school on applied micro-econometrics 2021, Cyprus, Lancaster, and Padova; Hans-Böckler 2021 conference on inequality; IIM Bangalore International Conference on Public Policy 2021; Italian Econometric Association (SIdE-IEA) Workshop in Econometrics and Empirical Economics 2021 for helpful and healthy discussions. Last but not the least, I am thankful to my friend Dr. Daphné Reguiessé who reviewed the initial draft of my paper and provided her valuable insights. All remaining errors are my own. The data used in this paper comes from the Young Lives survey and are publicly available on website https://www.younglives-india.org/index.php/. Young Lives is funded by UK aid from the Department for International Development (DFID). The views expressed here are solely mine. They are not necessarily those of Young Lives, the University of Oxford, DFID or other funders. I declare that there are no relevant or material financial interests that relate to the research described in this paper.

† Department of Economics & Management "Marco Fanno", University of Padova, Italy. Email- isha.gupta@studenti.unipd.it

1. Introduction

Ethnic fragmentation and traditional hierarchies based on social markers have been characteristic of several societies: race in the United States, castes in India, and ethnicity in Nigeria. These social stratifications are critical in shaping the social identities that influence individuals' behaviors and decision-making. In this paper, by focusing on a particular setting in India where social identities are defined along the lines of castes and religions, I study social identities' dynamic role in human capital development over time and how the perception of social discrimination contributes to gaps in human capital across social groups. This research question is particularly relevant since the notion of confronting economic inequality in societies has been gaining significant attention recently[1].

In India, caste is a historically defined social identity inherited at birth. By defining individuals' access to resources and opportunities at every stage of their economic and social lives, it plays a significant role in shaping economic mobility. There exists wide heterogeneity across castes in terms of socioeconomic status, access to education, and access to the labor market. Lower castes, in general, have been subjected to deep-rooted prejudices and social stigmatization. For example, *Dalits*—considered to be at the bottom of the Hindu caste hierarchy—were regarded as 'untouchables' and are often subjected to discriminatory and unequal treatment.[2]

Growing up, lower caste children are exposed to numerous disadvantaged environments such as lower income levels, access to lower quality health facilities, among others. It is well established in the literature that early life experiences and living conditions may have lasting consequences on children's overall development, later life outcomes, and even inter-generational transmission of human capital (Attanasio et al., 2015; Cunha & Heckman, 2007; Rubio-Codina et al., 2016). A growing body of empirical evidence suggests that compared to upper caste children regarding physical and human capital outcomes—health, educational attainment, and cognitive outcomes—lower caste children are disadvantaged, with most of the studies identifying children's economic disadvantages as the major reason for their poor development and later life outcomes (Banerjee & Knight, 1985; Borooah, 2012; Deshpande, 2001; Munshi, 2019).

A question then arises: at what age does the gaps in human capital start to emerge across castes, and how do these gaps evolve as children age? This question is yet to be answered, and answering it would bring to light crucial insights into the right age that public policies aiming to bridge caste-based differences should target. This paper attempts to bridge this this gap in the literature. To this end, it analyzes the dynamics of castes' role in human capital formation during mid-childhood and early adolescence—between ages 5 and 15—and focuses on human capital accumulation from two different perspectives: children's cognitive outcomes and parental investment in children's human capital. A wide range of indicators of children's human capital stock and parental investment over time, including performance in various tests and material investment in education, are examined.

Furthermore, an understanding of the factors that contribute to the unexplained differences in human capital across social groups is critical to moderate these caste-based inequalities. This paper focuses on one such factor: perceived social discrimination. First, I analyze the association between parents' perception of social discrimination and their investment decisions in children's education. Further, I examine how children's cognitive outcomes are associated with parents' perception of social discrimination over time. Parental perceptions of discrimination are likely to shape their children's understanding and perceptions of discrimination. For example, if parents concur that a

[1] The rallying cry of the 2030 Agenda for Sustainable Development is 'Leave no one behind' (United Nations Committee for Development Policy, 2018).

[2] Dalits, also called scheduled castes, are the lowest caste in the Indian caste system hierarchy and were previously characterized as 'untouchables' and considered outcasts. Section 2 describes the Indian caste system in detail.

store manager is being discriminatory toward them, then their children are likelier to make an attribution to discrimination than when the parents disagree with this perception.[3] Second, I examine whether a differential association exists between perceived social discrimination and human capital development across social groups. Such an association would offer suggestive evidence if perceived discrimination can partly explain the gaps in human capital observed across social groups. If parents perceive that they are discriminated against, then this, I hypothesize, lowers their expected returns to education, thus lowering their investment in their children's education—a channel that is expected to be more prominent for the backward castes since the perception of social discrimination could aggravate parents' existing beliefs about discriminatory practices in higher education and the labor market, thus disincentivizing them from investing in their children's education. Owing to the richness of my data, I use the self-reported measure of perceived social discrimination to assess the contribution of perceived discrimination to the observed gaps in human capital between social groups.[4]

Several studies in the US have analyzed the racial gaps in the cognitive outcomes of children. For instance, Fryer and Levitt (2004, 2005) analyzed the racial gaps in math and reading test scores during early school years, finding that the Black–White test score gap among incoming kindergartners disappeared after controlling for a small number of socioeconomic covariates. However, the Blacks, as they grow older, lose substantial ground compared to other races. Many studies in the US have also extensively investigated the age at which children start perceiving racial discrimination and its effect on their cognitive outcomes (see Brown & Bigler, 2005; Simons et al., 2002; Theimer et al., 2001). This study is the first to explore these patterns across Indian social groups.

I utilize the Young Lives (YL) survey data, which is particularly well suited for this study.[5] This longitudinal survey followed 2011 children for five rounds—at the ages of 1, 5, 8, 12, and 15—between 2002 and 2016. Its data contains various pieces of information key for this analysis: children's scores in the Peabody Picture Vocabulary Test (PPVT)[6], math, English, and reading tests conducted at different ages; the various measures of parental investment, such as expenditure on education-related activities and the type of school attended; a rich set of household and individual characteristics; and measures of perceived social discrimination.

Test scores, recent literature has affirmed, are measured on ordinal scales.[7] As with utility functions, any monotonic transformation of the test score scale is also potentially valid. Bond and Lang (2013) conducted a bounding exercise for Fryer and Levitt (2005), demonstrating that order-preserving scale transformations of test scores can offer contradictory conclusions about the growth in gaps between Blacks and Whites, starting from increasing gaps over time to decreasing gaps over time; some transformations have also suggested that Blacks outperform Whites over time. This study

[3] Understanding the relationship between perceptions of discrimination and children's development has been identified as a key priority in both development and education economics literature (see Brown, 2015; Brown & Bigler, 2005; Cheng et al., 2015; Stone & Han, 2005). Perceiving oneself to be the target of discrimination is likely to affect individuals' identity formation, peer relations, academic competencies, occupational goals, and mental and physical wellbeing (Brown & Bigler, 2005; Stone & Han, 2005).

[4] Unlike others in the literature, I do not estimate the contribution of discrimination to the observed gaps in human capital between social groups using the Blinder-Oaxaca decomposition (for instance, see Arouri et al., 2019; Banerjee & Knight, 1985; Borooah, 2012). This approach has been criticized for understating or overstating the effect of discrimination (see Madden, 2000; Munshi, 2019; Ospino et al., 2010). Using self-reported perceived discrimination also has some potential concerns, such as social-desirability bias, which I discuss in detail in section 4.3.

[5] Young Lives survey has detailed information on children's cognitive outcomes, parental investment in children's education, measured at various ages, socioeconomic background and measure of perceived social discrimination.

[6] The Peabody Picture Vocabulary Test (PPVT), a widely used test to measure receptive vocabulary, is administered individually, orally, untimed, and norm-referenced, where the test-taker selects the picture that best represents the meaning of a stimulus word orally presented by the examiner.

[7] See Cunha & Heckman (2008); Bond & Lang (2013)

deals with the issue of test score ordinality by employing the percentile rankings of test scores that are scale-invariant.[8]

The findings of this study suggest that substantial and persistent gaps exist in the test scores of lower and upper caste children in India throughout the 10 years of the study period. These gaps cannot be completely explained by the observed differences in socioeconomic status (SES) across castes. When compared to the racial test score gaps between Blacks and Whites in the US reported by Fryer and Levitt (2005), the caste-based test score gaps observed in this study are smaller in magnitude but persistent over time. Moreover, this study analyzes the gaps between Muslims and Hindus and finds that Muslims perform equally well in the PPVT as upper caste Hindus at the ages of 5 and 8. However, this gap becomes significant over time, and in math, English, and reading tests, Muslims perform worse than both upper and lower caste Hindus.[9]

This study also establishes that, compared to upper caste Hindus, lower caste Hindus and Muslims invest significantly less in the education of their children, even after controlling for socioeconomic background. These gaps in parental investments across social groups are consistent over time. These children are more likely to drop out of school early and are less likely to attend private and expensive schools.

Finally, this paper finds that perceived social discrimination plays a crucial role in human capital formation. It is negatively associated with parents' investment in the education of their children starting from a very early age, and the resulting gaps persist throughout the 10 years of the study period. On the contrary, the negative association between perceived discrimination and cognitive outcomes only appears as children age, a finding consistent with the hypothesis that children develop an awareness of discrimination only as they grow older (Brown & Bigler, 2005). The results on the differential association between perceived social discrimination and human capital across social groups suggest that the perception of discrimination disproportionately affects backward castes and, thus, is a significant contributor to the gaps in human capital observed across social groups.

From a policy perspective, the study findings have important implications. They emphasize that differences in cognitive outcomes observed across castes arise at an extremely early age and persist over time. Lower caste children are disadvantaged in terms of parental investments, are more likely to drop out of school early, and are less likely to attend private schools. To bridge these gaps, there is a need for public policies that promote quality investment in lower caste children at a very early age and discourage social discrimination. The Government of India, through its Right to Education (RTE) Act, offers free primary education to socially and economically disadvantaged children through public schools. However, as abundant evidence in the literature shows, these schools are poor in quality (Gouda et al., 2013; Muralidharan & Kremer, 2006; Muralidharan & Sundararaman, 2013). Moreover, although affirmative policies exist that offer representation to the backward communities in higher education and the labor market by reserving seats for them, according to the study findings, these policies may not help compensate for the gaps in human capital generated during the early and critical periods of development.

The rest of this paper is structured as follows. Section 2 offers a brief background of the Indian caste system. A brief discussion of the related literature in section 3 is followed by a discussion on the issue of test score ordinality in section 4. In section 5, the methodology, data, and descriptive

[8] Discussed in detail in section 4.2.

[9] For the analysis, I delineate the social groups by caste and religion. I construct four categories: lower caste Hindus (Hindus belonging to Scheduled Caste (SC), Scheduled Tribe (ST), and Other Backward Class (OBC); upper caste Hindus (Hindus belonging to forward/upper castes); Muslims; and other religions. This is described in detail in sections 2 and 4.

statistics are described. In section 6, the main results are presented, and, finally, in the concluding section, section 7, a few policy discussions are offered.

2. Background of the Indian Caste System

The Hindu caste system in India is a historical social stratification of people into various hierarchically ranked groups that were traditionally defined on the basis of professions. The existence of the caste system goes back to more than 2000 years ago. It comprises four hierarchical classes or varnas: *Brahmins* (priests and teachers), the *Kshatriyas* (rulers and soldiers), the *Vaishyas* (merchants and traders), and the *Shudras* (laborers and artisans). Certain population groups, known as *Dalits*, were historically excluded from the *varna* system and regarded as "untouchables." Each *varna*, based on their specific occupation, was further divided into hundreds of sub-castes called *jatis*. The contemporary manifestation of the caste system comprises 6,000 endogamous *jatis* (Coffey et al., 2019).

Even though people have shifted occupations over time, the Indian caste system has persisted and is inherited at birth. Throughout the life of an individual, their caste typically stays the same; there are some exceptions, such as women in inter-caste marriages taking on their husbands' caste, and their children also inheriting their fathers' caste. However, marriages in India are mostly endogamous. Thus, an implicit social status with limited mobility is attached to a person at birth.

For political purposes and to offer relief and support to backward social groups, in modern India, the Indian government introduced a new categorization scheme: the *Dalits* or the untouchable castes were clubbed together and categorized as scheduled castes (SCs); the socially and economically marginalized indigenous ethnic groups were categorized as scheduled tribes (STs); and another group of castes that were identified as socioeconomically disadvantaged was referred to as other backward castes (OBCs). The other higher caste groups were referred to as upper/general castes.[10]

While the SCs are exclusive to Hindus, Sikhs, and Buddhists, STs, and OBCs include some Muslim groups. ST Muslims are, however, relatively few in India, with most Muslims falling under OBC or the general category. According to the Mandal Commission, established in 1979 to identify India's socially or educationally backward classes, backward Muslims constituted 8.4 percent out of the total 11.2 percent of the Indian Muslim population (Mondal, 2003).

While these social groups are highly heterogeneous, on average, the lower castes—such as SCs, STs, and OBCs—have been disadvantaged in terms of income, education, and many other socioeconomic indicators. To address these disadvantages, untouchability was officially abolished in 1950,[11] and compensatory, affirmative actions in education and employment were introduced for the SCs and STs who have suffered cumulative, social, and economic disabilities (Sankaran et al., 2017a). More recently, in 1992, these benefits were extended to OBCs in response to their organized political assertions. However, even after various affirmative actions, significant gaps persist across these social groups. These gaps concern the level of schooling between the lower and upper castes, discriminatory

[10] Post-colonial India attempted to establish a democratic social order based on egalitarian ethos. Its constitutional policies aimed to offset the disabilities faced by its historically disadvantaged sections of population characterised as backward classes: the SCs, STs, and OBCs (Mondal, 2003).

[11] After several social movements emerged to abolish untouchability, laws were made in the constitution to accommodate the interests of the oppressed classes. Article 17 of the Indian Constitution abolished untouchability, declaring it as a punishable act—no one can now restrict the Dalits from entering temples, streets, or buses; they are free to use all public services with respect and dignity; and no one can refuse to sell anything to them. However, despite such constitutional amendments, untouchability and caste discrimination still prevail in India.

practices and preconceptions at school against the lower castes, pre-market discrimination against the lower castes, and labor market discrimination (Banerjee and Knight (1985), Munshi (2019), and Kijima (2006)).

3. Related Literature

As suggested by a growing body of literature globally, the economic differences observed among individuals and households may have ethnic and racial origins. A large stream of literature demonstrates racial differences in the cognitive ability of young children, the quality of schools attended, and school dropouts (Fryer & Levitt, 2004, 2005, 2013; Jones et al., 1984; Neal, 2006). Fryer and Levitt (2004, 2005) studied the racial gaps in math and reading tests in the US, finding that a substantial Black–White achievement gap exists at the entry to school, which can be completely explained by the differences in SES across races, and this gap increases with age and can no longer be explained by differences in SES. Arouri et al. (2019) found that, in developing countries like Vietnam, Peru, and Ethiopia, children from small ethnic groups have lower educational attainment and cognitive ability.

In India, the lower castes have been socially disadvantaged for centuries; they were originally assigned the lowest-status occupations, requiring little investment in human capital; and even when they managed to achieve occupational mobility over time, they typically ended up in low-skill industrial jobs (Munshi, 2019). These differences still exist in India, reinforcing inequality throughout generations.

Most studies in India have found significant differences in the cognitive outcomes of children across castes. Disparities in factors such as child health, parental education, and household socioeconomic status have been found to be important in explaining these gaps across castes in India. Even after controlling for the differences in initial endowments, such as socioeconomic endowments, other recent studies have found that the effects of these social institutions still persist (Borooah, 2012; Munshi, 2019). For example, Munshi (2019), reviewing the literature on caste and the Indian economy, reported that even after controlling for parental education, household wealth, measures of school quality, and teacher inputs, SC/ST and OBC children are significantly less likely to attend school.

Gangopadhyay and Sarkar (2014), even after controlling for all the available socioeconomic background variables, found that SC households invest significantly less than other households in the private coaching of their children. These differences, according to the researchers, are driven by the cultural paradigm: the lower castes, owing to historical reasons, may be inherently less motivated to invest in education.

Munshi and Rosenzweig (2006) analyzed the schooling and career choices of children across castes in India with the advent of liberalization in the 1990s. Not only was there a dramatic shift in the returns to different occupations but returns to learning English also increased, with greater access to white-collar jobs. While the lower caste children, the researchers found, continued attending local language schools, an increased number of upper caste children enrolled in English-medium schools.

The literature has also emphasized the contribution of caste-based discrimination, exclusion, and humiliation toward the underperformance of lower caste children (Borooah (2012); Rawal and Kingdon (2010)). Society's belief in the backwardness of certain communities, Singh and Husain (2016) argued, has resulted in discrimination against them in the labor market, lowering perceived returns to education for such communities. These communities, they posited, start behaving in a manner that justifies society's perceptions about them, reinforcing and perpetuating initial disparities

(Singh & Husain, 2016). Thorat and Attewell (2007) conducted a field experiment to document caste-based pre-market discrimination against lower castes in the labor market, finding that SC and Muslim candidates with identical educational qualifications and experience as upper caste candidates were significantly less likely to be called for interviews.

In addition to social identity-based discrimination, there is also evidence of discrimination due to parental economic status, which is also known as class discrimination (Schiller, 1971). Xiang et al. (2018) studied the influence of perceived discrimination on Chinese migrant children and found that children from lower-income families experienced greater discrimination than children from higher-income families. In India, caste-based discrimination overlaps with class-based discrimination. Lower social categories in India, especially SCs, STs, and Muslims, have been disproportionality poorer than others and, therefore, face relatively greater discrimination (Dhesi, 1998).

Perceived discrimination, studies from developed countries have also revealed, is strongly associated with physical, mental, and behavioral health outcomes, such as depression, anxiety, chronic stress, post-traumatic stress disorder, and low self-esteem (Cheng et al., 2015; Cooke et al., 2014; Stone & Han, 2005; Xiang et al., 2018). For example, as Brown and Bigler (2005) showed, children who experience discrimination from their teachers are more likely to have negative attitudes regarding school, lower academic motivation and performance, and are at increased risk of dropping out of high school.

To bridge the caste-based gaps and safeguard the economic and social interests of the lower castes—SC, ST, and OBC—and address disparities in wages, employment, education, and consumption, the Government of India has made several provisions: providing quotas of up to 50% of the total available seats in universities, government jobs, political positions, and so on, to the backward castes; introducing the *Sarva Shiksha Abhiyan* ("Universal Education Program"), which targets the education of lower caste children through incentives such as providing mid-day meals, establishing new schools, and offering scholarships. Despite these efforts, caste-based differences still persist in India. It is, thus, important to understand the role of caste in the formation of human capital, the mechanisms through which caste channelizes its impact, and how this relationship between castes and human capital evolves over time so that relevant policy measures can be taken to moderate caste-based differences.

4. Empirical Framework

4.1 Data and descriptive statistics

The main objective of this paper is to understand how the role of caste and perceived social discrimination in human capital formation evolves during childhood. To conduct the analysis, the ideal dataset would be one that provides detailed information on the following: children's cognitive outcomes and parental investment in children's education, measured at various ages; socioeconomic background, such as caste, religion, and household wealth; and the measure of perceived social discrimination.

This paper employs the younger cohort data from the YL survey. The survey started in 2002 with two cohorts: the younger cohort was aged between 6 and 18 months, and the older cohort was aged between 7.5 and 8.5 years. The sample of the younger cohort contained 2011 children, and the data were collected in five rounds at ages 1, 5, 8, 12, and 15. The children were selected from 7 districts: the Hyderabad district and a 'poor' and a 'non-poor' district from each of the 3 distinct agro-

climatic regions in Andhra Pradesh: Coastal Andhra, Rayalaseema, and Telangana.[12] Since the YL survey aimed to document child poverty, it deliberately oversampled poor communities. As a result, although households from different socioeconomic backgrounds were included, the sample is not representative of the whole population.

In each round, an extensive effort was made to find and interview children who had moved from their locations in the previous survey round. As a result, attrition between rounds was extremely low. The total attrition from round 1 to round 5 was 5.9%; these figures include attrition due to mortality, with 2.14% of children dying between the ages of 1 and 15.

For analysis, I delineate the social groups by caste and religion. I construct four categories: 1) lower caste Hindus (Hindus belonging to SC, ST, and OBC); 2) upper caste Hindus (Hindus belonging to forward/upper castes); 3) Muslims; and 4) people from other religions.[13]

Using two survey questions asked to parents in round 2, I then construct an index for parents' perceived social discrimination. The parents were asked to rate how much they agreed with the following two statements on a four-point Likert scale: "When I am at shops/market, I am usually treated with fairness and with respect by others" (called RESPECT henceforth); and "Other people in my street/village look down on me and my family" (called LOOKED DOWN henceforth). These two manifestations of discrimination are combined to form an index for parents' perceived discrimination called 'DISCR.' This index takes a value of 0 if parents report no discrimination and 1 if parents report facing either of the two manifestations of discrimination.[14]

I look at two key outcomes. The primary outcome variable is children's cognitive ability, which I proxy using the scores obtained in the PPVT, math, English, language, and reading tests conducted by the YL interviewer in various rounds. In the PPVT, a widely used receptive vocabulary test, the child is asked to select the picture that best represents the meaning of a stimulus word presented orally by the examiner. This test was administered in all rounds, starting from round 2, when the children were 5 years old. The math test scores, on the other hand, were administered in rounds 3, 4, and 5; these are, therefore, available at ages 8, 12, and 15. The English and reading test scores are available for rounds 4 and 5, respectively. The reading test was conducted in Telugu, the local and widely spoken language in Andhra Pradesh.

The second outcome variable is parental investment in education, an important part of human capital formation. Parents base their investment decisions on the returns to investments at different stages, the available resources, the prices of investment goods, and, importantly, on their beliefs about their children's development process (Attanasio et al., 2015). I use the following measures of parental investment: expenditure by the household on education-related commodities, such as books, school uniforms, private tuitions; school fees; enrollment at school; and the type of school enrolled in

[12] Andhra Pradesh is a state in the south-eastern coastal region of India. It is the seventh-largest state by area and tenth-most populous state (approx. 49 million residents with 70% rural population). It comprises of two major regions, namely Rayalaseema and Coastal Andhra. Telangana is a neighbouring state of Andhra Pradesh and is the eleventh-largest state and the twelfth-most populated state in India (approx. 35 million residents with 61% rural population). Before 2014, Andhra Pradesh and Telangana formed a single state, however, in 2014 Telangana was carved off to form a separate state.

[13] As described in section 2, Muslim groups are not supposed to be included in the SCs category, which is restricted to Hindus, Sikhs and Buddhists; Muslim STs are relatively few. According to Mondal (2003), the Mandal Commission framed in the 1980s treated 90 percent of India's Muslim population as OBCs. Thus, in this paper, I treat Muslims as a separate social group category and do not unpack the castes within Muslims. Similarly, other religions constitute a small share of population in India and are treated as a separate collective social group.

[14] The construction of this index is described in the appendix section A7.

(Private/Public) the last 12 months. All the investment variables have been deflated to 2002 prices and standardized with a mean of zero and a standard deviation of one.

