HCEO WORKING PAPER SERIES

Parental Beliefs and Investment in Children: The Distortionary Impact of Schools

Josh Kinsler
Ronni Pavan

Working Paper

2016-029

11/2016

Parental Beliefs and Investment in Children: The Distortionary Impact of Schools

Josh Kinsler, University of Georgia

Ronni Pavan, University of Rochester

May 2016

Abstract

Parental investments in early childhood have been shown to have a large impact on skill acquisition. In this paper, we examine how beliefs about a child's relative skill influences investment and how these beliefs are determined. Using data from the ECLS-K, we first show that parental beliefs about a child's skill relative to *children of the same age* is distorted by a child's skill relative to *children in the same school*. In other words, parents of children attending schools with high (low) average skills tend to believe their child is lower (higher) in the overall skill distribution. We then show that beliefs about a child's skill relative to children of the same age affects parental investments such as helping with homework or hiring a tutor. Thus, parents are making important investment decisions using inaccurate information. Building off our descriptive findings, we develop a model of parental investment that incorporates uncertainty about the average skill level of similarly aged children. We estimate the model using indirect inference and perform a set of counterfactuals where parents are fully informed about the average skill level in the population. We find that investment and achievement rise by a considerable amount for students at the bottom of the skill distribution. The mechanism behind this result is that parents of children in relatively low achieving schools revise upward their beliefs about the average child in the population, inducing an investment response.

1 Introduction

A recent and growing literature demonstrates that parental investment is an important input in the production of adolescent skills.[1] This research has sparked renewed interest in understanding the determinants of parental investment during childhood. Models of parental investment typically focus on the impact of credit constraints or the tradeoff between goods and time investments in children.[2] Relatively little attention has been given to the role of uncertainty, information, and learning in parental decision making.[3] This is in stark contrast to much of the recent literature on own human capital investments, where imperfect information about ability or returns plays a central role.[4]

In this paper, we use data from the Early Childhood Longitudinal Study-Kindergarten Class of 1999 (ECLS-K) to investigate the relationships between information, parental beliefs, investment, and the evolution of child skills. The primary advantage of using the ECLS-K to examine these issues is that parents are asked to report their beliefs regarding the relative skills of their child. Our initial exploration of the data revealed an interesting pattern in parental beliefs that serves as motivation for the current work. Parental beliefs about a child's skill relative to children of the same age is heavily influenced by a child's skill relative to children *in the same school*. In other words, parents of a student who is above average in their school are more likely to believe that their child is above average overall, even when we condition on a measure of skill relative to children of the same age.

[1]For example, Cunha *et al.* (2010) find that measured parental investment accounts for 15% of the variation in educational attainment. Carneiro & Heckman (2003), Cunha & Heckman (2008), and Todd & Wolpin (2007) provide additional supporting evidence on the importance of early childhood investment. Heckman & Mosso (2014) provide a nice summary of the literature.

[2]Caucutt & Lochner (2012) and Cunha (2013) estimate dynamic models of investment focusing on the role of credit constraints. These papers build on earlier work by Becker & Tomes (1986). Boca *et al.* (2014) and Bernal (2008) estimate dynamic models of investment focusing on the labor supply and time allocation problem facing parents.

[3]There are a few exceptions. Caucutt *et al.* (2015) explore theoretically whether uncertainty and parental bias can help explain key stylized facts regarding early investment, parental income, and achievement. Cunha *et al.* (2013) and Dizon-Ross (2013) also explore these ideas and are discussed further below.

[4]The role of learning and uncertainty has been explored in the following human capital contexts: college dropout (Stinebrickner & Stinebrickner (2009), Arcidiacono *et al.* (2014)), major choice (Kinsler & Pavan (2014), Wiswall & Zafar (2014)), and occupational choice (Antonovics & Golan (2012), Sanders (2014)). This list is not meant to be exhaustive, but simply to illustrate the prevalence of the subject.

To the extent that parents care about their child's rank in the overall skill distribution, say for attending college, the tendency to rely on local comparisons can have a significant distortionary impact on investment behavior.[5] Moreover, if parental investment is compensatory, then the distortion in parental beliefs has the potential to exacerbate gaps in student outcomes. Parents of children attending schools with relatively low average skills will invest less than the optimal level since they believe their children are higher in the overall skill distribution than they actually are. The opposite pattern would exist in relatively high skill schools.

Two recent papers also provide evidence that parental information regarding child skills and their production are not always accurate. Cunha *et al.* (2013) surveys a sample of socioeconomically disadvantaged, pregnant African American women and elicits their subjective expectation about the elasticity of child development with respect to investment. The median reported elasticity is between 4% and 19%, while estimates from the CNLSY/79 indicate an elasticity between 21% and 36%. If the median mother in the survey were given the objective elasticities, investment is estimated to increase between 4% and 24% with a subsequent increase in cognitive skills between 1% and 5%. Using data from a field experiment in Malawi, Dizon-Ross (2013) finds that parents' perceptions of their children's recent achievement diverges substantially from children's true recent achievement, with an average gap between the two being a full standard deviation. Providing parents with accurate information about child achievement causes them to alter their allocation of educational investments.

While we are also interested in understanding how investment behavior might change if parents are simply better informed, our approach is somewhat different. We seek to understand how parental beliefs develop and evolve, paying particular attention to the information that parents have available to them. Knowledge of the this process will allow policymakers to identify the most effective types of information interventions with respect to their content, timing, and frequency. Moreover, understanding the channels through

[5] Even if parents care about the level of skill a child ultimately attains, it is difficult for parents to understand what skills a child should have mastered at age six in order to achieve a particular adult skill level. The natural instinct would then be to assess your child relative to her peers.

which parental biases arise can suggest non-information interventions that can achieve similar goals.

Our investigation of parental beliefs and information begins in Section 2, where we present an illustrative example of how information regarding a child's local relative skill can influence beliefs about a child's overall relative skill. The link between local skills and global beliefs is generated through the different types of signals parents receive and by how parents interpret these signals.[6] We consider two types of signals: school average scores and teacher reports. Although the model is simple, it highlights the key mechanisms driving the paper and serves as motivation for our descriptive analysis and structural model.

Building off the illustrative example, we turn to our descriptive analysis and estimate the influence of a child's local relative skill on parental beliefs about a child's overall relative skill. We run a set of reduced-form regressions describing parents' global beliefs, where the key regressors are a student's standardized, survey-based test score and the difference between a student's test score and the average test score in a student's school. As previously mentioned, we find that parents are significantly more likely to respond that their child is above average relative to similarly aged children if the deviation from the school average is positive. We conduct a multitude of robustness checks to rule out concerns related to measurement error (both in eliciting parental beliefs and test scores), heterogenous global reference points, and unobserved school-level heterogeneity.

While the global belief regressions indicate a robust relationship with local relative skill, they fail to provide insight into the mechanisms behind this connection. To shed light on one of the possible mechanisms, we explore how teachers evaluate students and whether these evaluations influence parental beliefs. We find that teacher assessments of a child's skill relative to other children in the same grade are significantly impacted by the difference between a child's test score and the school average test score. Thus, teachers are also shading their evaluations according to the local distribution. Moreover, the evidence

[6]Throughout the paper we will use the term "global" beliefs to reflect parental comparisons of their own child's skill to the skill level of children of a similar age. "Local" beliefs will refer to skill comparisons with children in the same school. We will also use the terms global skills and local skills as the child's actual skill relative to the skills of similarly aged children and children in the same school respectively.

suggests that these teacher assessments influence parental beliefs about their child's overall relative skill.

The final section of our descriptive analysis focuses on the link between parental beliefs and investment. If there is no connection, then belief distortions induced by the local skill distribution would be inconsequential for skill development. However, we find that parental investment is strongly related to global beliefs. We focus on remedial types of investment, such as helping with homework and tutoring, since we show that these investments are most directly related to the academic achievement of children. We estimate that parents who believe their child is above average relative to children of the same age invest 10% to 20% of a standard deviation less than all other parents. Moreover, we document that remedial types of investment are more responsive to parental beliefs than other types of investments, such as reading, playing, or singing songs.

Although our descriptive analysis provides compelling evidence of the links between beliefs, information, and investment, it is not capable of connecting the various pieces in one coherent framework. In other words, it is not possible to investigate how investment behavior and outcomes might change if the information available to parents is altered. Moreover, teacher reports reflect just one observed source of information for parents and it is not clear that they alone drive the relationship between global beliefs and local relative skill. Finally, investment and learning are dynamic processes that are difficult to capture in reduced-form type regressions. As a result, we develop and estimate a dynamic model of parental beliefs, information, and investment using the illustrative example and descriptive results as guides.

The basic components of the structural model are as follows. Each child is characterized by a skill level in kindergarten, 1st, and 3rd grade. Parents are assumed to be fully informed about the level of their child's skill and the average skill level in the local school. What is unknown to parents is the average skill level in the population. Parents receive a series of signals that allow them to learn about the average skill level in the population. The types of signals parent receive, as well as how parents interpret these signals, will work to generate the reduced-form relationship between global beliefs and local relative skill. Using the available signals, parents update their beliefs about the average skills in the population, and then choose investment in order to maximize the present discounted value of utility. We assume that parents have direct utility over the their child's skill relative to both the local and overall averages. Following investment, child skill evolves according to a log-CES production function.

We solve the model through backward induction, however, this process is complicated by the fact that population and school average skill levels enter the utility function. Note that this feature combined with the dynamics of the investment problem mean that parents need to forecast average skills in future periods. We take an approach similar to Lee (2005) and Lee & Wolpin (2006) and assume that parents understand that population and school average skills evolve following a known law of motion. Thus, we need to find the appropriate parameters for this process, iterating until the implied averages evolve as close as possible to the ones predicted by the given law of motion.

We estimate the model using indirect inference, targeting moments based on regressions from our descriptive analysis. Estimates from the model indicate that fully correcting the distortions in parental beliefs would lead to significant changes in behavior. We predict that parents of students in the bottom 10% of the initial skill distribution would increase investment in 1st and 3rd grade by 20% of a standard deviation when they are fully informed about the skills of the average child in the population. Skills for this group subsequently increase by approximately 10% of a standard deviation in 3rd grade. In contrast, parents of students at the top of the initial skill distribution reduce investment slightly with a small negative impact on 3rd grade skill. We also investigate how skills and investment change in response to partial information interventions and reductions in sorting across schools.

The remainder of the paper is as follows. Section 2 provides an illustrative example of parental learning and beliefs about average skill. In Sections 3 and 4, we discuss the ECLS-K data in detail and present reduced-form evidence on the links between global beliefs, local skills, and investment. We develop a formal model of parental beliefs, information, and investment in Section 5. In Section 6, we discuss our estimation approach and provide parameter estimates. We perform counterfactual analyses in Section 7 and conclude in

Section 8.

2 Illustrative Example

The goal of this section is to illustrate how parental beliefs about a child's skill relative to the population can be influenced by a child's skill relative to the local population, as defined by the local school. The example presented below is a simplification of the dynamic information process we model more formally in Section 5. It also excludes other key components of the model, such as skill production and investment.

There are three skill levels relevant for parents of child i in school j: i's skill (A_{ij}), average skills in i's school (A_j), and average skills in the population (A). These skill levels are related in the following manner:

$$\begin{aligned} A_{ij} &= A_j + \epsilon_{ij} \\ A_j &= A + \epsilon_j \end{aligned}$$

where ϵ_{ij} and ϵ_j are mean-zero normally distributed shocks with standard deviations σ_{ij} and σ_j respectively. This structure allows schools to have heterogenous average skill levels, a feature that is consistent with the sorting observed in the ECLS-K.

The information parents have available to them regarding these three skill levels is given by the following. First, we assume that parents know both A_{ij} and A_j. Parents are likely familiar with their own child's ability to read, write, and perform basic math simply through daily interaction. Knowledge of A_j can come from direct interactions with other children in the community or through formal channels like pre-school.[7] However, parents do not know the overall skill average, A. Parents have a prior belief regarding this average, $\hat{A}$, which is distributed according to $\hat{A} \sim N(0, \sigma^2_{\hat{A}})$. Note we assume that the true overall skill level, A, is zero such that the prior is centered on the truth.

