Parents' Beliefs and Children's Education: Experimental Evidence from Malawi*

Rebecca Dizon-Ross

University of Chicago Booth School of Business

November 13, 2016

Abstract

Do parents' inaccurate beliefs about their children's academic performance cause them to misallocate their educational investments? I conduct a field experiment in Malawi and find that providing parents with academic performance information causes them to reallocate their investments, roughly tripling the correlation of investments with academic performance. For example, most parents believe that schooling is more valuable for higher performers; information thus increases retention in school among higher-performing students and decreases retention among lower-performing students. Parents' reallocations affect a broad range of outcomes, including textbook purchases, retention in primary school, and resources for secondary school. The evidence also suggests that poorer parents have less accurate beliefs than richer, more-educated parents, and often respond more to information. Inaccurate beliefs may thus exacerbate inequalities between richer and poorer households or societies.

*I am very grateful to Pascaline Dupas, Caroline Hoxby, and Seema Jayachandran for their guidance, and to Ran Abramitzky, Abhijit Banerjee, Jim Berry, Marianne Bertrand, Nick Bloom, Doug Bernheim, Manasi Deshpande, Celine Dizon, Elise Dizon-Ross, Natalie Douvos, Esther Duflo, Alex Eble, Liran Einav, Nick Hagerty, Rema Hanna, Johannes Haushofer, Yael Hochberg, Anil Jain, Asim Khwaja, Anjini Kochar, Dan Lee, Shirlee Lichtman, Matthew Lowe, Rachael Meager, Ben Olken, Arianna Ornaghi, Jonah Rockoff, Sheldon Ross, Ashish Shenoy, Fabiana Silva, Melanie Wasserman, Tom Wollmann, Jenny Ying, Owen Zidar and workshop participants at Stanford, MIT, and Harvard, and seminar participants at Princeton, Yale, Columbia, University of Chicago, Chicago Booth, Northwestern, Harvard Business School, Stanford GSB, UCLA, UCSD, Stanford, Columbia Teacher's College, NBER Development Fall 2014, NBER Education Spring 2014, PacDEV 2014, and NEUDC 2013 for helpful comments and discussions. I thank Bridget Hoffmann, Rachel Levenson, and Michael Roscitt for help with the fieldwork, and Christine Cai for excellent research assistance. I appreciate the generous support of the Endowment in Memory of B.F. Haley and E.S. Shaw, Innovations for Poverty Action, the National Science Foundation (DDRIG 1156155), the Russell Sage Foundation, the Shultz Graduate Student Fellowship, SIEPR, the Stanford Economics Department, and the DDRO and GRO Funds. All errors are my own. Contact: rdr@chicagobooth.edu.

1 Introduction

It is widely believed that the returns to education are heterogeneous across the population, and that the efficient allocation of schooling investments depends on individual traits, such as ability, that determine returns. This idea is embedded in a long line of human capital models dating back to Becker (1962), and implies that the correct individual education decisions (such as whether to go to college, whether to get a PhD, or whether to invest in remedial tutoring) vary across individuals.

In practice, however, schooling investments may not be allocated efficiently across the population. One potential reason, which is the focus of this paper, is information frictions: Since perceived rather than true traits govern educational investments, misinformation about the individual factors underlying returns could cause important misallocations. Many education investments are made by parents, who may misallocate if they have inaccurate beliefs about their children's academic ability. As a concrete example, consider a parent with two children, one who performs well in school and one who does not. The parent can only afford to send one child to secondary school, and wants to send her higher-performer. The parent has inaccurate beliefs about which of her children is higher-performing, and so accidentally chooses to send the lower performer, only to have that child fail out of secondary school. Inaccurate beliefs can also cause misallocations across types of investment within a given child; for example, a parent could mistakenly think her child is academically weak and enroll her in a remedial tutoring class when the advanced tutoring class would have had higher returns.

Inaccurate beliefs about children's academic ability may be particularly prevalent in developing countries because many parents are uneducated. Free primary schooling only became widely available in many developing countries in the last 10-20 years: The average adult in sub-Saharan Africa has fewer than 5 years of education (UNESCO, 2013). Limited education and illiteracy may make it difficult for parents to judge their children's academic performance, especially if their children go further in school than they did, as is common in developing countries. Banerjee et al. (2010) find that, in India, 55% of parents whose child can barely decipher letters mistakenly think the child can read paragraphs. These concerns are reinforced by data from the U.S. and Malawi indicating that both within and across countries, less educated parents have less accurate beliefs.[1] If inaccurate beliefs lead to misallocation of investments, this could help explain why educational outcomes in developing countries are both poor and unequal (EPDC, 2009). For example, in Malawi, the secondary school completion rate (conditional on starting) is below 50%, and is over twice as high for the richest quintile of households as for the poorest. For primary school, the completion rate

[1] U.S. data were provided by Alexander and Entwisle (2006); Malawi data are from this paper.

is below 60%, and, despite primary school's lower costs, the differences between rich and poor are even starker (World Bank, 2010). Researchers have examined many factors (e.g., credit constraints, school quality) to explain the patterns, but none fully do.