In Figure 1, I demonstrate the plausible pathways between the variables of interest, offering a quick sense of how these variables are related. Caste may channelize its impact on children's cognitive and educational outcomes (a) directly through the differences in initial endowments, such as socioeconomic status, parental education, and occupational status or differences in nurturing/parental time investment in children across social groups, and so on; and (b) indirectly via mediators such as perceived discrimination and parental investment. Similarly, caste can potentially affect parental investment in education (a) directly through the differences in inherent motivation to invest or perceived returns to education across castes; (b) indirectly via our mediator of interest, perceived discrimination; and (c) indirectly by affecting children's cognitive outcomes, which, in turn, affect parents' investment decisions.

Table 1 presents the descriptive statistics of the children's household characteristics at the baseline (round 1). Around 76% live in rural communities, 54% are male, and the average household size is 5.4. The mothers are relatively young, with an average age of 24, and their education is low, with an average of just 3 years. Lower caste Hindus comprise 72.6% of the surveyed households; 14% are upper caste Hindus; and the remaining are Muslims and other religions.

In Table 2, I report additional statistics that vary across rounds. A significant fraction of the children suffers from stunting, wasting, and being underweight. Together, these indicators suggest significant morbidity in this population. While the wealth index seems to be rising as the cohorts age (in part reflecting the economic growth in the area), health indicators (height, weight, and BMI for age z-scores) do not improve. Children spend minimal time working on household chores. By age 15, children spend approximately an hour a day helping out at home. Almost no children do paid work outside of their homes. School enrollment increases from 22% at age 5 to 96% at age 12 and then decreases to 83% at age 15.

In Table 3, I report the initial endowments of the children at the baseline[15] across the four social groups. The descriptive statistics provide suggestive evidence that children from backward castes are economically disadvantaged and, on average, exposed to challenging environments. For example, lower caste children belong to lower-income families and have lower parental education. They also have a lower height for the age z-score and are more likely to be stunted and underweight. Muslims, on average, are disadvantaged when compared to upper caste Hindus but are better off than lower caste Hindus. On average, Muslims start school late and have lower school enrollment at the age of 5 years when compared to other social groups. To check whether some social groups are more likely to report social discrimination than others, I run a regression of perceived discrimination on caste with and without controlling for income and wealth. The results are reported in Table 4. After controlling for wealth and income, lower Hindu castes (SC, ST, and OBC) are 7% more likely to face discrimination compared to upper caste Hindus. Other religions, on the other hand, are 15% more likely to face discrimination compared to upper caste Hindus. There is also evidence of class-based discrimination with higher perceived discrimination among low-income families.

4.2 Ordinality of test scores

As recent literature has affirmed, test scores, since they are monotonic transformations of some unobserved true measure of ability in a subject, are ordinal in nature and lack interval properties (Bond & Lang (2013), Lord (1975)). Though they offer a rank ordering of students, they cannot

[15] Table 3 reports the descriptive statistics in round 1 (age 1) for all variables except school enrollment, which is reported for round 2 (age 5) because children are not enrolled in schools at the age of 1.

measure by how much one student outperforms another. For example, the difference between a test score of 40 and 50 may be either more or less than the difference between a score of 70 and 80.

To understand how order-preserving monotonic transformations affect the evolution of gaps, Bond and Lang (2013) performed a bounding exercise on the Black–White test score gap. They used an algorithm to generate monotonic transformations of the original test score scale to maximize and then minimize the growth of the test score gap and also maximize the correlation between kindergarten and third-year scores. Order-preserving scale transformations of these test scores, it was found, can provide contradictory conclusions, starting from increasing gaps over time to decreasing gaps over time, and some transformations that also suggest that Blacks outperform Whites over time.

To deal with this issue of ordinality of scores, Cunha and Heckman (2008) and Bond and Lang (2018) proposed anchoring the test scores to adult outcomes such that the gaps are expressed in concrete units, such as completed years of education. As the adult outcomes are observed with significant delays, the primary challenge with this approach is the difficulty in the availability of relevant data.

In this paper, I transform the test scores in each round into percentile rankings that are scale-invariant. Percentile rankings are also invariant to changes in the scores that do not change ranks even though such modifications can change the relative means of the raw test scores. I estimate the gaps across castes at a given age as the difference in the mean percentile ranking. To maintain comparability with the literature on this subject, the test scores are standardized with a mean of zero and a standard deviation of one for the overall sample in each test and round.

Another solution to estimate the test score gaps that are invariant to monotonic transformations was recently proposed in a RES paper by Penney (2017). The proposed method employs the ordinary least squares variant of unconditional quantile regressions (UQR) as developed by Firpo, Fortin, and Lemieux (2009) to estimate the test score gap at the median and then normalizes the coefficients of interest by dividing them by the standard error of the regression. This is in contrast to the usual method, which normalizes the coefficients by dividing them by the standard deviation of the dependent variable. I use this solution as a robustness check to verify that the test score gaps across castes over time are not merely artifacts of test score scaling.

4.3 Estimation

To estimate the differences in children's cognitive outcomes and parental investment across social groups over time, I run the following regression model (1) separately for each YL round:

$$\boldsymbol{Y_{it} = \alpha_0 + \beta_t\, Social\ Group_i + \theta_t\, X_{it} + \gamma_i + \epsilon_{it}} \qquad \textbf{(1.1)}$$

where, Y_{it} denotes the two outcome variables of interest: a) child i's percentile ranking in round t in various tests, and b) child i's standardized parental investment in round t. $Social\ Group_i$ is a full set of social group dummies with upper caste Hindu as the base category. X_{it} represents a parsimonious set of the following socioeconomic controls: gender of the child; height by age score as a proxy for the child's health; birth order; sibling size; and socioeconomic status proxied by the wealth index[16] of the household, all at the baseline survey. γ_i represents the community-level fixed effects, which controls for unobserved heterogeneity among the communities, and ϵ_{it} is the random

[16] The data on wealth index are provided in the YL survey. The wealth index takes on a value between 0 and 1 and is constructed as an average of three indices: housing quality (quality of wall, roof, floor, and the number of rooms/person), access to services (drinking water source, sanitation, cooking fuel, and electricity), and the ownership of consumer durables (9 items, such as radio, television, motorbike, bicycle, automobile, landline, mobile phone, refrigerator, and fan) (Briones, 2017).

error term. The standard errors are clustered at the community level[17] to account for the correlation in outcomes within communities.

First, I estimate the model without including socioeconomic controls (X_{it}). Consequently, β_t is the coefficient of interest that captures the raw gaps in outcome variables between the named caste category and the upper Hindus in round t.

Next, I estimate the model with socioeconomic controls. Here, β_t captures the unexplained gaps in outcome variables across castes that cannot be explained by differences in SES. It represents the additional penalty of being born to a backward social group, as opposed to being born to an upper Hindu caste, that is not explained by other measures of socioeconomic disadvantage.

In Fryer and Levitt (2004), estimating the racial differences in children's test scores in the US, controls for parental education, parental employment, and household income are also included. However, one could argue that these controls are heritable and themselves outcomes of caste, which is the treatment in my analysis, making them bad controls. Hence, in my main specification, I do not include controls for these variables. However, I do conduct a robustness check by including these controls, and the results remain unchanged.[18]

Next, to estimate the association between perceived discrimination and the outcome variables, I run the following regression for each YL round.

$$\boldsymbol{Y_{it} = \alpha_0 + \beta_t\, Social\ Group_i + \delta_t Discr_i + \theta_t X_{it} + \gamma_i + \epsilon_{it}} \qquad \textbf{(1.2)}$$

where, $Discr_i$ is the parents' perceived social discrimination index, with 'no discrimination' as the omitted category. It takes a value of 1 if parents perceive any manifestation of social discrimination, and 0, otherwise.

I estimate equation (2), both with and without including socioeconomic controls (X_{it}). The coefficient of interest, δ_t, captures the association between parents' perceived discrimination and the outcome variables, holding the social group constant.

There are some concerns about using self-reported measures of perceived social discrimination. First, it may also pick up discrimination perceived due to one's class or economic status. Estimating equation (2), including controls for SES, addresses this concern as it captures the effect of discrimination holding one's economic status constant.

Second, there is a possible concern regarding 'social desirability bias.' Social desirability reflects the tendency of research subjects to claim or say things that are socially desirable and to place the speaker in a favorable light. For instance, in the context of this study, this may arise if the parents under-report social discrimination because they might feel ashamed in acknowledging it. While I cannot rule out this concern, I note that, in this case, the estimates on perceived social discrimination, controlling for caste, are attenuated. Another possibility is that the parents of children who perform worse on tests are more likely to report discrimination. It is important to note that to alleviate this concern, I do not use a contemporaneous measure of perceived discrimination. As described in section 4.1, I use the earliest available perceived discrimination measure from round 2.

Specification (2) assumes that the association of perceived discrimination on children's cognitive outcomes and parental investment is the same across social groups. As discussed in this paper before,

[17] Some children migrate over YL rounds to different communities, but I cluster the standard errors at round 1 community level to maintain comparison among the same children over time. I also conduct some robustness checks, which are described in detail in section 6.1.

[18] The results from the models controlling for household income and parental education are unchanged. I discuss robustness checks in detail in section 6.1.

social groups are likely to behave differently when faced with discrimination. I check whether the influence of perceived discrimination on human capital and parental investment varies across social groups by estimating the following equation—both without and with socioeconomic controls.

$$Y_{it} = \alpha_0 + \beta_t\, Social\ Group_i + \delta_t Discr_i + \mu_t\,(Social\ Group_i \times Discr_i) + \theta_t\, X_{it} + \gamma_i + \epsilon_{it} \quad (1.3)$$

In equation (4), μ_t, the coefficient of interest, captures the additive effect of belonging to the named social group and facing discrimination as compared to the upper Hindu castes. In other words, μ_t captures the differential effect of discrimination on human capital across social groups.

Finally, I test whether the estimates of the gaps in human capital across social groups vary over survey rounds by using the seemingly unrelated estimation (SUEST) method developed by Weesie (2000).

5. Results

5.1 Cognitive ability

Figure 2 presents a series of estimates of the gaps across social groups in percentile rankings in various tests over the YL survey rounds.[19] To analyze how the gaps evolve over time, I use the SUEST method to compare whether the coefficients across rounds are statistically different. The left panel presents the raw differences in mean percentile ranks between the stated social group and upper caste Hindus, not controlling for any socioeconomic covariates. These results simply reflect the raw percentile rank gaps across social groups. The right panel presents the estimates from the specifications, including SES controls.

At the age of 5 years, there is a significant raw percentile ranking difference of 0.31 standard deviations in the PPVT between lower caste Hindus and upper caste Hindus. This raw gap remains at 0.26 standard deviations at age 8, and 0.37 standard deviations at age 12 and 15. Once I introduce socioeconomic controls, significant gaps persist. At the age of 5 years, the gap in PPVT percentile ranking between lower and upper caste Hindus after controlling for the covariates is 0.18 standard deviations, statistically significant at the 1% level. By the age of 15 years, this gap is 0.268 standard deviations. The results are reported in Figure 2a.

Next, I test for the equality of the estimates of the PPVT percentile rank gaps across YL rounds, both with and without controls, and fail to reject the null hypothesis that the estimates are not significantly different across rounds, suggesting that the gap between lower and upper caste Hindus in the PPVT test remains constant over the 10 years of the study period.

Muslims, on the other hand, begin with a statistically insignificant deficit of 0.12 standard deviations, which disappears after controlling for socioeconomic factors. However, over time, they lose substantial ground relative to upper caste Hindus, and their PPVT percentile ranking deficit increases to 0.48 standard deviations by the age of 15 years, which is statistically significant at the 1% level. The unexplained gap, after controlling for socioeconomic background, remains at 0.39 standard deviations, significant at the 5% level, higher than that of lower caste Hindus.

Figures 2B and 2C are identical to Figure 2A but present estimates from math, English, and reading tests. Even after controlling for socioeconomic background, lower caste Hindus rank 0.20 standard deviations below upper caste Hindus in math at the age of 8. There is evidence of a slight

[19] The complete table of estimates is reported in the appendix (Table A1 to A6). The odd-numbered columns present the differences in mean percentile ranks in various tests, not controlling for any socioeconomic covariates. These results simply reflect the raw percentile rank gaps across social groups. The even-numbered columns present the estimates from specifications, including SES controls.

increase in this gap over time to 0.26 and 0.36 standard deviations by the ages of 12 and 15 years, respectively. These gaps are statistically significant at the 1% level. Muslims perform consistently lower than both upper and lower Hindu castes in math throughout the three YL rounds and rank 0.44, 0.47, and 0.40 standard deviations below upper caste Hindus, all significant at the 1% level. These estimates for Muslims are not significantly different across rounds, indicating that the gap persists over time.

As compared to the racial gaps in the math test scores between Blacks and Whites in the US, observed by Fryer and Levitt (2005), caste-based gaps in India are smaller in magnitude and are persistent over time. Fryer and Levitt (2005), on the other hand, found that the racial gap in kindergarten disappears after controlling for a small number of socioeconomic covariates, whereas over the three years of schooling, this gap increases significantly and can no longer be explained by the differences in SES across races.

In the English test conducted at the age of 12 years, lower caste Hindus rank 0.27 standard deviations below upper caste Hindus, and after controlling for SES, Muslims are 0.39 standard deviations below upper caste Hindus. In the reading test conducted at the age of 15 years, the gap is 0.26 and 0.38 standard deviations for lower caste Hindus and Muslims, respectively. All these estimates are statistically significant at the 1% level.

The results indicate that there are significant differences in the test scores of children across castes starting from a very early age and that these gaps persist over time. On average, lower caste Hindu children perform significantly worse than their upper caste counterparts, and Muslims perform worse than Hindus in general. These gaps cannot be completely explained by the differences in socioeconomic background, and when socioeconomic factors are controlled, significant gaps remain.

The controls in all specifications enter with the expected sign. An increase in wealth at the baseline survey is associated with an improvement in children's performance in all the tests. Better health proxied by height by the age z-score is also associated with improvement in the test scores of children. This suggests that socioeconomic background at an early age is a crucial determinant of human capital formation. I do not find any significant differences in the test scores of males and females.

Next, to capture the association between parents' perceived discrimination and children's test scores, I estimate equation (2) for each YL round. The results, reported in Figure 3, suggest that until the age of 8, there is no significant association between parents' perception of social discrimination and children's performance in the PPVT. However, starting from round 4, when children are aged 12 years, parents' perceived discrimination is significantly negatively associated with their children's performance. Controlling for the social group, the children of parents perceiving discrimination rank 0.19 standard deviations below when compared to children whose parents do not perceive discrimination in the PPVT. This gap remains at 0.12 in round 5. When the difference in SES background is accounted for, this gap is 0.18 standard deviations in round 4 and 0.10 in round 5. These estimates across rounds 4 and 5 are not statistically different from each other, suggesting that the negative association persists. These results are consistent with the fact that, as children age, they develop awareness about discrimination, which harms their academic performance. For example, as reported by Brown and Bigler (2005), in the US, most children (92%) are familiar with the meaning of discrimination by the age of 10.

Similarly, children whose parents report facing discrimination also score substantially lower on math and English tests. The raw math percentile ranking gaps between children whose parents perceive discrimination and children whose parents perceive no discrimination are 0.15, 0.13, and 0.17 standard deviations at ages 8, 12, and 15, respectively. When I control for the SES background, these deficits are 0.11, 0.09, and 0.14. These gaps across rounds are not statistically different from each other.

In the English test conducted at age 12, this deficit is 0.20 and 0.15 standard deviations, without and with SES controls, respectively. I do not find any significant correlation between perceived discrimination and reading test scores conducted at the age of 15 years, which could be because the reading test was conducted in Telugu. Overall, perceived social discrimination, the results confirm, is significantly negatively associated with the cognitive development of children as they grow older.

Next, to capture the association between discrimination and the cognitive outcomes of children within each social group, I estimate equation (3). The results are reported in Figure 4, and each estimate captures the average difference in standardized percentile rankings between children of parents perceiving discrimination and children of parents not perceiving discrimination belonging to the same social group. For upper caste Hindus, there is no significant association between perceived discrimination and PPVT scores. Similarly, for math, English, and reading test scores, the estimates are insignificant for upper caste Hindus, except for math test scores at age 8, which is significant only at the 10% level.

For lower caste Hindus, perceived discrimination is significantly negatively associated with children's performance in the PPVT at the age of 12; this effect statistically remains constant at age 15. Lower caste Hindus facing discrimination score 0.20 standard deviations below upper caste Hindus facing no discrimination at age 12, and this gap remains at 0.08 standard deviations at age 15. The effect of discrimination on lower caste Hindus is also significant in math, English, and reading tests. These results suggest that the perception of social discrimination disproportionately affects the test scores of children from lower Hindu castes. Perceived discrimination seems to be a significant contributor to the differences in cognitive outcomes observed between lower and upper caste Hindus. This may be because the lower castes have been subjected to discrimination when it comes to accessing skills, higher education, and the labor market. The perception of social discrimination among the lower castes may further aggravate these existing beliefs about discrimination, lowering their expected returns to education and the labor market and, thus, disproportionately affecting their outcomes. For Muslims, I do not find any significant gaps in children's cognitive outcomes by parents' perception of social discrimination.

5.2 Parental investment

Figure 5 presents the estimates of the differences in parental investment in children's education across social groups over time. The odd-numbered columns report the raw differences in the investments across social groups without controlling for socioeconomic factors, while the even-numbered columns control for socioeconomic covariates.

When compared to upper caste Hindus, lower caste Hindus, the estimates in Figure 5A suggest, invest significantly less in their children's education. These investment gaps remain constant over time. At the age of 5 years, the raw gap in annual expenditure in children's education between lower caste Hindus and upper caste Hindus is 0.52 standard deviations, and, at the age of 15, it is 0.39 standard deviations. These estimates across rounds are not statistically significantly different, suggesting that the gap in children's expenditures is persistent over time. After controlling for socioeconomic background, the gap is 0.34 standard deviations at age 5, remains statistically constant over time, and is 0.45, 0.30, and 0.30 in subsequent rounds. All estimates are statistically significant at the 1% level.

Muslim parents invest even less in the education of their children. The raw investment gap in Muslims and upper caste Hindus is 0.64 standard deviations at age 5 and remains statistically constant throughout. After controlling for the covariates, this investment gap is 0.54 standard deviations at the age of 5 and 0.48 standard deviations at the age of 15. Parents belonging to other religions, such as Christianity, Jainism, Sikhism, and so on, also invest significantly less in the education of their children relative to upper caste Hindus.

Figure 5B reports the differences in school enrollment for different social groups compared to upper caste Hindus. There are no significant differences in the enrollment at school across social groups until round 3. However, by the age of 15, compared to upper caste Hindus, children from lower Hindu castes and Muslims are 7% and 10% less likely to attend school, respectively. Muslims and lower caste Hindus, the results suggest, drop out of school early.

Figure 5C reports the estimates of the differences in expenditure on school fees across social groups in rounds 4 and 5, conditional on enrollment at school. Lower caste Hindus, the results indicate, spend significantly less on school fees compared to upper caste Hindus, suggesting that upper caste Hindus, in general, attend expensive schools. After controlling for SES, the difference in the annual school fees paid by the parents belonging to these two social groups is 0.38 and 0.40 standard deviations in rounds 4 and 5, respectively. The corresponding gap between Muslims and upper caste Hindus is 0.50 and 0.34 standard deviations in rounds 4 and 5, respectively.

The estimates of the differences in the type of school attended by children across castes, conditional on enrollment at school, are presented in Figure 5D. Even after controlling for socioeconomic differences, there are significant differences across social groups in the type of school children attend. The gaps are constant over time, and lower caste Hindus are 20% less likely to attend private schools—which are comparatively better and more expensive compared to public schools—than their upper caste Hindu counterparts. For Muslims, this difference is 32% at age 5 and 15% at age 15.

Next, I estimate equation (2) for parental investment. The results are reported in Figure 6. The estimates from panel A suggest that even after controlling for the social group and other observables, parents' perception of social discrimination is associated with lower investment in their children's education. Parents perceiving discrimination invest 0.16 standard deviations less compared to parents perceiving no discrimination in round 2. The estimate is statistically significant at the 1% level. In round 3, when children are 8 years old, the estimate on perceived discrimination is insignificant. In rounds 4 and 5, the gap in investment between parents perceiving discrimination and parents perceiving no discrimination is 0.23 and 0.09 standard deviations that are significant at the 1% and 5% level, respectively.

The results in panel B suggest that even after controlling for the observed SES covariates, parents perceiving discrimination spend around 0.10 standard deviations less on school fees compared to parents perceiving no discrimination in rounds 4 and 5. Both estimates are statistically significant. Parents' perceived discrimination exerts no significant effect on their children's school attendance until round 4. However, in round 5, discrimination reduces school enrollment by 5%, suggesting that children belonging to parents who perceive discrimination are significantly likelier to drop out of school by age 15. In other words, perceived discrimination leads to earlier school dropouts.

The estimates reported in Figure 6 C suggest that, conditional on enrollment at school, perceived social discrimination significantly affects the likelihood of attending private schools. Controlling for SES, at the age of 5, children are 15% less likely to attend private school if their parents perceive discrimination. This gap is 7%, 5%, and 7% in rounds 3, 4, and 5, respectively.

Finally, I estimate equation (3) to check if perceived discrimination has different significance for different social groups in terms of investment decisions in children's education. The results are reported in Figure 7. The estimates for upper caste Hindus capture the differences in parental investment within the upper Hindu castes between parents perceiving social discrimination and parents not perceiving any social discrimination. Similarly, the estimates for lower caste Hindus capture the gaps in parental investment within the lower Hindu castes between parents perceiving

discrimination and parents perceiving no discrimination. I find that there is no significant association between perceived discrimination and parental investment for upper caste Hindus. For lower caste Hindus, discrimination plays a crucial role and is associated with lower parental investment in the education of children. Compared to lower caste Hindus who do not perceive discrimination, lower caste Hindu parents who perceive discrimination invest significantly less in their children's education, spend less on school fees, and are less likely to send their children to private schools. For Muslims, I find that perceived discrimination is associated with lower school enrollment at all ages, except for round 3, where the coefficient is statistically insignificant.

This differential association between perceived social discrimination and parental investment across social groups could be because the perception of social discrimination among backward castes aggravates parents' existing belief about discriminatory practices in access to skills, higher education, and the labor market, thus disincentivizing parents to invest in their children's education and perpetuating social inequality over generations for backward castes.

5.3 Analysis of lower Hindu castes

I also analyze the lower Hindu castes to estimate the differences in cognitive outcomes and parental investment in education between SCs, STs, and OBCs. Tables 5 and 6 report the results for test scores and parental investment, respectively.

There are no differences in the percentile rankings in tests across SCs, STs, and OBCs. Concerning perceived discrimination, there exists no significant association between parents' perceived discrimination and children's PPVT scores at the ages of 5 and 8. However, at age 12, discrimination is associated with a lower PPVT percentile ranking by 0.20 standard deviations, significant at the 1% level. At the age of 15 years, the association is again insignificant.

Concerning parental investment, SCs and STs spend significantly less on children's education in the early years, but there are no significant differences in expenditure across castes in the last two rounds, at ages 12 and 15. SCs spend significantly less on school fees than STs. However, there is no significant difference in the school fees paid between OBCs and STs. There are also no differences in school enrollment across SCs, STs, and OBCs. Finally, SCs are significantly less likely to attend private schools than OBCs and STs. Thus, among the lower castes, parents' perception of discrimination significantly lowers expenditures on education, school enrollment, and private school attendance.