Parents update their prior regarding A using two pieces of information. First, the school

[7]The key reduced-form prediction does not hinge on this assumption, however, it does significantly simplify the model.

average acts as a direct signal of the overall average. Second, we assume that parents receive an additional piece of information, a teacher report given by

$$T_{ij} = (A_{ij} - A_j) + (A_{ij} - A) - \nu_{ij}$$

where $\nu_{ij} \sim N(0, \sigma_\nu^2)$. In our descriptive analysis, we provide evidence that teachers evaluate children according to how well the child is doing relative to both the local and overall averages. This motivates the structure of the above signal. While parents observe T_{ij}, they might not know the exact weights teachers give to each of relative skill measures. Thus, we allow parents to interpret the signal according to the following,

$$T_{ij} = (1+\gamma)(A_{ij} - A_j) + (1-\gamma)(A_{ij} - A) - \nu_{ij}.$$

In other words, parents may give too much or too little weight to the local component of the teacher report according to γ.

Parents use the school average and the teacher report to update their prior according to the procedure outlined in Appendix A. However, in the ECLS-K parents do not report their beliefs about the overall average. Instead, parents are asked how their child compares to the average child in the population. Define parental beliefs about their child's skill relative to the population average as $B_{ij} = A_{ij} - \tilde{A}$, where $\tilde{A}$ is the parent's posterior mean, $E(A|A_{ij}, A_j, T_{ij})$. We do not observe a continuous measure of beliefs, only an indicator for whether parents believe their child is above average, $1(B_{ij} > k)$, where k is a constant to allow for the fact that around 30-35% of parents think their child is above average. As detailed in Appendix A, we can write the probability that parents report their child is above average relative to the population in the following way

$$\Pr(B_{ij} > k) = \Pr\left(A_{ij} - \frac{\sigma_{\hat{A}}^2\left(\sigma_{\tilde{\nu}}^2 A_j + \frac{\sigma_j^2}{(1-\gamma)}(\gamma A_j + \nu_{ij})\right)}{\sigma_{\hat{A}}^2\sigma_j^2 + \sigma_{\hat{A}}^2\sigma_{\tilde{\nu}}^2 + \sigma_j^2\sigma_{\tilde{\nu}}^2} > k\right)$$

where A does not enter since we have assumed it is zero.

School (or local) average skill, A_j, enters the above probability through two channels. First, because parents use the school average as a signal for the overall average, parents will necessarily tilt their posterior towards the school average. This is captured by the first instance of A_j in the numerator. The size of the distortion will depend on the degree of confidence in the initial prior and the dispersion in school average scores. The second channel is slightly more subtle and depends on the degree to which parents misinterpret the teacher signal. If $\gamma \neq 0$, then the second A_j term in the numerator will also impact beliefs about skill relative to the population average. In practice, we estimate the degree to which parents misinterpret signals.

Both mechanisms discussed in the previous paragraph lead to a link between local relative skill and beliefs about overall relative skill. In Section 4 we present reduced-form evidence from the ECLS-K that is consistent with the model. However, in a context where parents receive multiple teacher signals across time, the dynamic influence of these two channels will differ, an issue we investigate more formally in the full model.[8] In the next section we introduce the data on parental beliefs and child skills.

3 Data

We use the Early Childhood Longitudinal Study, Kindergarten Class of 1999 (ECLS-K) to study parental beliefs, investment, and student outcomes. The ECLS-K is a longitudinal study that surveys a nationally representative sample of parents, children, teachers, and school administrators in the spring of kindergarten, 1st, 3rd, 5th, and 8th grades.[9] 21,409 children distributed across 1,018 schools are included in the initial kindergarten sample. Information about a child's home, school, and classroom environments is collected. We

[8]A reasonable alternative to the above setup is one where parents only know the skill of their own child. In this case, parents would first use the child's skill to update the local average skill level. This will also influence parent beliefs about the overall average according to the covariance matrix describing the relationship between the two priors. Parents would then use the mixed teacher report to update both the local and overall averages. The reduced-form prediction of this model is quite similar. The posterior mean for the overall average will be a function of the local average skill level.

[9]There is an additional survey in the fall of kindergarten and for a subsample of the original data a survey in the fall of 1st grade.

focus our analysis on data collected prior to 5th grade.[10] In the following paragraphs we discuss the key variables used in our analysis and describe the process by which we arrive at our estimation sample.

In each round of the survey, student math and reading skills are evaluated. The ECLS-K assesses skills that are typically taught and developmentally important, and the assessment frameworks are derived from national and state standards. The cognitive assessments are two-stage adaptive tests; all children begin a subject area test with a routing test, which is then followed by a second-stage form. The two-stage, adaptive assessment format helps ensure that children are tested with a set of items most appropriate for their level of achievement and minimized the potential for floor and ceiling effects. We standardize the Item Response Theory Scale Scores from the reading and math assessments and utilize these as unbiased measures of a child's skill.[11] Note that parents never observe the ECLS-K scores and thus cannot use them to learn about the average skill level in the population.

Parental beliefs about their child's skill relative to children of the same age are elicited in the fall of kindergarten and in the spring of 1st and 3rd grade. The precise wording of the question is as follows: "Does your child learn, think, and solve problems better, as well, slightly less well, or much less well than other children his/her age?" In the fall of kindergarten, 92% of parents respond that their child performs better or as well as other children of the same age. In the spring of 1st and 3rd grade, parents are also asked to compare the math and reading skills of their child to the math and reading skills of the other children in their child's class. Here parents are asked, "Compared to other children in your child's class, how well do you think he/she is doing in school this spring in math?

[10] We restrict the sample to earlier grades for a few reasons. First, attrition in the ECLS-K is considerable, an issue we discuss further below. Second, a key variable, how parents believe their child compares to their classmates is not available beginning in fifth grade. Third, limiting the number of periods facilitates estimation of the structural model. Finally, since standardized testing does not typically begin until the end of third grade, the timeframe we consider is one where parents are likely to be relatively uninformed. Note, however, that even in fifth grade beliefs about overall relative skill are influenced by measures of local relative skill.

[11] Item Response Theory uses the pattern of right, wrong, and omitted responses to the items actually administered in a test and the difficulty, discriminating ability, and "guess-ability" of each item to place each child on a continuous ability scale. The items in the routing test, plus a core set of items shared among the different second stage forms, made it possible to establish a common scale. It is then possible to estimate the score the child would have achieved if all of the items in all of the test forms had been administered.

Do you think he/she is doing much worse, a little worse, about the same, a little better, or much better?" A similar question is asked for reading. These classroom based questions are useful for demonstrating that parents understand the difference between local and global comparisons and utilize different reference points to assess their child's skills.

Kindergarten, 1st, and 3rd grade teachers are also asked to assess the math and reading skills of the surveyed children. Similar to the parent questions, teachers are asked to compare the child's math and reading skills to the skills of children at the same grade level. The choices available to the teacher are: far below average, below average, average, above average, and far above average.[12]

Along with beliefs, parental investment is the other essential variable for our model. We calculate "remedial" investment in 1st and 3rd grade as the primary factor of three underlying variables. The first variable is the number of times per week a parent helps their child with homework during the past school year. This variable captures the degree to which parents engage and assist their child in learning school material. The second variable is the ratio of the number of times per week the parent helps the child with homework to the number of times a child does homework at home.[13] Here we attempt to account for the fact that there may be variability in the amount of homework that children receive. The final variable is whether the child is tutored on a regular basis by someone other than a family member. The largest factor loadings are associated with the questions related to homework assistance.

The ECLS-K has additional variables that can be interpreted as parental investments. In the fall of kindergarten parents are asked how many times per week they tell stories, sing songs, read books, play games, play sports, do arts and crafts, and do science projects with their children. Starting in first grade, parents are also asked whether the child takes

[12]Teachers are also asked to rate specific math and reading skills of the the child on a five point scale. Examples of specific math skills in first grade are whether a child understands place value and uses a variety of strategies to solve math problems. The rating scale reflects the degree to which a child has acquired and/or chooses to demonstrate the targeted skills, knowledge, and behaviors. The ECLS-K combines these specific ratings using IRT to generate a single rating. These math and reading ratings correlate with the math and reading ratings based on comparisons to children in the same grade at a level of 0.72 and 0.82 respectively.

[13]In 3rd grade we use this ratio for both math and reading in our factor model.

music, art, or drama and how often the family eats together, goes to museums, and receives newspapers or magazines. We combine these measures of investment into a separate factor which we denote "activities" investments. In Section 4.3 we examine the sensitivity of both "remedial" and "activity" type investments to parental beliefs.

In addition to beliefs and investments, the survey also contains basic demographic and socioeconomic variables, such as race, gender, parental education, and family income in the fall of kindergarten. Another important feature of the data is the ability to group respondents together in schools. In the fall of kindergarten, we observe approximately twenty-one survey respondents per sampled school. This allows us to create proxies for the average math and reading skill in each school.

While the survey data is incredibly rich, the challenge in working with the ECLS-K is the high level of attrition. This is particularly problematic in our setting since we want to maintain a reasonable number of students in each school so that our proxies for school skill levels are informative. In the fall of kindergarten, there are 21,409 sampled children. Any kindergarten student who lacks a school identifier or is missing test scores, parental beliefs, and teacher assessments is excluded from the sample in kindergarten and all future grades. We pursue a similar strategy for 1st and 3rd grade observations, eliminating students who lack key information. This entails dropping a significant number of students since attrition is common between each round of interviews. Approximately 4,000 students attrit between kindergarten and 1st grade, with additional attrition of approximately 3,000 students between 1st and 3rd grade. Finally, we calculate the number of valid student observations available for each school-grade combination. The first time a student is associated with a school-grade combination with fewer than five students, we drop that observation and all subsequent observations associated with that student. This eliminates relatively few kindergarten students, but approximately 1,500 observations in both 1st and 3rd grade.

Our final sample contains 20,870, 15,239, and 11,100 students in the fall of kindergarten, spring of 1st grade, and spring of 3rd grade respectively. Table 1 provides means for the key variables discussed above for each grade in our sample. The first few rows of the table indicate that attrition is not entirely random since the sample becomes increasingly white and wealthier as measured by income in the fall of kindergarten. Also, the number of students per school declines considerably as a result of attrition. The parental belief variables indicate a significant skewness. More than 30% of the sample think their child thinks and solves problems better than other children his/her age, while only 7% think their child thinks and solves problems slightly less well or much less well than other children. Similar patterns are observed when parents are asked to compare their child to other children in the same class. The final few rows show the variability across households in helping with homework. Approximately 25% of parents help their child every day, while 5% never help their child with homework.

4 Descriptive Analysis

In this section we present reduced-form evidence consistent with our hypotheses regarding parental beliefs and investment. We first illustrate that parental beliefs about a child's skill relative to similarly aged children is impacted by a child's skill relative to his/her classmates. Next, we show that teacher beliefs about a child's skill is also related to a child's skill relative to his/her classmates and that teacher beliefs significantly influence parental beliefs. Finally, we demonstrate that parental investment is strongly related to beliefs about a child's skill relative to similarly aged children.

4.1 Parental Beliefs

In each round of the survey parents are asked to compare their child's ability to learn, think, and solve problems to children of a similar age. The answer to this question is our main outcome variable. Parents are given four options, however, almost all parents respond that their child is either better or as good as similarly aged children. Thus, for most of this section we treat parental beliefs as if they are binary, with a one indicating that they believe their child is above average.