In this paper, I conduct a field experiment in Malawi to test for the existence, magnitude, and implications of misallocations both across and within children. The hypothesis underlying the experiment is that parents' inaccurate beliefs, particularly among the uneducated, cause them to misallocate their investments and thus contribute to poor and unequal educational outcomes. The experiment is designed to reduce information frictions and assess the consequences for parents' educational investments. It delivers information to parents with children in primary school about the children's "academic performance" (which hereafter refers to performance on achievement tests administered by schools over the previous term). While the type of information delivered is very similar to the information already nominally given to parents through report cards in many countries, including Malawi, the official report cards are often hard for parents to understand, or do not reach them. The intervention in this experiment presented the information more clearly.

I assess the impact of accurate information on a broad range of investments and decisions, including book purchases, primary school retention, attendance, and resources allocated towards secondary school. The wide-ranging outcomes allow me to test both for misallocations of investments across children, and for misallocations across types of investment for a given child (i.e., whether inaccurate beliefs prevent parents from appropriately targeting their investments to their children's academic level). I find evidence of both types of misallocations, with the results suggesting that inaccuracies in parents' beliefs about their children may have large, negative impacts on children's education in developing countries.

I present three main sets of findings. The first finding is that beliefs are inaccurate, especially among the uneducated: On average, parents' beliefs about academic performance diverge from true performance by more than one standard deviation of the performance distribution. When comparing two of their children, one third of parents are mistaken about which child is higher-performing.

I then establish that these inaccurate beliefs can cause misallocations. This is the second finding: that due to inaccurate beliefs, investments are not as well tailored to academic performance as parents would like. I present parents with a series of investment options and decisions that are designed to have clear predictions for how the efficient investment depends on academic performance, allowing a clean test for misallocations. For example, I look at demand for books that are designed for students of different performance levels (e.g., a remedial book designed for the lower performers in the sample, an advanced book designed for higher performers). The prediction is that returns will be higher if the selected book matches children's performance. In the control group, parents try to match investments to the perceived performance of their child; for example, demand for the remedial book is over 12 times higher for parents who believe their child is in the bottom performance quintile relative to those who believe their child is in the top quintile. However, because many parents' beliefs are mistaken, the relationship between demand and true performance is much weaker: Demand is only 3 times higher among parents whose child is *truly* in the bottom quintile relative to those truly in the top.

The third finding is that providing information to parents about their children's academic performance causes them to reallocate. The findings are similar both for misallocations of the type of investment for a given child, and for misallocations across children. The magnitudes are economically significant, with the reallocations often more than tripling the targeting of parents' investments to their children's performance. For example, information quadruples the demand for the remedial book among parents whose children are truly in the bottom quintile relative to those whose children are truly in the top quintile, with demand going from 3 times higher for bottom quintile parents than top quintile parents in the control group, to 12 times higher in the treatment group. These types of parental decisions are likely more relevant now than ever in developing countries, as the use of supplementary inputs is growing rapidly (e.g., in Malawi, the share of 6th graders using tutoring services rose from 20% to over 50% between 2000 and 2005) (Paviot et al., 2008).

I find similar results when looking at parents allocating larger investments across children. Secondary school fees are the first high-cost educational investment that parents in Malawi make, and most parents cannot afford the fees for all of their children. I give parents the chance to win a secondary school scholarship and ask them to choose which of their children to give it to. In the control group, 60% of parents give it to their higher-performer. Provision of information increases this to 80%, thereby tripling targeting relative to the no-targeting benchmark of 50%. Similarly, primary school dropout rates for children whose parents find out they are above-median performance fall to nearly 0% relative to a control group mean of 2%, and they roughly double (to about 4%) for children whose parents are informed that they are below-median performance. This behavior suggests that parents believe years of schooling and academic performance are complements (i.e., that schooling is a higher-returns investment for higher-performing children), a belief consistent with the literature from other contexts (Pitt et al., 1990; Aizer and Cunha, 2012). It also highlights that academic performance information is not a panacea to increase education for all: it leads to reallocations, which may decrease education for some. Some parents of low-performers may decide that the returns to spending on education are lower than, say, the returns to spending on health. Parents are maximizing utility, not education.

In all of the above analyses, I also test for heterogeneity by parent education. Importantly, I find that less-educated parents have less accurate beliefs, and that, for some investments, they reallocate more in response to information than more-educated parents. For example, the treatment effects on demand for remedial textbooks are twice as large for parents without secondary education as for those who have secondary education. Since more-educated parents have more accurate beliefs, they appear to be better at targeting their investments at baseline.