6. Robustness Checks

6.1 Sensitivity of results

I conduct several robustness checks to test the sensitivity of my estimates to different specifications of the model. First, I test whether the estimates are sensitive to the inclusion of other SES controls, such as family income and parental education. This follows from the discussion before about these variables being themselves outcomes of caste, making them bad controls. Reassuringly, the estimates from the specifications, including the above controls, mirror the estimates from the main specification.

Next, since some children migrated over time to other places, I also tried specifications with round specific location fixed effects, and the results are robust. I also analyze a subsample excluding children who migrated outside their sentinel site over time to avoid comparing children who migrated with children who did not migrate, as migration could be an endogenous choice. The results are

reported in Table 7. The estimates are robust and mirror the results from the main specification reported in Figures 2 and 5.

Additionally, I check if gaps in cognitive outcomes and parental investment observed over time across social groups and the manifestation of perceived discrimination persist if controls for children's initial cognitive ability are included. The estimates, reported in Tables 8 and 9, capture the differences in children's test scores and parental investment in successive rounds across social groups conditional on lagged PPVT scores at the age of 5 years. All the regressions include socioeconomic controls as well as location fixed effects. The estimates reduce in size, but overall, the results mirror the findings from the main analysis.

6.2 Ordinality: Are the results simply an artifact of test score scaling?

As discussed in section 4.2, test scores are ordinal measures of achievement. To verify that the results are not mere artifacts of test score scaling, I use the solution proposed by Penney (2017), which employs UQR to estimate the test score gap at the median and then normalizes the coefficients of interest by dividing them by the standard error of the regression. Its invariance to monotonic transformations means that the same regression results will be obtained as if one had access to the "true" set of test scores. Table 10 displays the coefficient estimates for lower caste Hindus from estimating equation (1) using OLS, unconditional quantile regression at the median, and the method outlined in Penney (2017).

Comparing the z-UQR estimates at the median in columns (2) and (5) with those from the proposed method in columns (3) and (6) titled 'Normalized,' both without and with SES controls, respectively, the evolution pattern of the percentile ranking gaps over time is almost identical. The gaps in percentile rankings between lower and upper caste Hindus, the results reveal, are constant over time, suggesting that the gaps persist over time. The OLS results, also reported in the table in columns (1) and (4) for comparison purposes, exhibit similar evolution patterns as the normalized estimates. Overall, the above robustness check assuages the concerns of ordinality and test score scaling and raises confidence in the results.

6.3 Oster test

There are many unobserved individual or household characteristics, such as parents' preferences for children's education and children's innate ability, that may be correlated with caste or perceived social discrimination, and excluding them from the regression may lead to omitted variable bias. To check if the estimates of gaps in children's test scores and parental investment across castes and the effect of perceived discrimination are robust to omitted unobservable variable bias, I perform the Oster test (Oster, 2019).

Oster (2019) built on Altonji et al. (2005) to develop a novel method for assessing bias from unobservable factors and estimating the degree of selection on unobservable that would be required to drive the ATT to 0 (called $\tilde{\delta}$). It exploits information on coefficient movements and movements in R-squared values after the inclusion of controls to compute bounds for the treatment effect.

To identify the bias-adjusted β^*, one needs assumptions on (1) δ, the degree of selection on unobservables, and (2) $R_{\max}$ that indicates the maximum share of variance of the outcome that could be explained by any set of observable and unobservable covariates. Oster (2019) argued that $\delta \in [0, 1]$ is an appropriate bound because observed control variables are deliberately chosen as determinants of the outcome and must be at least as important as the unobservables. It is unlikely that $\delta > 1$ i.e., unobservables have a stronger impact on the outcome variable than the control variables.

For this analysis, I assume $R^2_{max} = 1.3 \times R^2_{controlled}$, a rule proposed by Oster, where $R^2_{controlled}$ is the R^2 of the model with all observables and assume δ = 0.80. In other words, I am assuming that selection on unobservable is 80% of selection on observable, and inclusion of these unobservables would have increased R^2 of the regression by 1.3 times. I also estimate the value of $\tilde{\delta}$ that will drive the estimate to 0.

I perform this test by estimating the following models:

$$\boldsymbol{Y_{it} = \alpha_0 + \beta_t\, UpperCaste_i + \theta_t X_{it} + \gamma_i + \epsilon_{it}} \quad \textbf{(1.4)}$$

$$\boldsymbol{Y_{it} = \alpha_0 + \beta_t\, Caste_i + \rho_t Discr_i + \theta_t X_{it} + \gamma_i + \epsilon_{it}} \quad \textbf{(1.5)}$$

where Y_{it} denotes child i's percentile ranking in round t in various tests and parental investment in child i's education in round t. $UpperCaste_{it}$ is an indicator of caste, which takes a value of 1 if the child belongs to the upper caste, and 0 if otherwise. $Caste_{it}$ is a full set of caste dummies with upper Hindu caste as the base category. $Discr_{it}$ is the index for parents' perceived discrimination. X_{it} represents an array of child-level social and economic variables

The results are reported in Table 11. The baseline effects (column 1) are the estimates from the regressions, including controls for the children's sex, birth order, and location fixed effect (variables are assumed to be unrelated to the set of proportionally related unobservables). The controlled effects are the estimates from the regressions, including the full control set. The bias-adjusted β^* are the estimates adjusting for the plausible bias due to the unobservables.

The estimates reported in Table 11a capture the differences in test scores and parental investment between upper Hindu castes and other social groups. Even when selection on unobservables is as high as 80% of selection on observables, the omitted variable bias does not change the direction of most of the estimates, except for PPVTs in rounds 2 and 3. The last column reports the minimum degree of selection on unobservables that would be required to drive the estimate to 0 (called $\tilde{\delta}$). For most of the significant estimates, the estimated $\tilde{\delta}$ is close to 1.

In Table 11b, I report the estimates of the effects of parents' perceived discrimination on children's test scores and parental investment. All the estimates are robust to omitted variable bias, except for school enrollment in round 5, and the estimated $\tilde{\delta}$ is greater than 1 for all significant estimates, thus raising confidence in the results.

6.4 Multiple hypothesis testing

Since this paper looks at multiple outcomes, this raises concerns about the over-rejection of null hypotheses unless the multiplicity of the testing framework is explicitly considered (Anderson, 2008; Conti et al., 2016).[20] To address this issue, I adjust for multiple hypotheses testing using the Romano & Wolf (2005) stepdown method.[21] The Romano-Wolf correction uses bootstrap to control for the familywise error rate (FWER) which captures the probability of rejecting at least one true null hypothesis in a family of hypotheses under test. As discussed in Conti et al. (2016) and Heckman et

[20] Suppose that a single-hypothesis test statistic rejects a true null hypothesis at significance level α. Thus, the probability of rejecting a single hypothesis out of K true hypotheses is given by $1 - (1 - \alpha)$ K. As the number of outcomes K increases, the likelihood of rejecting a true null hypothesis departs from α (Conti et al., 2016).

[21] Lehmann and Romano (2005) and Romano and Wolf (2005) discuss the stepdown procedure in depth.

al. (2010), I define two blocks of similar outcomes- test scores and parental investment and carry out 1000 bootstrap replications.

The estimates are reported in Table 12. Panel A reports the estimates for cognitive outcomes and Panel B reports the estimates for parental investment. Columns 1 and 3 report the p-values for estimates on gaps between lower and upper Hindu castes from specification 1.1 and estimates on discrimination from specification 1.2, respectively. Columns 2 and 4 report the p-values adjusted for multiple hypothesis testing associated with columns 1 and 3, respectively. The estimates on lower Hindu caste and discrimination are robust to multiple hypothesis testing as the Romano-Wolf adjusted p-values are always similar to the model p-values.

7. Discussion

The objective of this paper was twofold: 1) to investigate the age at which the gaps in human capital formation across social groups originate and how they evolve as children age from 5 to 15 years; and 2) to analyze the role of perception of social discrimination in human capital formation over this time and its contribution to the gaps in human capital observed across social groups. I look at two determinants of human capital: cognitive outcome, as measured by the performance on various tests, and parental investment in children's human capital. I focus on a particular setting in India where social identity is defined along the lines of caste and religion. This is the first study to analyze the dynamics of the role of the Indian caste system as well as perceived social discrimination in human capital formation over time. This paper speaks to the following two strands of literature: (1) the literature of human capital formation, and (2) the research in economics focusing on social institutions, social discrimination, and stereotypes.

The findings of this study suggest that there are substantial and persistent gaps in the test scores of children across social groups that cannot be completely explained by differences in SES. Lower caste Hindus and Muslims score significantly less on PPVT, math, English, and reading tests compared to upper caste Hindus. The gaps between lower caste Hindus and upper caste Hindus in the PPVT percentile rankings are constant over time. Whereas, there is evidence of a slight increase in percentile ranking gap on math tests over time. Muslims, on the other hand, perform equally well as upper caste Hindus at the age of 5 and 8 years in the PPVT; however, over time, they lose significant ground to both upper and lower caste Hindus. However, in math, English, and reading tests, Muslims persistently lag behind Hindus.

This paper also establishes that parents from backward social groups invest significantly less in the education of their children. These castes and religion-based gaps in parental investments also persist over time. Children from backward social groups are more likely to drop out early compared to children from upper Hindu castes; they are also less likely to attend private schools. Even after differences in socioeconomic backgrounds are accounted for, these gaps persist.

Consistent with the hypothesis that children develop an awareness of discrimination as they grow old, which harms their cognitive development, this study finds that parental perception of discrimination is associated with children's lower cognitive performance at a later age. While itis associated with lower investment in children's education throughout childhood. It is also associated with increased early dropouts from schools and lower likelihood of children attending private schools. A significant negative association between perceived social discrimination and children's cognitive development as well as parental investment is found only for lower Hindu castes, suggesting the higher vulnerability of this social group.

The study findings are informative for the enduring debates concerning social policies favoring children from disadvantaged social groups at an early age. The findings highlight that the social group to which children are born as well as the perceptions of social discrimination against oneself may shape their abilities and, thus, harm their life outcomes, such as cognitive development, educational attainment, and so on. Policies promoting investment in the education of children from backward castes at early ages and rooting out social discrimination against them are fundamental to improving human capital. Ensuring the social inclusiveness of backward castes, sensitizing teachers and youth, and convincing parents of first-generation students of the value of education can make a big difference.

The Government of India has introduced policies of affirmative action to address social inequality, such as reservations at government universities, government jobs, and legislative representation for socially backward groups. The literature has highlighted the success of these policies, including improved representation of marginalized communities in government jobs (Deshpande & Ramachandran, 2016) and improved diversity of social backgrounds in higher education (Bertrand et al., 2010). However, these affirmative action policies may be insufficient to adequately address the gaps in human capital generated early in life. These policies may not raise human capital for lower caste children but only provide them representations at higher educations and jobs by reserving seats for them and diluting the eligibility criterion.

The Government of India also provides free education to socially and economically disadvantaged children through public schools. However, there is abundant evidence in the literature that these schools are poor in quality, both in terms of infrastructure and quality of education. In 2009, the Right of Children to Free and Compulsory Education Act or Right to Education Act (RTE) was enacted, making education a fundamental right of every child between 6 and 14, necessitating all private schools to reserve 25% of their seats for children from poor and other categories. However, the act has been criticized for several reasons. In India, elementary education starts at the age of 2 and half years of age, but this policy excludes children below the age of 6. Moreover, according to the study's findings, significant differences in cognitive outcomes and parental investment develop across social groups by the age of 5, so this act fails to address the issue of the significant gaps generated before the age of 6. Moreover, the act excludes children aged above 14. In India, secondary education covers children aged 14 to 18, and the exclusion of children above 14 years would lead to a significant increase in school dropouts at this age, which is also one of the findings of this study. This study thus highlights the need for policies that can alleviate the consequences of being born in a particular social group—an important source of inequality not only in India but also globally.

8. References

Akerlof, G. A., & Kranton, R. E. (2000). Economics and Identity*. *Quarterly Journal of Economics*, *115*(3), 715–753. https://doi.org/10.1162/003355300554881

Altonji, J. G., Elder, T. E., & Taber, C. R. (n.d.). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 34.

Arouri, M., Ben-Youssef, A., & Nguyen, C. V. (2019). Ethnic and racial disparities in children's education: Comparative evidence from Ethiopia, India, Peru and Viet Nam. *Children and Youth Services Review*, *100*, 503–514. https://doi.org/10.1016/j.childyouth.2019.03.031

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., & Rubio-Codina, M. (n.d.). *Estimating the Production Function for Human Capital: Results from a Randomized Control Trial in Colombia*. 52.

Attanasio, O., Meghir, C., & Nix, E. (2015). *Human Capital Development and Parental Investment in India* (Working Paper No. 21740; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w21740

Attanasio, O., Meghir, C., Nix, E., & Salvati, F. (2017). Human capital growth and poverty: Evidence from Ethiopia and Peru. *Review of Economic Dynamics*, *25*, 234–259. https://doi.org/10.1016/j.red.2017.02.002

Banerjee, B., & Knight, J. (1985). Caste discrimination in the Indian urban labour market. *Journal of Development Economics*, *17*(3), 277–307.

Bertrand, M., Hanna, R., & Mullainathan, S. (2010). Affirmative action in education: Evidence from engineering college admissions in India. *Journal of Public Economics*, *94*(1–2), 16–29. https://doi.org/10.1016/j.jpubeco.2009.11.003

Bond, T. N., & Lang, K. (2013). THE EVOLUTION OF THE BLACK-WHITE TEST SCORE GAP IN GRADES K–3: THE FRAGILITY OF RESULTS. *THE REVIEW OF ECONOMICS AND STATISTICS*, 12.

Bond, T. N., & Lang, K. (2018). The Black-White Education-Scaled Test-Score Gap in Grades K-7. *The Jounal of Human Resources*, 65.

Borooah, V. K. (2012). Social Identity and Educational Attainment: The Role of Caste and Religion in Explaining Differences between Children in India. *Journal of Development Studies*, *48*(7), 887–903. https://doi.org/10.1080/00220388.2011.621945

Briones, K. (2017). How Many Rooms Are There in Your House? Constructing the Young Lives Wealth Index. *Young Lives*, 18.

Brown, C. S. (2015). The Educational, Psychological, and Social Impact of Discrimination on the Immigrant Child. *Migration Policy Institute*, 28.

Brown, C. S., & Bigler, R. S. (2005). Children's Perceptions of Discrimination: A Developmental Model. *Child Development*, *76*(3), 533–553.

Cheng, E. R., Cohen, A., & Goodman, E. (2015). The Role of Perceived Discrimination during Childhood and Adolescence in Understanding Racial and Socioeconomic Influences on Depression in Young Adulthood. *The Journal of Pediatrics*, *166*(2), 370-377.e1. https://doi.org/10.1016/j.jpeds.2014.10.010

Cunha, F., & Heckman, J. (2007). *The Technology of Skill Formation*. 38.

Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, *43*(4), 738–782. https://doi.org/10.3368/jhr.43.4.738

Dee, T. S. (n.d.). TEACHERS, RACE, AND STUDENT ACHIEVEMENT IN A RANDOMIZED EXPERIMENT. *THE REVIEW OF ECONOMICS AND STATISTICS*, 16.

Deshpande, A. (2001). Caste at Birth? Redefining Disparity in India. *Review of Development Economics*, *5*(1), 130–144. https://doi.org/10.1111/1467-9361.00112

Deshpande, A., & Ramachandran, R. (2016). The Changing Contours of Intergroup Disparities and the Role of Preferential Policies in a Globalizing World: Evidence from India. *CDE Working Paper No. 267*, 43.

Dommaraju, P., Agadjanian, V., & Yabiku, S. (2008). The Pervasive and Persistent Influence of Caste on Child Mortality in India. *Population Research and Policy Review*, *27*(4), 477–495. https://doi.org/10.1007/s11113-008-9070-0

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, *77*(3), 953–973. JSTOR.

Fryer, R. G., & Levitt, S. D. (2004). Understanding the Black-White Test Score Gap in the First Two Years of School. *THE REVIEW OF ECONOMICS AND STATISTICS*, 18.

Fryer, R. G., & Levitt, S. D. (2005). *The Black-White Test Score Gap Through Third Grade* (Working Paper No. 11049). National Bureau of Economic Research. https://doi.org/10.3386/w11049

Fryer, R. G., & Levitt, S. D. (2006). The Black-White Test Score Gap Through Third Grade. *American Law and Economics Review*, *8*(2), 249–281. JSTOR.

Fryer, R. G., & Levitt, S. D. (2013). Testing for Racial Differences in the Mental Ability of Young Children. *American Economic Review*, *103*(2), 981–1005. https://doi.org/10.1257/aer.103.2.981

Gangopadhyay, K., & Sarkar, A. (2014). *Private Investment in Education across Castes and Religion: Evidence from West Bengal*. 26.

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies. *American Political Science Review*, *105*(4), 765–789. https://doi.org/10.1017/S0003055411000414

Jones, L. V., Burton, N. W., & Davenport, E. C. (1984). Monitoring the Mathematics Achievement of Black Students. *Journal for Research in Mathematics Education*, *15*(2), 154–164. https://doi.org/10.2307/748891

Kijima, Y. (2006). Caste and Tribe Inequality: Evidence from India, 1983-1999. *Economic Development and Cultural Change*, *54*(2), 369–404.

Kulic, N., Skopek, J., Triventi, M., & Blossfeld, H.-P. (2019). Social Background and Children's Cognitive Skills: The Role of Early Childhood Education and Care in a Cross-National Perspective. *Annual Review of Sociology*, *45*(1), 557–579. https://doi.org/10.1146/annurev-soc-073018-022401

Lord, F. M. (1975). The 'ability' scale in item characteristic curve theory. *Psychometrika*, *40*(2), 205–217. https://doi.org/10.1007/BF02291567

Madden, D. (2000). Towards a broader explanation of male-female wage differences. *Applied Economics Letters*, *7*(12), 765–770. https://doi.org/10.1080/135048500444769

Mondal, S. R. (2003). Social Structure, OBCs and Muslims. *Economic and Political Weekly*, *38*(46), 4892–4897.

Munshi, K. (2019). Caste and the Indian Economy. *Journal of Economic Literature*, *57*(4), 781–834. https://doi.org/10.1257/jel.20171307

Munshi, K., & Rosenzweig, M. (2006). Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy. *American Economic Review*, *96*(4), 1225–1252. https://doi.org/10.1257/aer.96.4.1225

Nair, A. (2009). *Disadvantaged at Birth?*

Neal, D. (2006). Chapter 9 Why Has Black–White Skill Convergence Stopped? In *Handbook of the Economics of Education* (Vol. 1, pp. 511–576). Elsevier. https://doi.org/10.1016/S1574-0692(06)01009-9

Ospino, C. G., Roldan Vasquez, P., & Barraza Narvaez, N. (2010). Oaxaca-Blinder wage decomposition: Methods, critiques and applications. A literature review. *Revista de Economía Del Caribe*, *5*, 237–274.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204. https://doi.org/10.1080/07350015.2016.1227711

Penney, J. (2017). Test Score Measurement and the Black-White Test Score Gap. *The Review of Economics and Statistics*, *99*(4), 652–656. https://doi.org/10.1162/REST_a_00651

Rawal, S., & Kingdon, G. (2010). Akin to my teacher: Does caste, religious or gender distance between student and teacher matter? Some evidence from India. In *DoQSS Working Papers* (No. 10–18; DoQSS Working Papers). Department of Quantitative Social Science - UCL Institute of Education, University College London. https://ideas.repec.org/p/qss/dqsswp/1018.html

Rubio-Codina, M., Attanasio, O., & Grantham-McGregor, S. (2016). Mediating pathways in the socio-economic gradient of child development: Evidence from children 6–42 months in Bogota. *International Journal of Behavioral Development*, *40*(6), 483–491. https://doi.org/10.1177/0165025415626515

Sankaran, S., Sekerdej, M., & von Hecker, U. (2017). The Role of Indian Caste Identity and Caste Inconsistent Norms on Status Representation. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00487

Simons, R. L., Murry, V., Mcloyd, V., Lin, K.-H., Cutrona, C., & Conger, R. D. (2002). Discrimination, crime, ethnic identity, and parenting as correlates of depressive symptoms among African American children: A multilevel analysis. *Development and Psychopathology*, *14*(2), 371–393. https://doi.org/10.1017/S0954579402002109

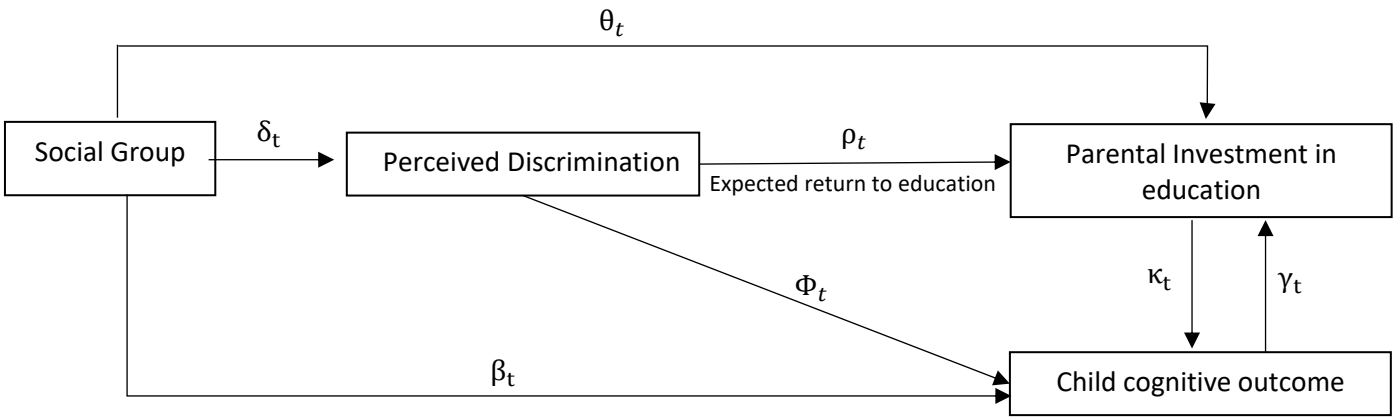
Singh, M., & Husain, Z. (2016). *Self-fulfilling Equilibrium and Social Disparities in Urban India*. *48*, 9.

Stone, S., & Han, M. (2005). Perceived school environments, perceived discrimination, and school performance among children of Mexican immigrants. *Children and Youth Services Review*, *27*(1), 51–66. https://doi.org/10.1016/j.childyouth.2004.08.011

Theimer, C. E., Killen, M., & Stangor, C. (2001). Young children's evaluations of exclusion in gender-stereotypic peer contexts. *Developmental Psychology*, *37*(1), 18. https://doi.org/10.1037/0012-1649.37.1.18

Thorat, S., & Newman, K. S. (2007). Caste and Economic Discrimination: Causes, Consequences and Remedies. *Economic and Political Weekly*, *42*(41), 4121–4124. JSTOR.