Table 2 illustrates the relationship between parents' beliefs about whether their child is above average and math test scores using a linear probability model. We focus on students who are in 1st and 3rd grade since this allows us to make direct comparisons to parental beliefs about their child's skill relative to children in the same class. The first column indicates that a one-standard deviation increase in the ECLS-K math assessment increases the probability that parents think their child is above average by 16 percentage points. In the second column we incorporate an additional regressor, the difference between a student's math score and the average math score in a student's school. The impact of the math score alone declines so that a one-standard deviation increase implies an 11 percentage point increase in the probability that parents respond that their child is above average. A one-standard deviation increase in the difference between the child's score and the school average increases the likelihood of an above average report by nearly 7 percentage points. Thus, local relative skill is exuding influence on parental beliefs about overall relative skill. If parents were fully aware of the overall average, then the local deviation would have no impact.[14]

A key concern regarding the above result is that parents may simply be misinterpreting the question. If parents believe they are being asked to compare their child to other kids in the child's class, we would expect the deviation from the school average to matter. However, we can investigate this directly since parents are also asked about their child's math and reading skills compared to other children in their class. The third and fourth columns of Table 2 change the dependent variable to an indicator for whether parents think their child is much better than the other children in his/her class in math and reading respectively. In these two regressions, it is only the difference between the test score and the school average score that has a meaningful impact on beliefs. The differential pattern of coefficients across the two dependent variables suggests that parents understand that they are being asked two different questions and utilize different reference points to assess their child's skills.

When parents are asked to compare their child to children of the same age, they do so according to how well the child learns, thinks, and solves problems. This question does not map exactly to the math skills being measured by the ECLS-K and maybe this

[14]In these regressions we treat attrition as if it were random. However, we can replicate these regressions using only students who remain in the sample through third grade with almost no change in the patterns. In the structural model we incorporate attrition explicitly.

is partly responsible for the impact of the test score deviation from the school average. However, Table 3 illustrates that the relative importance of the test score deviation from the school average score is nearly identical when we use reading scores instead of math scores. Moreover, parental beliefs about their child's skill relative to the child's class continue to be primarily a function of the local relative skill measure.

One potential problem affecting both the math and reading score regressions is measurement error. If a student's own test score is sufficiently noisy, then the school average might enter the belief regressions significantly since it may also act as a noisy measure of a student's underlying skill. Note that if this were the case we would expect the test score deviation from the school average to negatively influence beliefs. A student who attends a school with high average test scores would tend to have higher unobserved skills. Nevertheless, we investigate the role of measurement error in our initial regressions by instrumenting for both the own score and the deviation from the school average using lags of these same variables. The idea is that the instruments will only pick up true skills as opposed to any measurement error. The results in Table 4 indicate that our main findings are robust to concerns about measurement error. In fact, when we instrument the relative importance of the deviations from the school average increases.

The results thus far strongly suggest that parental beliefs about a child's skill relative to children of a similar age are influenced by comparisons between the child and his/her schoolmates. However, an alternative interpretation could be that parents have heterogenous reference points for similarly aged children. For example, parents in California don't compare their child to children in Texas since they won't compete for the same colleges or future jobs. If this is the case, then a child's standing in the local skill distribution may be more important for parents than the child's placement in the overall distribution. We investigate whether this type of behavior can explain the patterns we observe by including additional test score deviations based on groupings broader than the school. The two groupings we consider are based on socioeconomic characteristics and geography. For socioeconomic groupings, we find the average test score for students of the same race and gender with similar family incomes. For groupings based on geography, we construct the average test score by census region and whether the child lives in an urban, suburban, or rural community. Table 5 illustrates that when deviations from these averages are included in our baseline regressions, the coefficients on the own score and deviation from the school average are essentially unchanged. This suggests that the school deviation is not picking up a more “local” comparison then all children of a similar age.

As an additional check on the robustness of our result, we examine whether school level heterogeneity in beliefs that is correlated with average scores could be driving our findings. To do this we calculate the difference between each child’s test score and the average test score in his/her class. Note that this deviation is much noisier than the difference with the school average since we typically have only a handful of children in a particular class. Using this measure, however, allows us to examine the impact of test score deviations from the class average *within* schools. Table 6 shows that the classroom deviations are significant predictors of parental beliefs even when we control for school fixed effects. The coefficients are slightly smaller than the coefficients on the school deviations, but are not statistically different for math.

One robustness check we do not include is to estimate the influence of the difference between own skill and school average skill on global beliefs conditional on student level fixed effects. The challenge is that we have only two observations per student and our measure of the school average is extremely noisy. The concern is that we would be attempting to identify the effect of the local deviation based primarily on transitory variation in our school measure of average skill. One way to minimize the impact of transitory variation is to exploit variation in the school-level test score deviation for students who switch schools. We can accomplish this by estimating a model that includes fixed effects for the school where a child attends kindergarten and local test score deviations based on school averages calculated from all 1st and 3rd grade scores. Variation in the local deviation will then stem only from those students who switch schools, since for non-movers any variation in the local deviation is perfectly correlated with variation in the own score conditional on the school effects. The results, presented in Table 7, show that the effect of the difference between own skill and school average skill on global beliefs is very similar to our main OLS results for math and reading. The coefficients, however, are only marginally significant reflecting that fact that only about 5% of students in our sample switch schools and remain in the sample.

To summarize, we find that parents are significantly more likely to report that their child thinks, learns, and solves problems better than other similarly aged children if the child's math or reading test score is higher than the average test score in their school. This finding is robust to a number of plausible alternative hypotheses. The results suggest that parents are not fully informed about the average child in the population and that parents appear to use the local average as a proxy for the overall average. In the next section we provide evidence that teachers make it difficult to learn about the overall average since they send parents signals that mix overall and local averages.

4.2 Teacher Assessments

While parents may be able to observe the skills of their child through repeated interactions, it is difficult for parents to discern the skills of other children. This is precisely why parents may have difficulty comparing their child to other similarly aged children. However, parents of children in kindergarten through 3rd grade receive information on the relative performance of their child through teachers and other school personnel. Teachers typically send home multiple report cards over the course of a year and may meet with parents on various occasions. Through these interactions, teachers are able to convey to parents their opinions regarding a child. If teacher beliefs about a child's skill are also in part driven by where a child falls in the school distribution, then teachers may make it difficult for parents to gauge how well their child is doing relative to the general population.

As noted in the data section, teachers in the ECLS-K are also asked to compare each child to other children in the same grade. While we interpret "in the same grade" similarly to the parent comparison "of the same age", this is not relevant for our analysis since what ultimately matters is the type of information parents receive from teachers.[15] Our

[15] Note that teachers do not give child ratings that are relative to the school or class, so we are unable to do the same check as we did with parents.

assumption is that teachers provide parents with the same type of information they provide surveyors of the ECLS-K. Table 8 shows how teacher ratings about a child's relative skill in 1st and 3rd grade are influenced by test scores and test score deviations from the school average. Similar to parental beliefs, teacher ratings are also significantly influenced by a child's test score deviation from the school average. It is true for both math and reading, and is robust to controlling for lagged teacher ratings, test scores, and test score deviations from the school average.

The evidence from Table 8 indicates that teacher beliefs are also influenced by a student's standing in the school distribution. If parental beliefs are directly influenced by teacher beliefs than this could help explain the patterns illustrated in the previous section. Table 9 shows that higher teacher ratings in 1st and 3rd grade significantly increase the likelihood that parents will say their child is above average relative to similarly aged children. This is true even when we control for parental beliefs about how the child compares to his/her classmates, lag scores and teacher ratings, and school fixed effects. The fact that parent beliefs are affected by contemporaneous teacher ratings even when we control for lags suggest that parents are using new information to update beliefs about their child.

4.3 Beliefs and Investment

The fact that parental beliefs about a child's relative skill are tilted towards a child's relative skill in the local distribution will only matter for skill accumulation if parents act on these beliefs. In the following discussion we aim to show that parents respond to beliefs in terms of parental effort or investment choices. The first step is decide which type of parental investment to analyze. In Section 2 we introduced two types of investment, "remedial" investment and "activities" investment. "Remedial" investment is calculated as the principal factor of the number of times per week that parents help their child with homework, the fraction of the times per week that parents help their child with homework, and whether the child is tutored. "Activities" investment follows the broad investment types discussed in Cunha *et al.* (2010). We use the principal factor of how often the parent: reads, tells stories, plays games, does science projects, dines together, goes to concerts, and goes to museums with child. We also include variables indicating whether the child takes music classes, art classes, drama classes, and sporting classes. Finally, we check whether these investment measures in 1st and 3rd grade respond to parental beliefs. The main results are reported in Table 10.

The first thing to notice about Table 10 is the timing. Parents are asked about their investment behavior during the last school year, while teacher ratings and the math score are relative to the spring term. For this reason, we treat parental investment as having been decided before the contemporaneous year information has been released. Hence, when we look at the variables that influence investment we used lagged test scores and lagged teacher ratings. The first four columns pertain to remedial investment, while the last four columns pertain to "activities" investments.

Consider remedial investment first. The first column indicates that parents who believe their child is above average invest 0.17 standard deviations less than other parents conditional on grade effects, family income, race, gender, and parental education. To understand whether investment seems to respond to changes in beliefs we control in the second column for past variables. The coefficient on beliefs now indicates the impact on investment holding constant past beliefs and past investments. The coefficient is largely unchanged meaning that investment follows movements in beliefs. As shown in the previous section, we find that beliefs are driven in part by math test scores and teacher ratings. We check the consistency of the previous result by controlling explicitly for test scores and teacher ratings in the third column. As expected, both measures negatively impact investment, although math scores are more important (both measures are standardized). In the fourth column we again condition on past variables, and the coefficients on scores and teacher assessments do not change significantly.

In the next four columns we repeat the same regressions using "activities" type investment as the dependent variable. Activities type investment is strongly positively correlated with beliefs. This is more consistent with a story of dynamic complementarity where parents want to invest more in more able children. Once we control for past variables, however, we see that most of this effect is driven by ex-ante heterogeneity. Parents that start with positive beliefs tend to provide their children with more of this investment. This story is also supported by the fact that this type of investment is not strongly correlated, and does not strongly co-move, with observable measures of cognitive skill, as illustrated by the final two columns of Table 10. While these investments might be very important for the cognitive development of a child, as demonstrated in Cunha *et al.* (2010), parents are inelastic in their provision with respect to child scholastic performance.[16] Given that we want to focus on investments that are potentially affected by parental beliefs, in this paper we focus only on remedial investments as in the first four columns.

In Table 11, we check the robustness of our findings relating remedial investments to beliefs. One concern is that investment could be driven by different school policies. For example, in some schools children might be pushed harder and therefore require more help from their parents. This could be correlated with the average skill of the children, ex-ante and ex-post. To control for this, we incorporate school fixed effects into our regressions. While the point estimate is lower (-0.169), the difference is small and not statistically different from zero. In the third column, instead of including fixed effects, we control for the actual school homework policy. Teachers are asked how much time they expect their students to work on homework each day. The coefficient on beliefs in this case increases, although not significantly.

Another concern is whether parental investment responds to *global* beliefs or the child's *local* rank. This could be the case if, for example, parents care directly about their child's rank in the school or class. In the last three columns of Table 11 we try to address this concern. While we see that local relative skill does have an impact on investment, its presence does not reduce the impact of global beliefs on investment. Even in this case we might be concerned by the fact that our controls for beliefs are dichotomous. In the final column we control directly for continuous measures of cognitive skill. Conditional on the test score deviation from the school average, the level of the test directly impacts the level of investment. Note that in the last column we cannot interpret the math score as global

[16]This idea is also consistent with the findings in table A10-2 of Cunha *et al.* (2010) where current period cognitive skill is not a significant determinant of parental investment, at least in the ages that overlap with our analysis

beliefs and the test score deviation as local beliefs since the test score deviation impacts global beliefs. Yet, if the math score enters significantly, it strongly suggests that global beliefs enter significantly since the math score does not affect local beliefs.

5 A Model of Beliefs, Information, and Investment

The evidence reported in the previous section suggests that parent have distorted beliefs about the relative skill of their children, these beliefs significantly influence investment, and that teachers are one potential channel through which parent's receive information. However, it is difficult to discern the overall impact that locally distorted beliefs have on investment and skill accumulation through the reduced form regressions. In this section we develop a structural model of parental beliefs, information, and investment that can achieve this goal. We discuss estimation and the results from a series of counterfactuals in subsequent sections.