Because the results suggest that inaccurate beliefs may be more problematic among the less-educated, belief inaccuracy could contribute to the perpetuation of inequalities across generations. A back-of-the-envelope calculation suggests that providing information could close upwards of 15% of the gap in educational outcomes between less-educated and more-educated households in the sample. Information may therefore decrease the inter-generational persistence of inequalities, even as it may increase inequality within a given generation, as some parents invest more in their higher performers. In a context such as Malawi where parents believe the average returns to education are high, and their beliefs are in line with actual Mincerian returns (for both less-educated and more-educated households), misinformation about academic performance and individual-level returns may be a first-order information friction affecting investments and inequality.[2]

In addition to the main analysis, I perform a number of specification tests to rule out alternative interpretations, and I use baseline beliefs data to investigate whether the mechanism for information's effects is an effect on the first moment or on the second moment of parents' beliefs. I provide suggestive evidence that changes to the point estimate of beliefs is the primary driver of information's effects, but that for the larger investments, like dropouts and secondary school, the uncertainty of beliefs also matters.

It is important to note that I evaluate the effect of decreasing information frictions on investments, not welfare. The experiment takes as given parents' preferences and the perceived production function, so any "misallocation" identified is effectively defined as a wedge between how parents would like to allocate their investments given their children's academic performance, and how they allocate them in reality. The implications for welfare and the inter-generational persistence of inequalities thus depend on whether there are other interacting market failures. A first potential concern would be if parents were wrong about the education production function. I analyze and discuss this in detail throughout the paper; the evidence suggests this is not an issue in this context.[3] A second potential concern would

[2]Underestimating average returns is very important in some contexts (e.g., Jensen, 2010), but in others people overestimate returns, at least for post-primary (Hastings et al., 2013; Pekkala Kerr et al., 2015).

[3]Moreover, some of the investment choices presented to parents were designed to enter the production function in an obvious way (e.g., the remedial books are substitutes with ability), and across the parental education spectrum, parents' reallocations align with the predictions.

be if information affected the *level* of investments, and not just the allocation across the population. This could happen if beliefs are biased, or if parents asymmetrically update. In a world without other market failures, even if provision of information results in a decrease in the level of investments, this would still represent an unambiguous improvement to welfare. If we think parents are underinvesting due to other imperfections, however, an average decrease would be concerning. Reassuringly, I do not find effects on the average level of enrollment or expenditures.

My findings help advance our understanding of the causes of poor educational outcomes in developing countries and contribute to a number of strands of literature. First, I contribute to a growing literature on information constraints in education. Most of the existing literature has focused on misinformation about aggregate factors, such as the population-average returns to education, school quality, or other features of the education system (Jensen, 2010; Nguyen, 2008; Andrabi et al., 2016; Bettinger et al., 2012; Dinkelman and Martínez A, 2014; Hoxby and Turner, 2013; Wiswall and Zafar, 2015). These studies abstract away from the fact that correct individual education decisions (such as whether to go to college, or whether to invest in a remedial textbook) vary importantly across individuals. Here, I shift focus from aggregates to the heterogeneity within the population. Just as capital can be misallocated across firms with heterogeneous productivity, human capital can be misallocated across individuals with heterogeneous returns. Inaccurate beliefs about the individual-level factors underlying returns can therefore cause important misallocations. The few prior studies on beliefs about individual-level factors use observational data to show that students' beliefs about their own abilities predict their decisions, such as the choice of college major or college dropout (Chevalier et al., 2009; Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2012, 2014).

To the best of my knowledge, my experiment is the first to use exogenous variation in beliefs to establish a causal link between inaccurate beliefs about individual-level factors and the misallocation of investments. The findings complement recent information experiments by Bergman (2016), who shows that agency issues affect school performance by providing parents with information that allows them to monitor their children's effort and outcomes, and Bobba and Frisancho (2016), who test predictions about the differential roles of the mean and variance of beliefs on educational decisions.

This paper is also the first to document that education-related misinformation can be a more acute problem for parents with lower socio-economic status (SES). This heterogeneity is important as it may provide a channel for persistent educational inequalities if it causes less-educated parents to make sub-optimal investments in their children's schooling.

The paper also contributes to a large literature examining how parents' investments depend on their children's ability (e.g., Behrman et al., 1994; Griliches, 1979; Datar et al., 2010; Almond and Currie, 2011; Bharadwaj et al., 2013; Rosenzweig and Zhang, 2009). Identification in this setting is difficult, most notably because of potential reverse causality between investments and ability, and the resulting studies find mixed results. Most of these studies use either birthweight or twin comparisons, where there are concerns about endogeneity or external validity (as suggested in Bharadwaj et al. (2013)), with the recent exception of Leight (2014) and Adhvaryu and Nyshadham (2014), who use climatic shocks and policy-induced variation for identification. This paper contributes by using a new within-person identification method that exploits the exogenous "shock" to beliefs, and by examining a broader range of investments. The broad range is important since my results vary across investment types, which could help explain the mixed existing results.

Finally, this study adds to the literature documenting the positive influence of parents' education on children's education by highlighting one channel for effects: parents' beliefs (e.g., Rosenzweig and Wolpin, 1994; Andrabi et al., 2012; Banerji et al., 2013).

The paper proceeds as follows. Section 2 presents a conceptual framework. Section 3 describes the context and experimental design. Section 4 presents the results on shorter-run and longer-term investment outcomes. Section 5 examines robustness, and Section 6 concludes.

2 Conceptual framework

I begin by presenting a simple framework to generate predictions for how inaccurate beliefs affect investments. I then discuss how one can use an experiment to test the predictions.