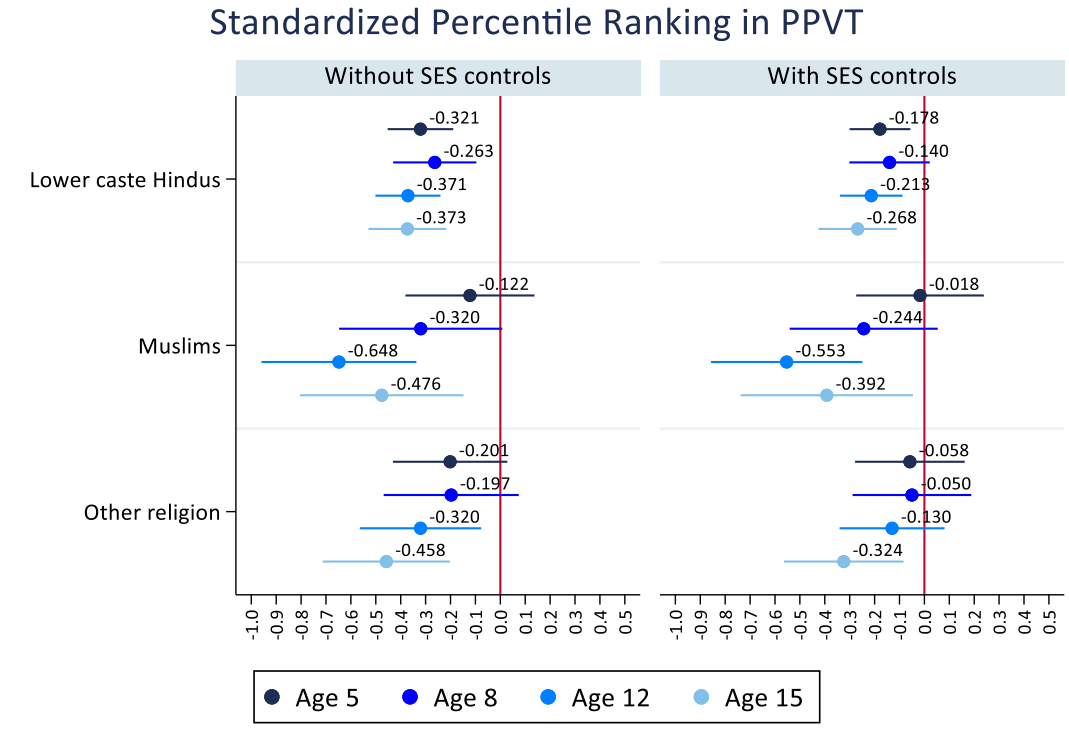
Figure 1: Plausible pathways between the variables of interest



Notes- This figure demonstrates the plausible pathways between the variables of interest.

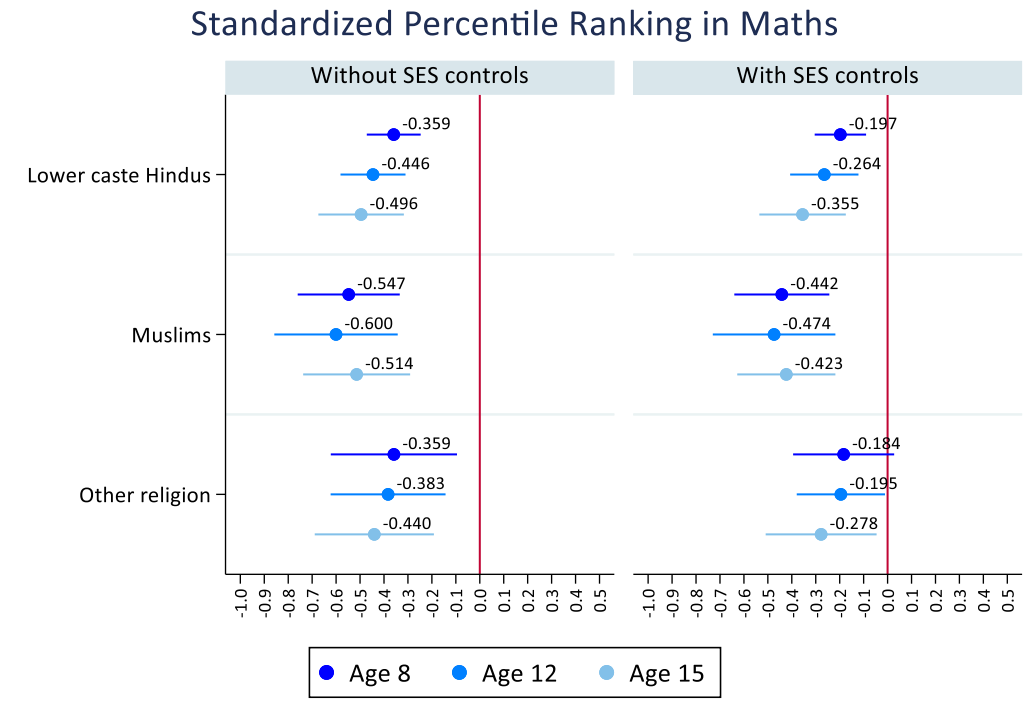
Figure 2: Gaps in test scores across social groups

Figure 2A: Gaps in standardized percentile ranking in PPVT test



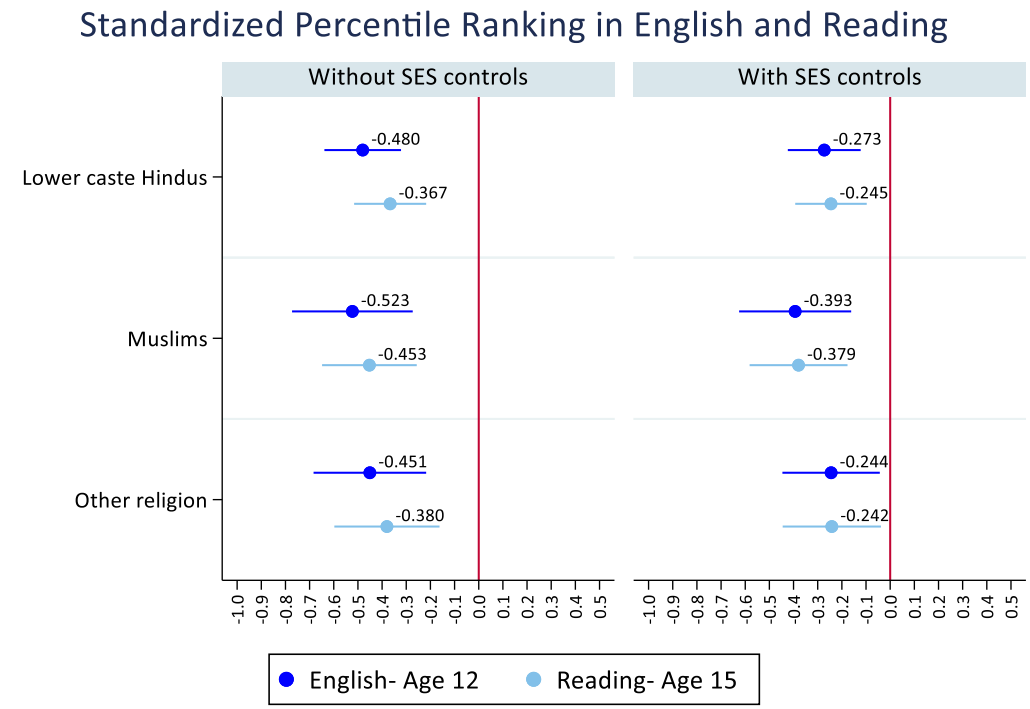
Omitted caste- Upper caste Hindus

Figure 2B: Gaps in standardized percentile ranking in Math test



Omitted caste- Upper caste Hindus

Figure 2C: Gaps in standardized percentile ranking in English and Reading tests



Omitted caste- Upper caste Hindus

Figure 3: Association Between Discrimination & Test Scores

Figure 3A: Standardized PPVT percentile ranking

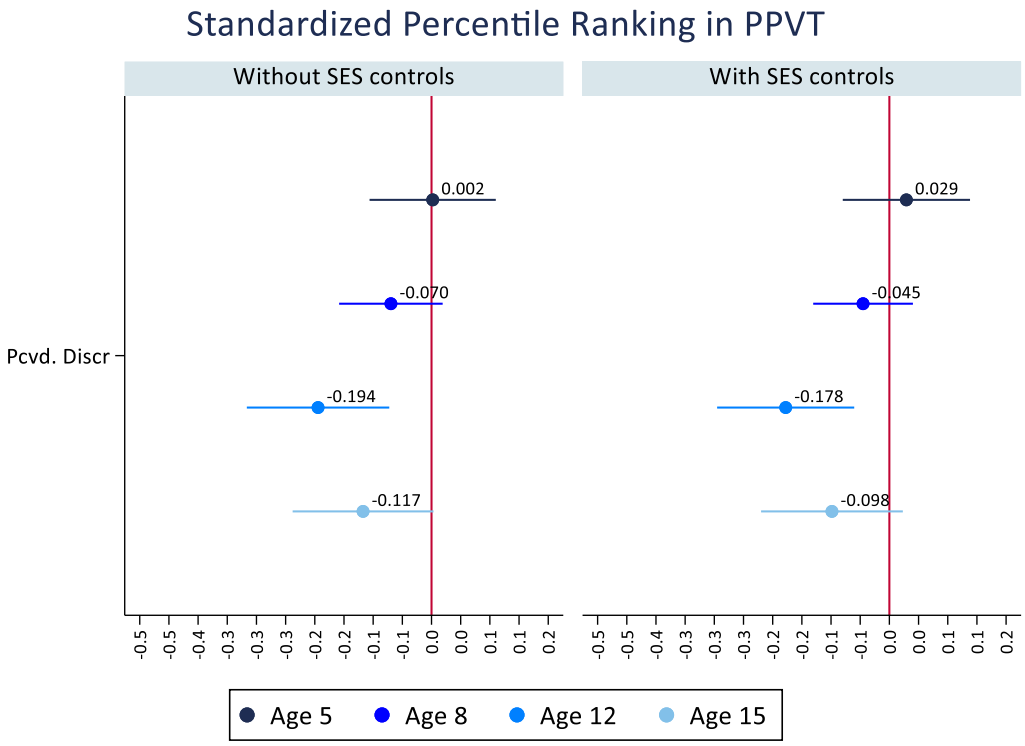


Figure 3B: Standardized Math percentile ranking

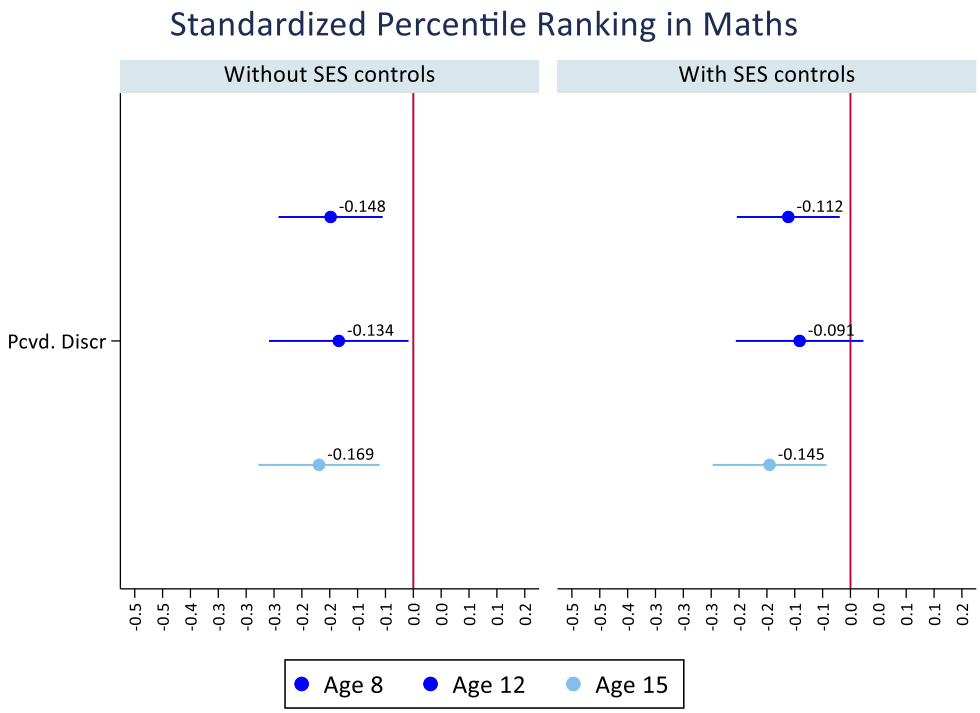


Figure 3C: Standardized English and Reading percentile ranking

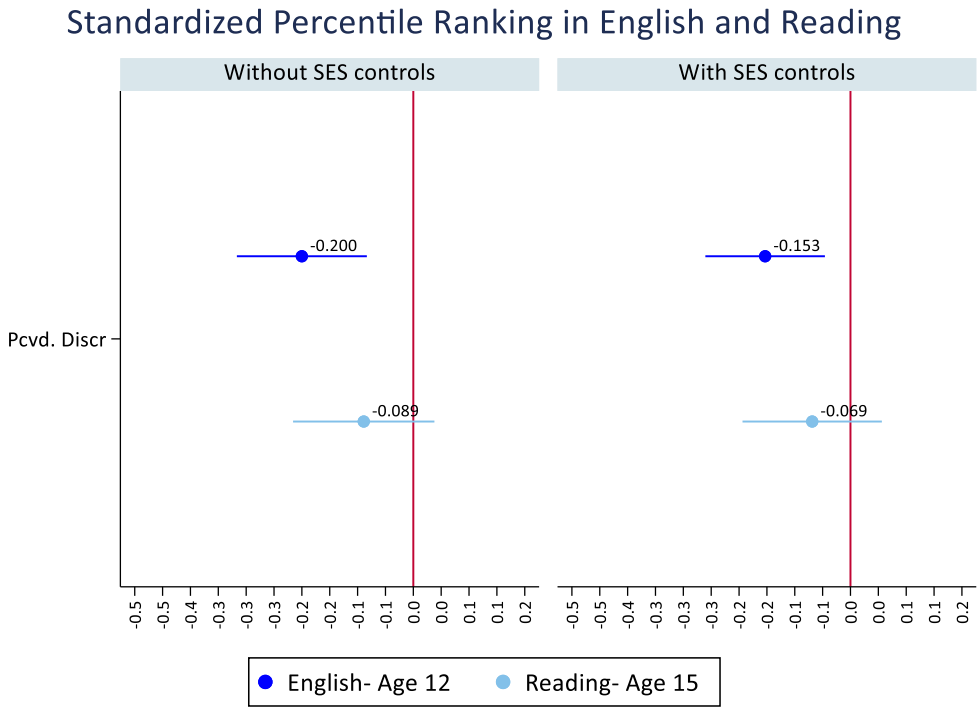


Figure 4: Differential association between perceived discrimination & test scores across social groups

Figure 4A: Standardized PPVT percentile ranking

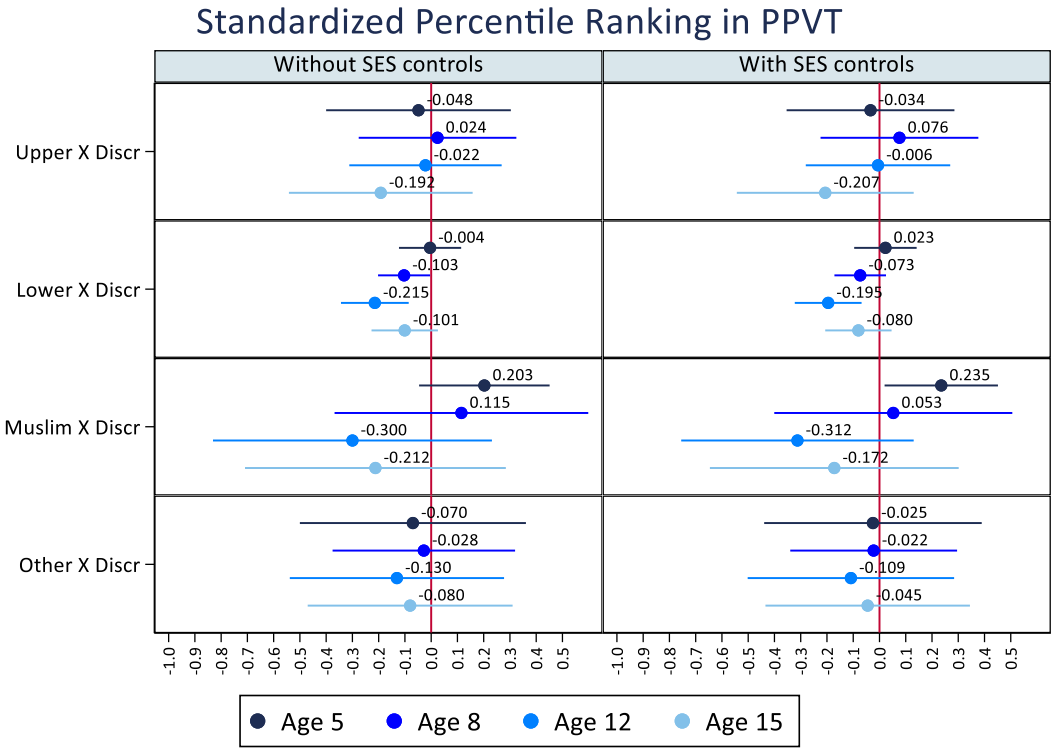


Figure 4B: Standardized Math percentile ranking

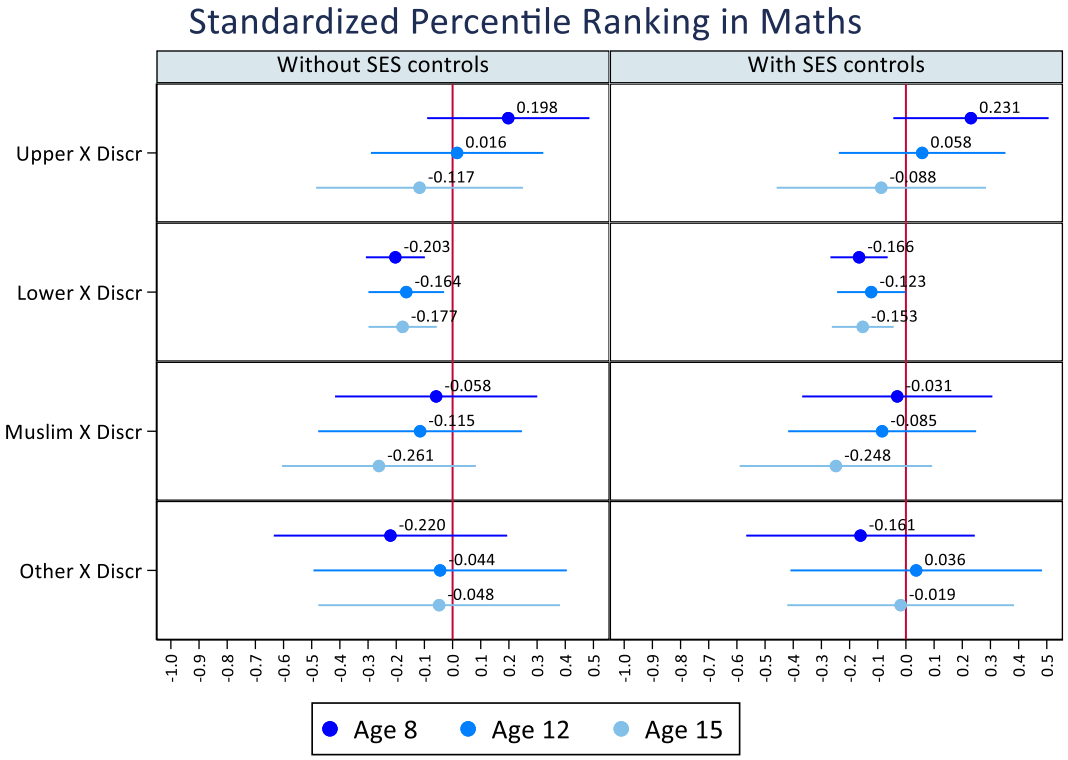


Figure 4C: Standardized English and Reading percentile ranking

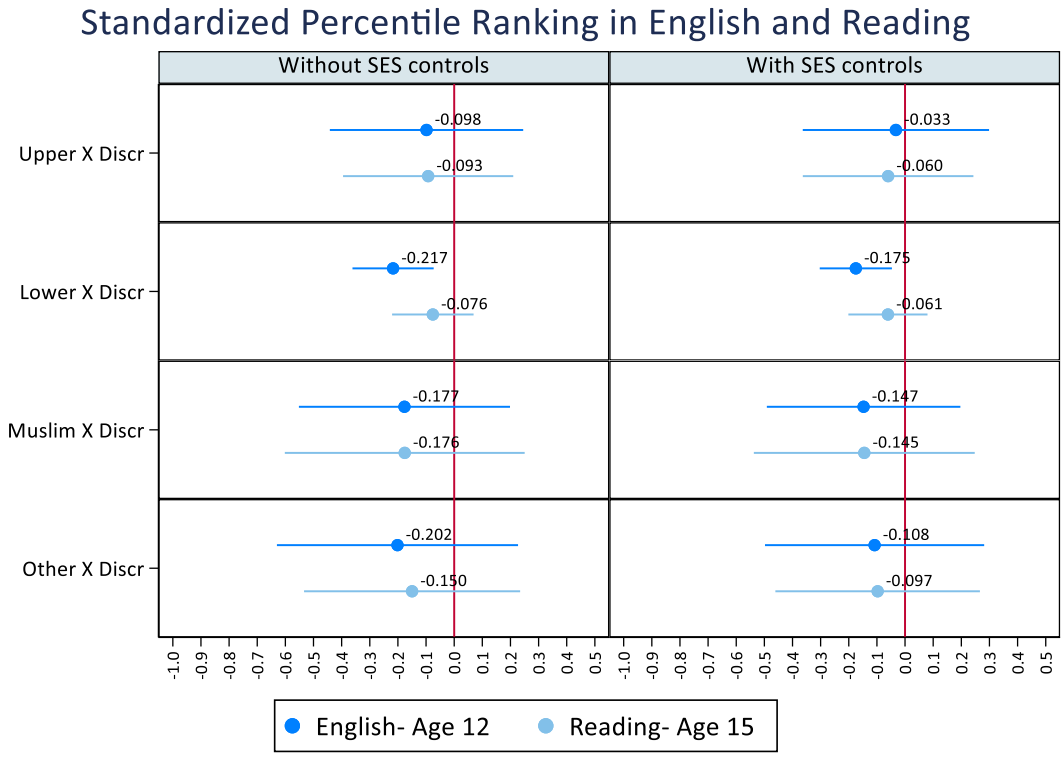
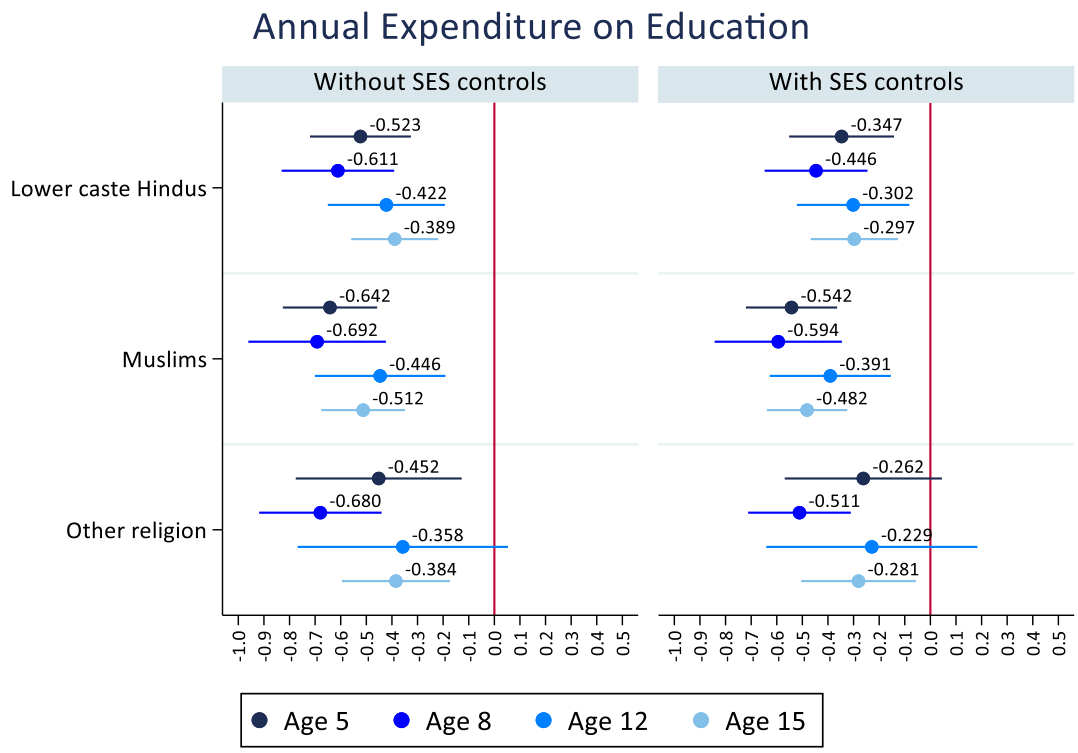