Environment. In our model children begin in kindergarten ($t = 0$) and attend primary school for T periods. Children attend different schools (indexed by j) and are assumed to never change school. We assume that childrens' skill A_{ijt} is unobserved by the econometrician but observed by the parents. Parents are able to directly observe their child and are thus aware of how well they add numbers, read, and write. The average skill in school j is $A_{jt} = \frac{1}{N_{jt}} \sum A_{ijt}$, where N_{jt} is the number of children in school j and grade t. We assume that parents observe this variable as well, likely through direct interaction with other children in the community.[17] In contrast, parents do not directly observe the overall average skill level, $A_t = \frac{1}{N_t} \sum A_{ijt}$, where N_t is the total number of children in a given grade. If all schools were identical, this imperfect information would be inconsequential.

[17]Recall that in a regression of parental beliefs about the skill of their child relative to their classmates, the coefficient associated with the test score of their child relative to their classmates dominated (Table 2 and Table 3). This is an indication that parents understand well the relative position of their child within the class. However, we have experimented with versions of the model where the school average skill level is also unknown to parents. In this setup, parents are learning about two unknown, potentially correlated quantities. Assuming the school average is known simplifies the estimation and doesn't reduce our ability to match the key moments from the descriptive models.

However, it is well known that parents cluster in local areas according to their tastes for amenities and willingness to pay, leading to a non-degenerate distribution of average skills across schools. Children are also differentiated by a time invariant measure of observable family characteristics X_{ij}. The average level of this variable in school j is $X_j = \frac{1}{N_j} \sum X_{ij}$. The overall average of observables is normalized to zero in the population.

Signals. While parents begin with a vague idea of the skill level of the average child, they update their beliefs using several signals. They receive a signal at $t = 0$ equal to school average skill, A_{j0}. At the end of each grade t, parents receive two additional signals related to average skill: a teacher report T_{ijt} and an unobserved signal L_{ijt}. While the parent observes both signals, the econometrician does not observe L_{ijt}. This signal allows the information set of the parents to be larger than the econometrician's information set and can be interpreted as parents gathering information from external sources not included in our data.

Teachers typically do not give information to parents about the skill of the average child in the population. Instead, teachers report how well a child is doing relative to some benchmark, which as the previous section illustrates is typically a mix of school and population averages. Given that parents know how well their child is doing at an absolute level, they can use the teacher signal to extract information on the average skill level in the population. We assume that the unobserved signal is similar in nature, and thus assume the two signals are generated according to:

$$M_{ijt} \quad = \quad \gamma_1^M (A_{ijt} - A_t) + \gamma_2^M (A_{ijt} - A_{jt}) + e_{ijt}^M \text{ for } M \in \{T, L\},$$

where each e_{ijt}^M is a mean zero independent random variable.

Parents interpret each signal as an unbiased measure of a child's relative skill. However, we allow the teacher signal and the unobserved signal to be misinterpreted by parents according to

$$\begin{aligned} M_{ijt} &= (\gamma_1^M + \alpha^M)(A_{ijt} - A_t) + (\gamma_2^M - \alpha^M)(A_{ijt} - A_{jt}) + e_{ijt}^M \\ &= \tilde{\gamma}_1^M (A_{ijt} - A_t) + \tilde{\gamma}_2^M (A_{ijt} - A_{jt}) + e_{ijt}^M \text{ for } M \in \{T, L\}. \end{aligned}$$

Although parents understand these signals are a combination of the relative performance of their child with respect to the school and with respect to the overall population, they may misinterpret the relative weights of the two measures. Notice that if $\alpha^M = 0$, parents interpret the signals correctly. In order to update their prior beliefs about A_t (which we will define later), they will therefore use the following measures:

$$M_{ijt} - \tilde{\gamma}_2^M (A_{ijt} - A_{jt}) - \tilde{\gamma}_1^M A_{ijt} = -\tilde{\gamma}_1^M A_t + e_{ijt}^M \text{ for } M \in \{T, L\},$$

where the left hand side is observed to parents.[18]

The complete parental information set in grade t is therefore $\Omega_{ij}^t = (\{\Omega_{ijn}\}_{n=0}^t, X_{ij}, X_j)$ where $\Omega_{ijt} = (A_{ijt}, A_{jt}, T_{ijt}, L_{ijt}, e_{ijt}^I)$ contains the signals received in grade t. Notice that this information set is larger than the econometrician's information set, since he does not observe L_{ijt}, the child's skill A_{ijt}, or the school average A_{jt}. The variable e_{ijt}^I is a shifter for the cost function for parental investments which will be introduced later.

The evolution of child skill. Although children are not allowed to change school after entering kindergarten, we allow for initial sorting across schools on observables. We assume that the initial average skill level in a child's school is $A_{j0} \sim N\left(\beta^s X_j + A_0, \sigma_s^2\right)$; and that the initial skill of the child is drawn from: $A_{ij0} \sim N\left(\beta(X_{ij} - X_j) + A_{j,0}, \sigma_A^2\right)$, where we allow parental characteristics to have different impacts on the initial school average and on the child's initial skill. Defining a_{ij0} as the mean zero stochastic deviation of the child's skill from the school average, and a_{j0} as the mean zero stochastic deviation of the school

[18]Because teacher signals are observed, it is possible to identify the variance of e_{ijt}^T directly. However, we allow parents to assign a precision to the teacher signal that is independent of the variance of e_{ijt}^T. This precision parameter is identified by the autocorrelation in beliefs across periods and we denote it $\frac{1}{\hat{\sigma}^T}$.

average skill from the overall average, we can write the child's initial skill level as

$$A_{ij0} = \beta X_{ij} + (\beta^s - \beta)X_j + A_0 + a_{j0} + a_{ij0}$$

where the last two components are mean zero and uncorrelated with the other arguments. Parents do not observe A_0, but have a prior belief, $\hat{A}_0$, distributed according to $\hat{A}_0 \sim N(A_0, \hat{\sigma}^2)$. The precision of the parent's prior belief is $1/\hat{\sigma}_P^2$. Notice that $A_{j0} - \beta^s X_j = A_0 + a_{j0}$, illustrating that school averages can be used by parents as measures of the overall average skill level.[19]

Skill varies over time as a result of unpredictable shocks and parental investment:

$$A_{ijt} = A\left(A_{ijt-1}, I_{ijt}, X_{ij}, X_j\right) + u_{ijt},$$

where I_{ijt} represents parental investment and u_{ijt} an idiosyncratic mean zero shock. We also let skill production depend on own parental characteristics and the school average parental characteristics. This latter component is a simple way to control for heterogeneity in school quality. We assume that skill production takes a log-CES functional form:

$$A_{ijt} = \left(\frac{1}{\rho}\right) \ln \Big(\pi_1 \exp(\rho A_{ijt-1}) + \pi_2 \exp(\rho I_{ijt}) + \pi_3 \exp(\rho X_{ij}) + \pi_4 \exp(\rho X_j)\Big) + u_{ijt}^A.$$

For this function, the marginal productivity of investments is an increasing function of past skill if $\rho < 1$. This is referred to as "dynamic complementarity" by Cunha & Heckman (2008). On the other hand, if $\rho > 1$, the marginal productivity of investment is greater insofar as lagged skill is lower. In contrast to Cunha *et al.* (2013), we assume that parents know this production function.[20]

Parental utility and investment. Each household seeks to maximize an increasing

[19]The variance of a_{j0} is identified in the data by the variance in average test scores across schools. However, we allow parents to assign a precision to the school signal independent of the variance of a_{j0}. This precision is identified by the change in beliefs after the first period and we denote it $\frac{1}{\hat{\sigma}^S}$

[20]In practice we allow for grade specific additive shifts in the production function to help match mean scores in each grade.

function of the child's skill relative to both the school average and the overall average skill level. We assume that parents do not directly care about the level of skill but only how a child compares to their peers. This type of untestable assumption is rationalized by envisioning future higher education and labor markets as tournaments, where only the best will succeed. Parents care not only about overall relative skills, but also about skills relative to the school average since high school rank can have a direct impact on college admission, for example. In practice, we assume final utility is given by the following power function

$$U_T = \frac{1}{\lambda}\Big(\chi \exp\big(A_{ijT} - A_T\big) + (1-\chi)\exp\big(A_{ijT} - A_{jT}\big)\Big)^{\lambda}.$$

As indicated above, parents can influence A_{ijT} through their investment choices. We assume that parental investment is costly and this cost depends on the characteristics of the family and of the school:

$$C_t = \exp(\alpha_0^I + \alpha_1^I X_{ij} + \alpha_2^I X_j + e_{ijt}^I) I_{ijt}^2,$$

where e_{ijt}^I is a *i.i.d.* cost shifter which is assumed to be observed by parents prior to the investment decision.

Solving the model. In the final period, parents choose investments such that:

$$I_{ijT} = \arg\max_I \left\{E\left[U_T(A_{ijT}(I) - A_T, A_{ijT}(I) - A_{jT})|\Omega_{ij}^{T-1}\right] - C_T(X_{ij}, X_j, I, e_{ijT}^I)\right\},$$

where the expectation is taken with respect to the shock to skill production and the average skill level, A_T. Finally, define the value of entering the last period as

$$V_T(\Omega_{ij}^{T-1}) = \max_{I_T}\{E[U_T(A_{ijT} - A_T, A_{ijT} - A_{jT})|\Omega_{ij}^{T-1}] - C_T(X_{ij}, X_j, I, e_{ijT}^I)\}.$$

Proceeding backward, at each grade parents choose investment to maximize:

$$V_t(\Omega_{ij}^{t-1}) = \max_{I_t}\{-C_t(X_{ij}, X_j, I_t, e_{ijt}^I) + \beta E(V_{t+1}(\Omega_{ij}^t(I_t)))\},$$

where we make explicit that next period's information set Ω_{ij}^t is affected by parental investments through its impact on the value of A_{ijt}.[21]

It is important to point out that we are assuming that at the time of choosing the optimal investment in grade t, parents have only received signals up to period $t-1$. Therefore, our model will not be able to capture patterns in investment that are generated by information that parents receive during the school year. For example, information received in September that leads to additional tutoring in November is excluded. However, this type of behavior is likely to occur and will make investment endogenous when estimating a skill production function like the one introduced earlier. To mimic this type of endogeneity in the model and relax the strict assumption on the timing of the investment decision, we allow e_{ijt}^I and u_{ijt}^A to be correlated in the model. This correlation allows us to capture the idea that parents might have additional unobserved information about their child's skill when making investment decisions.[22]

As part of the investment decision process described above, parents form beliefs about A_t given the information available at grade t, Ω^{ijt}. A key benefit of the ECLS-K data is that we have direct measures of these beliefs. In particular, parents are asked whether their child's skills are above the overall average. In our model, this would correspond to the following indicator function:

$$B_{ijt} = \mathbb{1}\left(Pr\left(A_{ijt} > A_t | \Omega_{ij}^t\right) > \tilde{K}_t\right)$$

where $\tilde{K}_t$ is some grade specific constant.

Finally, the parent's optimization problem requires them to understand how average skills evolve over time since both local and overall averages enter utility directly. A full solution method would require a guess as to how the averages evolve, solve for parental

[21]The timing of our data is such that the definition of a period is variable. Period 0, 1, 2, and 3 correspond to the fall of kindergarten, the spring of kindergarten, the spring of 1st grade, and the spring of 3rd grade. We adjust β to account for this timing.

[22]We operationalize this by including in the production function an additional term $\pi_5 e_{ijt}^I$ such that the investment cost shock can impact achievement directly. $\pi_5 > 0$ will imply higher scores for those with higher investment costs, all else equal. In the model, the only reason parents invest is because investments are productive. Thus, the reduced form investment moments we target will identify how strong this endogeneity factor is.

decisions, calculate the implied averages and then iterate until convergence. Instead, we use a method similar to Lee (2005) and Lee & Wolpin (2006) in which we assume that parents understand that averages evolve as a first-order linear difference equation with time varying intercepts.[23] Although we still need to find the appropriate coefficients for the linear difference equation, iterating until the implied averages evolve as close as possible to the ones predicted by the autoregressive process, the computational gains of this method are quite large. Three aspects of this simplification are worth noting. First, the evolution of the overall and school average skill levels is deterministic since individual productivity shocks are independent and identically distributed across students. Notice that this implies that the school average skill level is only useful as a signal in the initial period. Second, given the assumption of normality that we make for all skill shocks, initial values, and priors, our law of motion assumption allows us to utilize the Kalman filter to describe the evolution of parental beliefs regarding population average skills. Third, as the beliefs are normal, we can rewrite the indicator function for parental beliefs as $B_{ijt} = \mathbb{1}\left(A_{ijt} - E(A_t|\Omega^{ijt}) > K_t\right)$.