2.1 Setup

Parents are choosing investments in their children's schooling. There are various choices they make: the total amount to spend on education, the allocation of educational resources across the children in the household, and then, for each child, the specific bundle of educational resources – for example, what difficulty level of textbook or tutoring to choose for a given child. All of these choices may depend on the children's academic performance, since the returns to the various inputs may depend on performance.

For any of these decisions, the choice made by the parent can be described as follows. Denote the level or type of resource as s. A parent chooses s in order to maximize household utility subject to a budget constraint. The perceived returns to s depend upon the child's baseline ability, which is throughout the paper proxied for by academic performance, A. Thus, the utility-maximizing choice of s (i.e., "arg max") is a function of A. I denote this function as $s^*(A)$, and call it the parent's "preferred investment function"; it captures the inputs parents would opt to choose as a function of true performance if they knew it. A key assumption, which can later be tested for in the data, is that the perceived returns to s do in fact depend on A, and therefore that the preferred choice, s^*, depends on A as well – i.e., that the derivative of s with respect to A is not 0, $\frac{\partial s^*}{\partial A} \neq 0$. Much of the analysis thus centers around this derivative or slope.

Consider the various examples of parents' choices described above. Suppose s describes years of schooling. If schooling is a perceived complement with performance A (i.e., higher-returns for higher-performing children), then s^* would increase in A, $\frac{\partial s^*}{\partial A} > 0$; if it is a perceived substitute, then it would decrease $\frac{\partial s^*}{\partial A} < 0$. Alternatively, for a choice of allocation across two children, s can represent spending on child 1 relative to child 2, and A can represent the academic performance of child 1 relative to child 2. In that case, in addition to the production function, the derivative would also reflect parents' preferences for investing, such as whether they want to minimize inequality between their children. Third, for a parent's choice of which *type* of educational resource is best for a child, such as choosing the difficulty level of a textbook, s could represent the book's difficulty, and $\frac{\partial s^*}{\partial A} > 0$.

2.2 The effect of inaccurate beliefs

Assume that parents do not know true performance, A. Instead, they have a belief, denoted $\tilde{A}$. I define beliefs as inaccurate if $A \neq \tilde{A}$. Note that this is a statement about individual-level inaccuracies, not population-average. I initially assume there is no uncertainty in beliefs (i.e., $var(\tilde{A}) = 0$), and later discuss the effects of uncertainty. With no uncertainty in beliefs, instead of choosing the utility-maximizing investment $s^*(A)$, parents instead choose inputs as a function of beliefs $\tilde{A}$, and so choose $s^*(\tilde{A})$. If beliefs are inaccurate and preferred inputs vary with performance, then parents' chosen inputs, $s^*(\tilde{A})$, will not equal the utility-maximizing choice $s^*(A)$. As a result, utility will be inefficiently low.[4]

Since inaccurate beliefs decrease utility by causing actual chosen inputs to diverge from parents' preferred inputs, one way to test for the effects of inaccurate beliefs is to test for a divergence between actual and preferred inputs. Define the "actual investment function" as the average actual investments chosen as a function of *true* performance, $s(A) = E(s|A)$. To determine this function empirically, one can look at how investments depend on true performance. To determine the preferred investment function, one can look at how investments depend on beliefs. The form of the divergence between the two lines will depend on the statistical properties of the distributions of A and $\tilde{A}$. However, the same qualitative predictions hold whenever believed and true performance are positively correlated, and the

[4]Because the utility function incorporates the perceived education production function, this statement relies on the assumption that the perceived production function is correct. I discuss the appropriateness of this assumption – and what happens if it is relaxed – for each investment as I proceed through the analysis.

variance of beliefs is not too much larger than the variance of true performance. The data from my setting satisfies these conditions, as does most beliefs data, and so I sketch the intuition in that case, both verbally and in Figure 1.

In this setting, inaccuracies in beliefs cause beliefs to be "attenuated" relative to true performance, i.e., to have a slope less than 1 if plotted on true performance (Figure 1(a)).[5] The level of attenuation is driven by the correlation between believed and true performance: the lower the correlation, the more attenuated the slope. Attenuation implies that parents with children at the top of the true distribution underestimate their children on average, and parents with children at the bottom overestimate on average.

The attenuation of beliefs then causes actual chosen inputs to be attenuated as a function of true performance, i.e., flatter and not as responsive as parents would like.[6] The idea is that parents choose investments based on their (inaccurate) beliefs and the preferred investment function, and so investments are steeply sloped with beliefs, as depicted in Figure 1(b) for the case with $\frac{\partial s^*}{\partial A} > 0$. However, if we look at children who are *truly* at the top of the distribution, many of their parents think they are below the top, and so on average choose inputs appropriate for lower-performing children. Analogously, many parents of children at the bottom of the distribution choose inputs appropriate for higher-performing children. Together, this causes the slope of the actual investment function to be attenuated relative to the preferred slope, and decreases utility (Figure 1(c)). The attenuation captures the fact that investments are not as well tailored to performance as parents would like.

