# Parental Information and Investments in Children's Human Capital*

Jalnidh Kaur†

November 16, 2023

**Abstract**

This paper studies the relationship between parental perceptions about children's performance and parental investment in children's human capital, and how this relationship evolves over the course of schooling. Using rich longitudinal data on investments, test scores, and parental assessments, I implement alternative specifications for the parental investment function that allow investment to depend on the entire history of lagged investment and inputs, account for past parental beliefs to circumvent reverse causality, and use household fixed effects to account for fixed characteristics at the household level. I find that compared to children with poor perceived performance, children with better perceived performance are up to 16 percentage points more likely to be enrolled in private as opposed to public schools, and receive up to 40% higher investment in schooling. This relationship intensifies as children progress from primary to secondary school. Results are robust across specifications, with evidence of complementarity between perceived ability and schooling. Within a household, parents' behavior is reinforcing, with more spent on the child believed to be the better performer. These findings inform our understanding of parental investment response and intra-household allocation of human capital investment decisions.
*JEL codes: C23; D83; I21; I26; J24*

*The author is grateful for support and advice from Alex Eble, Abhijeet Singh, Liz Ananat, Ying Xu, and participants at Economics of Education colloquium at Teachers College, Columbia University.

†Ph.D. Candidate, Teachers College, Columbia University. E-mail: jk4203@tc.columbia.edu.

# 1 Introduction

Parental investments are one of the core determinants of the formation of human capital, however, there are stark disparities in the amount of resources that parents invest in their children (Carneiro et al., 2013; Kalil, 2015; Huttenlocher et al., 2010; Hoff, 2003; Putnam, 2015). These disparities can exacerbate educational and income inequalities at the societal level and, as a result, have profound consequences for intergenerational mobility, especially in settings with high rates of poverty. This raises an obvious question: what drives these differences in parental investment decisions?

One key factor driving these investments is the expected returns to investment. Parents' investment behavior responds to changes in the economic returns to human capital in a wide range of contexts (Munshi and Rosenzweig, 2006; Jensen, 2013; Attanasio and Kaufman, 2015). Crucially, however, parental knowledge of these returns is imperfect (Jensen, 2010; Jensen, 2013) which may lead to misallocation in child-specific investments (Dizon-Ross, 2019; Bergman, 2021). Data from surveys across developed and developing countries indicate that heterogeneity in parental beliefs about child skill and productivity of various skill inputs strongly correlates with heterogeneity in parental investments (Cunha et al., 2013; Boneva and Rauh, 2018; List et al., 2021, Attanasio, Boneva and Rauh, 2020). Information frictions are a common theme across these findings, highlighting the role of parental information sets in influencing investment decisions.

In this paper, I advance our understanding of how parental information about children's academic performance shapes investment decisions in a low-income context, and how this relation evolves over the course of schooling. Much of our knowledge to date on the belief-investment link comes from one-off experiments in particular settings. For example, Dizon-Ross (2019) examines the link for primary school children in Malawi, Bergman (2021) studies that for 6th to 11th graders in the US, and Gan (2021) does so for 12th graders in

China. There is scant evidence, however, on the link between beliefs and investments in non-experimental settings with broader and more representative samples. Using observational data to answer these questions is important if information flows work differently under status quo compared to experimental conditions, leading experimental effects to be short-term deviations from business-as-usual (Jayaraman, Ray and Vericourt, 2016).

I characterize how the belief-investment link varies across different age-groups, across children from different socioeconomic backgrounds, and across different domains of investment. Are parents more or less responsive to perceptions of children's academic ability at certain ages than others? Does the link operate differently for different types of investment? Additionally, parental investment response may differ within households across siblings (Gratz and Torche, 2016; Yi et al., 2015). Parents may choose to invest more in children perceived to be higher ability, reinforcing inequality in the distribution of child ability, or invest more in children perceived to be lower ability, thus attenuating inequality. Characterizing the nature and slope of the relationship between perceived ability and investment is helpful for predicting household responses to policy shocks, and estimating policy parameters as opposed to ceteris paribus effects (Todd and Wolpin, 2003).

Using observational data to answer these questions poses methodological challenges. In the absence of researcher-induced variation in parental information, it is difficult to distinguish a mere correlation between parental beliefs and investments from a true causal effect. In this context, endogeneity can come in three forms: (1) omitted variables, since we do not observe all variables that factor into parental investment decisions; (2) reverse causality, if investment impacts parental beliefs about child ability and not vice-versa; and (3) measurement error in parental information measure, if parental reports about child ability deviate from truth due to lack of information or lack of truthful reporting. I tackle these challenges by estimating the relationship between information and investments under alternative estimators that deal with endogeneity by relying on different assumptions. First, I condition on lagged achievement to account for individual-specific heterogeneity and history of past inputs. This specification builds on value-added frameworks used in prior literature. Second, to address concerns regarding reverse causality, I condition on past parental beliefs. Third, to account for unobserved factors at the household level (time-varying or time-invariant), I use a family fixed effect strategy and exploit within-household variation in parental information across siblings. I also test robustness to accounting for past investments as well as parental preferences. In the spirit of earlier literature that uses observational data to answer policy questions (Todd and Wolpin, 2003, 2007; Fiorini and Keane, 2014), I examine results that are robust across specifications.

Using rich longitudinal data on children and their families from the Young Lives panel dataset for India, I characterize parental beliefs about children's academic performance in this setting, documenting how much parents know about their children's academic performance. Across ages and types of assessments, parental beliefs positively correlate with independent test scores indicating that parental assessments contain an informative signal about children's cognitive ability. I establish that parental beliefs are more inaccurate for disadvantaged groups: poorer, and less-educated parents have less accurate beliefs compared to richer, and educated parents, respectively, suggesting the presence of information frictions.

Next, I examine the extent to which parents act on what they know. I find that parental assessments of their children's performance predict investments at the extensive margin of school choice and intensive margin of educational expenditures. Compared to children with poor perceived performance, children with better perceived performance are 9 to 22 percentage points more likely to be enrolled in private as opposed to public schools. Parents who perceive their children to be in high performance categories spend between 20 and 60 percent more on costs related to their child's education. These patterns are consistent with higher expected returns, and are most salient when children are in middle school and in the penultimate year of high school, at 12 and 15 years of age.

Results from intra-household comparisons across siblings show similar patterns. At 12 years of age, estimates are statistically significant for the extensive but not the intensive margin. At 15 years of age, estimates are statistically significant across both margins of investment. The robustness of results to the inclusion of family fixed effects suggests that fixed characteristics at the household level (such as deep-seated parental preferences or time-varying shocks) are unlikely to be unobserved determinants of parental investment at 15 years of age. In addition, I find evidence for complementarity between perceived performance and schooling decisions. In other words, parents invest more in children who are perceived to be better performing.

Further, I examine the robustness of the estimates. The richness of the data allow me to conduct placebo checks and test alternative explanations that may bias estimated effects. I show that accounting for past investment, inertia in school choice, apart from supply-side dynamics at the household level does not drastically alter the magnitude and significance of estimates. Also, results are robust to accounting for extended time lag between survey rounds.

My study advances prior work on parental beliefs using observational data from the US and UK (Kinsler and Pavan, 2021; Attanasio, Boneva and Rauh, 2020; Cunha, 2014). I advance this work by showing patterns for the belief-investment link in a developing country setting. Information frictions are more severe in developing countries due to low levels of parental education, which makes it critical to understand the link with investment. I also improve on measures used for capturing parental investment. Existing observational work uses granular measures of parental effort (e.g. taking a child to library or museum) for which there may not be a consensus among parents. In contrast, I use educational expenditure and school choice as revealed preference measures of investment that are more widely accepted indicators of investment. Additionally, my work complements experimental work on parental beliefs and investments (Dizon-Ross, 2019; Bergman, 2021; Gan, 2021).

More broadly, I add to the literature on intra-household allocation of resources and parental responses to disparities in child endowments (see Almond and Currie (2011) and Almond and Mazumder (2013) for reviews). There is mixed evidence on whether parents employ a reinforcing strategy (Aizer and Cunha, 2012; Rosenzweig and Zhang, 2009), or a compensating strategy (Leight, 2017; Kinsler and Pavan, 2021). I add to this literature by documenting complementarity in the relationship between parental perceptions of child skill and parental investment response.

The rest of the paper proceeds as follows: section 2 describes the data, sections 3 and 4 present the empirical approach and results, sections 5 and 6 discuss robustness and mechanisms, and section 7 concludes.

## 2 Data and Variables

Young Lives (YL) is a longitudinal study of childhood poverty that follows two cohorts of children across four low-income countries, over five rounds, starting in 2002. I use data for the younger cohort from India consisting of 2,011 children and their families who were surveyed at ages 1, 5, 8, 12, and 15.[1] The sample is representative of three regions in the Indian state of Andhra Pradesh spanning seven districts, 20 sub-districts (*mandals*), comprising 98 communities, with *mandals* being the primary sampling units.[2] Table 1 presents the timing of rounds and the average age for children in each round.

A number of distinctive features make the Young Lives dataset suitable for examining the research questions. First, the longitudinal dimension allows examination of link at different ages for the same cohort. Second, the data contain child-specific measures of parental

[1]Indian children usually start primary schools at 6 years, and graduate from high school at around 17 years.

[2]Within each region, households were selected from a 'poor' and a 'non-poor' district, in addition to the state capital Hyderabad.

information and investments for *all* children aged 5-18 years in sampled households. Third, survey attrition was low. Total attrition from round 1 to round 5 was 6%.

Table 2 shows descriptive statistics on key child and household characteristics. The sample consists of mostly rural households with less than 5 years of average parental education. Disadvantaged socioeconomic groups (including scheduled castes, scheduled tribes and backward castes) comprise a major share of the sample. Around 60% of households earn less than $2 a day.[3] Monthly per capita expenditure is around 17 USD in 2006 (age 8 years) and 40 USD in 2016 (age 15 years). Below, I explain the measurement of key variables used in this paper.

**Parental Information:** The household survey collected parents' assessments about their children's academic performance. The specific question asked "How would you say your child is performing in school?" The question is asked for all children in the ages 5-18 years in the household who are currently enrolled in school.[4] Respondents are asked to rate the performance on a five-point scale with 01= Excellent, 02=Good, 03=Reasonably well, 04=Poorly, and 05=Very bad. The same question is asked in each round and is answered by the primary caregiver, which is the biological parent in 90% of the cases.[5] I use the response to the question as a measure of parental information about children's academic performance.[6] Figure 1 shows the univariate distribution for parental beliefs. 'Good' is the modal response, followed by 'Reasonably well'. The distributions are skewed reflecting optimistic beliefs about child ability. This is the generally observed pattern across ages. For simplicity, I

[3]Based on calculations in Attanasio et al. (2020).

[4]The parental response is a school-based reference point. The fieldworker instruction handbook (p42.) states: "... determine how well the child is performing at school. You may need to prompt the respondent by asking him/her to compare to how well other children of the same age are doing at the same school."

[5]For the last round (R5), the question is asked for currently enrolled index children and one sibling.

[6]Given that the question is asked only for school-going children, and the cohort attended pre-school in round 2 (age 5 years), I use parental information and investment data from round 2 onwards.

combine ‘04=Poorly’ and ‘05=Very bad’ into a single category, and use dummies for each parental assessment category as explanatory variables in the analysis.