6 Estimation and Results

We estimate the parameters of the model using indirect inference, matching moments from a series of auxiliary models similar to the descriptive analysis of Section 4. As noted in the previous section, we do not observe child skill directly. However, each round of the ECLS-K survey includes a battery of standardized tests. The results of these tests are never shared with parents, and therefore, the tests are not part of the parent's information set. We estimate our auxiliary models using the ECLS-K math test scores as a noisy proxy for skill. To match the moments in the auxiliary regressions, we use our model to generate noisy measures of unobserved skill. Denote student level test scores by $S_{ijt} = \gamma_1^S A_{ijt} + e_{ijt}^S$. School and overall average scores are given by S_{jt} and S_t respectively.

Prior to estimation, we impose a few normalizations which are without loss of generality. Skill is never directly observed and therefore its scale is not identified. To set the scale we

[23] For the overall average we assume that the autoregressive coefficient is 1, as otherwise we would not have enough degrees of freedom to identify all the parameters.

fix the loading on skill for the noisy test score measure, γ_1^S, to one. We set the variance of the idiosyncratic component of the unobserved signal, e_{ijt}^L, to 1 since the measurement L_{ijt} is never directly observed. We also assume that parents interpret the unobserved signal as strictly a measure of the overall skill deviation, though in reality it can be a measure of both. Finally, to reduce the size of the state space we collapse observable family and child characteristics (income, race, gender, parental education) into a single index, X_{ij}, using weights from a regression where the dependent variable is the log of math test scores in the fall of kindergarten. School average observables, X_j, is then the average of this index. Both variables are demeaned using the individual-level mean.

In addition to these normalizations, we estimate a handful of ancillary parameters outside of the model. First, as described in the data section, there is significant attrition in the ECLS-K. To capture this attrition, we estimate a probit model of attrition as a function of test scores, observable family characteristics (X_{ij}), average family characteristics in the school (X_j), and a grade effect. In the full model we simulate attrition using the coefficients from this external model. Second, in the data we observe the proportion of parents who believe their child is above average, P_t^{Above}. To mimic this in the model, we rank households according to the difference between child skill and beliefs about the overall average. We then assign above average beliefs to any household whose rank is less than or equal to P_t^{Above}. Lastly, we replicate the distribution of pupils per school and the average observables in a school by randomly drawing from the data both the number of observed students for each school and average school characteristics.

There remain 33 parameters to estimate. These parameters describe the initial skill distributions, priors over average skill, signal equations and precisions, the production function, and the parent's utility function. For each candidate set of parameters, we simulate data for a large number of schools (10,000). Within each iteration, we find the coefficients of the first-order linear difference equations that describe the evolution of overall and school average skills and generate student level data on skills, test scores, signals, beliefs, and investment for kindergarten, 1st, and 3rd grades. After building the simulated data, we compute a set of moments to compare to the ones produced by the actual data.

We iterate over the parameter space until we minimize the Euclidean distance between the simulated and actual moments. The full list of data and simulated moments are presented in appendix Tables A-1 and A-2. Model parameters and standard errors are shown in Table A-3.

A few points regarding the auxiliary models and moments are worth mentioning. First, we use the natural log of the ECLS-K IRT math scores in our auxiliary models. The distribution of log scores is better approximated by a normal distribution relative to the standardized IRT scores. Second, the tables specify the grades covered by the particular auxiliary regression. The ECLS-K questionnaires occur in the fall and spring of kindergarten and the spring of 1st and 3rd grade. However, all of the data is not available each year. The most important omissions are parental beliefs in the spring of kindergarten, teacher reports in the spring of kindergarten, and investment during kindergarten. In the model we simulate teacher reports and investment for the kindergarten year but treat it as if it is unobserved. Note that test score measures in the spring of kindergarten are used as lag score controls in many of the regressions.

The auxiliary equations are chosen with the purpose of identifying the key parameters. Regressions for the production function (3) and investment function (4) speak directly to the production and utility functions embedded in the model. The teacher report regression (10) pins down the true content of the teacher signal, while the parental distortion in interpreting the teacher signal is identified by the bias in beliefs given the signals (5, 6, and 7) and the strength of the relationship between beliefs and signals (8). The initial distribution of skill within and across schools is identified by the relationship between test scores and observables (1 and 2), along with the initial test score variances. Parents initial prior regarding overall average skills and its precision is identified by both initial beliefs reported in the fall of kindergarten (5) and the degree to which beliefs change over time (9). The variances in Table A-2 aid in identifying the variances of the idiosyncratic shocks to production, investment, and teacher signals, while the means help to regulate the model.

The crucial moments are those related to skill production, investment, and parental beliefs. The skill production function (3) links test scores in 1st and 3rd grade to lagged scores and investments. Notice that the coefficient on the linear investment term in the data is negative. This reflects the obvious endogeneity of parental remedial investment. Parents are likely to help more when new information arrives indicating that their child is struggling. Our model accommodates these within-year changes in information through the correlation between the skill shock and investment cost shock. Additionally, because skill is unobserved, investment will act as a negative signal of skill which will help fit the negative impact of investment. In the investment function (4), the linear term for each skill deviation is negative, indicating that students who are above average both locally and globally invest less. The squared terms indicate non-linearities in these relationships that depend on whether the student is above or below average. For global skill deviations, students who are way behind tend to remediate less, while students who are way ahead reduce investment even more. We are able to match all these patterns in our model. Finally there are a series of belief regressions, the most important of which is the link between global beliefs and local and global test score deviations (5, 6, and 7). The model is able to match the distortionary effect school level test score deviations have on overall beliefs even when we condition on the global skill deviations. We are also able to reproduce the impact the teacher report has on parental beliefs.

The model parameters, listed in Table A-3, also provide insight into parental behavior. First, prior beliefs regarding the overall average skill level are rather diffuse ($\hat{\sigma} = 3.09$) and not particularly precise ($\hat{\sigma}^P = 0.32$). As a result, the initial school signal will have a strong impact on parental beliefs about the overall average. The teacher signals could overcome this deficiency if parents were able to correctly interpret these signals. However, parents weigh the global deviation and local deviation in the teacher signal according to $\gamma_1^T + \alpha_T = 1.89$ and $\gamma_2^T - \alpha_T = 0.40$ when the true weights are given by 0.90 and 1.40 respectively. In other words, parents mistakenly interpret the teacher signal as if it is primarily information about a child's skill relative to the overall population. The production function is consistent with dynamic complementarity between investment and past skill since $\rho = -0.30$ is less than 1. We find that own observables are productive, however, school average observables do not influence skill production. Utility function estimates indicate that parents care about how their child's skill compares to both the overall and school average skill levels ($\chi = 0.68$). Also, the curvature in the utility function ($\lambda = -1.55$) indicates that parents are particularly averse to having relatively low skill children. As a result, we would expect that if information distortions are eliminated the largest responses are likely to come from the lower portions of the skill distribution.

Before moving on to this type of counterfactual analysis, we provide some additional insight into our modeling choices and the robustness of our findings. In our baseline specification, we allow parents to potentially misinterpret the teacher signals. We have estimated alternative versions of the model that eliminate this possibility and the fit of these models is significantly worse. In particular, we are unable to match the effect of the local skill deviation on global beliefs, a crucial parameter in our model. The baseline model also allows for flexibility in the confidence that parents have in the school and teacher signals. Rather than tie these precision parameters to the data on the signals themselves, we let the variability in beliefs over time identify the precision of these signals from the parent's point of view. We have estimated alternative versions where we eliminate this flexibility. The fit of the model declines somewhat, but the main predictions of the model remain unchanged. Finally, to help match the production function auxiliary parameters, we allow the investment cost shock and skill shock to be correlated in the model. Eliminating this correlation reduces the fit of the model, primarily as it relates to the investment related production parameters. In fact, investment is more productive in this version (due to the interaction term increasing), leading to even larger impacts in the counterfactual exercises discussed below. However, we believe the model with endogeneity is more appropriate and provides a more conservative prediction regarding the impact of information.

7 Simulations and Counterfactuals

While the parameter estimates are informative, an easier way to illustrate the predictions of the model is to examine the simulated data and run a series of counterfactual exercises. In this section, we briefly present some additional findings from the simulated data. We then explore how student skill and parental investment might change when we alter the environment. In particular, we consider altering parental information as well as changing the degree of student sorting across schools.

7.1 Simulations

The main idea behind our model is that parents are unaware of the overall skill level when making investment decisions for their children. Parents will use information available at the local level to try and determine the overall average. Thus, in schools where the local average is high (low), parents will tend to overstate (understate) the overall average. These patterns are evident in our simulated moments discussed in the previous section. Below we provide some additional details.

While all parents begin with an unbiased prior, upon receiving school and teacher signals in the first period, distortions in beliefs arise. For example, students who begin kindergarten in the bottom 10% of the skill distribution, update their beliefs about the overall average skill level such that they are biased downwards by about 30% of a standard deviation. Students in the top 10% of the initial skill distribution update their priors based on the first round of local signals such that they overstate the overall average by approximately 30% of a standard deviation. These distortions would be even larger if we split the sample according to the initial distribution of school average scores. Interestingly, as parents accumulate additional teacher and unobserved signals, their beliefs do not improve all that much. The reason for this is that they continue to significantly misinterpret the signals that teachers provide. By the final period, the distortion in parental beliefs for both the top and bottom of the initial skill distribution has only closed by about 10%.

Because parents care about the skills of their child relative to the overall average, the distortions highlighted above can significantly impact investment behavior. Still, parents of students in the bottom 10% of the initial skill distribution consistently make more compensatory investments than other parents. For example, the gap in parental investment between the bottom 10% and top 10% of the initial skill distribution is approximately 20-25% of a standard deviation. The first panel of Table 12 shows the raw gaps in investment across students in the top and bottom of the initial skill distribution for the baseline model. This gap is driven by the fact that parents care about relative skill and children at the bottom of the distribution are lagging significantly. The gap in investment would likely be even larger absent any distortion in beliefs.

The fact that parents of low skill children tend to invest more than high skill children aids in closing any initial gaps in skill over time. Random skill shocks will also tend to close these gaps. However, skill gaps might close even further if parents of children in schools where the average child is of relatively low skill had correct beliefs regarding overall average skill. The next section investigates how skill and investment gaps might change under various counterfactual scenarios.

7.2 Counterfactuals

The distortion in parental beliefs about the average skill level in the population arises as a result of two features in the model. First, parents are unaware of the overall average and use local information to learn. Second, because households sort into localities based in part on skill, local information is typically not a good signal for the overall average. In the following sections we explore how parental investment and child skill would change if either information or sorting is altered.

7.3 Information Interventions

The type of counterfactual we explore in this section is related to the information available to parents regarding average skill. The model as written seems a good approximation to the types of information parents might have available to them. While standardized testing in public schools has expanded recently, it is still rare that children are tested prior to 3rd grade.[24] Thus, parents are left to ascertain population averages based on information they receive from teachers and other local sources.

[24] Even in grades where standardized testing is available, parents may receive only categorical information regarding the proficiency of their child. This would make it extremely difficult to infer what the average child can do.

Our first exercise is to ameliorate this information deficiency by assuming that parents are fully aware of the average skill level in the population. We assume parents have this information available to them at the beginning of the child's life, or kindergarten in our model. All the parameters unrelated to beliefs are held fixed relative to the baseline.[25] The results of this exercise are presented in the second panel of Table 12. The first panel presents identical statistics for the baseline model for comparison purposes.