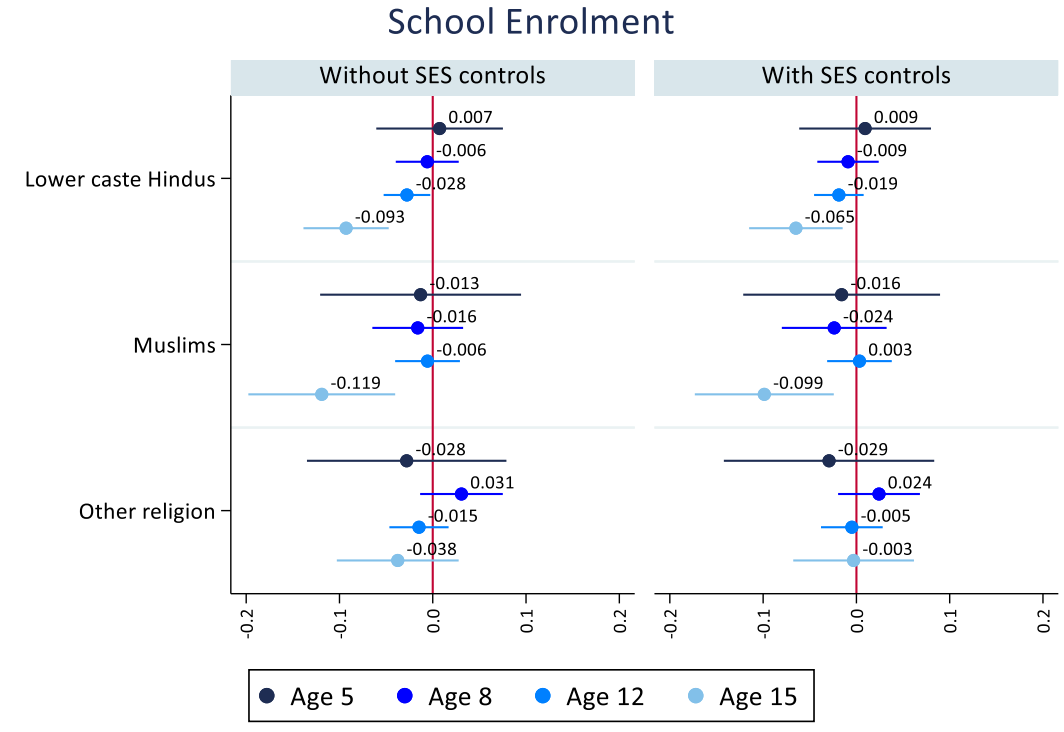
Figure 5: Gaps in parental investment across social groups

Figure 5A: Annual expenditure on children's education



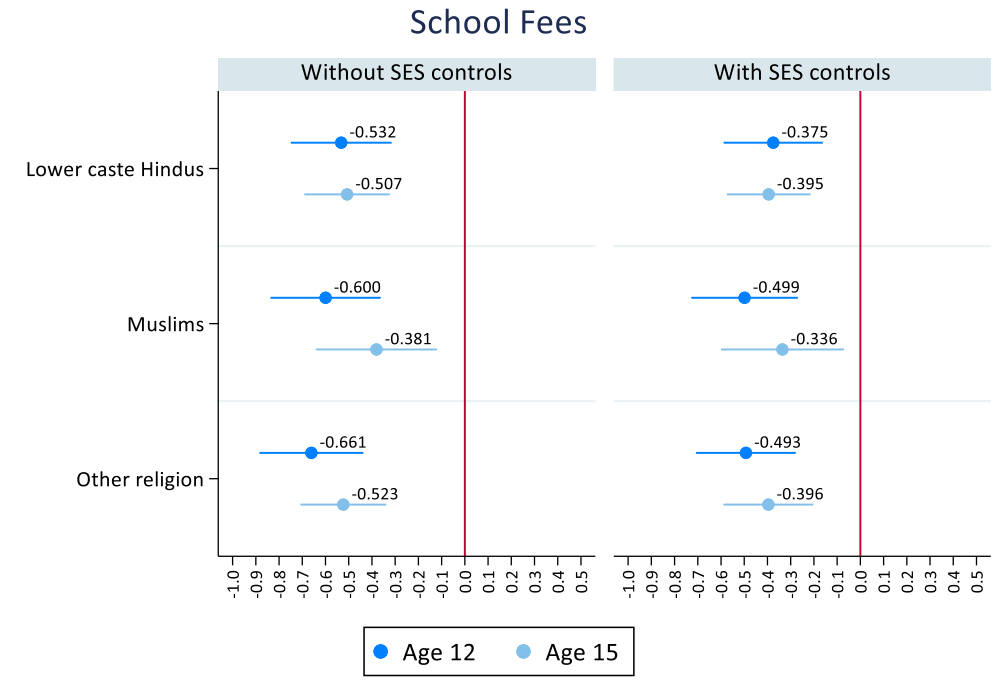
Omitted caste- Upper caste Hindus

Figure 5B: Enrolment at school



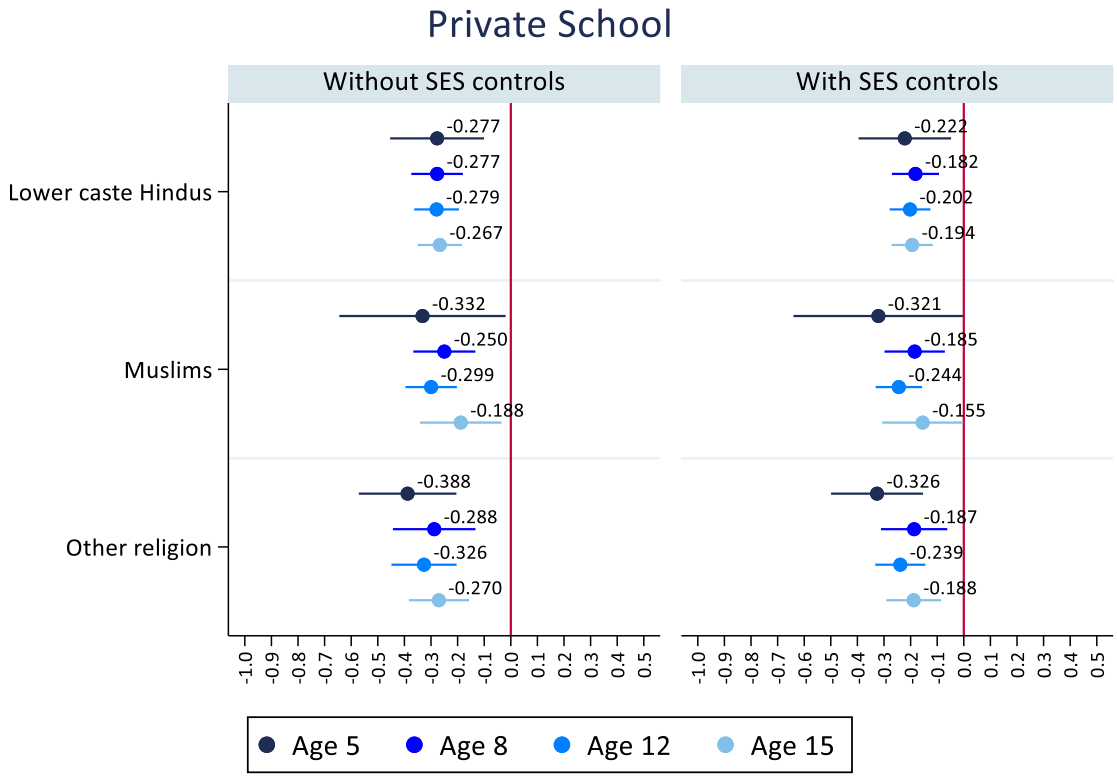
Omitted caste- Upper caste Hindus

Figure 5C: Expenditure on school fees



Omitted caste- Upper caste Hindus

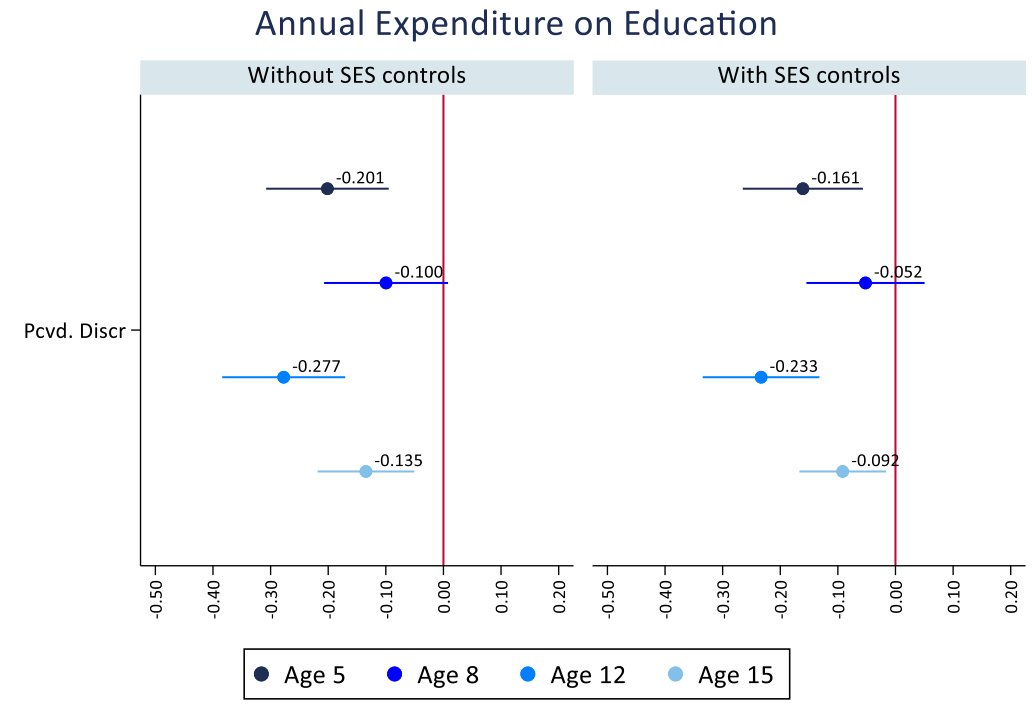
Figure 5D: Attendance at private school



Omitted caste- Upper caste Hindus

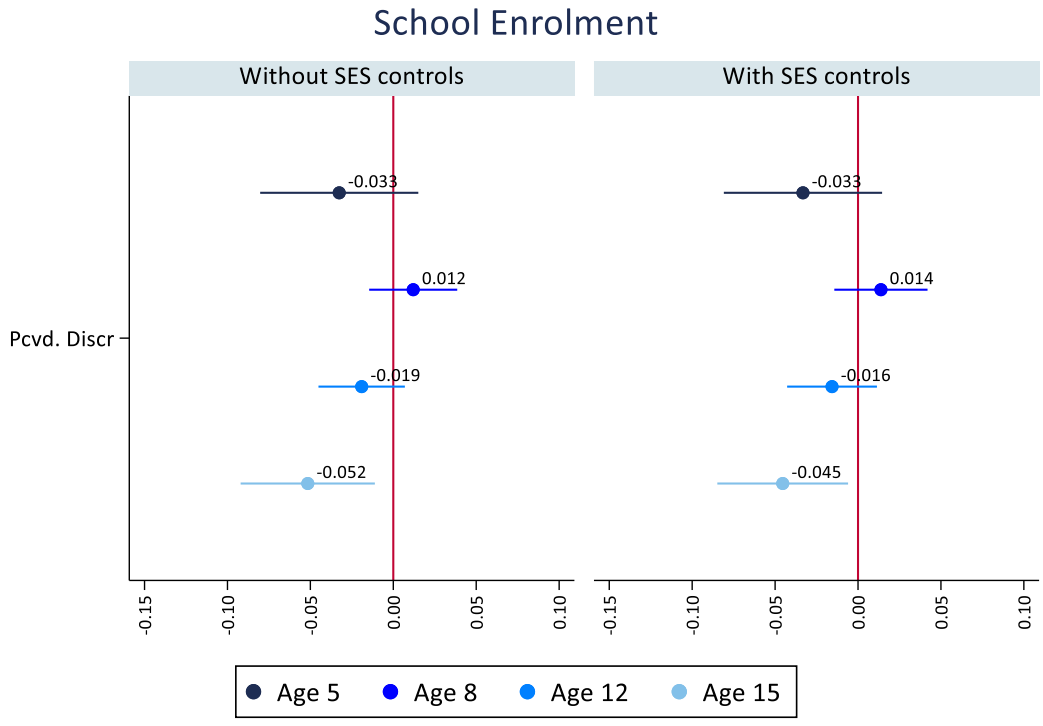
Figure 6: Association between perceived discrimination and parental investment

Figure 6A: : Annual expenditure on children's education



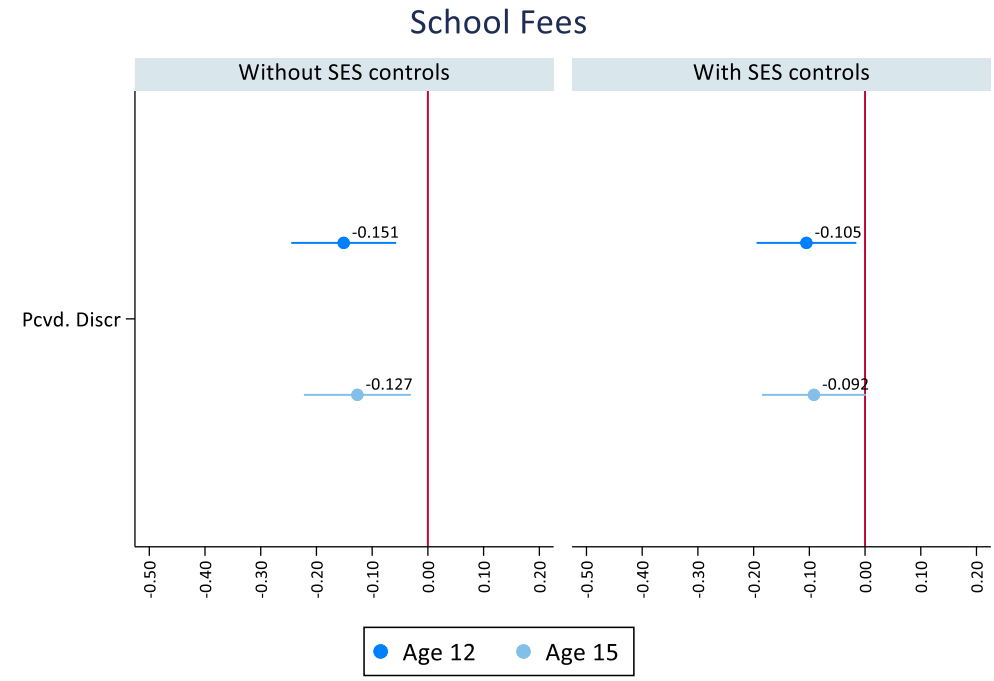
Omitted caste- Upper caste Hindus

Figure 6B: Enrolment at school



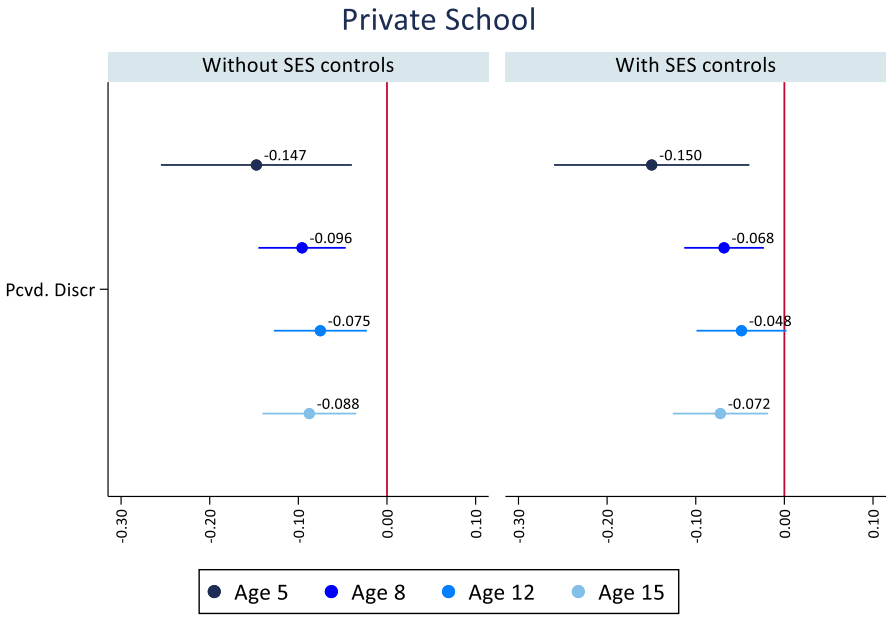
Omitted caste- Upper caste Hindus

Figure 6C: Expenditure on school fees



Omitted caste- Upper caste Hindus

Figure 6D: Attendance at private school



Omitted caste- Upper caste Hindus

Figure 7: Differential association between perceived discrimination and parental investment across social groups