**Parental Investment:** The main outcome variable is parental investment. I use two key measures of parental investment in children’s human capital: (1) private school enrolment, and (2) expenditure on school fees and extra tuition. Choice of private as opposed to public school is reflective of investment for two reasons: First, public schools are free, while private schools charge a fee. Second, there is a widespread perception that private schools provide better quality education (Central Square Foundation, 2020). To that extent, school choice is reflective of parental investment at the extensive margin. As a second measure of investment, I use expenditure incurred by parents on school fees and after-school tuition. Conditional on enrollment, this reflects investment at the intensive margin since private schools typically display wide heterogeneity in terms of price (Kingdon, 2020). The data contain disaggregated expenses on school fees and after-school tuition for rounds 4 and 5, while round 3 contains a composite measure combining both types of expenses. For the sake of comparison across rounds, I look at the aggregate expenses in all rounds.

**Test Scores:** Independent tests of achievement for math and language were administered in each round. These tests were administered individually during household visits and captured a wide variety of cognitive domains. It is important to note that parents cannot base their perceptions about children’s performance on these achievement measures for the following reasons. First, the survey administration protocols mandate that the sequencing of the questionnaire is adhered to, the test is administered to children after the parental module (covering belief questions) is completed. Second, the test is conducted in a private distraction-free setting within the home where parents may not be present. Third, no feedback is provided to the child or parent after the test. Thus, test score measures can be considered as independent and objective measures of student learning.

**Individual and Household controls:** Additionally, the survey also collected information on a range of individual and household characteristics which I use as controls in my analysis. These include child gender, birth order, maternal and paternal years of education, caste dummies, household size and wealth index. The wealth index is a measure of household wealth, constructed as an average of measures of housing quality, consumer durables, and access to services. I standardize the index for each round.

# 3 Empirical Strategy

Parental choices about investments in children's human capital are likely to depend on a series of factors including parental preferences, budget constraints, and beliefs about the effectiveness of investments. A fundamental challenge in examining the link between parental beliefs and investments using observational data is the endogeneity of parental assessments. In the absence of an exogenous source of variation in parental assessments about children's performance, parents' subjective beliefs may be picking up unobserved heterogeneity that is correlated with educational investments. I estimate the investment function using different empirical strategies that allow different parts of the data to come into play. Each estimation method attempts to handle endogeneity in a different way, relying on different assumptions. In this section, I present the empirical framework for estimation highlighting the assumptions and threats to identification for each strategy.

## 3.1 Conditional Exogeneity using Lagged Test Scores

The first specification borrows from the literature on value-added frameworks to account for lagged achievement. I estimate the following equation for each age-group using an OLS regression of outcome on parental assessment dummies, conditioning on child- and household-level background controls, as well as lagged test-score. Identification is reliant on the assumption that lagged achievement is sufficient to account for individual-specific heterogeneity and the full history of past investment and inputs. This assumption underlies most value-added models where results typically agree well with independent lottery-estimates and quasi-experimental designs (Chetty et al., 2014; Singh, 2015; Muralidharan and Sundararaman, 2013). I estimate the following equation:

$$Inv_{iha} = \beta_{0a} + \beta_{1a} PA_{iha} + \beta_{2a} T_{ih,a-1} + X_{iha} + \epsilon_{iha} \tag{1}$$

where, $PA_{iha}$ is a vector of dummies for parental assessments with "poor" as the reference category, corresponding to child $i$ in household $h$ at age $a$, $Inv_{iha}$ is an indicator variable equalling 1 if the child is enrolled in a private school and 0 if the child is enrolled in a government school. For the intensive margin investment outcome, $Inv_{iha}$ is a log transformation of expenditure on school fees and tuition incurred by the household. $T_{iha-1}$ is the standardized test-score on math cognitive test in the previous round, $X_{iha}$ is a vector of controls including child gender, parental years of education, household wealth, household size and caste dummies. Standard errors are clustered at the sub-district level which is the primary sampling unit. The specification exploits variation in parental assessments across students who are similar on observable characteristics, and past inputs captured by lagged test-scores. Estimation relies on an ignorability assumption - conditional on background controls and lagged test scores, parental assessments are orthogonal to other determinants of investment.

However, the direction of causality is not clear. It is plausible that parents update their beliefs after making investments, leading to bias in the estimation of $\beta_1$. To tackle the concern of reverse causality, I estimate a second specification regressing parental investment in current period on parental assessments from the *previous* round. I estimate the following lagged specification:

$$Inv_{iha} = \alpha_{0a} + \alpha_{1a} PA_{ih,a-1} + \alpha_{2a} T_{ih,a-1} + X_{iha} + \epsilon_{iha} \qquad (2)$$

where, $PA_{ih,a-1}$ is a vector of dummies for parental assessments with poor as the reference category, corresponding to the previous age at which the child is observed. I present estimates of $\hat{\alpha}_{1a}$ at 8, 12, and 15 years. The parental assessment measure is also available for the age of 5 years, when the child attends a pre-school.

The advantage of both contemporaneous and lagged specifications in equations (1) and (2) is that it is possible to estimate effects separately at each age i.e. the relation between parental beliefs and investments at 8, 12 and 15 years. This is important if the effects on investments vary at particular ages. The downside is that the potential for bias remains due to omitted variables that correlate with both parental information and investments, conditional on past scores and observed controls. For example, children with motivated parents may be more likely to attend private schools, and if this unobserved trait correlates with child skill, it will be reflected in parental assessments about the child. In this case, estimates of $\hat{\alpha}_{1a}$ will be upwardly biased on account of this omission. The following strategies attempt to deal with some of these sources of bias.

## 3.2 Household fixed effects

The second strategy exploits within-household variation in parental information across siblings. Given that households have multiple children, and parental assessments and investments are available for each child, I implement this strategy by employing a household fixed effects specification. I estimate both contemporaneous and lagged specifications.

$$Inv_{iha} = \beta_0 + \beta_{1a} PA_{ih,a} + \beta_{2a} Score_{ih,a-1} + X_{iha} + \gamma_h + \epsilon_{iha} \tag{3}$$

$$Inv_{iha} = \alpha_0 + \alpha_{1a} PA_{ih,a-1} + \alpha_{2a} Score_{ih,a-1} + X_{iha} + \gamma_h + \epsilon_{iha} \tag{4}$$

where, $\gamma_h$ is the family fixed effect, and $X_{iha}$ is a vector of sibling-varying controls including gender, birth-weight and BMI-for-age. In the previous framework, if households differ on unobservables that are correlated with both parental perceptions and investments, then a simple comparison to children across households will be biased. Including family fixed effect ensures that factors at the family/household level (either time-varying or time-invariant) are differenced out. The identifying assumption is that within the same household, parental perceptions of school performance across siblings are uncorrelated with unobserved determinants of investments.

This framework has two main advantages: first, it directly ties the estimation to intra-household allocation of resources which is the core channel that I want to estimate; second, it increases the sample sizes for the estimation of the main effects as the sample now includes both index children and their siblings. The limitation is that siblings, unlike the index children, are of different ages, due to which it is not possible to account for a comparable test-score from the previous round.

With a household fixed effects specification, the potential for bias stems from more specific type of confounders that are at the individual-level and differ across siblings within households.

For instance, if siblings differ from each other in terms of ability that is observed by parents but unobserved by the econometrician, and parents engage in favoritism, systematically investing more on abler kids, then estimates will have an upward bias. On the other hand, if parents compensate for disparities on unobserved trait by investing on less able kids, then estimates will be biased downwards.

## 3.3 Individual fixed effects

The third strategy exploits variation in parental information within children, over time. Adding individual fixed effects allows us to remove individual-specific fixed heterogeneity by pooling repeated observations for the same individual together. This enables a within-child comparison across waves. I estimate both contemporaneous and lagged specifications, conditioning on time-varying controls including wealth index, household size and residence location (urban/rural). The identifying assumption is that within individuals, parental perceptions of performance over time are uncorrelated with unobserved determinants of investments. The identifying variation comes from shifts in parental perceptions over time for children. While the advantage of this model is that it allows us to tackle omitted variables and unobservables at the level of the individual, it comes at a cost to power. The design has limited power given limited time dimensions (T = 4). Another downside is that we are able to estimate only one parameter for the cohort and cannot allow the effect to vary over age.

# 4 Results

## 4.1 Informativeness of parental assessments

To what extent are parental beliefs reflective of children's performance, as measured by objective measures of learning? The data contain measures of cognitive achievement based on math and language tests administered to children. Figure 2 plots normalized test scores against parental beliefs. I find that parental beliefs systematically relate to objective measures of achievement. Across rounds and types of assessments, average test score in both math and language increases incrementally for each point on the parental assessment scale, indicating meaningful variation in the belief measure that is informative about child's cognitive ability. This is corroborated by regression results in Table A1. Despite the fact that cognitive test scores are unobserved to parents, a one standard deviation increase in a child's math score is associated with 6 to 15 percentage point increase in the likelihood that a parent thinks the child is good or excellent.

An important limitation in this setting is that unlike related work in this literature, our parental belief measure is categorical, and not continuous. This has two implications. First, the belief measure amounts to being a coarse measure capturing a broad assessment about child's school performance and not a specific estimate about performance on a test. Second, it is not possible to measure belief accuracy or divergence in the traditional sense, given that there is no true score for comparison. While it is hard to characterize the absolute amount by which parents over- or under-estimate their children's ability, it is possible to compare relative levels of belief inaccuracies across subgroups. I examine this in the next section.

## 4.2 Heterogeneity in parental assessments

I test whether parental beliefs vary across socioeconomic status (as defined by being above or below the median on the wealth index) and maternal education (defined as the mother having some or no education).[7]

Figure 3 plots the non-parametric relationship between parental belief that the child is perceived as excellent and independent measure of achievement (test scores). To construct the figure, I divide math test scores into twenty equal size bins (vingtiles) and plot the mean value of parental belief in each bin, for each subgroup. This binned scatterplot represents the conditional expectation function for parental beliefs, across SES and maternal education. Table A2 shows the corresponding regression coefficients from specifications regressing the probability that the parental belief is excellent on an indicator for SES, standardized test score and the interaction between the two. The coefficient on the interaction indicates the extent to which the belief-score relationship is different across subgroups.

I find that across ages, the slope for the belief-score correspondence is flatter for children from low SES backgrounds as well as those with uneducated mothers. This is reflected in the regression results in Table A2, where the coefficient on the interaction is positive and statistically significant across ages, until 12 years. These patterns suggest that beliefs are more *attenuated* for disadvantaged groups compared to advantaged groups. Attenuation is a form of belief inaccuracy that occurs when beliefs are positively but imperfectly correlated with true performance. In this context, the results show that (a) uneducated parents have less accurate beliefs compared to educated parents, and (b) lower-SES parents have less accurate beliefs comapred to higher-SES parents. These results align with experimental findings from Dizon-Ross (2019).

[7] Note that 62% of children have mothers with no education, while the remaining 38% have mothers with some education, with the median child having 8 years of maternal education.

## 4.3 Parental information and investments

Given that parental assessments are informative about child ability, the next question is whether parents act on this information. I examine the link between parental information and investments using alternative specifications.

Figure 4 plots the raw private school enrolment gap by age between children with "poor" and "excellent" parental perceptions across ages. This gap is substantial at the age of 5 years when children are in pre-school, around 6 percentage points, and decreases over age, to around 4 percentage points, when children are 15 years old. The decline appears to be primarily driven by lowering share of private school enrolment for top performing children. However, decomposing by gender, panel B shows that poorly perceived male and female children experience different enrollment trends. While poorly perceived females see a fall in probability of private school enrollment over time, males who are poorly perceived see a rise, indicating a differential pattern of investment-belief relationship, by gender over time.