More broadly, for any set of distributions of $\tilde{A}$ and A, the prediction to test is:

Prediction 1. *Inaccurate beliefs can cause the slope of the actual investment function to differ from the slope of the preferred investment function.*

2.3 Estimation

It is difficult to empirically estimate the difference between the slopes of the actual and preferred investment functions because neither regression line will in general be causal. Assume parents invest according to the model above plus a noise term due to heterogeneous tastes (ε). Consider comparing the slope estimated from regressing investments on beliefs, $\tilde{A}$, to the slope estimated from regressing investments on true performance, A. The estimated slopes could differ from the true causal slopes as a result of correlation between ε and $\tilde{A}$ or

[5]To see that this holds when (i) A and $\tilde{A}$ are positively correlated, and (ii) the variance of $\tilde{A}$ is not "too much larger" than the variance of A, use the standard formula for the OLS slope to express the slope as $corr(\tilde{A}, A)\frac{SD(\tilde{A})}{SD(A)}$, where *corr* is correlation and SD is the standard deviation. There is thus attenuation when $corr(\tilde{A}, A)\frac{SD(\tilde{A})}{SD(A)} < 1$. Since correlations are bounded above by 1, this means that there is attenuation whenever $SD(\tilde{A}) < SD(A)$, and more broadly that there will be attenuation when $\frac{SD(\tilde{A})}{SD(A)} < \frac{1}{corr(\tilde{A},A)}$.

[6]See Appendix C for more formal discussion.

ε and A, which would cause omitted variable bias (OVB). In particular, if $\tilde{A}$ and A have *different* correlations with ε, then this could cause the slopes of the estimated lines to differ, but that difference might only reflect OVB, not inefficiencies due to inaccurate beliefs.

We can use an experiment to overcome this estimation challenge. Consider an information intervention that tells parents the true A and changes their beliefs from $\tilde{A}$ to A. If inaccurate beliefs were causing attenuation at baseline, then providing information would allow parents to choose their preferred investment, $s^*(A)$, instead of $s^*(\tilde{A})$;[7] that is, information would allow parents to invest along the preferred investment function. In the case of baseline attenuation, this would increase the slope of actual investments on true performance. This leads to the following testable prediction:

Prediction 2. *If inaccurate beliefs distort parents' baseline investments, then information will change the slope of the actual investment function. If baseline investments were attenuated, then information will increase the slope.*

The intuition is that information allows parents to correct their baseline mistakes and make their preferred investment choice (see Appendix C for more rigorous discussion).

Note that the predictions and empirical test focus on information's impact on the slope of investments, not the average treatment effect (ATE) of information. ATE will tend to understate information inefficiencies. For example, if beliefs are inaccurate at the individual level but have the correct population-level mean (i.e., $E[\tilde{A}] = E[A]$), and investments are linear in A, there would be no ATE, but providing information could still increase welfare.

An alternate empirical approach is to examine how treatment effects vary by the "beliefs shock" (A-$\tilde{A}$). The approaches are similar and results largely consistent, with beliefs shock results included in the robustness section. I focus on the slope approach since (a) it can provide nice insight into parents' objectives, as I show in Section 4.2; and (b) the "beliefs shock" specification relies on an assumption that can fail in the presence of uncertainty: that the heterogeneity in the treatment effects by A and $\tilde{A}$ are equal and opposite.[8]

2.4 Extending the model: Uncertainty

The statement above that parents with beliefs $\tilde{A}$ choose inputs $s^*(\tilde{A})$ depends on the assumption that the certainty of beliefs does not matter, e.g., that parents would choose the same investment regardless of whether they knew for sure their child had true performance of $\tilde{A}$ or just had some uncertain beliefs with mean $\tilde{A}$. In reality, beliefs uncertainty could cause attenuation in the preferred slope of investments on (mean) beliefs, since, if beliefs are

[7]This assumes that parents fully update their beliefs in response to information. If they only partially update, then their choice would be a weighted combination of $s^*(A)$ and $s^*(\tilde{A})$.

[8]This assumption can fail because uncertainty can affect the slope of preferred investments as a function of beliefs, as described in the next subsection. See the second approach in Section 4.4 for further discussion.

uncertain, parents may hesitate to make their investments depend as strongly on their beliefs (see App. C.1 for one potential framework producing this prediction). One can think of this as a second channel for baseline attenuation which I test for in the empirical analysis. Under some but not all models, uncertainty could also affect the average level of investments, akin to underinvestment in risky financial assets (Levhari and Weiss, 1974; Olson et al., 1979).

3 Context and experimental design

The setting for the experiment is Malawi. Primary school in Malawi covers grades 1-8. It has technically been free since 1994, but it does involve expenditures. Parents in the study sample spent an average of 1,750 Malawi Kwacha (MWK), roughly 10.6 USD or 1% of annual household income, annually per child. The main expenditures are uniforms (33% of total), informal but required school fees (22%), and supplemental investments such as school supplies, tutoring, and books (45%). The access rate to the first grade of primary school is above 95%, but dropouts are common in primary school. Sources vary, but all suggest the completion rate (conditional on enrolling) is less than 60% (World Bank, 2010). Secondary school, covering grades 9-12, is not free, and costs significantly more than primary school. Annual fees for government secondary schools range from 5,000 - 10,000 MWK per year (30 - 60 USD, over 4 times the median primary-school expenditures in the sample). Uniforms and supplementary supplies are additional expenses. Many children do not attend because of the high fees (World Bank, 2010). Secondary slots are also limited, with admissions governed by an achievement test administered at the end of primary school.