Figure 7A: : Annual expenditure on children's education

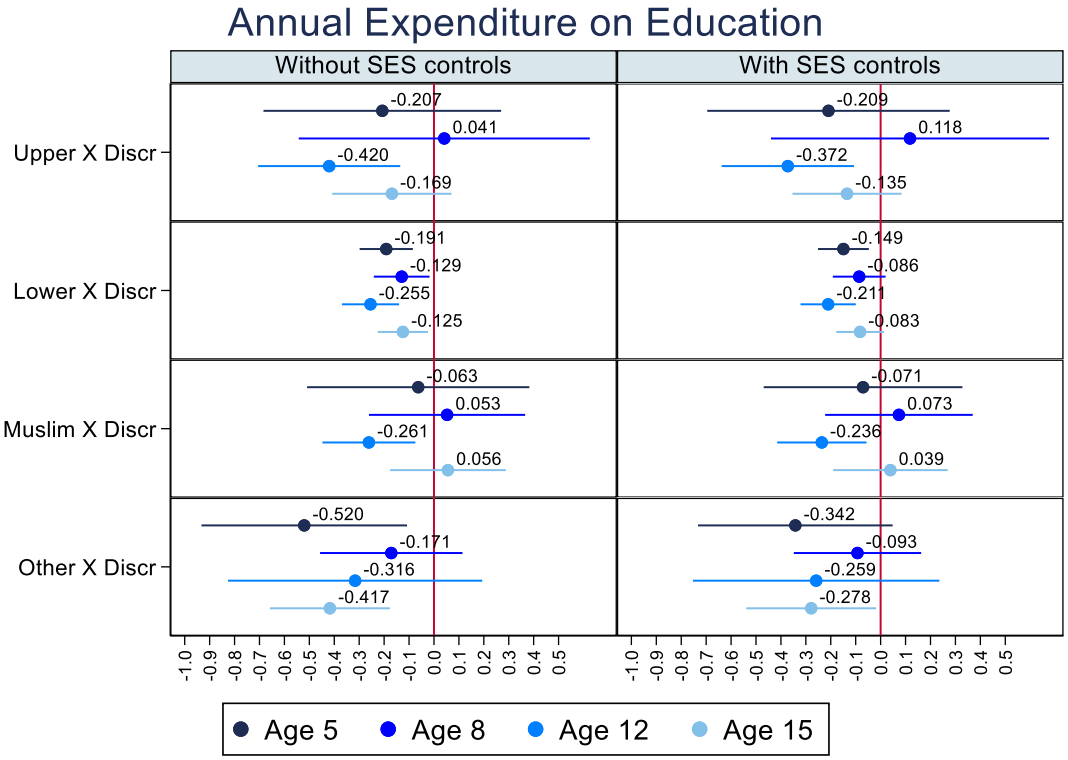


Figure 7B: Enrolment at school

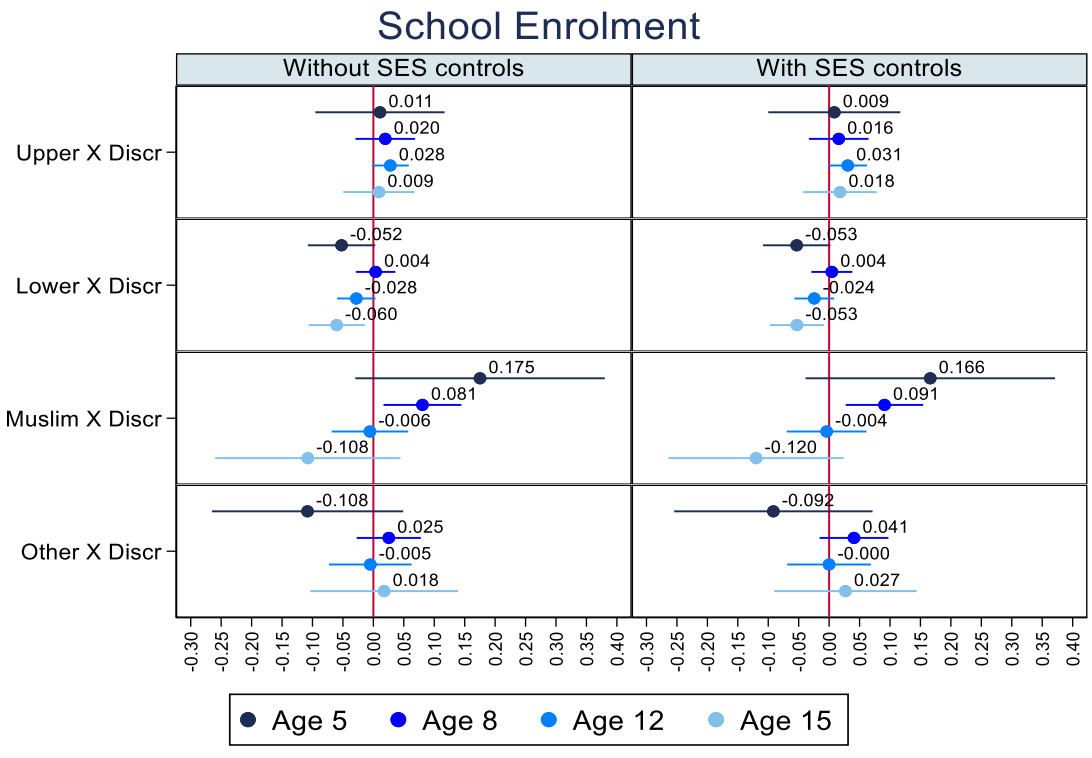


Figure 7C: Expenditure on school fees

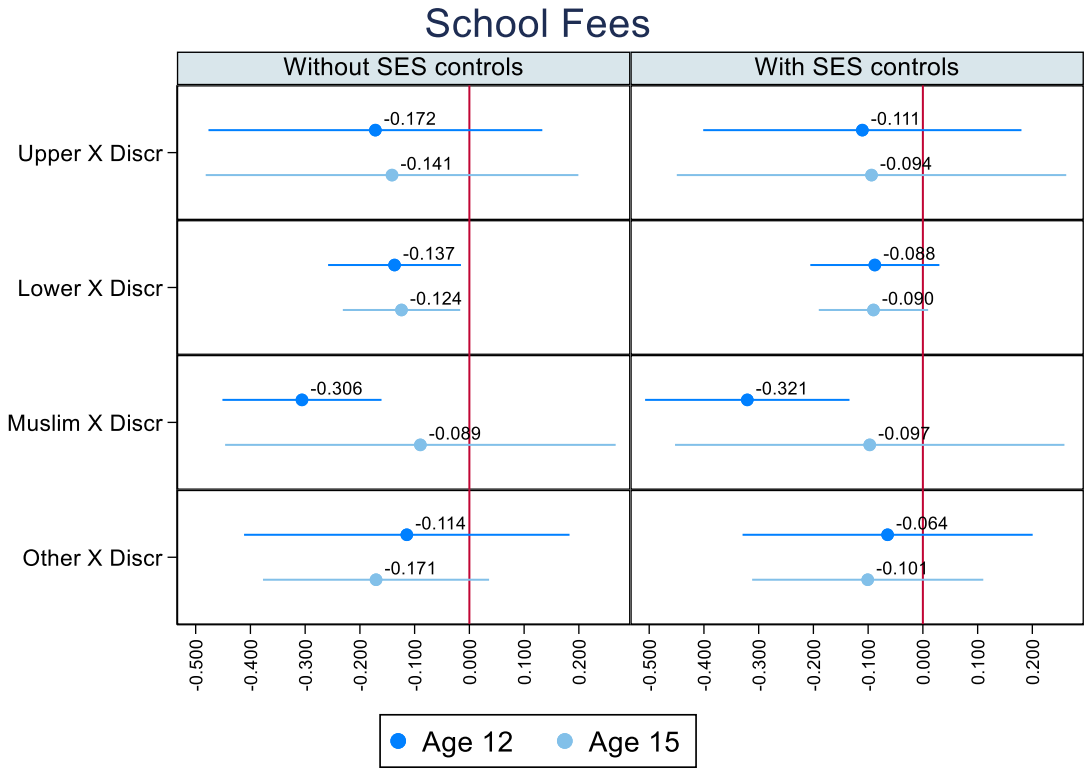


Figure 7D: Attendance at private school

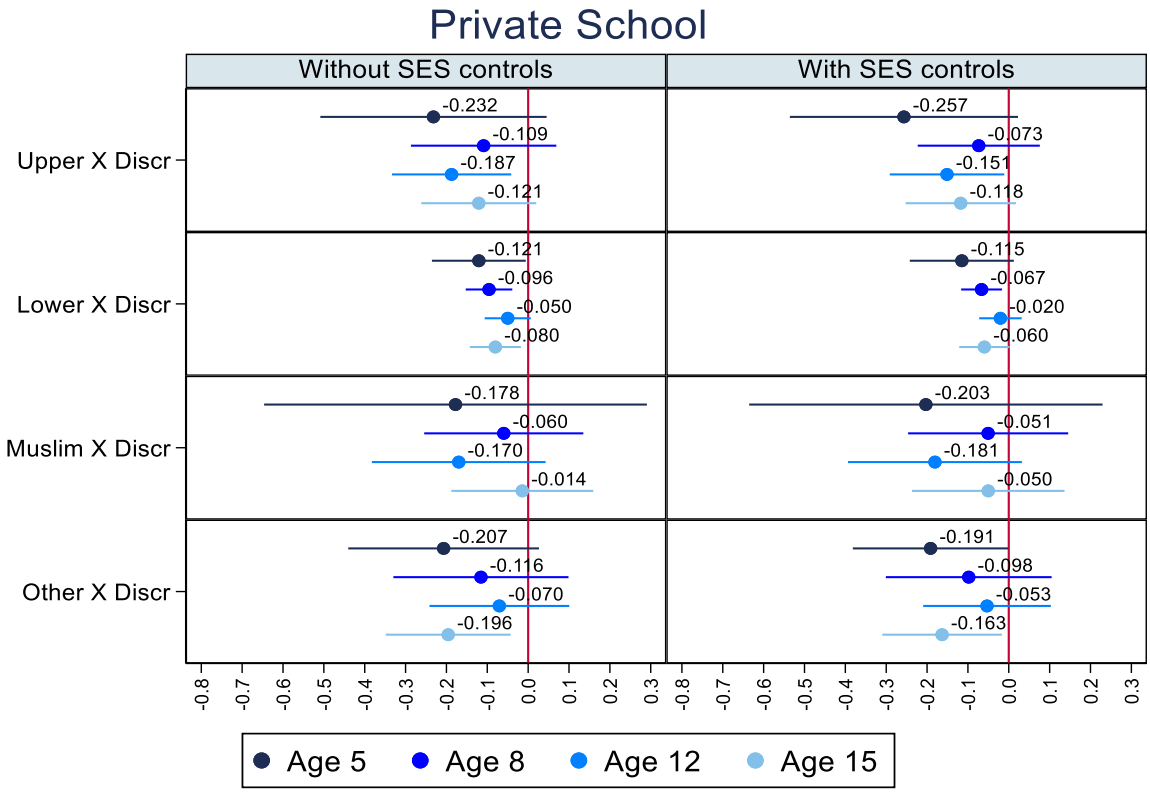


Table 1: Descriptive statistics of the sample at baseline (round 1)

Variable	Mean	Std. Dev.	Min	Max
Gender= boy child	0.537	0.499	0	1
Region= urban	0.253	0.434	0	1
Lower-caste Hindu (SC, ST and BC)	0.726	0.446	0	1
Upper-caste Hindu	0.143	0.349	0	1
Muslim	0.074	0.262	0	1
Other religions	0.057	0.231	0	1
Age in months	11.822	3.492	5	21
Household size	5.422	2.356	2	22
Mother's age	23.681	4.331	12	48
Mother's years of education	2.982	4.195	0	14
No. of children alive	1.892	0.999	1	8
Order of birth	1.979	1.116	1	11

Notes- This table presents descriptive statistics of the sample at baseline.

Table 2: Descriptive statistics across rounds

Variables	R1	R2	R3	R4	R5
Underweight	0.329	0.444	0.459	.	.
Stunting	0.307	0.357	0.289	0.292	0.278
Wasted/Thinness	0.188	0.187	0.273	0.330	0.255
Height for age Z-score	-1.337	-1.644	-1.425	-1.462	-1.468
Weight for age Z-score	-1.548	-1.866	-1.869	.	.
BMI for age Z-score	-1.038	-1.176	-1.414	-1.354	-1.138
Wealth index	0.408	0.459	0.514	0.585	0.633
Housing quality index	0.494	0.541	0.582	0.679	0.693
Access to services index	0.551	0.611	0.645	0.700	0.789
Consumer durables index	0.178	0.226	0.315	0.376	0.417
Access to electricity	0.821	0.898	0.965	0.976	0.980
Access to sanitation	0.299	0.327	0.347	0.407	0.497
Access to safe drinking water	0.838	0.949	0.967	0.989	0.992
Access to adequate fuels for cooking	0.245	0.270	0.299	0.428	0.686
Hours spent in paid activity per day	.	0.001	0.009	0.060	0.484
Hours of household chore per day	.	0.056	0.334	0.859	1.190
Hours at school per day	.	5.743	7.667	7.996	7.824
Hours of study per day	.	1.041	1.833	1.916	2.113
School Enrolment	.	0.218	0.935	0.963	0.878

Notes: This table provides descriptive statistics of the sample across YL rounds.

Table 3: Summary statistics for children belonging to various social groups

Sample Baseline	Full	Lower Hindu	Other Hindu	Muslim	Other religion
Perceived Discrimination	0.329	0.349	0.237	0.296	0.339
Height for age Z-score	-1.337	-1.458	-0.861	-1.133	-1.272
Wealth Index	0.408	0.364	0.538	0.586	0.405
Mother's years of education	2.982	2.169	6.049	5.013	3.009
Father's years of education	4.439	3.731	7.445	5.705	4.291
Stunting	0.307	0.343	0.168	0.216	0.327
Underweight	0.329	0.358	0.210	0.295	0.304
School enrolment at age 5	0.219	0.223	0.232	0.134	0.226

Notes: This table provides descriptive statistics of variables at the baseline, for all social groups. All the variables are from round 1, except for school enrolment which is reported for round 2.

Table 4: Perceived discrimination across social groups

VARIABLES	(1)	(2)
Caste (Base- Upper Hindu)		
SC/ST/BC Hindu	0.079**	0.054*
	(0.028)	(0.027)
Muslim	0.055	0.043
	(0.063)	(0.063)
Other religion	0.128**	0.102*
	(0.060)	(0.058)
Observations	1,867	1,867
R-squared	0.178	0.186
Income and Wealth controls	No	Yes
Location FE	Yes	Yes

*Notes: This table reports the estimates from regression of discrimination on castes. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table 5: Test score percentile ranking gaps across SCs, STs, and OBCs

	PPVT				Maths			English	Reading
	Age 5	Age 8	Age 12	Age 15	Age 8	Age 12	Age 15	Age 12	Age 15
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Test score percentile gaps within lower Hindu castes									
Lower Caste (Base- OBC)									
SC Hindu	-0.053	-0.025	0.084	-0.051	-0.048	-0.051	-0.130	-0.044	-0.091
	(0.075)	(0.076)	(0.091)	(0.082)	(0.083)	(0.092)	(0.083)	(0.085)	(0.084)
ST Hindu	-0.099	-0.059	-0.078	0.020	-0.137	-0.098	-0.090	-0.047	-0.161
	(0.144)	(0.115)	(0.104)	(0.115)	(0.093)	(0.121)	(0.121)	(0.119)	(0.114)
Observations	1,348	1,382	1,385	1,373	1,384	1,345	1,331	1,348	1,323
SES controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel b: Effect of perceived discrimination on test score percentiles within lower Hindu castes									
Lower Caste (Base- OBC)									
SC Hindu	-0.045	-0.035	0.096	-0.047	-0.042	-0.052	-0.120	-0.056	-0.094
	(0.076)	(0.078)	(0.088)	(0.080)	(0.082)	(0.091)	(0.080)	(0.086)	(0.084)
ST Hindu	-0.111	-0.054	-0.047	0.034	-0.133	-0.080	-0.061	-0.049	-0.149
	(0.144)	(0.117)	(0.101)	(0.118)	(0.091)	(0.123)	(0.117)	(0.125)	(0.117)
Perceived social discrimination	0.020	-0.081	-0.206***	-0.082	-0.162***	-0.108*	-0.144***	-0.154**	-0.056
	(0.063)	(0.052)	(0.064)	(0.065)	(0.052)	(0.058)	(0.054)	(0.065)	(0.072)
Observations	1,337	1,369	1,372	1,361	1,371	1,332	1,320	1,336	1,312
SES controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table 6: Parental investment within lower Hindu castes

Table 6a: Differences in parental investment across SCs, STs, and OBCs

	Expenditure on education				School Fees	
	Age 5	Age 8	Age 12	Age 15	Age 12	Age 15
	(1)	(2)	(3)	(4)	(5)	(6)
Lower Caste (Base- OBC)						
SC Hindu	-0.104*	-0.209***	-0.068	-0.063	-0.163***	-0.163***
	(0.059)	(0.057)	(0.089)	(0.058)	(0.056)	(0.049)
ST Hindu	-0.106	-0.171**	0.020	0.209	-0.019	-0.052
	(0.078)	(0.081)	(0.080)	(0.199)	(0.078)	(0.058)
Observations	1,082	1,370	1,299	1,263	1,329	1,178
SES controls	Yes	Yes	Yes	Yes	Yes	Yes

	Enrolment at school				Private School			
	Age 5	Age 8	Age 12	Age 15	Age 5	Age 8	Age 12	Age 15
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lower Caste (Base- OBC)								
SC Hindu	-0.033	0.053**	0.005	0.042	-0.124*	-0.073**	-0.129***	-0.175***
	(0.034)	(0.025)	(0.017)	(0.030)	(0.063)	(0.034)	(0.030)	(0.039)
ST Hindu	-0.021	-0.016	-0.005	-0.018	0.019	-0.094	-0.049	-0.104*
	(0.045)	(0.024)	(0.025)	(0.033)	(0.123)	(0.058)	(0.039)	(0.055)
Observations	1,404	1,404	1,397	1,381	313	1,302	1,329	1,179
SES controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects and SES. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table 6b: Association between discrimination and parental investment for lower caste Hindus

	Expenditure on education				School Fees	
	Age 5	Age 8	Age 12	Age 15	Age 12	Age 15
	(1)	(2)	(3)	(4)	(5)	(6)
Lower Caste (Base- OBC)						
SC Hindu	-0.106*	-0.217***	-0.068	-0.060	-0.166***	-0.160***
	(0.058)	(0.057)	(0.086)	(0.060)	(0.057)	(0.050)
ST Hindu	-0.102	-0.154*	0.034	0.211	-0.017	-0.051
	(0.077)	(0.081)	(0.082)	(0.205)	(0.083)	(0.057)
Perceived social discrimination	-0.136***	-0.088*	-0.189***	-0.079	-0.075	-0.078
	(0.050)	(0.051)	(0.058)	(0.049)	(0.059)	(0.051)
Observations	1,071	1,358	1,289	1,253	1,317	1,169
SES controls	Yes	Yes	Yes	Yes	Yes	Yes

	Enrolment at school				Private School			
	Age 5	Age 8	Age 12	Age 15	Age 5	Age 8	Age 12	Age 15
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lower Caste (Base- OBC)								
SC Hindu	-0.034	0.050**	0.008	0.040	-0.113**	-0.072**	-0.131***	-0.176***
	(0.034)	(0.025)	(0.017)	(0.029)	(0.056)	(0.035)	(0.030)	(0.039)
ST Hindu	-0.025	-0.027	0.006	-0.013	0.023	-0.091	-0.050	-0.100*
	(0.042)	(0.026)	(0.021)	(0.032)	(0.123)	(0.062)	(0.044)	(0.053)
Perceived social discrimination	-0.054*	0.006	-0.026	-0.055**	-0.120*	-0.061**	-0.018	-0.055*
	(0.029)	(0.018)	(0.017)	(0.024)	(0.061)	(0.025)	(0.027)	(0.030)
Observations	1,391	1,391	1,384	1,369	310	1,292	1,317	1,170
SES controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects and SES. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table 7: Analysis on the sub-sample of children who did not migrate over time

Panel A: Gaps in test scores across castes

	Standardized PPVT percentiles							
	Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Caste (Base- Upper Hindu)								
SC/ST/BC Hindu	-0.389***	-0.229***	-0.298***	-0.171**	-0.387***	-0.199**	-0.403***	-0.273***
	(0.062)	(0.057)	(0.084)	(0.084)	(0.086)	(0.079)	(0.087)	(0.087)
Muslim	-0.243**	-0.048	-0.350*	-0.237	-0.756***	-0.581***	-0.505***	-0.349**
	(0.121)	(0.114)	(0.178)	(0.166)	(0.155)	(0.161)	(0.168)	(0.169)
Other religion	-0.195*	-0.063	-0.179	-0.065	-0.278**	-0.097	-0.442***	-0.313**
	(0.115)	(0.112)	(0.127)	(0.118)	(0.135)	(0.124)	(0.124)	(0.120)
Observations	1,479	1,464	1,531	1,514	1,548	1,531	1,543	1,526
SES controls	No	Yes	No	Yes	No	Yes	No	Yes

Standardized Percentiles	**Maths**						**English**		**Reading**	
	Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Caste (Base- Upper Hindu)										
SC/ST/BC Hindu	-0.413***	-0.231***	-0.441***	-0.245***	-0.496***	-0.336***	-0.491***	-0.270***	-0.382***	-0.244***
	(0.066)	(0.065)	(0.075)	(0.075)	(0.084)	(0.082)	(0.080)	(0.070)	(0.079)	(0.077)
Muslim	-0.595***	-0.406***	-0.705***	-0.497***	-0.667***	-0.513***	-0.628***	-0.392***	-0.670***	-0.529***
	(0.118)	(0.113)	(0.115)	(0.114)	(0.147)	(0.133)	(0.144)	(0.127)	(0.134)	(0.127)
Other religion	-0.356**	-0.204*	-0.346***	-0.176	-0.410***	-0.263**	-0.371***	-0.185	-0.401***	-0.290**
	(0.137)	(0.117)	(0.128)	(0.108)	(0.127)	(0.118)	(0.123)	(0.111)	(0.123)	(0.117)
Observations	1,536	1,519	1,510	1,493	1,503	1,486	1,514	1,497	1,498	1,482
SES controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Panel B: Gaps in parental investment across castes

	Expenditure on Education							
	Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Caste (Base- Upper Hindu)								
SC/ST/BC Hindu	-0.387***	-0.211*	-0.561***	-0.409***	-0.307**	-0.187	-0.348***	-0.262***
	(0.103)	(0.108)	(0.124)	(0.114)	(0.122)	(0.122)	(0.093)	(0.096)
Muslim	-0.592***	-0.467***	-0.720***	-0.592***	-0.338*	-0.248	-0.521***	-0.455***
	(0.149)	(0.133)	(0.149)	(0.132)	(0.182)	(0.169)	(0.143)	(0.137)
Other religion	-0.380***	-0.193	-0.564***	-0.407***	-0.183	-0.052	-0.325***	-0.218*
	(0.128)	(0.127)	(0.129)	(0.120)	(0.237)	(0.236)	(0.110)	(0.116)
Observations	1,202	1,185	1,531	1,513	1,460	1,442	1,449	1,431
SES controls	No	Yes	No	Yes	No	Yes	No	Yes

	School Fees				Enrolment at school							
	Age 12 (Round 4)		Age 15 (Round 5)		Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Caste (Base- Upper Hindu)												
SC/ST/BC Hindu	-0.407***	-0.246*	-0.412***	-0.306***	0.022	0.024	0.002	-0.007	-0.023	-0.017	-0.095***	-0.069**
	(0.135)	(0.135)	(0.099)	(0.098)	(0.037)	(0.038)	(0.018)	(0.019)	(0.016)	(0.017)	(0.028)	(0.029)
Muslim	-0.537***	-0.402***	-0.235	-0.158	0.002	-0.002	-0.009	-0.022	-0.005	0.002	-0.132**	-0.111**
	(0.150)	(0.142)	(0.277)	(0.280)	(0.055)	(0.053)	(0.025)	(0.031)	(0.024)	(0.025)	(0.051)	(0.051)
Other religion	-0.549***	-0.378***	-0.386***	-0.270***	-0.024	-0.032	0.029	0.017	-0.008	-0.002	-0.031	-0.009
	(0.116)	(0.114)	(0.102)	(0.103)	(0.056)	(0.059)	(0.025)	(0.024)	(0.016)	(0.017)	(0.034)	(0.035)
Observations	1,496	1,478	1,360	1,345	1,559	1,541	1,559	1,541	1,559	1,541	1,553	1,535
SES controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

	Type of School (Private School)							
	Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Caste (Base- Upper Hindu)								
SC/ST/BC Hindu	-0.270***	-0.233***	-0.269***	-0.180***	-0.253***	-0.183***	-0.279***	-0.209***
	(0.087)	(0.080)	(0.054)	(0.048)	(0.050)	(0.046)	(0.045)	(0.041)
Muslim	-0.414**	-0.414**	-0.298***	-0.220***	-0.336***	-0.277***	-0.260***	-0.222**
	(0.188)	(0.198)	(0.072)	(0.068)	(0.067)	(0.059)	(0.084)	(0.089)
Other religion	-0.357***	-0.315***	-0.274***	-0.182***	-0.280***	-0.200***	-0.257***	-0.184***
	(0.095)	(0.086)	(0.079)	(0.067)	(0.065)	(0.053)	(0.063)	(0.059)
Constant	1.270***	1.249***	1.490***	1.508***	1.360***	1.336***	1.404***	1.320***
	(0.087)	(0.130)	(0.056)	(0.066)	(0.051)	(0.061)	(0.045)	(0.073)
Observations	344	338	1,465	1,448	1,496	1,478	1,361	1,346
R-squared	0.637	0.655	0.428	0.500	0.349	0.407	0.317	0.366
SES controls	No	Yes	No	Yes	No	Yes	No	Yes

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table 8: Robustness Check- Controlling for baseline PPVT test score

Panel A: Differences in test scores across social groups- controlling for baseline PPVT test scores

	PPVT				Maths		
VARIABLES	PPVT R2	PPVT R3	PPVT R4	PPVT R5	Maths R3	Maths R4	Maths R5
Base- Upper Hindu							
SC/ST/BC Hindu	-0.178***	-0.117	-0.174***	-0.231***	-0.167***	-0.236***	-0.361***
	(0.061)	(0.080)	(0.066)	(0.088)	(0.050)	(0.073)	(0.091)
Muslim	-0.018	-0.258	-0.577***	-0.402**	-0.438***	-0.476***	-0.464***
	(0.129)	(0.168)	(0.156)	(0.184)	(0.095)	(0.115)	(0.106)
Other religion	-0.058	-0.041	-0.102	-0.329***	-0.196*	-0.156*	-0.276**
	(0.111)	(0.123)	(0.106)	(0.119)	(0.103)	(0.090)	(0.113)
Observations	1,832	1,796	1,797	1,783	1,883	1,828	1,805
R-squared	0.301	0.320	0.333	0.247	0.395	0.394	0.399

Panel B: Effect of discrimination on test scores- controlling for baseline PPVT test scores

	PPVT				Maths		
	PPVT R2	PPVT R3	PPVT R4	PPVT R5	Maths R3	Maths R4	Maths R5
Perceived social discrimination	0.029	-0.040	-0.185***	-0.100	-0.112**	-0.086	-0.127**
	(0.055)	(0.041)	(0.056)	(0.062)	(0.046)	(0.057)	(0.052)
Base- Upper Hindu							
SC/ST/BC Hindu	-0.170***	-0.111	-0.181***	-0.229**	-0.160***	-0.226***	-0.359***
	(0.062)	(0.082)	(0.066)	(0.089)	(0.051)	(0.075)	(0.092)
Muslim	-0.012	-0.244	-0.590***	-0.406**	-0.425***	-0.466***	-0.463***
	(0.122)	(0.164)	(0.154)	(0.186)	(0.095)	(0.115)	(0.105)
Other religion	-0.053	-0.035	-0.097	-0.322***	-0.180*	-0.142	-0.272**
	(0.110)	(0.124)	(0.107)	(0.122)	(0.101)	(0.091)	(0.117)
Observations	1,815	1,780	1,781	1,768	1,784	1,741	1,725
R-squared	0.303	0.322	0.343	0.248	0.394	0.333	0.299