In Table 3, I present estimates of the belief-investment link for private school enrollment from across-household comparisons using contemporaneous and lagged specifications across 8, 12, and 15 years of age, corresponding to equations (1) and (2). Standard errors are clustered at the sub-district level. Contemporaneous parental assessments are highly predictive of private school enrollment across all observed ages, as seen in columns 1, 3, and 5. The likelihood of private school enrolment increases in perceived performance as seen in positive coefficients on parental belief dummies. Is this evidence of sorting into private/public schools based on perceived ability? The direction of causality is not clear. If students are systematically more likely to perform better at private schools compared to public schools, and this is accurately reflected in parental information sets, we would see a similar correlation between parental assessment and school type. Alternatively, parental beliefs about child performance and investment might be jointly determined.

To assuage these concerns, columns 2, 4, and 6 show results for investment regressed on parental assessment from previous round, while conditioning on lagged achievement, child and household background controls. I find that parental assessments about child performance matter at 12 years and 15 years of age, but not as much when children are younger. Compared to children with poor perceived performance, children with better perceived performance are 9 to 22 percentage points more likely to be enrolled in private school, from a reference group mean of 30 percent enrollment. The highest effects are seen for middle school – children who are perceived to be in the top category (excellent) are 66% more likely to be enrolled in private schools compared to the reference category (poor). At the bottom of the table, I report the F-test p-values testing for the equality of coefficients across parental belief categories. I find that the estimates are statistically distinguishable for the majority of pairwise comparisons in the last two rounds (ages 12 and 15 years).

Table 4 presents results for the intensive margin of investment. The dependent variable is child-specific educational expenditure incurred on school fees and after-school tuition. The expenditure measure (for all rounds) is right-skewed with a concentration of zeroes. Following Bellemare and Wichman (2020), I use inverse hyberbolic sine transformation since it approximates the natural logarithm and thus, (a) reduces the influence of outliers, and (b) allows retaining zero-valued observations. I model selection out of zero by conditioning on positive expenditure. A similar pattern holds for results on educational expenditures. Parental beliefs about child performance strongly predict educational expenses at 12 years and 15 years of age. Parents who perceive their children to be in high performance categories spend between 20 and 60 percent more on costs related to child's education. Educational expenditures increase monotonically in perceived performance of the child, consistent with higher expected returns. Decomposing the results for expenditure in Appendix Table A3, I find that the results are driven exclusively by the school fee component of total educational expenses, rather than extra tuition after school.

Next, I examine the parental investment response to changes in parental assessments in an intra-household setting comparing siblings within the same household. The average household has 2.57 children, the number of children ranging from from 1 to 10. The specification with household fixed effects takes into account all households with at least two children. Investment measures are not available for siblings at 8 years, so I present results at 12 and 15 years of age. Given that siblings are of different ages, it is not possible to normalize scores, so I use percentage correct scores instead. I present results without controlling for lagged test scores at 12 years, and with lagged test score control at 15 years.[8]

Table 5 reports estimation results using household fixed effects for both investment measures. The findings are qualitatively aligned with across-household comparisons, although the magnitudes are attenuated across ages and margins of investment. Within households, conditional on lagged test scores and gender, better parental beliefs correspond to a 6-16 percentage point higher likelihood of private school enrolment, and around 40% higher expenditure on education, compared to the reference category. At 12 years, estimates are statistically significant for the extensive but not the intensive margin.[9] At 15 years, estimates are statistically significant across both margins of investment. One potential reason for difference in results across ages is that parental investment choices are more high-stakes and consequential for entry into college, at the age of 15 years. Parental investment is likely to be more sensitive to parental perceptions at this age. On the other hand, middle school investment choices are not as high-stakes, so parental investment-information link is significant only for the extensive margin of school choice.

The robustness of results to the inclusion of family fixed effects suggests that fixed characteristics at the household level (such as deep-seated parental preferences or time-varying shocks) are unlikely to be unobserved determinants of parental investment at 15 years of

[8]Sibling-specific test scores are available for rounds 4 and 5, for math only.

[9]Note that the specification at 12 years is not directly comparable with results in Tables 3 and 4, as lagged test-score is not accounted for due to lack of data availability.

age. The positive sign of estimates points to a reinforcing strategy at play within households, where parents invest more in children perceived to be better performing. Another salient fact from Table 5 is that girls are less likely to be enrolled in private school, and also likely to receive lower educational spending, reflecting a gender bias in the setting.

I also examine results for individual fixed effects (Table A4) for the pooled sample including both index children and siblings. All specifications include time-varying controls including wealth index, household size, an indicator for whether the location is urban and round fixed effects. Estimates are lower in magnitude, around one-tenth the size of the raw gap in private school enrollment, and are not statistically significant. In making sense of these results, it is helpful to explore the source of the identifying variation, which comes from the sub-sample for whom parental perception shifted across rounds. Table A5 shows the matrices of parental perceptions for consecutive rounds, with the diagonal elements showing the percentage of children with unchanged parental perceptions across successive rounds. Most variation occurs in the adjacent categories (good/reasonably well) rather than the extremes (excellent/poor). For more than 40% of the sample, parental perceptions remain unchanged across consecutive rounds.[10] These stable differences across children may in part be attributable to systematic behaviors on the part of children (e.g. fixed ability of the child) or to features of parents that make them more likely to assess children in a particular way (e.g. reference points shaped by school or siblings).

Given lack of substantive variation in parental perceptions over time for children, the individual fixed effect model is underpowered to detect precise correlations. Placing these results in the context of earlier results, there are two ways to reconcile the differences. First, it is possible that the investment function operates differently across settings. While parental assessments of children's ability may not matter for an individual over time,

[10] 47% of children have the same parental perception across rounds 2 and 3, 45% across rounds 3 and 4, and 43% across rounds 4 and 5.

these matter for explaining differences in investments across siblings within households, and for children across similar households. Alternatively, it is possible that unobserved fixed traits about children which are observed by parents, and not necessarily by the econometrician drive changes in perceptions and investments. These results might be plausibly driven by confounding variables that follow a particular pattern - individual-specific, time-varying, sibling-varying unobservables that are correlated with both parental perceptions and investments.

# 5 Robustness Checks

While the results are consistent with our hypothesis that parental investments are determined by parental beliefs about child performance, there are alternative explanations for a link between the two. Here, I pursue a number of these alternative theories.

**Extended time-lag:** Given that time lag between successive survey rounds is around 3 years, the estimates at best identify the 3-year effect of parental assessments. A plausible concern is that the estimates are picking up the effect of factors during the intervening period that co-determine both parental information and investments. For example, if positive income shocks lead parents to move their children to private schools, following which parents adjust their beliefs upwards in the next period, then estimates in the current specifications will be picking up these effects operating via omitted variables and reverse causal pathway from school choice to parental beliefs. To deal with this concern, I make use of data on private school enrollment, on which year-level information is available. I regress outcome at year $r+1$ on parental assessment in year $r$, where $r$ is the year of survey round. The school choice information for year $r+1$ is available for years when children are 6, 8, and 13 years old respectively. Results from lagged specifications are in Table 6. The relation between parental assessments and school choice continues to be statistically significant, with slightly reduced magnitudes for coefficients on parental assessment indicators.

**Past Investment:** A second potential concern is that parental taste for investment might be correlated with parental skill, which in turn is correlated with child skill. Some parents might have a deep-seated taste or preference for schooling and investing in education. If heterogeneity in parental tastes is not accounted for, then our estimate of parental beliefs may be overestimated. In line with recommendation from Kinsler and Pavan (2021), I control for past investment to account for heterogeneous tastes. Tables 7 and 8 show results for private school enrollment and expenditures respectively. The proxy used for past investment is educational expenditure in the previous round. The patterns for significance of estimates are similar, though there is a slight reduction in magnitudes.

**Inertia in School Choice:** It is possible that school choice exhibits path dependence over time if parents who decide to send their child to a public school, stay in the public school in the absence of a major shock. As a stronger test of my hypothesis, I generate an indicator for switching (that takes the value 1 if the child switches across sectors, and 0 otherwise), and regress on RHS variables in (1). Table 9 shows age-wise regression results, for the full sample, subsample of students in public schools, and for subsample of students in private schools. First, I find that there is a significant amount of churning across years, as opposed to path dependence. Between 8% to 23% of children migrate across sectors in consecutive years. Second, results from age-wise regressions show a stark correlation between parental beliefs in current year and probability of switching in the following year, though the nature of correlations exhibits slightly different trends at each age. At 8 years, bad performance is strongly associated with switching next year, the direction of switching is driven by movement from public to private schools. At 12 years, poorly performing children in either school type are equally likely to switch in the following year. At 15 years, good performance is associated with switching, driven mainly by migration from public to private schools.

**Placebo Checks:** Investments are likely to be serially correlated over time within individuals. At the same time, it is possible that parents update their beliefs after observing investments. The combination of serial correlation of investments with dynamic belief updating could generate a spurious link between past beliefs and future investments. To test the seriousness of this concern, I employ a falsification test in which investments are regressed on the full vector of past, present, and future beliefs of parents. Table 10 shows results. I find that current investment is predicted by current and past beliefs, but not by future beliefs. Across all columns, the coefficient on future beliefs is statistically indistinguishable from zero indicating the absence of evidence against exogeneity of past beliefs with respect to the current investment.

**Supply-side factors:** Another concern is that school supply is likely to be a dominant factor affecting investment choices of parents and likely to be a key omitted variable. If parents who face a rich and diverse choice set of schools are more likely to engage in children's education, thereby having better information as well as making higher investments, then the correlation between parental beliefs and investments might be driven by omitted school supply. To address this concern, I include sub-district fixed effects in the main regressions.[11] These fixed effects control for dimensions of economic status, educational supply and infrastructure, as well as unobserved conditions within a smaller geographic unit that is also used as an entity for governance at the local level, thus capturing larger supply-side dynamics that are reflected

[11] The state of Andhra Pradesh is administratively divided into 23 districts and 1,125 sub-districts, with the sub-districts being the primary sampling units for the YL sample. Each sub-district contains between 20 and 40 villages.

in parental choice sets. I find that the main results are robust to the inclusion of these fixed effects. Results are in Tables A6 and A7.

# 6 Mechanisms and Discussion

There can be multiple pathways through which belief formation takes place. First, home and school environments might affect parental beliefs by providing information on relative ranking in the school/ neighbourhood distribution, or through comparison with siblings. Second, information directly shared by schools and children with parents is most likely to influence parental perceptions. If children choose to withhold information or frictions prevent parents from accessing this information, parental information may reflect distorted beliefs. Third, measurement error can affect parental perceptions if parents cannot reliably assess child's progress. Regardless of the origin of parental perceptions, these reflect a close proxy of parental information sets on which parental investment decisions are likely to be based.

While experimental studies better identify the causal effect of parental information on investments, they limit the spectrum of results to certain age-groups. Dizon-Ross (2019) shows in a different context (Malawi) that while school enrolment responds to changes in parental beliefs, there is no effect on school expenditures. In contrast, results in this paper for India show that parental perceptions predict a higher likelihood of enrollment in private schools as well as higher expenditure at 12 and 15 years of age. Within households, parents are more likely to invest more in children that are perceived to be higher-performing suggesting complementarity between perceived ability and schooling. In the intra-household setting, parental investment responds to parental perceptions at the school choice margin when children are in middle school, and responds at both the extensive and intensive margins at 15 years, when children are in high school and investments matter for college entry and labor market outcomes.