Schools are required to send report cards home each term with achievement test scores. They vary by school, but all are supposed to have average absolute test scores, and the corresponding grade on the standard Malawian grading scale of 1-4. (Online App. H contains an example from the study sample.) However, according to baseline survey data, 65% of parents state that they do not know their child's performance from the last report. The main reasons are that the parents (a) were unable to read or understand the report, or (b) did not receive it in the first place. Students are supposed to deliver the reports, so children could either lose or choose not to deliver them: Parents of students who performed badly were much less likely to receive the report. Overall, the education system in Malawi is similar to that in many countries in sub-Saharan Africa and beyond in terms of the information given to parents as well as its overall structure.

3.1 Experimental design

The experiment delivers academic performance information to randomly selected parents and measures the effects on educational investments and decisions. To fit the framework presented in section 2, the experiment should provide information about the individual-level trait on which parents' educational investments depend. According to qualitative interviews, parents in Malawi think academic performance (i.e., scores on school-administered exams) is the most important determinant of both their investments and the returns to educational investments.[9]

Sample selection: The study worked with 39 schools in two districts (Machinga and Balaka) in Malawi. Schools were selected randomly from the universe of primary schools, oversampling schools with high and low expected levels of parent education to increase heterogeneity in parent education within the sample. The study team first conducted a census at schools, mapping the sibling structures for all students in grades 2-6, which were chosen because they span most of primary school. Since one of the outcomes to be examined is inter-sibling tradeoffs, multiple-sibling households were used as the sampling frame.[10] The team also gathered achievement test data from the most recent term (term 2 of the 2011-2012 school year) for use in the intervention, focused on the most recent term since (i) it was the most relevant to parents, and (ii) schools often do not keep good historical records.

Based on the test score and sibling data, a sample of 3,464 households with at least two children enrolled in grades 2-6 with test score data was drawn. For households with more than two children, two were randomly selected. The criterion that children needed test score data means students who have the highest absence rates (and whose parents might have the largest information problems) are under-represented in my sample.

Randomization: I randomly assigned half the households in the sample to a treatment group that received information about their children's test scores, and half to a control group, which did not.[11] The randomization was stratified on a test score measure (between-sibling score gap), and a proxy for parent education (the estimated literacy rate in the household's village), since a key *ex ante* goal was to look at heterogeneity by parent education.

Eligibility interviews: Sample selection and randomization were based on data gath-

[9] If parents were wrong about the education production function, a second objective relevant for a scale-up would be to use the trait most correlated with *actual* individual-level returns. (See Section 4.5 for further discussion.) Academic performance also meets this second objective: It determines progression through school and selection into secondary school, thereby almost surely affecting the returns to investment. "Innate" ability is another possible determinant of returns, but, as has been extensively documented, it is difficult to measure "innate" ability; any measure would represent some combination of innate ability and past inputs.

[10] Fewer than 3% of households in Malawi with children have only one child, so the potential external validity concern with this sampling approach is that households with tighter birth spacing are over-represented in the sample. Reassuringly, birth spacing is uncorrelated with belief accuracy in the sample.

[11] Half the treatment group was also assigned to receive an add-on intervention designed to test a hypothesis intended for study in separate work: that providing more detailed information would increase parental engagement, as measured through non-monetary and monetary investments. This group received additional skills information (e.g., whether their child could add 3-digit numbers, see Online Appendix F for sample). In this paper, I ignore this add-on treatment and pool the treatment households. I do not find that this treatment had an effect on the pre-specified outcomes.

ered from students at school and on school administrative data. Household eligibility (i.e., whether both siblings lived in the household and were still enrolled in school) was then verified through an eligibility questionnaire with parents.[12] Among the 3,464 sampled households, 21% of households were found to be ineligible during the parent interviews, leaving a sample of 2,716 eligible households. Of the 2,716 sampled and eligible households, 97% (2,634 households) were located at their homes, available, and consented to participate in the baseline survey. Thus, the final experimental sample comprises 5,268 parent-child pairs. Both eligibility and baseline survey completion are unrelated to treatment assignment.

Baseline survey visit: Surveyors visited all sampled households and asked to speak with the parent who is the primary decision-maker about education.[13] Surveyors then conducted a baseline survey, which included a module on education spending and beliefs about children's test scores. While eliciting baseline beliefs about test scores, surveyors explained the grading scale used by schools to parents, including reviewing a sample report card with the same format as those later delivered to the treatment group. This was done to aid the elicitation of beliefs and to hold knowledge of the grading scale and report card format constant across treatment groups. After the survey, during the same visit, surveyors conducted the information intervention for the treatment group.