Panel C: Differential effect of discrimination on test scores across social groups- controlling for baseline PPVT test scores

	PPVT				Maths		
VARIABLES	PPVT R2	PPVT R3	PPVT R4	PPVT R5	Maths R3	Maths R4	Maths R5
Base- Upper Hindus × No Discr							
Lower Hindu × No Discr	-0.181***	-0.055	-0.126*	-0.244***	-0.045	-0.186**	-0.354***
	(0.069)	(0.091)	(0.074)	(0.086)	(0.057)	(0.082)	(0.095)
Muslim × No Discr	-0.080	-0.234	-0.501***	-0.389*	-0.366***	-0.429***	-0.418***
	(0.128)	(0.164)	(0.180)	(0.213)	(0.116)	(0.123)	(0.108)
Other religion × No Discr	-0.048	0.006	-0.081	-0.337**	-0.047	-0.161	-0.306***
	(0.125)	(0.139)	(0.131)	(0.150)	(0.103)	(0.117)	(0.113)
Upper Hindu × Discr	-0.034	0.164	0.038	-0.149	0.323**	0.060	-0.100
	(0.161)	(0.125)	(0.119)	(0.166)	(0.133)	(0.127)	(0.181)
Lower Hindu × Discr	-0.159*	-0.134	-0.334***	-0.330***	-0.218***	-0.301***	-0.481***
	(0.086)	(0.093)	(0.096)	(0.090)	(0.067)	(0.092)	(0.098)
Muslim × Discr	0.155	-0.165	-0.847***	-0.583**	-0.382**	-0.546***	-0.686***
	(0.153)	(0.263)	(0.198)	(0.242)	(0.147)	(0.160)	(0.167)
Other religion × Discr	-0.073	-0.024	-0.169	-0.425***	-0.266	-0.086	-0.313
	(0.183)	(0.166)	(0.158)	(0.143)	(0.191)	(0.174)	(0.204)
Observations	1,815	1,780	1,781	1,768	1,784	1,741	1,725
R-squared	0.304	0.323	0.345	0.248	0.399	0.334	0.300

*Notes: Regressions for rounds 3, 4, and 5 control for PPVT scores obtained at age 2 (round 2). All regressions include socioeconomic controls and location fixed effects. Standard errors are adjusted for clustering at the community level. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table 9: Robustness Check- Controlling for baseline PPVT test scores

Panel A: Differences in parental investment across social groups- controlling for baseline PPVT test scores

	Expenditure on education				School fees		Enrolment at school			
	EXP R2	EXP R3	EXP R4	EXP R5	FEES R4	FEES R5	ENROL R2	ENROL R3	ENROL R4	ENROL R5
Base- Upper Hindu										
SC/ST/BC Hindu	-0.347***	-0.396***	-0.288**	-0.288***	-0.375***	-0.391***	0.009	-0.001	-0.015	-0.060**
	(0.103)	(0.096)	(0.120)	(0.090)	(0.118)	(0.091)	(0.036)	(0.016)	(0.013)	(0.026)
Muslim	-0.542***	-0.578***	-0.411***	-0.502***	-0.539***	-0.353**	-0.016	-0.027	0.005	-0.099**
	(0.090)	(0.132)	(0.127)	(0.085)	(0.124)	(0.140)	(0.053)	(0.028)	(0.017)	(0.038)
Other religion	-0.262*	-0.466***	-0.215	-0.257**	-0.489***	-0.403***	-0.029	0.027	-0.004	-0.005
	(0.155)	(0.100)	(0.219)	(0.119)	(0.118)	(0.102)	(0.057)	(0.023)	(0.017)	(0.035)
Observations	1,463	1,781	1,682	1,656	1,754	1,586	1,908	1,815	1,804	1,788
R-squared	0.343	0.265	0.272	0.233	0.285	0.307	0.129	0.109	0.104	0.157

	Private school			
	R2	R3	R4	R5
Base- Upper Hindu				
SC/ST/BC Hindu	-0.222**	-0.167***	-0.225***	-0.203***
	(0.088)	(0.043)	(0.039)	(0.041)
Muslim	-0.321**	-0.189***	-0.279***	-0.178**
	(0.161)	(0.055)	(0.044)	(0.076)
Other religion	-0.326***	-0.167***	-0.251***	-0.185***
	(0.087)	(0.063)	(0.053)	(0.057)
Observations	416	1,704	1,736	1,565
R-squared	0.575	0.452	0.366	0.314

Panel B: Effect of discrimination on parental investment- controlling for baseline PPVT test scores

	Expenditure on education				School fees		Enrolment at school			
	EXP R2	EXP R3	EXP R4	EXP R5	FEES R4	FEES R5	ENROL R2	ENROL R3	ENROL R4	ENROL R5
Perceived social discrimination	-0.161***	-0.068	-0.256***	-0.086**	-0.106**	-0.088*	-0.033	0.015	-0.013	-0.044**
	(0.053)	(0.054)	(0.056)	(0.038)	(0.046)	(0.046)	(0.024)	(0.014)	(0.013)	(0.019)
Base- Upper Hindu										
SC/ST/BC Hindu	-0.350***	-0.397***	-0.279**	-0.227**	-0.337***	-0.320***	0.015	0.000	-0.014	-0.056**
	(0.105)	(0.097)	(0.120)	(0.087)	(0.119)	(0.100)	(0.036)	(0.016)	(0.013)	(0.026)
Muslim	-0.550***	-0.585***	-0.404***	-0.422***	-0.501***	-0.258	-0.009	-0.027	0.005	-0.096**
	(0.093)	(0.135)	(0.130)	(0.109)	(0.120)	(0.226)	(0.056)	(0.028)	(0.017)	(0.038)
Other religion	-0.263*	-0.465***	-0.195	-0.201*	-0.456***	-0.337***	-0.018	0.027	-0.002	0.002
	(0.157)	(0.100)	(0.219)	(0.119)	(0.119)	(0.114)	(0.057)	(0.023)	(0.017)	(0.036)
Observations	1,447	1,766	1,669	1,643	1,738	1,573	1,888	1,799	1,788	1,773
R-squared	0.351	0.270	0.273	0.230	0.278	0.301	0.130	0.107	0.108	0.160

	Private school			
	R2	R3	R4	R5
Perceived social discrimination	-0.150***	-0.082***	-0.045	-0.069**
	(0.055)	(0.023)	(0.028)	(0.027)
Base- Upper Hindu				
SC/ST/BC Hindu	-0.247***	-0.157***	-0.220***	-0.198***
	(0.090)	(0.042)	(0.040)	(0.042)
Muslim	-0.282*	-0.181***	-0.279***	-0.176**
	(0.169)	(0.055)	(0.044)	(0.073)
Other religion	-0.343***	-0.152**	-0.247***	-0.178***
	(0.088)	(0.064)	(0.053)	(0.057)
Observations	411	1,690	1,720	1,552
R-squared	0.602	0.458	0.367	0.317

Panel C: Differential effect of discrimination on parental investment across social groups- controlling for baseline PPVT test scores

	Expenditure on education				School fees		Enrolment at school			
	EXP R2	EXP R3	EXP R4	EXP R5	FEES R4	FEES R5	ENROL R2	ENROL R3	ENROL R4	ENROL R5
Base- Upper Hindus × No Discr										
Lower Hindu × No Discr	-0.365**	-0.387***	-0.314**	-0.230**	-0.326***	-0.309***	0.031	0.007	-0.002	-0.040
	(0.148)	(0.105)	(0.139)	(0.104)	(0.123)	(0.116)	(0.038)	(0.020)	(0.015)	(0.026)
Muslim × No Discr	-0.586***	-0.621***	-0.424***	-0.454***	-0.423***	-0.236	-0.055	-0.043	0.013	-0.059
	(0.129)	(0.157)	(0.150)	(0.122)	(0.127)	(0.273)	(0.043)	(0.033)	(0.013)	(0.038)
Other religion × No Discr	-0.225	-0.447***	-0.212	-0.137	-0.447***	-0.331**	0.012	0.020	0.002	-0.011
	(0.219)	(0.122)	(0.287)	(0.154)	(0.129)	(0.133)	(0.065)	(0.029)	(0.021)	(0.048)
Upper Hindu × Discr	-0.209	-0.043	-0.388***	-0.091	-0.034	-0.039	0.009	0.029	0.032*	0.020
	(0.245)	(0.246)	(0.138)	(0.124)	(0.193)	(0.196)	(0.055)	(0.023)	(0.017)	(0.031)
Lower Hindu × Discr	-0.515***	-0.469***	-0.547***	-0.311***	-0.421***	-0.401***	-0.022	0.011	-0.023	-0.091***
	(0.146)	(0.126)	(0.128)	(0.096)	(0.131)	(0.118)	(0.042)	(0.021)	(0.018)	(0.031)
Muslim × Discr	-0.656***	-0.543***	-0.698***	-0.423***	-0.762***	-0.376**	0.111	0.044*	0.003	-0.189**
	(0.205)	(0.151)	(0.143)	(0.152)	(0.139)	(0.163)	(0.119)	(0.026)	(0.039)	(0.074)
Other religion × Discr	-0.567***	-0.549***	-0.504***	-0.414***	-0.534***	-0.405***	-0.079	0.063***	0.006	0.023
	(0.148)	(0.138)	(0.180)	(0.118)	(0.147)	(0.145)	(0.073)	(0.023)	(0.031)	(0.046)
Observations	1,447	1,766	1,669	1,643	1,738	1,573	1,888	1,799	1,788	1,773
R-squared	0.352	0.271	0.273	0.230	0.279	0.302	0.134	0.108	0.110	0.163

	Private school			
VARIABLES	R2	R3	R4	R5
Base- Upper Hindus × No Discr				
Lower Hindu × No Discr	-0.278***	-0.159***	-0.243***	-0.202***
	(0.101)	(0.040)	(0.042)	(0.045)
Muslim × No Discr	-0.277	-0.188***	-0.251***	-0.179**
	(0.235)	(0.051)	(0.044)	(0.080)
Other religion × No Discr	-0.355***	-0.150**	-0.259***	-0.155**
	(0.105)	(0.073)	(0.064)	(0.064)
Upper Hindu × Discr	-0.257*	-0.090	-0.113	-0.077
	(0.141)	(0.070)	(0.084)	(0.083)
Lower Hindu × Discr	-0.393***	-0.240***	-0.263***	-0.265***
	(0.119)	(0.045)	(0.046)	(0.046)
Muslim × Discr	-0.480***	-0.252**	-0.450***	-0.242**
	(0.160)	(0.101)	(0.106)	(0.103)
Other religion × Discr	-0.546***	-0.243***	-0.310***	-0.297***
	(0.101)	(0.087)	(0.070)	(0.073)
Observations	411	1,690	1,720	1,552
R-squared	0.605	0.458	0.369	0.318

Notes: Regressions for rounds 3, 4, and 5 control for PPVT scores obtained in round 2. All regressions include socioeconomic controls and location fixed effects. Standard errors are adjusted for clustering at the community level. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 10: Jeffrey Penney

Table 10a: Differences in test scores between lower and upper-caste Hindus

	PPVT					
	z-OLS	**z-UQR**	**Normalized**	**z-OLS**	**z-UQR**	**Normalized**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Round 2	-0.321***	-0.576***	-0.374***	-0.178***	-0.381***	-0.251***
	(0.066)	(0.119)	(0.069)	(0.061)	(0.099)	(0.063)
Round 3	-0.278***	-0.372**	-0.224***	-0.158*	-0.190	-0.116
	(0.086)	(0.146)	(0.085)	(0.084)	(0.151)	(0.089)
Round 4	-0.368***	-0.455***	-0.285***	-0.218***	-0.246*	-0.158**
	(0.070)	(0.129)	(0.078)	(0.066)	(0.126)	(0.078)
Round 5	-0.381***	-0.390***	-0.248***	-0.267***	-0.240*	-0.155**
	(0.086)	(0.123)	(0.076)	(0.084)	(0.125)	(0.078)
	Maths					
Round 3	-0.373***	-0.514***	-0.332***	-0.213***	-0.274***	-0.182***
	(0.057)	(0.100)	(0.063)	(0.053)	(0.097)	(0.062)
Round 4	-0.454***	-0.678***	-0.422***	-0.277***	-0.436***	-0.277***
	(0.073)	(0.124)	(0.075)	(0.076)	(0.131)	(0.081)
Round 5	-0.548***	-0.715***	-0.429***	-0.401***	-0.521***	-0.315***
	(0.093)	(0.162)	(0.095)	(0.093)	(0.166)	(0.098)
	English					
Round 4	-0.476***	-0.682***	-0.430***	-0.274***	-0.367***	-0.241***
	(0.084)	(0.134)	(0.082)	(0.079)	(0.134)	(0.086)
	Reading					
Round 5	-0.349***	-0.437***	-0.273***	-0.235***	-0.293**	-0.185**
	(0.079)	(0.117)	(0.071)	(0.077)	(0.121)	(0.074)
SES controls	No	No		Yes	Yes	

Table 10b: Effect of discrimination on test scores

	PPVT					
	z-OLS	**z-UQR**	**Normalized**	**z-OLS**	**z-UQR**	**Normalized**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Round 2	0.002	0.125	0.081	0.029	0.164*	0.108*
	(0.055)	(0.086)	(0.055)	(0.055)	(0.090)	(0.058)
Round 3	-0.058	-0.055	-0.033	-0.035	-0.018	-0.011
	(0.046)	(0.084)	(0.049)	(0.044)	(0.081)	(0.048)
Round 4	-0.193***	-0.294***	-0.186***	-0.178***	-0.266***	-0.173***
	(0.060)	(0.097)	(0.060)	(0.058)	(0.097)	(0.061)
Round 5	-0.112*	-0.119	-0.076	-0.094	-0.099	-0.064
	(0.062)	(0.092)	(0.057)	(0.062)	(0.090)	(0.057)
	Maths					
Round 3	-0.141***	-0.262***	-0.170***	-0.105**	-0.211***	-0.140***
	(0.049)	(0.079)	(0.050)	(0.047)	(0.080)	(0.051)
Round 4	-0.118*	-0.105	-0.065	-0.078	-0.039	-0.024
	(0.065)	(0.104)	(0.062)	(0.059)	(0.096)	(0.059)
Round 5	-0.149**	-0.182**	-0.110**	-0.126**	-0.168**	-0.103**
	(0.057)	(0.082)	(0.048)	(0.054)	(0.078)	(0.046)
	English R4					
Round 4	-0.186***	-0.231**	-0.146**	-0.145**	-0.168*	-0.111*
	(0.061)	(0.092)	(0.057)	(0.056)	(0.089)	(0.057)
	Reading R5					
Round 5	-0.087	-0.098	-0.061	-0.068	-0.077	-0.048
	(0.066)	(0.104)	(0.063)	(0.065)	(0.105)	(0.064)
SES controls	No	No		Yes	Yes	

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table 11: Oster test
Table 11a: Effect of caste

Treatment variable- Lower Hindu castes; Rmax=1.3 ˜ R; delta=0.8				
Dependent variable	**Baseline effect (std dev), [R2]**	**Controlled effect (std dev), [R2]**	**Bias adjusted β**	$\tilde{\delta}$ **for** β**=0 given** R_{max}
		Test Scores		
z-score PPVT pctile-R2	0.264*** (0.065) [0.266]	0.156** (0.061) [0.300]	-0.049	0.61
z-score PPVT pctile-R3	0.223*** (0.082) [0.263]	0.142* (0.080) [0.283]	-0.137	0.41
z-score PPVT pctile-R4	0.352*** (0.063) [0.257]	0.236*** (0.059) [0.291]	0.034	0.92
z-score PPVT pctile-R5	0.357*** (0.079) [0.203]	0.283*** (0.080) [0.223]	0.069	1.03
z-score Maths pctile-R3	0.344*** (0.056) [0.301]	0.273*** (0.057) [0.289]	0.004	0.78
z-score Maths pctile-R4	0.414*** (0.062) [0.241]	0.277*** (0.066) [0.291]	0.105	1.26
z-score Maths pctile-R5	0.454*** (0.085) [0.225]	0.355*** (0.087) [0.261]	0.167	1.43
z-score English pctile-R4	0.445*** (0.078) [0.265]	0.281*** (0.073) [0.340]	0.113	1.31
z-score Reading pctile-R5	0.352*** (0.072) [0.190]	0.256*** (0.071) [0.216]	0.088	1.18
		Parental Investment		
z-Investment R2	0.517*** (0.093) [0.293]	0.360*** (0.098) [0.341]	0.098	1.07
z-Investment R3	0.591*** (0.107) [0.209]	0.463*** (0.100) [0.261]	0.312	2.14
z-Investment R4	0.406*** (0.116) [0.229]	0.305*** (0.110) [0.263]	0.124	1.30
z-Investment R5	0.404*** (0.083) [0.186]	0.311*** (0.083) [0.221]	0.163	1.58
z-School Fees R4	0.512*** (0.104) [0.234]	0.395*** (0.106) [0.274]	0.216	1.64
z-School Fees R5	0.488*** (0.088) [0.257]	0.390*** (0.085) [0.290]	0.187	1.44
School Enrolment R2	-0.005 (0.033) [0.124]	-0.005 (0.035) [0.129]	0.008	0.34
School Enrolment R3	0.007 (0.017) [0.104]	0.008 (0.016) [0.106]	-0.050	0.14
School Enrolment R4	0.022* (0.012) [0.093]	0.016 (0.013) [0.095]	-0.027	0.30
School Enrolment R5	0.081*** (0.021) [0.133]	0.063*** (0.023) [0.143]	-0.002	0.78
Private School R2	0.284*** (0.090) [0.544]	0.058(0.051) [0.106]	0.016	1.02
Private School R3	0.251*** (0.045) [0.384]	0.183*** (0.048) [0.320]	0.054	1.11
Private School R4	0.266*** (0.042) [0.306]	0.225*** (0.041) [0.317]	0.127	1.68
Private School R5	0.246*** (0.040) [0.262]	0.269*** (0.053) [0.275]	0.155	1.73

Table 11b: Effect of discrimination

Treatment variable- Discrimination; Rmax=1.3*R; delta=0.8				
Dependent variable	**Baseline effect (std dev), [R2]**	**Controlled effect (std dev), [R2]**	**Bias adjusted β**	$\tilde{\delta}$ **for** β **=0 given** R_{max}
		Test Scores		
z-score PPVT pctile-R2	0.009 (0.055) [0.262]	0.029 (0.055) [0.303]	0.094	-0.36
z-score PPVT pctile-R3	-0.078* (0.045) [0.261]	-0.045 (0.043) [0.286]	0.047	0.40
z-score PPVT pctile-R4	-0.211*** (0.061) [0.255]	-0.178*** (0.059) [0.303]	-0.012	2.21
z-score PPVT pctile-R5	-0.136*** (0.060) [0.193]	-0.098 (0.061) [0.225]	-0.025	1.07
z-score Maths pctile-R3	-0.161*** (0.047) [0.293]	-0.112** (0.047) [0.346]	-0.036	1.17
z-score Maths pctile-R4	-0.155** (0.062) [0.228]	-0.091 (0.058) [0.293]	-0.026	1.12
z-score Maths pctile-R5	-0.195*** (0.054) [0.212]	-0.145*** (0.051) [0.264]	-0.071	1.52
z-score English pctile-R4	-0.222*** (0.058) [0.257]	-0.153*** (0.054) [0.295]	-0.084	1.74
z-score Reading pctile-R5	-0.107 (0.065) [0.177]	-0.069 (0.063) [0.213]	-0.010	0.93
		Parental Investment		
z-Investment R2	-0.212*** (0.060) [0.281]	-0.161*** (0.053) [0.351]	-0.093	1.86
z-Investment R3	-0.127** (0.054) [0.181]	-0.052 (0.052) [0.267]	0.003	0.75
z-Investment R4	-0.289*** (0.053) [0.217]	-0.233*** (0.051) [0.263]	-0.159	2.33
z-Investment R5	-0.141*** (0.039) [0.169]	-0.092** (0.038) [0.222]	-0.044	1.52
z-School Fees R4	-0.175*** (0.048) [0.207]	-0.105** (0.045) [0.268]	-0.038	1.24
z-School Fees R5	-0.151*** (0.050) [0.221]	-0.095** (0.047) [0.283]	-0.038	1.33
School Enrolment R2	-0.032 (0.024) [0.125]	-0.033 (0.024) [0.130]	-0.054	-1.47
School Enrolment R3	0.013 (0.013) [0.103]	-0.014 (0.014) [0.106]	0.018	-4.12
School Enrolment R4	-0.020 (0.013) [0.096]	-0.016 (0.014) [0.098]	0.000	0.80
School Enrolment R5	-0.054*** (0.020) [0.131]	-0.045** (0.020) [0.147]	-0.024	1.61
Private School R2	-0.153*** (0.053) [0.521]	-0.150*** (0.055) [0.602]	-0.136	3.90
Private School R3	-0.108** (0.023) [0.369]	-0.068*** (0.023) [0.448]	-0.015	1.03
Private School R4	-0.088*** (0.025) [0.283]	-0.048* (0.026) [0.357]	-0.007	0.94
Private School R5	-0.102*** (0.027) [0.248]	-0.072*** (0.027) [0.310]	-0.034	1.51

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table 12: Multiple hypothesis testing

Treatment Variable-	Lower Hindu Caste		Discrimination	
	Model p-value	Romano-Wolf p-value	Model p-value	Romano-Wolf p-value
Outcome Variables	(1)	(2)	(3)	(4)
		Panel A: Cognitive Outcomes		
z-score PPVT pctile-R2	0.005	0.002	0.597	0.445
z-score PPVT pctile-R3	0.088	0.007	0.297	0.327
z-score PPVT pctile-R4	0.001	0.001	0.003	0.001
z-score PPVT pctile-R5	0.001	0.001	0.111	0.099
z-score Maths pctile-R3	0.0004	0.001	0.018	0.004
z-score Maths pctile-R4	0.0004	0.001	0.117	0.099
z-score Maths pctile-R5	0.0002	0.001	0.006	0.001
z-score English pctile-R4	0.0005	0.001	0.006	0.001
z-score Reading pctile-R5	0.0014	0.001	0.278	0.327
		Panel B: Parental investment		
z-Investment R2	0.001	0.002	0.003	0.001
z-Investment R3	0.000	0.001	0.317	0.290
z-Investment R4	0.008	0.003	0.000	0.001
z-Investment R5	0.001	0.001	0.017	0.003
z-School Fees R4	0.001	0.001	0.021	0.007
z-School Fees R5	0.000	0.001	0.044	0.024
School Enrolment R2	0.794	0.696	0.171	0.171
School Enrolment R3	0.589	0.689	0.332	0.290
School Enrolment R4	0.164	0.126	0.254	0.251
School Enrolment R5	0.012	0.003	0.024	0.008
Private School R2	0.361	0.388	0.035	0.013
Private School R3	0.001	0.001	0.125	0.130
Private School R4	0.000	0.001	0.058	0.028
Private School R5	0.000	0.001	0.003	0.001