Note that the results have a causal interpretation as long as the key identifying assumptions are satisfied - ignorability of parental perceptions conditional on lagged test-scores and other observed determinants of investments. However, selection on unobservables cannot be ruled out completely. The results using household fixed effects suggest that while fixed characteristics at the household level are unlikely to confound the estimates, the possibility for individual-level fixed traits to confound the link between parental perceptions and investments, remains. For example, if parents perceive certain children to be more gifted than others and engage differently with children by being more strict or disciplined with them, then these individual-level parenting behaviors might influence both child skill and parental perceptions, as well as investments. Such a story is relatively less plausible in light of results from Berry et al. (2020) that suggests that parents exhibit an intra-household preference for equity in inputs between children. Other fixed traits at the level of the child which are not reflected in test-scores are more plausible confounders.

# 7 Conclusion

In this paper, I present evidence that parental beliefs about children's performance predict parental investment in children's human capital. While earlier literature examines this link using experiments for specific age-groups, I use observational data to uncover results for different age-groups, and add to our understanding of how the link operates across the full range of schooling. Parents who believe their child is a high performer invest more in child's education, and are more likely to enroll them in private, as opposed to public schools. This relationship intensifies as children progress from primary to secondary school. Additionaly, I find that parents view investments as complementary to child skill, investing more in children who are perceived to be better performing. These findings inform our understanding of parental investment response and intra-household allocation of human capital investment decisions, which has direct implications for inequality and inter-generational mobility at the societal level.

In sum, this work highlights the importance of parental beliefs in the household context, and also opens avenues for future research to examine the causal link between parental perceptions and investments by designing interventions that impact parental beliefs. This is directly relevant for policy; understanding how the link evolves over age is informative about critical education stages where providing information to parents can improve the demand for education.

# References


Attanasio, O., F. Cunha, and P. Jervis (2019). Subjective parental beliefs. their measurement and role. Technical report, *National Bureau of Economic Research.*

Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2020). Estimating the production function for human capital: results from a randomized controlled trial in Colombia. *American Economic Review* 110 (1), 48–85.

Attanasio, O. P., Kaufmann, K. M. (2014). Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics*, 109, 203-216.

Attanasio, O., Boneva, T., Rauh, C. (2020). Parental Beliefs about Returns to Different Types of Investments in School Children. *Journal of Human Resources*, 0719-10299R1.

Bellemare, M. F., Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, 82(1), 50-61.

Bergman, P. (2021). Parent-child information frictions and human capital investment: Evidence from a field experiment. *Journal of Political Economy*, 129(1), 286-322.

Boneva, T., and Christopher R. (2018). "Parental Beliefs about Returns to Educational Investments—The Later the Better?" *J. European Econ. Assoc.* 16 (6): 1669–711.

Bergman, P. (2021). Parent-child information frictions and human capital investment: Evidence from a field experiment. *Journal of Political Economy*, 129(1), 286-322.

Central Square Foundation. (2020). State of the Sector Report on Private Schools in India.

Cunha, F., I. Elo, and J. Culhane. (2013). "Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation." Working Paper no. 19144, NBER, Cambridge, MA.

Cunha, F. (2014). Gaps in early investments in children. Manuscript, Dept. Econ., Univ. Pennsylvania.

Dizon-Ross, R. (2019). Parents' beliefs about their children's academic ability: Implications for educational investments. American Economic Review, 109(8), 2728-65.

Gan, Tianqi (2021) How Parents' Beliefs About Their Children's Academic Ability Affect Educational Investments. Job Market Paper.

Grätz, M., Torche, F. (2016). Compensation or reinforcement? The stratification of parental responses to children's early ability. *Demography*, 53(6), 1883-1904.

Huttenlocher, J., H. Waterfall, M. Vasilyeva, J. Vevea, and L. V. Hedges (2010). Sources of variability in children's language growth. *Cognitive psychology* 61 (4), 343–365.

Hoff, E. (2003b). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development* 74 (5), 1368–1378.

Jayaraman, R., Ray, D., De Véricourt, F. (2016). Anatomy of a contract change. *American Economic Review*, 106(2), 316-58.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, 125(2), 515-548.

Jensen, R. (2012). Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *The Quarterly Journal of Economics*, 127(2), 753-792.

Jung, E. (2016). The development of reading skills in kindergarten influence of parental beliefs about school readiness, family activities, and children's attitudes to school. *International Journal of Early Childhood*, 48(1), 61-78.

Kalil, A. (2015). Inequality begins at home: The role of parenting in the diverging destinies of rich and poor children. In Families in an era of increasing inequality, pp. 63–82. Springer.

Kingdon, G. G. (2020). The private schooling phenomenon in India: A review. *The Journal of Development Studies*, 56(10), 1795-1817.

Kinsler, J., Pavan, R. (2021). Local distortions in parental beliefs over child skill. *Journal of Political Economy*, 129(1), 81-100.

List, J. A., Pernaudet, J., Suskind, D. (2021). It All Starts with Beliefs: Addressing the Roots of Educational Inequities by Shifting Parental Beliefs (No. w29394). *National Bureau of Economic Research.*

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.

Muralidharan, K., Sundararaman, V. (2015). The aggregate effect of school choice: Evidence from a two-stage experiment in India. *The Quarterly Journal of Economics*, 130(3), 1011-1066.

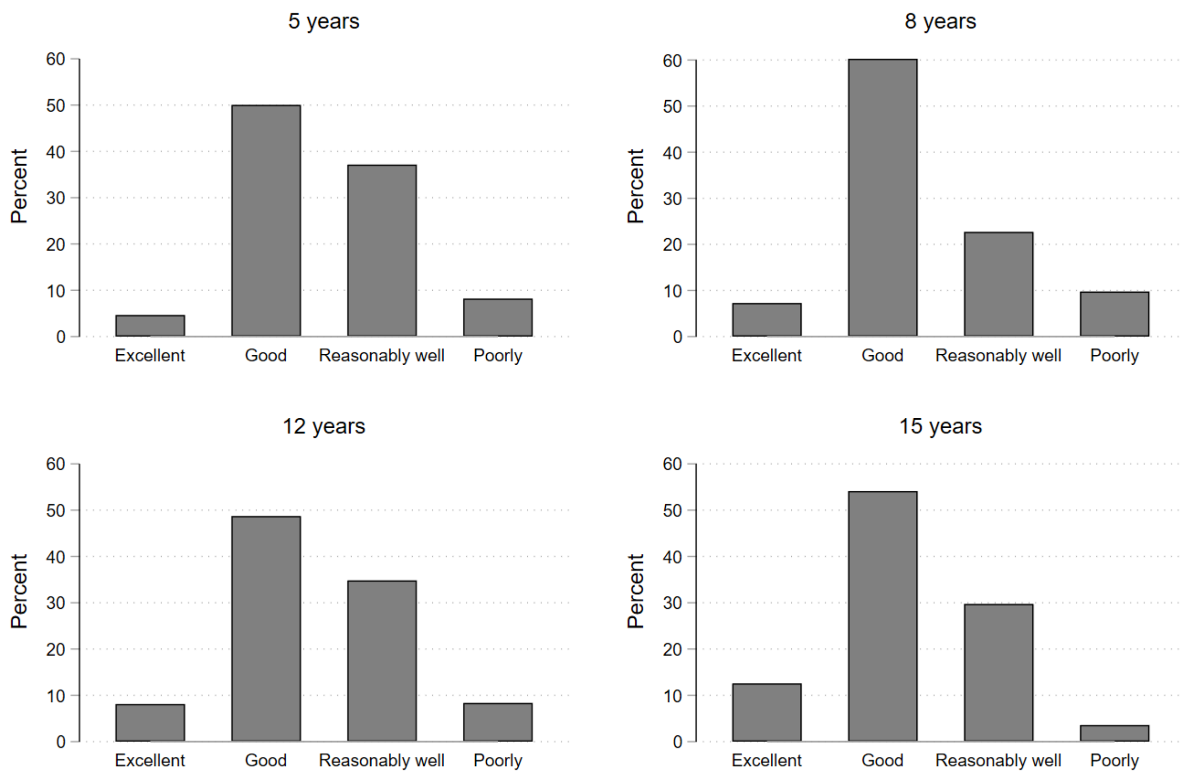
Rosenzweig, M. and J. Zhang, (2009). "Do population control policies induce more human capital investment? Twins, birth weight and China's 'one-child policy'," *The Review of Economic Studies*, 76, 1149–1174.

Singh, A. (2015). Private school effects in urban and rural India: Panel estimates at primary and secondary school ages. *Journal of Development Economics*, 113, 16-32.

Yi, J., Heckman, J. J., Zhang, J., Conti, G. (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal*, 125(588), F347-F371.

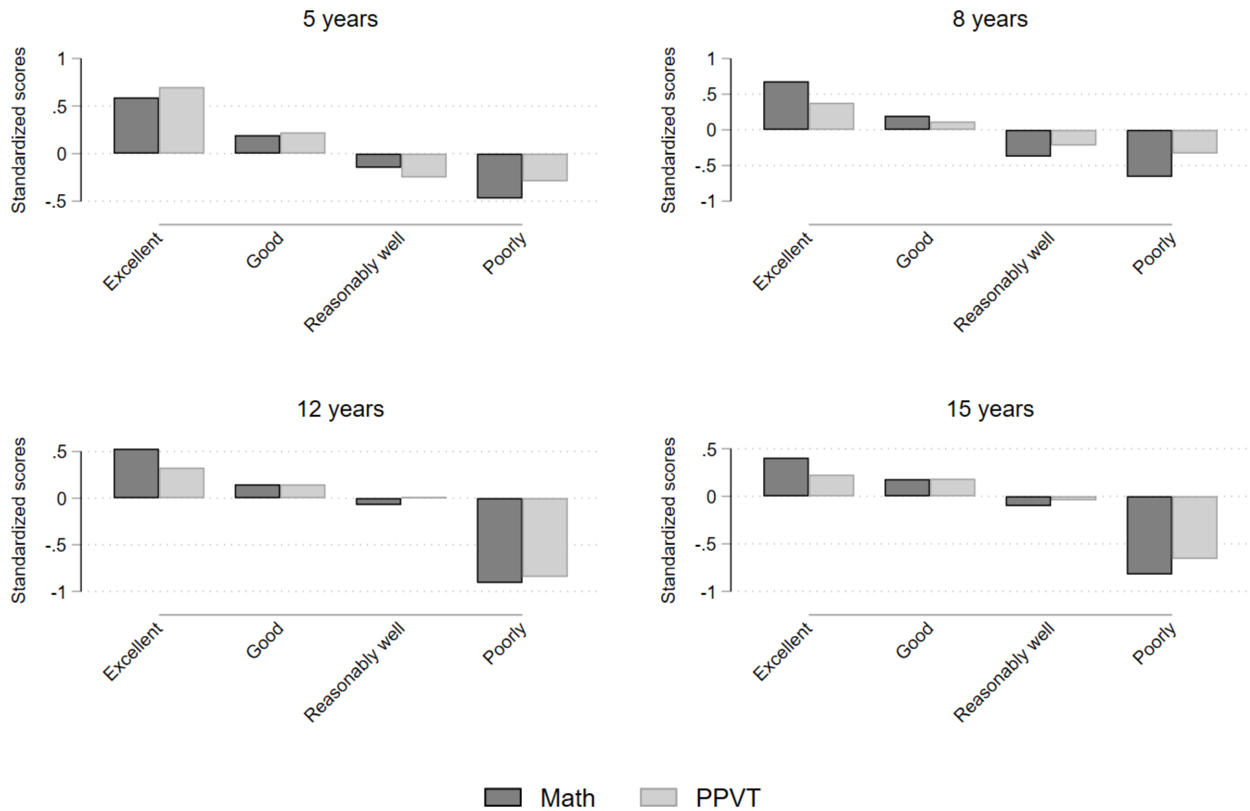
## Figures

### Figure 1: Parental belief distributions across ages




Notes: The figures show the univariate distribution of parental beliefs across ages using responses to the question 'How would you say the YL child is performing in school?' The primary caregiver is asked to rate the performance on a five-point scale with 01 = Excellent, 02 = Good, 03 = Reasonably well, 04 = Poorly, and 05 = Very Bad. The last two points on the scale are combined into a single category. There are 781, 1857, 1855, and 1683 responses in rounds 2, 3, 4, and 5 respectively. Given that the question is asked for only school-going children, and not all children are enrolled in school at age 5 years, N is low in the first panel.