The first two columns of the table show the overall means and standard deviations of skill measured at the end of 3rd grade and investments during 1st and 3rd grade. Relative to the baseline, overall skill has increased mildly, as has investment. These general patterns mask important heterogeneity in responses to information. The third and fourth columns of Table 12 show how students who begin kindergarten at different points of the skill distribution respond differently to the information about average skill. The skill level of children in the bottom 10% of the initial skill distribution increases by 0.039 (4.586-4.547) relative to the baseline. This represents a 9.4% of standard deviation increase in skill. The skill level of children in the top 10% of the initial skill distribution see a decrease of 0.01 relative to the baseline, or a 2.4% of standard deviation decline.[26] These differential changes in skill are a result of the differential investment responses of the two groups. Children at the bottom of the initial skill distribution increase investment by approximately 20% of a standard deviation, while children at the top decrease investment by about 10% of a standard deviation. By the end of third grade, the skill gap between students starting at the top and bottom of the initial skill distribution closes by about 12% of a standard deviation when parents are fully informed.

The intuition for the changes in behavior is clear. Students initially at the bottom of

[25] In each of the counterfactual scenarios considered in this section, we solve for a new rational expectations equilibrium for the first-order linear difference equations governing school and overall average ability. As an example, when we give parents full information, the autoregressive coefficient in the school average regression declines. This occurs since low and high skill schools move closer to the mean as a result of changing investment patterns.

[26] When discussing counterfactuals, we focus on students at the very bottom and very top of the skill distribution. However, responses occur throughout the distribution. For example, students between the 10th and 25th percentiles of the initial skill distribution see a 7.0% increase in skill by the end of 3rd grade when parents are fully informed. Not surprisingly, the counterfactual predictions become attenuated the more central the child, or school, is in the initial distribution.

the distribution are likely to attend schools where the local average is low, and thus infer that the overall average is low. If parents are given correct information, they realize their child's true place in the overall distribution and boost compensatory investment to improve their child's relative position. The reverse pattern occurs at the top of the distribution, but given the curvature of the utility function the responses are considerably smaller.

To the extent that sorting into schools is not perfect, looking at just the top and bottom of the initial skill distribution of *students* may actually understate the impact of the information intervention. Even average students at a below average school may be underinvesting relative to the full information benchmark. To explore this idea, the final two columns of Table 12 show how student skill and investment change for students that attend *schools* initially in the bottom and top of the initial skill distribution. Here we see that the responses are even larger. The average 3rd grade skill level among students that attend a school at the bottom of the initial skill distribution increases by 11.1% of a standard deviation. At the top of the distribution, 3rd grade skill levels drop by 4.8% of a standard deviation. Thus, the overall gap between students attending the very best and very worst schools closes by about 16% of a standard deviation when parents are fully informed.

As noted earlier, many states have already implemented statewide standardized testing. To the extent that these tests ultimately reveal the true average skill level in the population, then parents will eventually be fully informed. However, in a dynamic investment model, the timing of information can have an important impact. In the bottom panel of Table 12 we illustrate how skill and investment respond to an unexpected late information intervention, in this case at the end 1st grade.[27] The first thing to note is that because parents make investment choices in 1st grade under the baseline regime, there is no difference in 1st grade investment for all students and schools across the baseline and late intervention counterfactual. While not shown in the table, average achievement in 1st grade is also unchanged. However, once parents receive the information at the end of 1st grade, investment in the next period, or 3rd grade in our model, changes. In particular,

[27] An alternative version would allow parents to know that they will ultimately learn the true average in a future period. The results are similar across the two types of late interventions.

students at the bottom of the initial skill distribution (or attending schools at the bottom of the initial skill distribution) increase investment. The reverse is true at the top of the skill distribution.

While the direction of the change in investment is similar to the full information counterfactual considered earlier, the magnitudes are slightly different. Both low and high skill children invest less in 3rd grade compared with the full information case. The reason for this is the dynamic complementarity in skill production. In the full information case, the low skill students have already received increased levels of investment in 1st grade, and thus start 3rd grade at a higher skill level. As a result, 3rd grade investment is more productive than it is in the late intervention case. High skill children are even further ahead at the start of 3rd grade under the late information intervention, and thus cut investment by even more to compensate. Ultimately, the late intervention is not as successful in closing skill gaps in the short-run.

7.4 Sorting Interventions

Providing parents with the right information, at the right time, with enough frequency can be difficult. However, the belief distortions that arise in our model can potentially be ameliorated through another channel - sorting. Because parents use the local average to infer something about the overall average, bringing these two objects closer together can mitigate suboptimal investment choices.

We explore this idea in Table 13. The top panel is identical to the baseline results of the information intervention, where we repeat it for convenience. The second panel examines how investment and skill would change if sorting into schools based on observables was eliminated. This mimics a busing type system where there are quotas at each school. However, within a quota type system unobservably high skill kids may still sort into certain schools. The third panel eliminates all sorting across schools such that the only variation in skill is within school. This can be seen in the last two columns of the table in the sense that schools at the top and bottom of the distribution look almost identical. For both sorting counterfactuals, we maintain the underlying variability in individual observables and individual skill levels.

When examining the sorting impacts on skill and investment it is important to keep in mind that there are potentially three mechanisms at work. First, variability in school average skill levels is diminished, so that the biases in parental beliefs are lessened. In contrast to the information intervention, uncertainty about the population average skill level remains. The other two mechanisms arise because we allow school average observables to enter both the production and cost functions. We included these to proxy for school quality, which will likely also change if sorting changes.

Both sorting interventions yield increases in investment among children who are initially in the bottom 10% of the skill distribution. As expected, when sorting is completely eliminated the increase in investment is even larger since now parental beliefs are, on average, not distorted. However, investment does not increase nearly as much as in the full information counterfactual. This reflects the fact that parents are still unsure about the population average skill level and that investment costs are higher when classmate observables are larger. For students who are at the top of the initial skill distribution, eliminating sorting across schools leads to less investment. The drop in investment is even larger than in the perfect information case, reflecting the importance of uncertainty on investment choices. Overall, the 3rd grade skill gap between children initially at the top and bottom of the skill distribution closes by 11.8% of a standard deviation. This is similar to the full information case, but results from skill at the top of the distribution dropping by more than in the baseline case.

8 Conclusion

In this paper we present evidence that parental beliefs about a child's skill relative to similarly aged children are distorted by a child's skill relative to children in the same school. This distortion in beliefs has important consequences for parental investment and the evolution of childrens' skill. Parents of low skill children who attend schools where average skill is also low will perform fewer remedial type investments than parents of similarly able children who attend schools where average skill is higher. Because of the tendency for students and families to sort into schools and neighborhoods, low skill children are more likely to attend schools where average skill is also low. As a result, the distortion in parental beliefs generated by local skill comparisons leads to underinvestment for low skill children.

While we find that an information intervention increases skill by about 10% of a standard deviation at the bottom of the skill distribution, there are reasons to believe that the true effect is larger. First, our measures of beliefs are quite coarse and likely introduce a significant amount of noise into the model. Second, our remedial investment measures are quite limited and may miss key avenues through which parents help their children. Third, it is difficult for us to pin down the true impact of parental compensatory investment on skill development. Remedial investments are inherently endogenous and thus require either an experimental framework or a strong instrument to confidently identify their effect. Finally, we suspect that local information distortions like the ones we illustrate are ubiquitous. In a world with multidimensional cognitive and non-cognitive skills, these types of information distortions likely compound.

Our paper complements recent work illustrating parental misinformation about child skill and development. Cunha *et al.* (2013) finds that socioeconomically disadvantaged, pregnant African American women have biased beliefs regarding the productivity of parental investment. Closer to our paper, Dizon-Ross (2013) finds that parents in Malawi significantly overstate their child's skill and when given more accurate information choose more remedial type investments to help their children. The elicitation of parental beliefs in these papers is cleaner than in our setup since the survey/experiments employed were designed precisely for this reason. However, by using the ECLS-K we are able to explore in greater detail the nature and source of parental distortions. As a result we are able to gain additional insight into policies capable of ameliorating these distortions.

The finding that parent beliefs about a child's relative skill are distorted by the local distribution also connects our paper to the broader peer effects literature. In an effort to estimate the impact of peers, researchers often estimate the impact average classroom skill has on individual test score outcomes. It is generally not clear the channel through which average peer skill operates, but the typical interpretation is that it works through in-school behaviors of either the teacher or students themselves. Our paper suggests that average peer skill also matters for individual outcomes through its impact on parental investment.

Parental investment in children, particularly at young ages, has been shown to a be a key input into skill development. As a result, it is imperative that we understand the key determinants of these investment decisions. Our paper suggests that one important factor are parental beliefs about the cognitive skill of their child. However, significant work remains. In particular, parental beliefs about the returns to investment and beliefs about non-cognitive skills are likely to significantly influence investment decisions. Moreover, embedding parental beliefs into a broader model of investment that accounts for borrowing constraints and the trade-offs between goods and time investments would be extremely informative. These additional constraints may temper the impact of beliefs or exacerbate them depending on the relationships between beliefs and other family characteristics.

References

Antonovics, Kate, & Golan, Limor. 2012. Experimentation and Job Choice. *Journal of Labor Economics*, **30**(2), 333–366.

Arcidiacono, Peter, Aucejo, Esteban, Maurel, Arnaud, & Ransom, Tyler. 2014. *College Attrition and the Dynamics of Information Revelation*. Working Paper.

Becker, Gary S., & Tomes, Nigel. 1986. Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, **4**(3), S1–S39.

Bernal, Raquel. 2008. The Effect of Maternal Employment and Child Care on Children's Cognitive Development. *International Economics Review*, **49**(4), 1173–1209.

Boca, Daniela Del, Flinn, Christopher, & Wiswall, Matthew. 2014. Household Choices and Child Development. *Review of Economic Studies*, **81**(1), 137–185.

Carneiro, Pedro, & Heckman, James J. 2003. Human Capital Policy. *In:* Heckman, James J., Krueger, Alan B., & Friedman, Benjamin M. (eds), *Inequality in America: What Role for Human Capital Policies*. Cambridge, MA: MIT Press.

Caucutt, Elizabeth M., & Lochner, Lance. 2012. *Early and Late Human Capital Investments, Borrowing Constraints, and the Family*. NBER Working Paper 18493.

Caucutt, Elizabeth M., Lochner, Lance, & Park, Youngmin. 2015. *Correlation, Consumption, Confusion, or Constraints: Why do Poor Children Perform so Poorly?* NBER Working Paper 21023.

Cunha, Flavio. 2013. *Investments in Children When Markets are Incomplete*. Working Paper.

Cunha, Flavio, & Heckman, James J. 2008. Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, **43**(4), 738–782.

Cunha, Flavio, Heckman, James J., & Schennach, Susanne M. 2010. Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, **78**(3), 883–931.

Cunha, Flavio, Elo, Irma, & Culhane, Jennifer. 2013. *Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation*. NBER Working Paper 19144.

Dizon-Ross, Rebecca. 2013. *Parents' Perceptions and Children's Education: Experimental Evidence from Malawi*. Working Paper.

Heckman, James J., & Mosso, Stefano. 2014. The Economics of Human Development and Social Mobility. *Annual Review of Economics*, **6**, 689–733.

Kinsler, Josh, & Pavan, Ronni. 2014. *The Specificity of General Human Capital: Evidence from College Major Choice*. forthcoming Journal of Labor Economics.

Lee, Donghoon. 2005. An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice. *International Economic Review*, **46**(1), 1–34.

Lee, Donghoon, & Wolpin, Kenneth I. 2006. Intersectoral Labor Mobility and the Growth of the Service Sector. *International Economic Review*, **74**(1), 1–46.

Sanders, Carl. 2014. *Skill Accumulation, Skill Uncertainty, and Occupational Choice*. Working Paper.

Stinebrickner, Todd R., & Stinebrickner, Ralph. 2009. *Learning about Academic Ability and the College Drop-Out Decision*. NBER Working Paper 14810.

Todd, Petra E., & Wolpin, Kenneth I. 2007. The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, **1**(1), 91–136.

Wiswall, Matthew, & Zafar, Basit. 2014. *Determinants of College Major Choice: Identification using an Information Experiment*. forthcoming Review of Economic Studies.