*Notes- Columns 1 and 3 report the p-values for estimates on lower Hindu castes from specification 1.1 and discrimination from specification 1.2, respectively. Columns 2 and 4 report the p-values adjusted for multiple hypothesis testing associated with columns 1 and 3, respectively, using 1000 bootstrap replications. Standard errors are adjusted for clustering at the community level. All regressions include SES controls. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

9. Appendix

Table A1: Test score gaps across social groups over time

Panel A:	**Standardized PPVT percentiles**							
	Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Caste (Base- Upper Hindu)								
SC/ST/BC Hindu	-0.321***	-0.178***	-0.263***	-0.140*	-0.371***	-0.213***	-0.373***	-0.268***
	(0.066)	(0.061)	(0.084)	(0.081)	(0.066)	(0.063)	(0.079)	(0.079)
Muslim	-0.122	-0.018	-0.320*	-0.244	-0.648***	-0.553***	-0.476***	-0.392**
	(0.131)	(0.129)	(0.165)	(0.150)	(0.157)	(0.153)	(0.165)	(0.174)
Other religion	-0.201*	-0.058	-0.197	-0.050	-0.320**	-0.130	-0.458***	-0.324***
	(0.116)	(0.111)	(0.137)	(0.120)	(0.123)	(0.106)	(0.129)	(0.121)
Observations	1,851	1,832	1,901	1,879	1,903	1,881	1,890	1,867
SES controls	No	Yes	No	Yes	No	Yes	No	Yes

Panel B: Standardized Percentiles	**Maths**						**English**		**Reading**	
	Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Caste (Base- Upper Hindu)										
SC/ST/BC Hindu	-0.359***	-0.197***	-0.446***	-0.264***	-0.496***	-0.355***	-0.480***	-0.273***	-0.367***	-0.245***
	(0.057)	(0.054)	(0.068)	(0.072)	(0.090)	(0.091)	(0.080)	(0.076)	(0.075)	(0.075)
Muslim	-0.547***	-0.442***	-0.600***	-0.474***	-0.514***	-0.423***	-0.523***	-0.393***	-0.453***	-0.379***
	(0.107)	(0.100)	(0.130)	(0.129)	(0.113)	(0.103)	(0.126)	(0.117)	(0.099)	(0.102)
Other religion	-0.359***	-0.184*	-0.383***	-0.195**	-0.440***	-0.278**	-0.451***	-0.244**	-0.380***	-0.242**
	(0.133)	(0.106)	(0.121)	(0.093)	(0.125)	(0.117)	(0.117)	(0.102)	(0.110)	(0.103)
Observations	1,904	1,883	1,858	1,838	1,840	1,818	1,862	1,842	1,831	1,810
SES controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table A2: Association between perceived discrimination and children's test scores

Panel A:	**Standardized PPVT percentiles**							
	Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Caste (Base- Upper Hindu)								
SC/ST/BC Hindu	-0.308***	-0.170***	-0.250***	-0.132	-0.368***	-0.221***	-0.370***	-0.271***
	(0.067)	(0.062)	(0.086)	(0.083)	(0.064)	(0.061)	(0.078)	(0.079)
Muslim	-0.110	-0.012	-0.302*	-0.229	-0.644***	-0.558***	-0.470***	-0.394**
	(0.125)	(0.122)	(0.164)	(0.148)	(0.153)	(0.150)	(0.168)	(0.175)
Other religion	-0.190	-0.053	-0.170	-0.037	-0.305**	-0.131	-0.454***	-0.332***
	(0.115)	(0.110)	(0.143)	(0.123)	(0.122)	(0.107)	(0.130)	(0.122)
Perceived social discrimination	0.002	0.029	-0.070	-0.045	-0.194***	-0.178***	-0.117*	-0.098
	(0.055)	(0.055)	(0.045)	(0.043)	(0.062)	(0.059)	(0.061)	(0.061)
Observations	1,833	1,815	1,879	1,859	1,881	1,861	1,868	1,848
SES Controls	No	Yes	No	Yes	No	Yes	No	Yes

Panel B: Standardized Percentiles	**Maths**						**English**		**Reading**	
	Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Caste (Base- Upper Hindu)										
SC/ST/BC Hindu	-0.342***	-0.191***	-0.425***	-0.253***	-0.482***	-0.351***	-0.449***	-0.251***	-0.342***	-0.230***
	(0.058)	(0.055)	(0.070)	(0.074)	(0.091)	(0.092)	(0.078)	(0.076)	(0.075)	(0.075)
Muslim	-0.524***	-0.428***	-0.581***	-0.462***	-0.507***	-0.423***	-0.492***	-0.371***	-0.435***	-0.370***
	(0.107)	(0.100)	(0.128)	(0.127)	(0.111)	(0.104)	(0.123)	(0.116)	(0.098)	(0.101)
Other religion	-0.323**	-0.169	-0.361***	-0.177*	-0.427***	-0.272**	-0.404***	-0.200*	-0.337***	-0.213**
	(0.135)	(0.107)	(0.121)	(0.095)	(0.130)	(0.121)	(0.120)	(0.107)	(0.108)	(0.099)
Perceived social discrimination	-0.148***	-0.112**	-0.134**	-0.091	-0.169***	-0.145***	-0.200***	-0.153***	-0.089	-0.069
	(0.047)	(0.047)	(0.063)	(0.058)	(0.055)	(0.052)	(0.059)	(0.054)	(0.064)	(0.063)
Observations	1,882	1,863	1,837	1,818	1,819	1,800	1,842	1,823	1,810	1,792
SES Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects.

Table A3: Differential association between discrimination and test scores across social groups over time

	Standardized Test Scores								
	PPVT				Maths			English	Reading
	Age 5	Age 8	Age 12	Age 15	Age 8	Age 12	Age 15	Age 12	Age 15
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Base- Upper Hindus × No Discr									
Lower Hindu × No Discr	-0.181***	-0.094	-0.175**	-0.302***	-0.093	-0.209***	-0.335***	-0.217***	-0.230***
	(0.069)	(0.091)	(0.072)	(0.076)	(0.062)	(0.079)	(0.097)	(0.076)	(0.081)
Muslim × No Discr	-0.080	-0.230	-0.484***	-0.396*	-0.376***	-0.431***	-0.381***	-0.347***	-0.348***
	(0.128)	(0.156)	(0.180)	(0.201)	(0.128)	(0.137)	(0.111)	(0.120)	(0.124)
Other religion × No Discr	-0.048	-0.017	-0.115	-0.376**	-0.069	-0.183	-0.301***	-0.186	-0.202*
	(0.125)	(0.140)	(0.140)	(0.151)	(0.117)	(0.119)	(0.114)	(0.134)	(0.106)
Upper Hindu × Discr	-0.034	0.076	-0.006	-0.207	0.231*	0.058	-0.088	-0.033	-0.060
	(0.161)	(0.151)	(0.139)	(0.170)	(0.139)	(0.149)	(0.187)	(0.167)	(0.153)
Lower Hindu × Discr	-0.159*	-0.168*	-0.371***	-0.382***	-0.259***	-0.332***	-0.489***	-0.391***	-0.291***
	(0.086)	(0.093)	(0.098)	(0.085)	(0.072)	(0.091)	(0.100)	(0.094)	(0.106)
Muslim × Discr	0.155	-0.177	-0.796***	-0.568**	-0.407***	-0.516***	-0.630***	-0.494***	-0.493***
	(0.153)	(0.239)	(0.186)	(0.224)	(0.135)	(0.163)	(0.170)	(0.187)	(0.172)
Other religion × Discr	-0.073	-0.039	-0.224	-0.421***	-0.230	-0.147	-0.320	-0.294*	-0.299
	(0.183)	(0.157)	(0.157)	(0.145)	(0.192)	(0.174)	(0.206)	(0.157)	(0.182)
Observations	1,815	1,859	1,861	1,848	1,863	1,818	1,800	1,823	1,792
SES controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table A4: Gaps in parental investment across social groups

Panel A:	**Annual Expenditure on Education**							
	Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Caste (Base- Upper Hindu)								
SC/ST/BC Hindu	-0.523***	-0.347***	-0.611***	-0.446***	-0.422***	-0.302***	-0.389***	-0.297***
	(0.099)	(0.103)	(0.111)	(0.101)	(0.115)	(0.111)	(0.086)	(0.086)
Muslim	-0.642***	-0.542***	-0.692***	-0.594***	-0.446***	-0.391***	-0.512***	-0.482***
	(0.093)	(0.090)	(0.135)	(0.125)	(0.128)	(0.119)	(0.083)	(0.079)
Other religion	-0.452***	-0.262*	-0.680***	-0.511***	-0.358*	-0.229	-0.384***	-0.281**
	(0.163)	(0.155)	(0.121)	(0.101)	(0.207)	(0.208)	(0.106)	(0.113)
Observations	1,484	1,463	1,893	1,870	1,788	1,767	1,763	1,742
SES controls	No	Yes	No	Yes	No	Yes	No	Yes

Panel B:	**School Fees**				**Enrolment at school**							
	Age 8 (Round 3)		Age 12 (Round 4)		Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Caste (Base- Upper Hindu)												
SC/ST/BC Hindu	-0.532***	-0.375***	-0.507***	-0.395***	0.007	0.009	-0.006	-0.009	-0.028**	-0.019	-0.093***	-0.065**
	(0.109)	(0.108)	(0.093)	(0.091)	(0.034)	(0.036)	(0.017)	(0.017)	(0.013)	(0.013)	(0.023)	(0.025)
Muslim	-0.600***	-0.499***	-0.381***	-0.336**	-0.013	-0.016	-0.016	-0.024	-0.006	0.003	-0.119***	-0.099***
	(0.120)	(0.116)	(0.132)	(0.133)	(0.054)	(0.053)	(0.025)	(0.028)	(0.017)	(0.017)	(0.040)	(0.038)
Other religion	-0.661***	-0.493***	-0.523***	-0.396***	-0.028	-0.029	0.031	0.024	-0.015	-0.005	-0.038	-0.003
	(0.113)	(0.108)	(0.093)	(0.097)	(0.054)	(0.057)	(0.022)	(0.022)	(0.016)	(0.017)	(0.033)	(0.033)
Observations	1,839	1,818	1,657	1,638	1,931	1,908	1,931	1,908	1,920	1,896	1,900	1,876
SES controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Panel C:	**Type of School (Private School)**							
	Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Caste (Base- Upper Hindu)								
SC/ST/BC Hindu	-0.277***	-0.222**	-0.277***	-0.182***	-0.279***	-0.202***	-0.267***	-0.194***
	(0.089)	(0.088)	(0.049)	(0.045)	(0.042)	(0.039)	(0.042)	(0.039)
Muslim	-0.332**	-0.321**	-0.250***	-0.185***	-0.299***	-0.244***	-0.188**	-0.155**
	(0.157)	(0.161)	(0.059)	(0.057)	(0.049)	(0.044)	(0.077)	(0.077)
Other religion	-0.388***	-0.326***	-0.288***	-0.187***	-0.326***	-0.239***	-0.270***	-0.188***
	(0.092)	(0.087)	(0.078)	(0.063)	(0.062)	(0.047)	(0.057)	(0.052)
Observations	422	416	1,806	1,784	1,839	1,818	1,658	1,639
SES controls	No	Yes	No	Yes	No	Yes	No	Yes

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table A5: Association between perceived discrimination and parental investment

Panel A:	**Annual Expenditure on Education**							
	Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Caste (Base- Upper Hindu)								
SC/ST/BC Hindu	-0.514***	-0.350***	-0.614***	-0.450***	-0.410***	-0.294***	-0.326***	-0.241***
	(0.100)	(0.105)	(0.112)	(0.102)	(0.115)	(0.111)	(0.080)	(0.080)
Muslim	-0.639***	-0.550***	-0.700***	-0.599***	-0.441***	-0.388***	-0.436***	-0.413***
	(0.095)	(0.093)	(0.140)	(0.129)	(0.129)	(0.120)	(0.106)	(0.103)
Other religion	-0.444***	-0.263*	-0.674***	-0.502***	-0.326	-0.202	-0.316***	-0.220*
	(0.164)	(0.157)	(0.123)	(0.103)	(0.211)	(0.211)	(0.107)	(0.114)
Perceived social discrimination	-0.201***	-0.161***	-0.100*	-0.052	-0.277***	-0.233***	-0.135***	-0.092**
	(0.054)	(0.053)	(0.054)	(0.052)	(0.054)	(0.051)	(0.042)	(0.038)
Observations	1,467	1,447	1,872	1,851	1,770	1,751	1,744	1,725
SES controls	No	Yes	No	Yes	No	Yes	No	Yes

Panel B:	**School Fees**				**Enrolment at school**							
	Age 12 (Round 4)		Age 15 (Round 5)		Age 5 (Round 2)		Age 85 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Caste (Base- Upper Hindu)												
SC/ST/BC Hindu	-0.486***	-0.337***	-0.424***	-0.319***	0.011	0.015	-0.003	-0.006	-0.026**	-0.018	-0.088***	-0.061**
	(0.109)	(0.107)	(0.094)	(0.092)	(0.035)	(0.036)	(0.017)	(0.017)	(0.013)	(0.014)	(0.023)	(0.025)
Muslim	-0.561***	-0.462***	-0.283	-0.243	-0.007	-0.009	-0.015	-0.023	-0.004	0.004	-0.114***	-0.095**
	(0.115)	(0.111)	(0.220)	(0.218)	(0.057)	(0.056)	(0.025)	(0.028)	(0.018)	(0.018)	(0.039)	(0.037)
Other religion	-0.613***	-0.449***	-0.438***	-0.318***	-0.015	-0.018	0.031	0.026	-0.013	-0.003	-0.023	0.004
	(0.114)	(0.110)	(0.103)	(0.107)	(0.055)	(0.057)	(0.023)	(0.022)	(0.016)	(0.017)	(0.033)	(0.034)
Perceived social discr	-0.151***	-0.105**	-0.127**	-0.092*	-0.033	-0.033	0.012	0.014	-0.019	-0.016	-0.052**	-0.045**
	(0.047)	(0.045)	(0.048)	(0.047)	(0.024)	(0.024)	(0.013)	(0.014)	(0.013)	(0.014)	(0.020)	(0.020)
Observations	1,818	1,799	1,639	1,622	1,909	1,888	1,909	1,888	1,897	1,876	1,878	1,857
SES Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Panel C:	**Type of School (Private School)**							
	Age 5 (Round 2)		Age 8 (Round 3)		Age 12 (Round 4)		Age 15 (Round 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Caste (Base- Upper Hindu)								
SC/ST/BC Hindu	-0.301***	-0.247***	-0.263***	-0.173***	-0.270***	-0.196***	-0.255***	-0.187***
	(0.090)	(0.090)	(0.049)	(0.044)	(0.043)	(0.040)	(0.042)	(0.040)
Muslim	-0.301*	-0.282*	-0.238***	-0.178***	-0.295***	-0.243***	-0.185**	-0.157**
	(0.164)	(0.169)	(0.058)	(0.057)	(0.049)	(0.045)	(0.075)	(0.075)
Other religion	-0.404***	-0.343***	-0.262***	-0.164**	-0.309***	-0.224***	-0.249***	-0.171***
	(0.091)	(0.088)	(0.080)	(0.065)	(0.065)	(0.052)	(0.057)	(0.054)
Perceived social discrimination	-0.147***	-0.150***	-0.096***	-0.068***	-0.075***	-0.048*	-0.088***	-0.072***
	(0.054)	(0.055)	(0.025)	(0.023)	(0.026)	(0.026)	(0.027)	(0.027)
Observations	417	411	1,787	1,767	1,818	1,799	1,640	1,623
R-squared	0.575	0.602	0.361	0.448	0.288	0.357	0.258	0.310

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

Table A6: Differential association between discrimination and parental investment across social groups

Panel A:	**Expenditure on education**				**School Fees**	
	Age 5	Age 8	Age 12	Age 15	Age 12	Age 15
	(1)	(2)	(3)	(4)	(5)	(6)
Base- Upper Hindus × No Discr						
Lower Hindu × No Discr	-0.365**	-0.400***	-0.333**	-0.253**	-0.343***	-0.320***
	(0.148)	(0.101)	(0.132)	(0.098)	(0.116)	(0.108)
Muslim × No Discr	-0.586***	-0.598***	-0.417***	-0.457***	-0.406***	-0.242
	(0.129)	(0.147)	(0.142)	(0.115)	(0.122)	(0.262)
Other religion × No Discr	-0.225	-0.447***	-0.227	-0.166	-0.464***	-0.316**
	(0.219)	(0.118)	(0.280)	(0.149)	(0.124)	(0.127)
Upper Hindu × Discr	-0.209	0.118	-0.372***	-0.135	-0.111	-0.094
	(0.245)	(0.281)	(0.134)	(0.110)	(0.147)	(0.179)
Lower Hindu × Discr	-0.515***	-0.485***	-0.544***	-0.336***	-0.431***	-0.410***
	(0.146)	(0.123)	(0.126)	(0.091)	(0.125)	(0.111)
Muslim × Discr	-0.656***	-0.524***	-0.653***	-0.418***	-0.727***	-0.340**
	(0.205)	(0.148)	(0.143)	(0.146)	(0.132)	(0.162)
Other religion × Discr	-0.567***	-0.540***	-0.485***	-0.444***	-0.528***	-0.416***
	(0.148)	(0.136)	(0.177)	(0.115)	(0.147)	(0.141)
Observations	1,447	1,851	1,751	1,725	1,799	1,622
SES controls	Yes	Yes	Yes	Yes	Yes	Yes

Panel B:	**Enrolment at school**				**Private School**			
	Age 5	Age 8	Age 12	Age 15	Age 5	Age 8	Age 12	Age 15
	(1)	(2)	(3)	(4)	(7)	(8)	(9)	(10)
Base- Upper Hindus × No Discr								
Lower Hindu × No Discr	0.031	-0.002	-0.005	-0.045*	-0.278***	-0.175***	-0.229***	-0.201***
	(0.038)	(0.020)	(0.015)	(0.026)	(0.101)	(0.041)	(0.042)	(0.045)
Muslim × No Discr	-0.055	-0.044	0.011	-0.061	-0.277	-0.184***	-0.230***	-0.172**
	(0.043)	(0.034)	(0.014)	(0.038)	(0.235)	(0.053)	(0.043)	(0.081)
Other religion × No Discr	0.012	0.017	0.002	-0.007	-0.355***	-0.154**	-0.247***	-0.150**
	(0.065)	(0.029)	(0.021)	(0.045)	(0.105)	(0.074)	(0.062)	(0.063)
Upper Hindu × Discr	0.009	0.016	0.031*	0.018	-0.257*	-0.073	-0.151**	-0.118*
	(0.055)	(0.025)	(0.016)	(0.030)	(0.141)	(0.075)	(0.071)	(0.068)
Lower Hindu × Discr	-0.022	0.002	-0.029	-0.098***	-0.393***	-0.241***	-0.250***	-0.261***
	(0.042)	(0.021)	(0.020)	(0.032)	(0.119)	(0.046)	(0.045)	(0.047)
Muslim × Discr	0.111	0.047*	0.007	-0.181**	-0.480***	-0.235**	-0.411***	-0.222**
	(0.119)	(0.024)	(0.039)	(0.072)	(0.160)	(0.102)	(0.106)	(0.103)
Other religion × Discr	-0.079	0.058**	0.002	0.020	-0.546***	-0.252***	-0.300***	-0.313***
	(0.073)	(0.023)	(0.029)	(0.042)	(0.101)	(0.086)	(0.068)	(0.068)
Observations	1,888	1,888	1,876	1,857	411	1,767	1,799	1,623
SES controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Notes- Standard errors are adjusted for clustering at the community level. All regressions control for location fixed effects and SES. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.*

A7: Index of parents' perceived discrimination

I construct an index for parents' perceived social discrimination using two survey questions asked to parents in round 2. Parents were asked to rate how much they agreed with the following two statements on a four-point Likert scale (from strongly agree to strongly disagree): 'When I am at shops/market I am usually treated with fairness and with respect by others'; and 'Other people in my street/village look down on me and my family'. The descriptive statistics of the two survey questions are provided below.

a. RESPECT: When I am at shops/market I am usually treated by others with fairness and with respect.

RESPECT CASTE	strongly agree	agree	disagree	strongly disagree	Total
SC/ST/BC Hindu	1,064	301	29	26	1,420
	0.749	0.212	0.020	0.018	
Other Hindu	226	35	3	6	270
	0.837	0.130	0.011	0.022	
Muslim	106	26	2	1	135
	0.785	0.193	0.015	0.007	
Other religion	81	19	3	3	106
	0.764	0.179	0.028	0.028	
Total	1,438	367	35	34	1,874
	0.767	0.196	0.019	0.018	

Almost 76 percent of the respondents strongly agree to be treated with respect. For upper-caste Hindus, this share is 84 percent. I create a binary variable "*Respect*" which takes a value of 1 if parents strongly agree to feel respected and 0 otherwise. Table A7.1 below reports the descriptive statistics.

RESPECT CASTE	0	1	Total
SC/ST/BC Hindu	356	1,064	1,420
	0.251	0.749	
Other Hindu	44	226	270
	0.163	0.837	
Muslim	29	106	135
	0.215	0.785	
Other religion	25	81	106
	0.236	0.764	
Total	454	1,477	1,931

Table A7.1: Descriptive statistics of RESPECT

b. LOOKED DOWN: Other people in my STREET/VILLAGE look down on me and my family

LOOKED DOWN CASTE	strongly agree	agree	disagree	strongly disagree	Total
SC/ST/BC Hindu	62	100	118	1,145	1,425
	0.044	0.070	0.083	0.804	
Other Hindu	6	11	13	244	274
	0.022	0.040	0.047	0.891	
Muslim	6	8	4	119	137
	0.044	0.058	0.029	0.869	

Other religion	7	6	10	84	107
	0.065	0.056	0.093	0.785	
Total	81	121	141	1,543	1,886
	0.043	0.064	0.075	0.818	

Overall, 81% strongly disagree to feel looked down upon in the community. This share is 80% and 89% for the lower and upper-caste Hindus, respectively. I construct a binary variable "*Looked down*" which takes a value of 1 if parents strongly disagreed with being looked down and 0 otherwise.

LOOKED DOWN CASTE	0	1	Total
SC/ST/BC Hindu	1,145	280	1,425
	0.804	0.196	
Other Hindu	244	30	274
	0.891	0.109	
Muslim	119	18	137
	0.869	0.131	
Other religion	84	23	107
	0.785	0.215	
Total	454	1,477	1,931

Table A7.2: Descriptive statistics of LOOKED DOWN

c. DISR: Combining the two variables- *Respect* and *Looked down*, I construct an index for discrimination "*DISCR*". The index takes a value of 0 if parents perceive no manifestation of discrimination (i.e. if parents strongly agree with being respected and strongly disagree to being looked down upon the community) and 1 if parents perceive any manifestation of discrimination. Table A7.3 below reports the descriptive statistics.

DISCR CASTE	0	1	Total
SC/ST/BC Hindu	923	494	1,417
	0.651	0.349	
Other Hindu	206	64	270
	0.763	0.237	
Muslim	95	40	135
	0.704	0.296	
Other religion	70	36	106
	0.660	0.340	
Total	1,294	634	1,928

Table A7.3: Descriptive statistics of DISCR