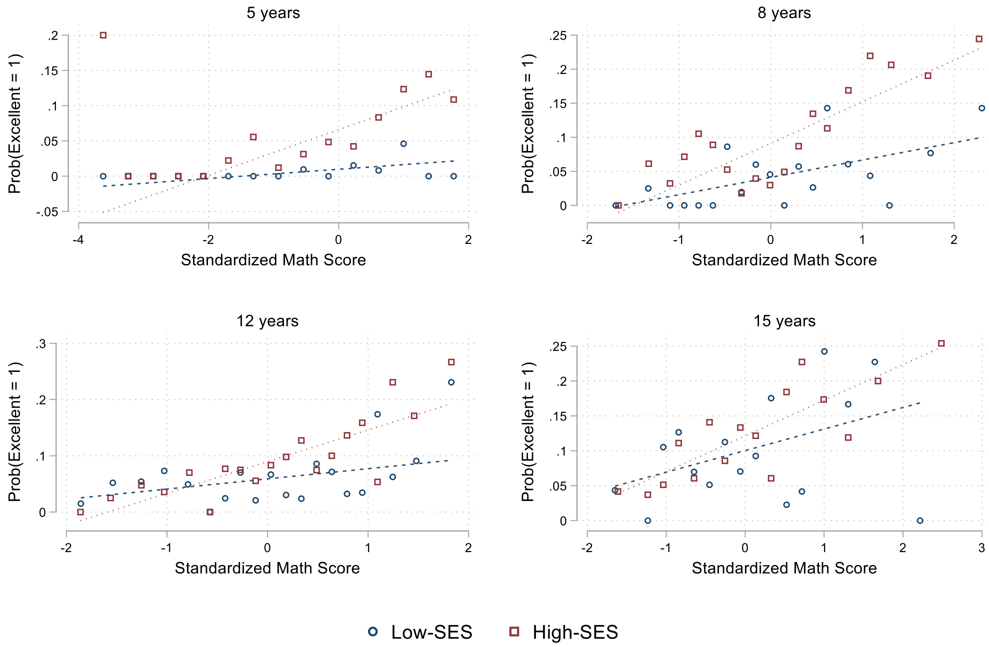
**Figure 2: Distribution of scores across parental belief categories, by age**



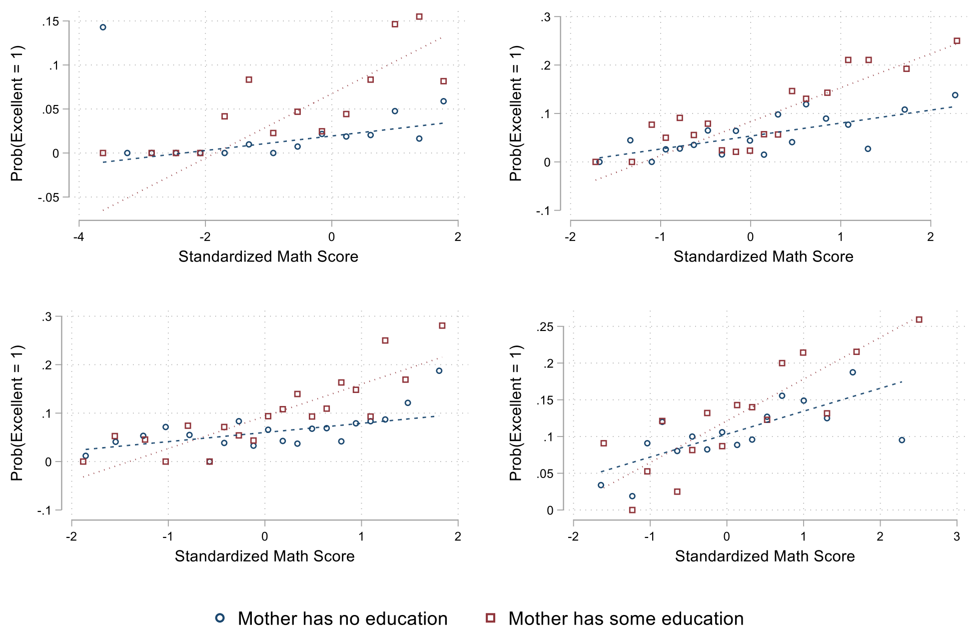

Notes: The figures plot standardized scores for Math and PPVT (Peabody Picture Vocabulary Test) against parental belief responses for all index children across ages 5, 8, 12, and 15 years. Scores are normalized within round (age) and correspond to performance on independently administered tests.

**Figure 3: Heterogeneity of parental beliefs across subgroups**

**Panel A: SES**



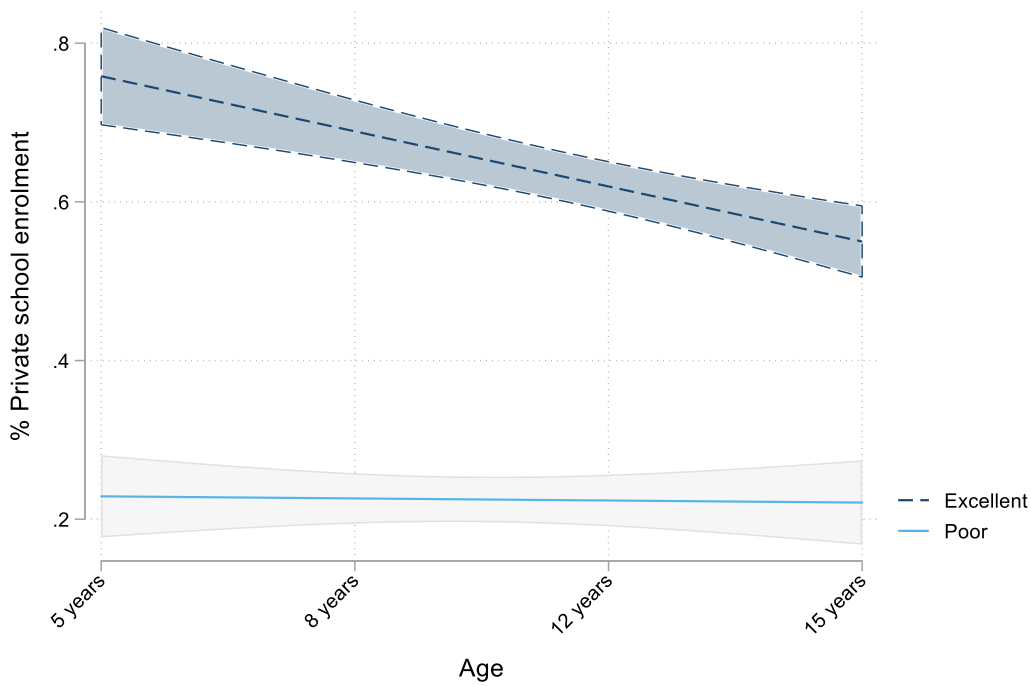

**Panel B: Maternal Education**



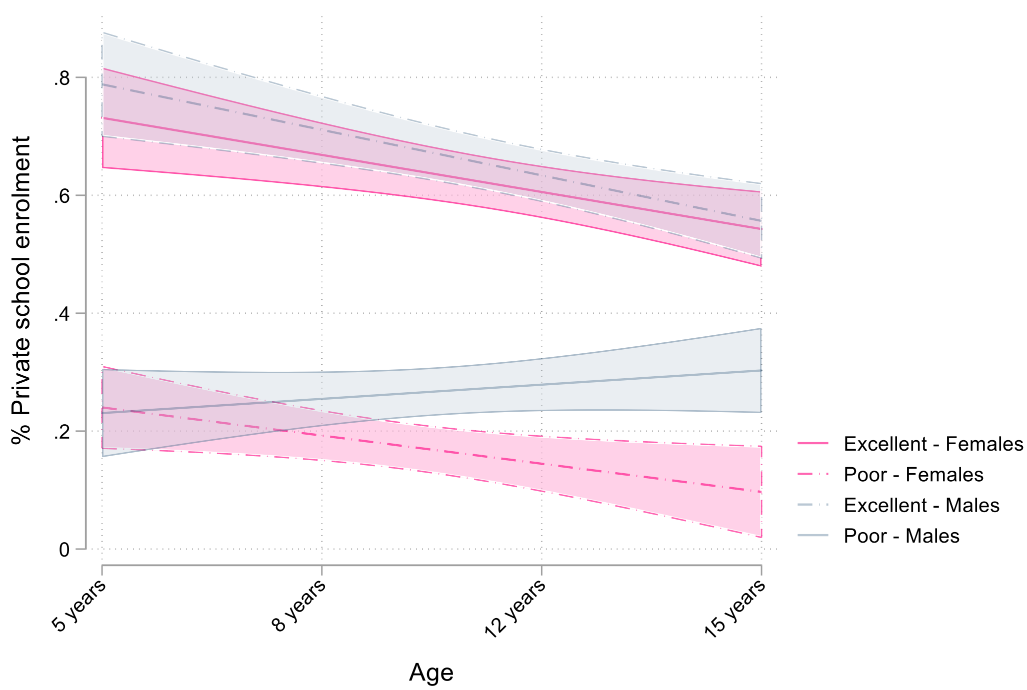

Notes: The figures show binned scatter plots showing the non-parametric conditional expectation function for parental belief against standardized math test scores, across subgroups. The y-axis plots the average probability that the parental belief is "excellent" as opposed to other values on the scale (good/well/poor), for each equal-sized bin of the standardized test score. The dotted lines show the best linear fit estimated using OLS. Panel A shows belief-score correspondence for high and low socioeconomic status (SES) groups. SES is defined as an indicator of being above or below the median on the household wealth index. Panel B shows the belief-score correspondence across children with educated and non-educated mothers. The plots correspond to regressions in columns (1), (3), (5) and (7) of Table A2, with standard errors clustered at the sub-district level.

**Figure 4: Private school enrollment by age and parental perception**

**Panel A: Pooled Sample**




**Panel B: Trends by gender**




Notes: Panel A plots the private school enrolment gap by age between children with poor and excellent parental perceptions. The shaded area shows 95% confidence intervals. Panel B shows private school enrolment trends by age across children with poor and excellent parental perceptions, disaggregated by gender.

## Tables

**Table 1: Average age in successive rounds**

| Year | Round | Average age at interview |
|---|---|---|
| 2002 | 1 | 1 year |
| 2006 | 2 | 5 years |
| 2009 | 3 | 8 years |
| 2013 | 4 | 12 years |
| 2016 | 5 | 15 years |

Notes: The table shows the calendar year and the average age at interview corresponding to each round (wave) of Young Lives survey for the younger cohort.