Table 1: Summary Statistics

	K	1st	3rd
% White	0.55	0.57	0.60
Log Income	10.51	10.54	10.60
Mom has BA	0.28	0.28	0.29
Children per School	21.0	16.1	12.5
Comparisons to children of same age			
Above Average	0.34	0.31	0.34
Below Average	0.07	0.07	0.09
Comparisons to children in same class			
Above Average, Math		0.36	0.35
Below Average, Math		0.05	0.08
Above Average, Reading		0.40	0.36
Below Average, Reading		0.09	0.09
Parents help with HW, 5+ times per week		0.28	0.22
Parents help with HW, 3-4 times per week		0.36	0.31
Parents help with HW, 1-2 times per week		0.21	0.30
Parents help with HW, Never		0.05	0.06
N	20,870	15,239	11,100

Data include students in kindergarten, 1st, and 3rd grade from the Early Childhood Longitudinal Study, Kindergarten Class of 1999. Data cleaning and sample restrictions are described in Section 3. Text for the questions pertaining to how a child compares to others can also be found in Section 3.

Table 2: Parental Beliefs and Math Scores

	Above Average Relative to ...			
	Similarly Aged Children		Children in Same Class	
			Math	Reading
Math	0.159*	0.110*	0.021*	0.006
	(0.003)	(0.006)	(0.006)	(0.007)
Math - School Avg. Math		0.067*	0.134*	0.136*
		(0.007)	(0.007)	(0.007)
Grade Effects	Y	Y	Y	Y
N	23,372	23,372	23,418	23,433

*,** Indicates coefficients significant at a 1% and 5% level respectively. The unit of observation is 1st and 3rd grade students from the Early Childhood Longitudinal Study, Kindergarten Class of 1999 (ECLS-K). The dependent variable is an indicator for whether parents believe their child is above average relative to a particular reference group. The first two columns use similarly aged children as the reference group, while the final two columns use children from the child's class as the reference group. The precise survey questions are presented in Section 3. The regressors are standardized IRT math scores from the ECLS-K. School averages are calculated within sample.

Table 3: Parental Beliefs and Reading Scores

	Above Average Relative to ...			
	Similarly Aged Children		Children in Same Class	
			Math	Reading
Reading	0.162*	0.114*	0.013**	0.025*
	(0.003)	(0.006)	(0.006)	(0.006)
Reading - School Avg. Reading		0.066*	0.100*	0.174*
		(0.007)	(0.007)	(0.007)
Grade Effects	Y	Y	Y	Y
N	23,092	23,092	23,136	23,151

*,** Indicates coefficients significant at a 1% and 5% level respectively. The unit of observation is 1st and 3rd grade students from the ECLS-K. The dependent variable is an indicator for whether parents believe their child is above average relative to a particular reference group. The first two columns use similarly aged children as the reference group, while the final two columns use children from the child's class as the reference group. The precise survey questions are presented in Section 3. The regressors are standardized IRT reading scores from the ECLS-K. School averages are calculated within sample.

Table 4: Parental Beliefs, Robustness to ME

	Above Avg. Relative to Similarly Aged Children			
	OLS	IV	OLS	IV
Math	0.110*	0.127*		
	(0.006)	(0.007)		
Math - School Avg. Math	0.067*	0.115*		
	(0.007)	(0.009)		
Reading			0.114*	0.139*
			(0.006)	(0.007)
Reading - School Avg. Reading			0.066*	0.099*
			(0.007)	(0.009)
Grade Effects	Y	Y	Y	Y
N	23,372	23,129	23,092	22,476

*,** Indicates coefficients significant at a 1% and 5% level respectively. The unit of observation is 1st and 3rd grade students from the ECLS-K. The dependent variable is an indicator for whether parents believe their child is above average relative to similarly aged children. The precise survey question is presented in Section 3. The regressors are standardized IRT math and reading scores from the ECLS-K. The IV regressions use lagged scores to instrument for both the contemporaneous individual and school deviation measures. For 1st grade score we use Kindergarten scores, while for 3rd grade we use the 1st grade scores. School averages are calculated within sample.

Table 5: Parental Beliefs, Robustness to Varying Reference Points

	Above Average Relative to Similarly Aged Children					
	using math score controls			*using reading score controls*		
Own Score	0.110*	0.114*	0.090*	0.114*	0.128*	0.131*
	(0.006)	(0.009)	(0.021)	(0.006)	(0.009)	(0.024)
School Deviation	0.067*	0.069*	0.065*	0.066*	0.073*	0.069*
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.007)
Socioeconomic Deviation		-0.007			-0.023*	
		(0.009)			(0.010)	
Geographic Deviation			0.022			-0.018
			(0.022)			(0.025)
Grade Effects	Y	Y	Y	Y	Y	Y
N	23,372	23,372	23,372	23,092	23,092	23,092

*,** Indicates coefficients significant at a 1% and 5% level respectively. The unit of observation is 1st and 3rd grade students from the ECLS-K. The dependent variable is an indicator for whether parents believe their child is above average relative to similarly aged children. The precise survey question is presented in Section 3. Own scores are standardized IRT math and reading scores from the ECLS-K. School deviation is the difference between the own score and the school average score in the same subject. Socioeconomic deviation measures the difference between the own score and the average score among test takers from the same income quartile, race, and gender. Geographic deviation measures the difference between the own score and the average score among test takers from the same census region and population density (central city, large town, rural). All averages are constructed within sample.

Table 6: Parental Beliefs, Robustness to School Heterogeneity

	Above Average Relative to Similarly Aged Children					
	OLS	OLS	IV	OLS	OLS	IV
	using math score controls			*using reading score controls*		
Own Score	0.131*	0.149*	0.205*	0.137*	0.155*	0.213*
	(0.004)	(0.006)	(0.007)	(0.004)	(0.006)	(0.009)
Classroom Deviation	0.055*	0.037*	0.055*	0.049*	0.031*	0.038*
	(0.006)	(0.007)	(0.010)	(0.006)	(0.007)	(0.010)
School Effects	N	Y	Y	N	Y	Y
Grade Effects	Y	Y	Y	Y	Y	Y
N	23,372	23,372	23,129	23,092	23,092	22,476

*,** Indicates coefficients significant at a 1% and 5% level respectively. The unit of observation is 1st and 3rd grade students from the ECLS-K. The dependent variable is an indicator for whether parents believe their child is above average relative to similarly aged children. The precise survey question is presented in Section 3. Own scores are standardized IRT math and reading scores from the ECLS-K. Classroom deviation is the difference between the own score and the class average score in the same subject. Class averages are constructed within sample. The IV regressions use lagged scores to instrument for both the contemporaneous individual and school deviation measures.

Table 7: Parental Beliefs, Identification from School Switchers

	Above Avg. Relative to Similarly Aged Children	
	using math controls	*using reading controls*
Own score	0.127*	0.098**
	(0.045)	(0.041)
Own Score - Fixed School Average	0.047	0.078***
	(0.045)	(0.043)
Kindergarten School Effects	Y	Y
Grade Effects	Y	Y
N	23,372	23,092

*,**, *** Indicates coefficients significant at a 1%, 5%, and 10% level respectively. The unit of observation is 1st and 3rd grade students from the ECLS-K. The dependent variable is an indicator for whether parents believe their child is above average relative to similarly aged children. The precise survey question is presented in Section 3. The regressors are standardized IRT math and reading scores from the ECLS-K. Fixed school average scores are based on averages across 1st and 3rd grade. Thus, the school average score only varies for those who students who switch schools. Kindergarten school effects are indicators for the initial school a student attends.

Table 8: Teacher Assessments and Student Test Scores

	Math Skills			Reading Skills		
Own Score	0.615*	0.421*	0.229*	0.699*	0.444*	0.281*
	(0.006)	(0.011)	(0.018)	(0.005)	(0.010)	(0.015)
School Deviation		0.264*	0.152*		0.350*	0.236*
		(0.012)	(0.019)		(0.011)	(0.016)
Lag Teacher Rating			0.249*			0.306*
			(0.007)			(0.007)
Lag Test Score			0.106*			0.055*
			(0.018)			(0.016)
Lag School Deviation			0.099*			0.047*
			(0.019)			(0.018)
Grade Effects	Y	Y	Y	Y	Y	Y
N	23,169	23,169	21,700	22,950	22,950	21,244

*,** Indicates coefficients significant at a 1% and 5% level respectively.The unit of observation is 1st and 3rd grade students from the ECLS-K. The dependent variable is a standardized teacher assessment of the the child's skill in math (first three columns) or reading (final three columns). Own scores are standardized IRT math and reading scores from the ECLS-K. School deviation is the difference between the own score and the school average score in the same subject. School averages are constructed within sample.

Table 9: Parental Beliefs and Teacher Assessments

	Above Average Relative to Similarly Aged Children					
Math	0.159*	0.106*	0.095*	0.040*	0.046*	0.052*
	(0.003)	(0.004)	(0.004)	(0.005)	(0.009)	(0.006)
Teacher Assessed Math		0.091*	0.060*	0.058*	0.041*	0.053*
		(0.004)	(0.004)	(0.004)	(0.007)	(0.004)
Beliefs, Comparison to Class	N	N	Y	N	Y	N
Lagged Controls	N	N	N	Y	Y	Y
School Effects	N	N	N	N	N	Y
Grade Effects	Y	Y	Y	Y	N	Y
N	23,372	20,809	20,607	17,988	7,397	17,988

*,** Indicates coefficients significant at a 1% and 5% level respectively. The unit of observation is 1st and 3rd grade students from the ECLS-K. The dependent variable is an indicator for whether parents believe their child is above average relative to similarly aged children. The precise survey question is presented in Section 3. Math score is a standardized IRT score from the ECLS-K. The teacher assessment is a standardized measure of the child's skill as reported by the teacher. Beliefs, Comparison to Class indicate controls for whether the parent believes their child is above average relative to their classmates. Lagged controls indicate that lags of the dependent variable and all contemporaneous regressors are included.

Table 10: Parental Investment

	Remedial				Activities			
Parental Beliefs (Lagged)	-0.173*	-0.164*			0.131*	0.007		
	(0.015)	(0.025)			(0.015)	(0.021)		
Math Score (Lagged)			-0.134*	-0.120*			0.004	0.000
			(0.009)	(0.018)			(0.009)	(0.015)
Teacher Assessment (Lagged)			-0.063*	-0.064*			-0.007	-0.020
			(0.009)	(0.014)			(0.009)	(0.013)
Investment (Lagged)		0.172*		0.154*		0.493*		0.492*
		(0.012)		(0.012)		(0.010)		(0.010)
Twice Lagged Beliefs	N	Y	N	N	N	Y	N	N
Twice Lagged Math and Teacher	N	N	N	Y	N	N	N	Y
Grade Effects	Y	Y	Y	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y	Y	Y	Y
N	21,668	8,496	22,107	8,378	21,538	8,404	21,963	8,276

*,** Indicates coefficients significant at a 1% and 5% level respectively. The unit of observation is 1st and 3rd grade students from the ECLS-K. The dependent variables are a measure of remedial and activities investment. The remedial investment factor includes help with homework and tutoring while the activity investment factor includes playing games, reading books, and going to the museum. Parental beliefs is an indicator that parents believe their child is above average relative to similarly aged children. Math score is a standardized IRT score from the ECLS-K. The teacher assessment is a standardized measure of the child's skill as reported by the teacher. Parental beliefs, math scores, and teacher assessments are lagged since investment during 1st and 3rd grade is based on information available at the end of the previous grade. Lagged investment is a lag of the dependent variable.