Information intervention and report cards (Treatment group only): Surveyors walked treatment parents through two report cards describing the academic performance of their two children. The order was randomized. The reports showed children's performance on all tests administered in the most recent school term, specifically: the percent score (an absolute measure), the corresponding grade on the Malawian grading scale, and the within-class percentile ranking (see Online Appendix I for more details). The statistics were listed for the three subjects that Malawian educators deem most important – math, English, and Chichewa, the local language – and for "overall" (the average of the three). The report card also showed the number of individual tests included in the averages; the sample average is 4.5 tests. The correlation between students' scores on the different individual tests is roughly 0.8 for overall performance, and 0.6 - 0.7 within subjects.

A sample report card is presented in Appendix B. The format was chosen based on a series of focus groups, with the primary selection criterion being how well uneducated parents could understand it. Surveyors walked treatment parents through every number on their report cards. They previously received training on how to explain the information clearly. Much of the information in the intervention report card overlapped with the information

[12] Eligibility for the initial sample was based on children's reports. Most ineligibility reflects children's misreports, with the most common source being that the sampled children were not siblings, just friends.

[13] If that parent was unavailable, the surveyor spoke with the second parent if there was one and he or she was knowledgeable about educational decisions. If not, the surveyor returned later.

already nominally provided by school report cards, since all school report cards are required to include absolute scores and grades. However, by using a clear report card format (there is no standard school report card format) and having a surveyor walk the parents through each number on the cards, the intervention presented this information in a clearer way.

3.2 Data and outcomes

The analysis uses several data sources, including data from surveys with parents and administrative data from schools.

(1) Baseline survey data: The baseline survey was rolled out immediately after term 2 of the school year ended in March 2012, and ran from April to June of 2012. The survey included modules on demographics, baseline spending on education, beliefs about the returns to education, and beliefs about academic performance, described more in Section 3.3.

(2) First endline survey - Immediate investments and endline beliefs: During this survey, surveyors presented parents with three real-stakes investment decisions described in detail in Section 4.2: remedial textbooks, workbooks of different difficulty levels, and secondary school lottery tickets. Surveyors also measured parents' beliefs, as described in Section 3.3. This survey was conducted immediately after the baseline survey and information intervention, as laid out in the data collection timeline in Figure 2. This was done primarily for budgetary reasons, but also has the advantage that the outcomes were measured before parents had a chance to speak with others, allowing the outcomes to more cleanly reflect parents' preferences, as opposed to the preferences of the people they talk to, such as their children. I thus refer to the real-stakes investment decisions as the "immediate" outcomes.

(3) Longer-term data: I also collected two types of data in the year following the intervention: a second endline survey of parents 1 year after the intervention (June-July 2013), which I use to examine treatment effects on dropouts and expenditures, and administrative data on attendance gathered roughly 1 month after the intervention (July 2012).

For the 1-year endline data collection, given the very limited budget available, I focused on outcomes where (a) I expected results, and (b) data collection costs were lower. I thus focused on dropouts and expenditures, rather than academic performance. Dropouts and expenditures are parental decisions that are easy to adjust, whereas academic performance reflects many other factors.[14] There was sufficient budget to include 900 households in the endline survey sample.

[14] The cost of measuring dropouts and expenditures was also much lower, since they could be assessed through surveys with parents, and a household visit with a subset of parents was already being conducted for other reasons (see Online Appendix I.2). Measuring academic performance would have been considerably more difficult, likely requiring the direct administration of a test to students at home: availability of school data would be endogenous to dropouts, and very costly to gather and match to sample data. Schools often also do not keep good historical records.

The administrative attendance data was gathered by giving schools templates to record the data for the month following the intervention, and was collected from schools covering 35% of the sample. During the attendance data collection, we were able to collect data on endline exams for 7% of the sample, which allows me to validate the accuracy of the baseline academic performance measure but does not give sufficient statistical power to look at treatment effects. Online Appendix I provides more detail on the endline sample and data.

3.3 Measurement of baseline and endline beliefs

During the baseline survey, surveyors measured baseline beliefs by asking parents about the same performance metrics later delivered in the intervention report cards – average scores and percentile rankings on previous-term school exams in math, English, Chichewa, and overall. We used the same measure used in the intervention so that any gap between believed and true performance represents belief inaccuracies, not differences between measures, and because parents identify it as the most important measure for decision-making. Beliefs uncertainty was measured by asking parents to distribute tokens across bins representing score ranges (e.g., 0-20).

During the first endline survey, I assess beliefs updating. Because I wish to assess whether information affects the beliefs underlying parents' behavior, I want to know both whether (a) parents understood and believed the information presented in the intervention, and (b) the information is relevant for their decisions going forward. As a result, parents were asked what score they thought their child would receive if he took an exam *that same day*: Asking about the previous-term scores as done in the baseline survey would only have measured (a), since those exams happened in the past, and parents may attribute those scores to idiosyncratic events such as illness; in contrast, asking about current performance proxies for both (a) and (b). In the analysis, I refer to these beliefs as endline beliefs.