**Table 2: Descriptive Statistics**

| | Mean | SD |
|---|---|---|
| Mother's years of education | 2.99 | 4.20 |
| Father's years of education | 4.49 | 4.87 |
| Female | 0.46 | 0.50 |
| First-born | 0.39 | 0.49 |
| Scheduled caste | 0.18 | 0.39 |
| Scheduled tribe | 0.15 | 0.35 |
| Backward caste | 0.46 | 0.50 |
| Household size (2006) | 5.52 | 2.23 |
| Urban (2006) | 0.25 | 0.44 |
| Number of siblings (2016) | 2.35 | 1.36 |
| Monthly per capita expenditure (2006) | 16.94 | 12.05 |
| Monthly per capita expenditure (2016) | 40.39 | 36.94 |

Notes: The table shows summary statistics for children in the analytic sample (N = 1,942). Monthly per capita expenditure is reported in USD, adjusted for the year-specific exchange rate: at age 5 (2006), 1 USD ~= 45 INR, at age 15 (2016), 1 USD ~= 67 INR.

**Table 3: Parental beliefs and private school investment**

| | Round 3 (8 years) | | Round 4 (12 years) | | Round 5 (15 years) | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Parent Belief in current round* | | | | | | |
| Excellent | 0.17** | | 0.20*** | | 0.22*** | |
| | (0.063) | | (0.051) | | (0.060) | |
| Good | 0.09* | | 0.16*** | | 0.17*** | |
| | (0.044) | | (0.034) | | (0.051) | |
| Reasonably well | 0.10** | | 0.13*** | | 0.16** | |
| | (0.039) | | (0.032) | | (0.060) | |
| *Parent Belief in previous round* | | | | | | |
| Excellent | | 0.04 | | 0.22*** | | 0.12* |
| | | (0.042) | | (0.042) | | (0.059) |
| Good | | 0.01 | | 0.10*** | | 0.12** |
| | | (0.034) | | (0.036) | | (0.052) |
| Reasonably well | | -0.03 | | 0.11*** | | 0.09* |
| | | (0.042) | | (0.037) | | (0.044) |
| | | | | | | |
| Lagged test score (normalized) | 0.03*** | 0.04*** | -0.03* | -0.03* | 0.02 | 0.02 |
| | (0.010) | (0.012) | (0.016) | (0.015) | (0.018) | (0.017) |
| Constant | 0.39*** | 0.50*** | 0.30*** | 0.33*** | 0.24*** | 0.30*** |
| | (0.061) | (0.060) | (0.051) | (0.046) | (0.075) | (0.070) |
| | | | | | | |
| Observations | 1,836 | 1,665 | 1,806 | 1,756 | 1,640 | 1,647 |
| R-squared | 0.403 | 0.388 | 0.365 | 0.377 | 0.322 | 0.317 |
| | | | | | | |
| F-test p-values | | | | | | |
| Excellent = Good | 0.060 | 0.257 | 0.334 | 0.019 | 0.220 | 0.960 |
| Excellent = Well | 0.199 | 0.083 | 0.070 | 0.046 | 0.136 | 0.402 |
| Excellent = Poor | 0.016 | 0.305 | 0.001 | 0.000 | 0.002 | 0.057 |
| Good = Well | 0.481 | 0.196 | 0.182 | 0.819 | 0.421 | 0.308 |
| Good = Poor | 0.061 | 0.770 | 0.000 | 0.008 | 0.003 | 0.029 |
| Well = Poor | 0.017 | 0.427 | 0.001 | 0.008 | 0.019 | 0.060 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 The dependent variable is an indicator for private school enrollment of index children in the respective round. Parental beliefs are on a 4-point scale (excellent/good/well/poor), parental belief dummies are used an explanatory variables with "poor" as the reference category. All regressions control for wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. Standardized value of math test score in the previous round is used as proxy for lagged achievement. Scores are standardized for each age (round). Standard errors are clustered at the mandal (sub-district) level. The bottom of the table presents p-values from F-test of equality of coefficients.

**Table 4: Parental beliefs and school expenditures**

| | Round 3 (8 years) | | Round 4 (12 years) | | Round 5 (15 years) | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Parent Belief in current round* | | | | | | |
| Excellent | 0.34* | | 0.61*** | | 0.90*** | |
| | (0.193) | | (0.165) | | (0.221) | |
| Good | 0.14 | | 0.44** | | 0.79*** | |
| | (0.118) | | (0.165) | | (0.175) | |
| Reasonably well | 0.12 | | 0.27** | | 0.77*** | |
| | (0.118) | | (0.118) | | (0.198) | |
| *Parent Belief in previous round* | | | | | | |
| Excellent | | 0.12 | | 0.49*** | | 0.61*** |
| | | (0.204) | | (0.143) | | (0.212) |
| Good | | 0.02 | | 0.11 | | 0.53*** |
| | | (0.161) | | (0.111) | | (0.160) |
| Reasonably well | | -0.13 | | 0.19* | | 0.49*** |
| | | (0.178) | | (0.095) | | (0.132) |
| Lagged test score (normalized) | 0.05 | 0.04 | -0.08* | -0.06 | 0.18** | 0.18** |
| | (0.042) | (0.050) | (0.045) | (0.047) | (0.074) | (0.075) |
| Constant | 0.56*** | 0.76*** | -0.05 | 0.15 | -0.56** | -0.27 |
| | (0.157) | (0.197) | (0.183) | (0.147) | (0.237) | (0.193) |
| Observations | 1,712 | 1,523 | 1,815 | 1,765 | 1,630 | 1,625 |
| R-squared | 0.901 | 0.903 | 0.952 | 0.952 | 0.889 | 0.889 |
| F-test p-values | | | | | | |
| Excellent = Good | 0.201 | 0.613 | 0.150 | 0.004 | 0.297 | 0.509 |
| Excellent = Well | 0.264 | 0.285 | 0.005 | 0.022 | 0.210 | 0.463 |
| Excellent = Poor | 0.097 | 0.558 | 0.001 | 0.003 | 0.001 | 0.009 |
| Good = Well | 0.831 | 0.141 | 0.028 | 0.200 | 0.777 | 0.692 |
| Good = Poor | 0.265 | 0.899 | 0.015 | 0.341 | 0.000 | 0.004 |
| Well = Poor | 0.337 | 0.490 | 0.033 | 0.062 | 0.001 | 0.001 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. All regressions model selection out of zero by conditioning on positive expenditure. The dependent variable in all columns is the inverse hyperbolic sine transformation of the reported expenditure on education of the child that is composed of school fees and extra tuition fees, for the respective round. Parental beliefs are on a 4-point scale (excellent/good/well/poor), parental belief dummies are used an explanatory variables with "poor" as the reference category. All regressions control for wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. Standardized value of math test score in the previous round is used as proxy for lagged achievement. Scores are standardized for each age (round). Standard errors are clustered at the mandal (sub-district) level. The bottom of the table presents p-values from F-test of equality of coefficients.

## Table 5: Household Fixed Effects

| | Private School Enrollment (12 years) | | Educational Expenditure (12 years) | | Private School Enrollment (15 years) | | Educational Expenditure (15 years) | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Parent Belief in current round* | | | | | | | | |
| Excellent | 0.13*** | | 0.32*** | | 0.02 | | 0.17 | |
| | (0.033) | | (0.086) | | (0.073) | | (0.211) | |
| Good | 0.07*** | | 0.19*** | | 0.06 | | 0.23 | |
| | (0.024) | | (0.071) | | (0.063) | | (0.170) | |
| Reasonably well | 0.03 | | 0.04 | | 0.04 | | 0.03 | |
| | (0.023) | | (0.063) | | (0.064) | | (0.169) | |
| *Parent Belief in previous round* | | | | | | | | |
| Excellent | | 0.16*** | | 0.14 | | 0.09 | | 0.42** |
| | | (0.060) | | (0.138) | | (0.062) | | (0.172) |
| Good | | 0.08* | | 0.04 | | 0.10** | | 0.39*** |
| | | (0.045) | | (0.117) | | (0.050) | | (0.143) |
| Reasonably well | | 0.06 | | 0.00 | | 0.11** | | 0.42*** |
| | | (0.047) | | (0.114) | | (0.047) | | (0.136) |
| Female | -0.12*** | -0.12*** | -0.20*** | -0.20*** | -0.08*** | -0.08*** | -0.35*** | -0.31*** |
| | (0.015) | (0.018) | (0.037) | (0.046) | (0.022) | (0.022) | (0.072) | (0.073) |
| Constant | 0.41*** | 0.39*** | 0.16 | 0.24 | 0.38*** | 0.35*** | -0.24 | -0.42 |
| | (0.021) | (0.042) | (0.122) | (0.168) | (0.064) | (0.047) | (0.518) | (0.480) |
| Observations | 3,678 | 3,074 | 3,705 | 3,101 | 2,573 | 2,550 | 2,578 | 2,543 |
| R-squared | 0.050 | 0.050 | 0.920 | 0.920 | 0.024 | 0.029 | 0.885 | 0.884 |
| No. of households | 1,873 | 1,797 | 1,883 | 1,811 | 1,721 | 1,724 | 1,725 | 1,717 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 The dependent variable is columns 1, 2, 5, and 6 is an indicator for private school enrollment. The dependent variable in columns 3, 4, 7, and 8 is the inverse hyperbolic sine transformation of educational expenditure on school fees and extra tuition. All specifications control for sibling-varying observables including birthweight, gender and BMI-for-age. Sibling scores are available at ages 12 and 15 years only and not at 8 years. Specifications in columns (1)-(4) do not use lagged test score as test score is not available for sibling at 8 years. Cols (5)-(8) control for percentage correct version of lagged test score which is available for both index children and siblings. Scores are not standardized as siblings are of different ages. All expenditure regressions condition on enrolment and positive fees. Standard errors are clustered at the household level.

**Table 6: Robustness for consecutive years**

| | Year 2007-08 (6 years old) | Year 2010-11 (9 years old) | Year 2014-15 (13 years old) |
|---|---|---|---|
| | (1) | (2) | (3) |
| *Parent Belief in previous year (same as round-year)* | | | |
| Excellent | 0.03 | 0.13*** | 0.10** |
| | (0.054) | (0.045) | (0.046) |
| Good | 0.05 | 0.06** | 0.10*** |
| | (0.035) | (0.030) | (0.032) |
| Reasonably well | -0.02 | 0.05 | 0.09*** |
| | (0.034) | (0.031) | (0.031) |
| Lagged test score (normalized) | 0.03*** | 0.00 | 0.04*** |
| | (0.010) | (0.011) | (0.011) |
| Constant | 0.46*** | 0.43*** | 0.31*** |
| | (0.048) | (0.042) | (0.046) |
| Observations | 1,636 | 1,838 | 1,816 |
| R-squared | 0.469 | 0.431 | 0.420 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is an indicator for private school enrolment of index children. All regressions control for wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. Standard errors are clustered at the mandal (sub-district) level.

**Table 7: Robustness to past investment (Private School Enrollment)**

| | Round 4 (12 years old) | | Round 5 (16 years old) | |
|---|---|---|---|---|
| *Parent Belief in current round* | | | | |
| Excellent | 0.13*** | | 0.22*** | |
| | (0.045) | | (0.049) | |
| Good | 0.10*** | | 0.15*** | |
| | (0.034) | | (0.044) | |
| Reasonably well | 0.09*** | | 0.14*** | |
| | (0.034) | | (0.044) | |
| *Parent Belief in previous round* | | | | |
| Excellent | | 0.13*** | | 0.08* |
| | | (0.042) | | (0.044) |
| Good | | 0.06** | | 0.11*** |
| | | (0.028) | | (0.036) |
| Reasonably well | | 0.08*** | | 0.09** |
| | | (0.030) | | (0.035) |
| Education expenditure in previous round | -0.02** | -0.02** | 0.01 | 0.01 |
| | (0.010) | (0.010) | (0.010) | (0.011) |
| Lagged test score (normalized) | 0.05*** | 0.05*** | 0.05*** | 0.05*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Constant | 0.10** | 0.13*** | 0.03 | 0.08* |
| | (0.047) | (0.043) | (0.056) | (0.049) |
| Observations | 1,690 | 1,688 | 1,636 | 1,647 |
| R-squared | 0.500 | 0.498 | 0.474 | 0.471 |

Notes: Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The dependent variable is an indicator for private school enrollment. All regressions control for wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. For round 3, lagged expenditure (corresponding to round 2) is not available. Standard errors are clustered at the mandal (sub-district) level.