Table 11: Parental Remedial Investment, Robustness

Parental Beliefs (Lagged)	-0.173*	-0.169*	-0.184*	-0.180*	-0.178*	
Similarly Aged Children	(0.015)	(0.015)	(0.016)	(0.023)	(0.025)	
Parental Beliefs (Lagged)				-0.160*	-0.164*	
Class Comparison				(0.023)	(0.025)	
Math Score (Lagged)						-0.127*
						(0.017)
School Deviation (Lagged)						-0.061*
						(0.018)
School FE	N	Y	N	Y	N	N
HW Policy	N	N	Y	N	Y	Y
Grade and Demographics	Y	Y	Y	Y	Y	Y
N	21,668	21,668	18,916	9,477	8,103	20,468

*,** Indicates coefficients significant at a 1% and 5% level respectively. The unit of observation is 1st and 3rd grade students from the ECLS-K. The dependent variable is a measure of remedial investment. The remedial investment factor includes help with homework and tutoring. Parental beliefs is an indicator that parents believe their child is above average relative to similarly aged children or relative to children in the same class. Math score is a standardized IRT score from the ECLS-K. School deviation is the own math score minus the school average math score. Parental beliefs and math scores are lagged since investment during 1st and 3rd grade is based on information available at the end of the previous grade.

Table 12: Information Interventions

	Overall		Student Initial Skill		School Initial Average Skill	
	Mean	SD	Bottom 10%	Top 10%	Bottom 10%	Top 10%
Baseline						
Skill, 3rd grade	4.817	0.413	4.547	5.062	4.602	5.019
Invest, 1st grade	0.289	0.820	0.375	0.146	0.295	0.270
Invest, 3rd grade	-0.060	0.747	0.000	-0.144	-0.078	-0.056
Full Information						
Skill, 3rd grade	4.831	0.396	4.586	5.052	4.648	4.999
Δ Skill from Baseline (SDs)	*0.033*		*0.094*	*-0.024*	*0.111*	*-0.048*
Invest, 1st grade	0.347	0.717	0.533	0.087	0.489	0.168
Invest, 3rd grade	-0.027	0.703	0.165	-0.201	0.110	-0.155
Late Intervention						
Skill, 3rd grade	4.817	0.402	4.561	5.050	4.621	5.003
Δ Skill from Baseline (SDs)	*0.000*		*0.034*	*-0.029*	*0.046*	*-0.039*
Invest, 1st grade	0.289	0.820	0.375	0.146	0.295	0.270
Invest, 3rd grade	-0.053	0.697	0.133	-0.245	0.089	-0.199

All calculations based on data simulated from structural model. Placement in initial student (school) distribution is based on initial skill draw. **Baseline** results are generated using estimates presented in Table A-3. **Full Information** model assumes parents know population average skill in all periods. **Late Intervention** model assumes parents are informed about population average skill at the end of 1st grade.

Table 13: School Sorting Interventions

	Overall		Student Initial Skill		School Initial Average Skill	
	Mean	SD	Bottom 10%	Top 10%	Bottom 10%	Top 10%
Baseline						
Skill, 3rd grade	4.817	0.413	4.547	5.062	4.602	5.019
Invest, 1st grade	0.289	0.820	0.375	0.146	0.295	0.270
Invest, 3rd grade	-0.060	0.747	0.000	-0.144	-0.078	-0.056
Sorting on Unobservables Only						
Skill, 3rd grade	4.818	0.412	4.555	5.057	4.684	4.942
Δ Skill from Baseline (SDs)	*0.002*		*0.019*	*-0.012*	*0.199*	*-0.186*
Invest, 1st grade	0.287	0.825	0.393	0.129	0.252	0.293
Invest, 3rd grade	-0.061	0.747	0.025	-0.173	-0.092	-0.040
No Sorting						
Skill, 3rd grade	4.817	0.409	4.574	5.040	4.817	4.820
Δ Skill from Baseline (SDs)	*0.000*		*0.065*	*-0.053*	*0.521*	*-0.481*
Invest, 1st grade	0.298	0.818	0.498	0.062	0.295	0.300
Invest, 3rd grade	-0.051	0.744	0.119	-0.227	-0.054	-0.051

All calculations based on data simulated from structural model. Placement in initial student (school) distribution is based on initial skill draw. **Baseline** results are generated using estimates presented in Table A-3. **Sorting on Unobservables Only** model assumes that the average observable, X_j, is zero for all j. **No Sorting** model assumes that the average observable, X_j, is zero for all j and that $\sigma_s^2 = 0$.

Appendix A

Parents use the school average and the teacher report to update their prior regarding the average skill level in the population. To make the signal component of the teacher report a bit clearer, we rewrite the parent's interpretation of the teacher report in the following manner

$$\tilde{T}_{ij} = \frac{2A_{ij} - T_{ij} - (1+\gamma)A_j}{(1-\gamma)} = A + \tilde{\nu}_{ij}$$

where $\tilde{\nu}_{ij} = \frac{\nu_{ij}}{(1-\gamma)}$. Using this modified teacher signal, it is straightforward to show that the posterior mean for the overall average skill level, denoted $\tilde{A}$, is given by:

$$\tilde{A} = E(A|A_{ij}, A_j, T_{ij}) = \frac{\sigma^2_{\hat{A}}\left(\sigma^2_{\tilde{\nu}}A_j + \sigma^2_j\tilde{T}_{ij}\right)}{\sigma^2_{\hat{A}}\sigma^2_j + \sigma^2_{\hat{A}}\sigma^2_{\tilde{\nu}} + \sigma^2_j\sigma^2_{\tilde{\nu}}}.$$

The final step is to show how this posterior mean affects parental beliefs. In the ECLS-K, parents are asked how their child compares to the average child in the population. Define parental beliefs about their child's skill relative to the population average as $B_{ij} = A_{ij} - \tilde{A}$. We do not observe a continuous measure of beliefs, only an indicator for whether parents believe their child is above average, $1(B_{ij} > k)$, where k is a constant such that around 30-35% of parents think their children are above the average. Thus, the probability that parents report that their child is above average relative to the population is given by

$$\Pr(B_{ij} > k) = \Pr(A_{ij} - \tilde{A} > k) = \Pr\left(A_{ij} - \frac{\sigma^2_{\hat{A}}\left(\sigma^2_{\tilde{\nu}}A_j + \sigma^2_j\tilde{T}_{ij}\right)}{\sigma^2_{\hat{A}}\sigma^2_j + \sigma^2_{\hat{A}}\sigma^2_{\tilde{\nu}} + \sigma^2_j\sigma^2_{\tilde{\nu}}} > k\right).$$

We can replace $\tilde{T}_{ij}$ with $\frac{2A_{ij} - T_{ij} - (1+\gamma)A_j}{(1-\gamma)}$, and substitute for T_{ij} with the true signal process. This yields

$$\Pr(B_{ij} > k) = \Pr(A_{ij} - \tilde{A} > k) = \Pr\left(A_{ij} - \frac{\sigma^2_{\hat{A}}\left(\sigma^2_{\tilde{\nu}}A_j + \frac{\sigma^2_j}{(1-\gamma)}(\gamma A_j + \nu_{ij})\right)}{\sigma^2_{\hat{A}}\sigma^2_j + \sigma^2_{\hat{A}}\sigma^2_{\tilde{\nu}} + \sigma^2_j\sigma^2_{\tilde{\nu}}} > k\right)$$

where A does not enter since we have assumed it is zero.

Table A-1: Auxiliary Regression Moments

	Dependent Variable	Regressor	Data Coefficient	SE	Model Coefficient	t
(1)	S_{j0}	X_j	1.257	0.031	1.259	F-K
(2)	$S_{ij0} - S_{j0}$	$X_{ij} - X_j$	0.724	0.018	0.717	F-K
(3)	S_{ijt}	S_{ijt-1}	0.684	0.005	0.722	S-1&3
		I_{ijt}	-0.033	0.015	-0.086	
		X_{ij}	0.156	0.011	0.146	
		X_j	0.040	0.015	0.017	
		$I_{ijt} \times S_{ijt-1}$	0.005	0.004	0.022	
		$\mathbb{1}(t = 3)$	0.119	0.003	0.111	
(4)	I_{ijt}	$S_{ijt-1} - S_{jt-1}$	-0.358	0.067	-0.450	S-1&3
		$(S_{ijt-1} - S_{jt-1})^2 \times \mathbb{1}(S_{ijt-1} > S_{jt-1})$	0.301	0.134	0.157	
		$(S_{ijt-1} - S_{jt-1})^2 \times \mathbb{1}(S_{ijt-1} \leq S_{jt-1})$	-0.177	0.128	-0.069	
		$S_{ijt-1} - S_{t-1}$	-0.218	0.054	-0.154	
		$(S_{ijt-1} - S_{t-1})^2 \times \mathbb{1}(S_{ijt-1} > S_{t-1})$	-0.263	0.091	-0.158	
		$(S_{ijt-1} - S_{t-1})^2 \times \mathbb{1}(S_{ijt-1} \leq S_{t-1})$	-0.344	0.092	-0.252	
		X_{ij}	0.203	0.048	0.217	
		X_j	-0.332	0.084	-0.272	
		$\mathbb{1}(t = 3)$	-0.346	0.011	-0.345	
(5)	B_{ij0}	$S_{ij0} - S_{j0}$	0.153	0.023	0.155	F-K
		$S_{ij0} - S_0$	0.339	0.020	0.294	
(6)	B_{ij1}	$S_{ij1} - S_{j1}$	0.158	0.023	0.170	S-1
		$S_{ij1} - S_1$	0.369	0.020	0.409	
(7)	B_{ij3}	$S_{ij3} - S_{j3}$	0.298	0.023	0.185	S-3
		$S_{ij3} - S_3$	0.382	0.020	0.489	
(8)	B_{ijt}	T_{ijt}	0.154	0.003	0.223	S-1&3
		$\mathbb{1}(t = 3)$	0.020	0.006	0.022	
(9)	B_{ijt}	B_{ijt-1}	0.357	0.007	0.461	S-1&3
(10)	T_{ijt}	$S_{ijt} - S_{jt}$	1.007	0.029	1.020	S-1&3
		$S_{ijt} - S_t$	1.310	0.025	1.271	

Auxiliary regressions are labeled (1)-(10). All regressions include an intercept which we repress for ease of presentation. Individual test scores S_{ijt} are logs of the ECLS-K IRT measures. F-K (S-1&3) indicates that the dependent variables are measured in the fall kindergarten (spring 1st and 3rd) questionnaire.

Table A-2: Auxiliary Means and Variances

	Data Coefficient	SE	Model Coefficient	Grades Covered
Mean(S_{ij1})	3.539	0.002	3.677	S-K
Mean(S_{ij2})	4.073	0.004	4.288	S-1
Mean(S_{ij3})	4.572	0.004	4.844	S-3
Var(S_{j0})	0.032	0.001	0.043	F-K
Var(S_{ij0})	0.109	0.001	0.077	F-K
Var(S_{ijt})	0.092	0.001	0.153	S-1&3
Var(T_{ijt})	0.992	0.007	1.012	S-1&3
Var(I_{ijt})	0.643	0.007	0.650	S-1&3

Individual test scores S_{ijt} are logs of the ECLS-K IRT measures. School average scores are created using the logged individual test scores.

Table A-3: Model Parameters

		Coefficient	SE
Initial Skill Distributions	A_0	3.254	0.004
	σ_s	0.145	0.019
	σ_A	0.168	0.021
	β^s	1.246	0.033
	β	0.744	0.019
Priors	$\hat{\sigma}$	3.093	0.103
	$\hat{\sigma}^P$	0.321	0.024
Signals	γ_0^T	-0.033	0.008
	γ_1^T	0.895	0.064
	γ_2^T	1.397	0.072
	St. Dev. e_{ijt}^T	0.554	0.020
	γ_1^L	0.993	2.300
	α_T	0.993	0.076
	α_L	0.073	2.319
	$\hat{\sigma}^T$	0.419	0.015
	$\hat{\sigma}^S$	0.335	0.089
Production Function	ρ	-0.303	0.005
	π_1	0.662	0.002
	π_2	0.027	0.001
	π_3	0.031	0.004
	π_4	0.000	0.004
	St. Dev. u_{ijt}^A	0.284	0.003
	π_5	0.142	0.012
Utility Function	χ	0.675	0.049
	λ	-1.546	0.040
	$\alpha_{0,1}$	-4.616	0.028
	$\alpha_{0,2}$	-3.923	0.024
	α_1	-0.632	0.086
	α_2	0.558	0.308
	St. Dev. e_{ijt}^I	0.672	0.031
Test Score Measure	St. Dev. e_{ijt}^S	0.003	0.112