**Table 8: Robustness to past investment (Educational Expenditures)**

| | Round 4 (12 years old) | | Round 5 (16 years old) | |
|---|---|---|---|---|
| *Parent Belief in current round* | | | | |
| Excellent | 0.46*** | | 0.89*** | |
| | (0.133) | | (0.196) | |
| Good | 0.31*** | | 0.67*** | |
| | (0.112) | | (0.177) | |
| Reasonably well | 0.18 | | 0.64*** | |
| | (0.115) | | (0.181) | |
| *Parent Belief in previous round* | | | | |
| Excellent | | 0.32*** | | 0.47*** |
| | | (0.110) | | (0.179) |
| Good | | 0.03 | | 0.47*** |
| | | (0.072) | | (0.144) |
| Reasonably well | | 0.11 | | 0.44*** |
| | | (0.081) | | (0.146) |
| Education expenditure in previous round | 0.10*** | 0.11*** | 0.18*** | 0.18*** |
| | (0.010) | (0.010) | (0.011) | (0.011) |
| Lagged test score (normalized) | -0.05** | -0.04 | 0.14*** | 0.15*** |
| | (0.026) | (0.027) | (0.040) | (0.041) |
| Constant | -0.27** | -0.12 | -0.90*** | -0.67*** |
| | (0.135) | (0.112) | (0.214) | (0.185) |
| Observations | 1,700 | 1,698 | 1,625 | 1,625 |
| R-squared | 0.960 | 0.960 | 0.910 | 0.910 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the arc sine transformation of educational expenditure on school fees and extra tuition for index children. All regressions control for wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. Standard errors are clustered at the mandal (sub-district) level.

**Table 9: Switching across sectors and parental beliefs**

| | 5 years | | | 8 years | | | 12 years | | |
|---|---|---|---|---|---|---|---|---|---|
| | All | Public | Private | All | Public | Private | All | Public | Private |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| *Parent Belief in previous round* | | | | | | | | | |
| Excellent | -0.06** | -0.09 | -0.02 | -0.00 | 0.02 | -0.05 | 0.03 | -0.00 | 0.01 |
| | (0.027) | (0.075) | (0.059) | (0.031) | (0.049) | (0.045) | (0.027) | (0.022) | (0.073) |
| Good | -0.05** | -0.06* | -0.03 | 0.01 | -0.01 | -0.02 | 0.05*** | 0.02* | 0.03 |
| | (0.024) | (0.028) | (0.058) | (0.022) | (0.028) | (0.042) | (0.015) | (0.012) | (0.070) |
| Reasonably well | -0.05** | -0.07** | -0.03 | 0.03 | 0.01 | 0.01 | 0.04** | 0.02* | 0.04 |
| | (0.023) | (0.025) | (0.062) | (0.020) | (0.026) | (0.034) | (0.015) | (0.009) | (0.074) |
| Constant | 0.12** | 0.21*** | 0.07 | 0.05* | 0.11** | 0.12** | 0.04 | 0.03 | 0.19** |
| | (0.044) | (0.059) | (0.064) | (0.030) | (0.042) | (0.048) | (0.029) | (0.032) | (0.092) |
| Observations | 1,476 | 894 | 582 | 1,838 | 1,023 | 815 | 1,790 | 1,050 | 740 |
| R-squared | 0.017 | 0.069 | 0.070 | 0.013 | 0.027 | 0.084 | 0.008 | 0.017 | 0.130 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is an indicator for switching that takes the value 1 if a child switches from private to public school or from public to private school. Columns (2), (5), and (8) restrict to children who were in public school in given round, while columns (3), (6), and (9) restrict to children in private schools in given round. All regressions control for wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. For round 3, lagged expenditure (corresponding to round 2) is not available. Standard errors are clustered at the mandal (sub-district) level.

**Table 10: Placebo Checks**

| | Private School Enrollment | | | Educational Expenses | | |
|---|---|---|---|---|---|---|
| | 8 years | 12 years | 15 years | 8 years | 12 years | 15 years |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Parental belief at 5 years* | | | | | | |
| Excellent | 0.07 | 0.09 | 0.21*** | 0.12 | 0.17 | 0.28 |
| | (0.045) | (0.053) | (0.056) | (0.202) | (0.157) | (0.194) |
| Good | 0.05 | 0.05 | 0.15*** | 0.07 | -0.10 | 0.05 |
| | (0.030) | (0.028) | (0.032) | (0.173) | (0.133) | (0.152) |
| Reasonably well | -0.01 | -0.02 | 0.04 | -0.09 | -0.10 | -0.33** |
| | (0.040) | (0.032) | (0.035) | (0.196) | (0.125) | (0.131) |
| *Parental belief at 8 years* | | | | | | |
| Excellent | 0.15*** | 0.18*** | 0.13* | 0.35 | 0.42** | 0.53** |
| | (0.054) | (0.053) | (0.067) | (0.227) | (0.147) | (0.205) |
| Good | 0.08* | 0.08** | 0.06 | 0.16 | 0.07 | 0.06 |
| | (0.042) | (0.037) | (0.053) | (0.159) | (0.129) | (0.158) |
| Reasonably well | 0.09* | 0.09* | 0.04 | 0.01 | 0.15 | 0.15 |
| | (0.046) | (0.046) | (0.055) | (0.166) | (0.106) | (0.179) |
| *Parental belief at 12 years* | | | | | | |
| Excellent | 0.06 | 0.17*** | 0.10* | 0.19 | 0.53** | 0.56* |
| | (0.043) | (0.041) | (0.050) | (0.205) | (0.201) | (0.272) |
| Good | 0.06 | 0.12*** | 0.09 | 0.24 | 0.35* | 0.44** |
| | (0.041) | (0.034) | (0.055) | (0.183) | (0.170) | (0.209) |
| Reasonably well | 0.03 | 0.10** | 0.08** | 0.26 | 0.18 | 0.48*** |
| | (0.040) | (0.036) | (0.029) | (0.182) | (0.150) | (0.151) |
| *Parental belief at 15 years* | | | | | | |
| Excellent | 0.02 | 0.05 | 0.12 | -0.40 | 0.14 | 0.48 |
| | (0.056) | (0.071) | (0.079) | (0.277) | (0.136) | (0.288) |
| Good | 0.03 | 0.08 | 0.08 | -0.28 | 0.17 | 0.36 |
| | (0.048) | (0.075) | (0.067) | (0.288) | (0.150) | (0.256) |
| Reasonably well | 0.02 | 0.09 | 0.10 | -0.21 | 0.07 | 0.43 |
| | (0.053) | (0.074) | (0.071) | (0.284) | (0.138) | (0.253) |
| Constant | 0.32*** | 0.10 | 0.06 | 0.47 | -0.22 | -0.91*** |
| | (0.071) | (0.074) | (0.083) | (0.301) | (0.262) | (0.299) |
| Observations | 1,424 | 1,415 | 1,408 | 1,334 | 1,423 | 1,397 |
| R-squared | 0.358 | 0.305 | 0.263 | 0.894 | 0.952 | 0.882 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is indicator for private school enrollment, for columns (1) to (3), and aggregate educational expenses for columns (4) to (6). All regressions control for wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. Standard errors are clustered at the mandal (sub-district) level.

**Table A1: Parental Beliefs and Math Scores**

| | Parental belief about child performance | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 5 years | | | 8 years | | | 12 years | | | 15 years | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Math score | 0.065*** | 0.049*** | 0.060*** | 0.149*** | 0.139*** | 0.123*** | 0.092*** | 0.082*** | 0.097*** | 0.078*** | 0.078*** | 0.094*** |
| | (0.016) | (0.016) | (0.013) | (0.012) | (0.014) | (0.014) | (0.016) | (0.015) | (0.016) | (0.014) | (0.013) | (0.015) |
| Observations | 1,626 | 1,605 | 1,605 | 1,824 | 1,803 | 1,803 | 1,815 | 1,795 | 1,795 | 1,666 | 1,652 | 1,652 |
| R-squared | 0.077 | 0.127 | 0.196 | 0.110 | 0.125 | 0.168 | 0.046 | 0.062 | 0.127 | 0.044 | 0.060 | 0.105 |
| Demographics | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Cluster FE | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |
| Mean belief | 0.546*** | 0.486*** | 0.487*** | 0.674*** | 0.648*** | 0.638*** | 0.572*** | 0.519*** | 0.489*** | 0.657*** | 0.573*** | 0.579*** |
| | (0.035) | (0.051) | (0.034) | (0.021) | (0.041) | (0.032) | (0.028) | (0.058) | (0.049) | (0.022) | (0.051) | (0.038) |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is an indicator that parents believe that their child's school performance is good or excellent. Demographic controls include wealth index, parental years of education, child gender, household size, caste dummies, and an indicator for whether the child is the eldest child. All columns control for language test score. Child math score is a standardized score from each round of Young Lives survey. Standard errors are clustered at the mandal (sub-district) level.

**Table A2: Heterogeneity in parental beliefs across subgroups**

**Panel A: SES (Wealth Index)**

| | 5 years | | 8 years | | 12 years | | 15 years | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| High-SES*Math score | 0.03** | 0.02* | 0.04** | 0.03** | 0.04*** | 0.03** | 0.02 | 0.01 |
| | (0.010) | (0.010) | (0.016) | (0.014) | (0.013) | (0.012) | (0.017) | (0.018) |
| Math score | 0.01* | 0.00 | 0.03** | 0.02** | 0.02 | 0.02 | 0.03** | 0.03** |
| | (0.004) | (0.004) | (0.010) | (0.011) | (0.012) | (0.012) | (0.012) | (0.013) |
| High-SES (Wealth index > median) | 0.06*** | 0.03*** | 0.05** | 0.04** | 0.03* | 0.01 | 0.02 | 0.01 |
| | (0.013) | (0.011) | (0.020) | (0.013) | (0.014) | (0.014) | (0.019) | (0.021) |
| Constant | 0.03*** | 0.03 | 0.05** | 0.07** | 0.01 | 0.02 | 0.04 | 0.03 |
| | (0.010) | (0.016) | (0.017) | (0.029) | (0.018) | (0.028) | (0.028) | (0.028) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 1,918 | 1,918 | 1,897 | 1,897 | 1,850 | 1,850 | 1,833 | 1,833 |
| R-squared | 0.040 | 0.054 | 0.049 | 0.056 | 0.037 | 0.045 | 0.028 | 0.038 |

**Panel B: Mother's Education**

| | 5 years | | 8 years | | 12 years | | 15 years | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Educated-mother*Math score | 0.03** | 0.02** | 0.04** | 0.04** | 0.05*** | 0.05*** | 0.03* | 0.02 |
| | (0.010) | (0.009) | (0.016) | (0.016) | (0.013) | (0.014) | (0.015) | (0.016) |
| Math score | 0.01** | 0.01** | 0.03*** | 0.02** | 0.02* | 0.02* | 0.03*** | 0.03** |
| | (0.003) | (0.003) | (0.009) | (0.009) | (0.011) | (0.010) | (0.010) | (0.011) |
| Educated-mother | 0.05*** | -0.05** | 0.03 | -0.05 | 0.03* | 0.02 | 0.02 | -0.03 |
| | (0.010) | (0.019) | (0.025) | (0.039) | (0.018) | (0.036) | (0.016) | (0.052) |
| Constant | 0.04*** | 0.04** | 0.06*** | 0.09*** | 0.01 | 0.03 | 0.04 | 0.03 |
| | (0.011) | (0.018) | (0.015) | (0.030) | (0.018) | (0.029) | (0.032) | (0.030) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 1,918 | 1,918 | 1,897 | 1,897 | 1,850 | 1,850 | 1,833 | 1,833 |
| R-squared | 0.036 | 0.052 | 0.046 | 0.056 | 0.041 | 0.047 | 0.029 | 0.039 |

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is an indicator for whether the parent reports their child's school performance to be excellent. High-SES is an indicator for whether the household's wealth index is above the median. Educated-mother is an indicator for whether the mother has some education as opposed to no education. Math test score is standardized within each age (round). All regressions control for wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. Standard errors are clustered at the mandal (sub-district) level.

**Table A3: Parental perceptions and expenditure (decomposed into school fees and tuition)**

| | Round 4 (12 years) | | | Round 5 (15 years) | | |
|---|---|---|---|---|---|---|
| | Combined Expenses | School Fees | Extra tuition | Combined Expenses | School Fees | Extra tuition |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Parent Belief in Previous Round* | | | | | | |
| Excellent | 0.49*** | 0.99*** | -0.09 | 0.61*** | 0.89** | -0.58 |
| | (0.143) | (0.291) | (0.221) | (0.212) | (0.338) | (0.355) |
| Good | 0.11 | 0.30 | 0.13 | 0.53*** | 0.73** | -0.52 |
| | (0.111) | (0.213) | (0.203) | (0.160) | (0.309) | (0.367) |
| Reasonably well | 0.19* | 0.54** | -0.04 | 0.49*** | 0.68** | -0.30 |
| | (0.095) | (0.199) | (0.143) | (0.132) | (0.255) | (0.284) |
| Constant | 0.15 | -0.04 | 0.06 | -0.27 | -0.55* | 0.54 |
| | (0.147) | (0.223) | (0.250) | (0.193) | (0.281) | (0.412) |
| Observations | 1,765 | 1,765 | 1,765 | 1,625 | 1,626 | 1,626 |
| R-squared | 0.952 | 0.821 | 0.187 | 0.889 | 0.808 | 0.088 |

Notes: The dependent variable in all columns is the inverse hyperbolic sine transformation of the reported expenditure, columns (1) and (4) report aggregate expenses on school fees and extra tuition. No decomposed expenditure measures are available for round 3 (8 years). All regressions control for lagged test score, wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

## Table A4: Individual Fixed Effects

| | Private school enrollment | | Educational Expenditure | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Parent Belief in current round* | | | | |
| Excellent | 0.06** | | 0.23** | |
| | (0.027) | | (0.102) | |
| Good | 0.02 | | 0.08 | |
| | (0.021) | | (0.083) | |
| Reasonably well | 0.03 | | 0.01 | |
| | (0.021) | | (0.081) | |
| *Parent Belief in previous round* | | | | |
| Excellent | | 0.01 | | 0.04 |
| | | (0.027) | | (0.103) |
| Good | | 0.02 | | 0.02 |
| | | (0.019) | | (0.076) |
| Reasonably well | | 0.02 | | 0.11 |
| | | (0.019) | | (0.077) |
| Lagged test score (percentage correct) | 0.00** | 0.00** | 0.01*** | 0.01*** |
| | (0.000) | (0.000) | (0.002) | (0.002) |
| Constant | 0.35*** | 0.41*** | 0.70*** | 0.72*** |
| | (0.046) | (0.047) | (0.234) | (0.240) |
| Observations | 6,218 | 5,974 | 6,176 | 5,897 |
| R-squared | 0.014 | 0.015 | 0.870 | 0.873 |
| Number of individuals | 2,840 | 2,793 | 2,837 | 2,793 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The sample includes all children including index children and their siblings. Expenditure regressions condition on positive fees to model selection out of zero. All specifications include time-varying controls including wealth index, household size, an indicator for whether the location is urban, and round fixed effects.

**Table A5: Parental perceptions across consecutive rounds**

| | | | 8 years | | | | |
|---|---|---|---|---|---|---|---|
| | | | Poor | Reasonably well | Good | Excellent | Total |
| 5 years | Poor | N | 19 | 61 | 97 | 4 | 181 |
| | | Row % | 10.5 | 33.7 | 53.59 | 2.21 | 100 |
| | Reasonably well | N | 85 | 215 | 596 | 62 | 958 |
| | | Row % | 8.87 | 22.44 | 62.21 | 6.47 | 100 |
| | Good | N | 98 | 256 | 1,048 | 113 | 1,515 |
| | | Row % | 6.47 | 16.9 | 69.17 | 7.46 | 100 |
| | Excellent | N | 2 | 22 | 101 | 32 | 157 |
| | | Row % | 1.27 | 14.01 | 64.33 | 20.38 | 100 |
| | Total | N | 204 | 554 | 1,842 | 211 | 2,811 |
| | | Row % | 7.26 | 19.71 | 65.53 | 7.51 | 100 |

| | | | 12 years | | | | |
|---|---|---|---|---|---|---|---|
| | | | Poor | Reasonably well | Good | Excellent | Total |
| 8 years | Poor | N | 47 | 88 | 107 | 14 | 256 |
| | | Row % | 18.36 | 34.38 | 41.8 | 5.47 | 100 |
| | Reasonably well | N | 87 | 290 | 293 | 39 | 709 |
| | | Row % | 12.27 | 40.9 | 41.33 | 5.5 | 100 |
| | Good | N | 114 | 646 | 995 | 163 | 1,918 |
| | | Row % | 5.94 | 33.68 | 51.88 | 8.5 | 100 |
| | Excellent | N | 5 | 46 | 115 | 46 | 212 |
| | | Row % | 2.36 | 21.7 | 54.25 | 21.7 | 100 |
| | Total | N | 253 | 1,070 | 1,510 | 262 | 3,095 |
| | | Row % | 8.17 | 34.57 | 48.79 | 8.47 | 100 |

| | | | 16 years | | | | |
|---|---|---|---|---|---|---|---|
| | | | Poor | Reasonably well | Good | Excellent | Total |
| 12 years | Poor | N | 47 | 94 | 78 | 10 | 181 |
| | | Row % | 20.52 | 41.05 | 34.06 | 4.37 | 100 |
| | Reasonably well | N | 44 | 321 | 518 | 80 | 958 |
| | | Row % | 4.57 | 33.33 | 53.79 | 8.31 | 100 |
| | Good | N | 13 | 325 | 729 | 178 | 1,515 |
| | | Row % | 1.04 | 26.1 | 58.55 | 14.3 | 100 |
| | Excellent | N | 0 | 49 | 117 | 49 | 157 |
| | | Row % | 0 | 22.79 | 54.42 | 22.79 | 100 |
| | Total | N | 104 | 789 | 1,442 | 317 | 2,811 |
| | | Row % | 3.92 | 29.75 | 54.37 | 11.95 | 100 |

Notes: The tables show matrices of parental perceptions for consecutive rounds, with the diagonal elements showing the percentage of children with unchanged parental perception across successive rounds.

### Table A6: Sensitivity of main results on private enrollment, to sub-district fixed effects

| | Round 3 (8 years) | | Round 4 (12 years) | | Round 5 (16 years) | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Parent Belief in Current Round* | | | | | | |
| Excellent | 0.12* | | 0.17*** | | 0.20*** | |
| | (0.061) | | (0.051) | | (0.053) | |
| Good | 0.07 | | 0.15*** | | 0.16*** | |
| | (0.046) | | (0.028) | | (0.047) | |
| Reasonably well | 0.07* | | 0.12*** | | 0.14** | |
| | (0.036) | | (0.030) | | (0.057) | |
| *Parent Belief in Previous Round* | | | | | | |
| Excellent | | 0.05 | | 0.15*** | | 0.11* |
| | | (0.042) | | (0.041) | | (0.059) |
| Good | | 0.09** | | 0.08** | | 0.11** |
| | | (0.033) | | (0.038) | | (0.047) |
| Reasonably well | | 0.02 | | 0.07* | | 0.09** |
| | | (0.039) | | (0.037) | | (0.041) |
| Lagged test score (normalized) | 0.03*** | 0.03*** | -0.01 | -0.01 | 0.04*** | 0.04*** |
| | (0.009) | (0.010) | (0.010) | (0.009) | (0.014) | (0.013) |
| Constant | 0.38*** | 0.42*** | 0.30*** | 0.35*** | 0.25*** | 0.29*** |
| | (0.058) | (0.045) | (0.042) | (0.047) | (0.061) | (0.062) |
| Observations | 1,836 | 1,665 | 1,806 | 1,756 | 1,640 | 1,647 |
| R-squared | 0.473 | 0.469 | 0.430 | 0.436 | 0.368 | 0.365 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 The dependent variable is an indicator for private school enrollment of index children in the respective round. All regressions control for sub-district fixed effects along with wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. Standard errors are clustered at the mandal (sub-district) level.

**Table A7: Sensitivity of main results on educational expenditure, to sub-district fixed effects**

| | Round 3 (8 years) | | Round 4 (12 years) | | Round 5 (16 years) | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Parent Belief in Current Round* | | | | | | |
| Excellent | 0.26 | | 0.38*** | | 1.01*** | |
| | (0.152) | | (0.112) | | (0.192) | |
| Good | 0.07 | | 0.30** | | 0.81*** | |
| | (0.117) | | (0.135) | | (0.159) | |
| Reasonably well | 0.02 | | 0.22** | | 0.75*** | |
| | (0.143) | | (0.093) | | (0.169) | |
| *Parent Belief in Previous Round* | | | | | | |
| Excellent | | 0.06 | | 0.25* | | 0.49** |
| | | (0.167) | | (0.131) | | (0.214) |
| Good | | 0.08 | | 0.06 | | 0.42*** |
| | | (0.135) | | (0.104) | | (0.131) |
| Reasonably well | | -0.10 | | 0.10 | | 0.46*** |
| | | (0.150) | | (0.090) | | (0.125) |
| Lagged test score (normalized) | 0.08** | 0.07* | 0.02 | 0.04 | 0.18*** | 0.20*** |
| | (0.035) | (0.036) | (0.024) | (0.025) | (0.044) | (0.047) |
| Constant | 0.53** | 0.60*** | 0.04 | 0.21 | -0.88*** | -0.51** |
| | (0.217) | (0.172) | (0.167) | (0.173) | (0.214) | (0.190) |
| Observations | 1,712 | 1,523 | 1,815 | 1,765 | 1,630 | 1,625 |
| R-squared | 0.911 | 0.914 | 0.958 | 0.958 | 0.903 | 0.902 |

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 . All regressions model selection out of zero by conditioning on positive expenditure. The dependent variable in all columns is the inverse hyperbolic sine transformation of the reported expenditure on education of the child that is composed of school fees and extra tuition fees, for the respective round. All regressions control for sub-district fixed effects along with wealth index, parental years of education, household size, child gender, parental years of education, an indicator for whether the child is the eldest and caste dummies. Standard errors are clustered at the mandal (sub-district) level.