3.4 Summary statistics and balance

Table 1 presents summary statistics and tests for balance across the treatment and control groups. 77% of respondents are female, and 92% are the primary education decision maker in the household. Average levels of parental education are low, at 4.7 years. Households are large, with an average of 5 children. Sampled children were 12 years old on average, primarily aged 8 to 16, with 51% female. To test balance, I regress each variable on a dummy for being in the treatment group. The differences between the treatment and control groups are never large, only one of the 39 variables tested is statistically significant at the 5% level: baseline math scores. This is unlikely to confound the treatment effect estimates since the analysis focuses on heterogeneity in treatment effects by performance. However, to ensure this does not affect the results, all regressions control for an academic performance measure, although the results are robust to omitting this control.

4 Empirical results

In this section, I begin by demonstrating that parents have inaccurate beliefs about their children's academic ability. I then show that their belief inaccuracies affect their investments, both immediate and longer-term ones. In the remainder of the section, I discuss the role of uncertainty in beliefs, and then the implications of these information frictions for welfare, the average level of investment in education, and the intergenerational persistence of inequalities.

4.1 Beliefs

Result 1: Parents' beliefs about academic performance appear to be inaccurate.

Data from the baseline survey can be used to assess the accuracy of parents' beliefs about their children's "academic performance," i.e., scores on school-administered exams the previous term. Panel F of Table 1 presents the average absolute value of the gap between parents' beliefs about their children's academic performance and their children's true academic performance. Scores are absolute percentages, expressed on a scale from 0 to 100.[15]

The average gap is large: 20 points, or 1.2 standard deviations of the performance distribution for overall performance. Beliefs about between-subject (math vs. English) and between-sibling (child 1 vs. child 2) performance are also inaccurate, with beliefs about the inter-sibling gap incorrect by an average of 1.1 std. dev., and 31% of parents wrong about which of their own children is higher-scoring. While parents overestimate on average, 21% of parents underestimate.

As described in the conceptual framework, because the true and believed performance distributions have similar variances, the inaccuracies in beliefs should cause beliefs to be an attenuated function of true scores. Figure 3(a) substantiates this with a local linear regression of beliefs on true performance: the slope is visually less than 1. The lines are shown separately for the treatment and control groups to show that there is baseline balance. The attenuation in the slope captures the fact that the correlations between believed and true performance are low: 0.3 for overall performance, as depicted in the graph, and 0.2-0.3 for performance in the individual subjects like math. Since these tests determine progression through school, these inaccuracies are likely relevant for a broad range of investments.[16]

[15] Online Appendix Tables E.2, E.3, and E.16 show the main results in the paper using relative performance (i.e., within-class percentiles) instead of absolute performance and show that they are robust. (The first two tables show beliefs, the third shows other outcomes.) The correlation between absolute and relative performance is 0.8. Online Appendix I.2 further discusses the reasons for using absolute performance for the analysis and shows results when both performance measures are analyzed simultaneously; among other reasons, parents seem to respond more to absolute than relative (see Online Appendix Table E.17).

[16] A reasonable question is whether these patterns simply reflect noise in the performance measure, or

Result 1A: Less-educated parents have less accurate beliefs.

Table 2 presents the results of the following regression testing for heterogeneity in the attenuation of beliefs by parental education:

$$\tilde{A_{ij}} = d_0 + d_1 A_{ij} + d_2 A_{ij} \times Educ_i + d_3 Educ_i + \varepsilon_{ij} \tag{1}$$

where A_{ij} is child j's academic performance, $\tilde{A_{ij}}$ is parent i's baseline beliefs about child j's academic performance, and $Educ_i$ is household-average years of parental education.[17] The prediction is $d_2 > 0$: more-educated parents have less attenuated beliefs. The table shows that d_2 is strongly positive. The magnitudes of the estimates suggest that going from 0 years of education to 8 years, which would mean finishing primary school, increases the slope by 43% (column 1). Note that parental education could be correlated with other dimensions of SES; throughout the paper I use education as a proxy for SES.

An alternate way to look at belief accuracy is to test whether the absolute value of the gap between beliefs and true scores is larger for less-educated parents. Appendix Table A.1 presents the results of this test with consistent results. The table also shows that less-educated parents have more uncertain beliefs, but are not significantly more overconfident.

Result 1B: Providing information affects beliefs.

Having demonstrated that parents have inaccurate beliefs, I next ask whether information changes beliefs and decreases attenuation. Figure 3(b) graphs parents' endline beliefs relative to true baseline scores for both the control and treatment groups. Information decreases attenuation: the slope of endline beliefs on true baseline scores is substantially steeper for the treatment group than for the control. This is consistent with Bayesian updating where parents' posterior beliefs move in the direction of the signal.[18]

4.2 Immediate investment results

I now turn to the "immediate investments," i.e., the investment decisions presented to parents in the first endline survey. A key advantage of these investments is that they were designed to have clear predictions for how parents would like to tailor their investments to performance – the "preferred investment function." This allows a cleaner test for the effects of inaccurate beliefs. I begin by using data from the control group to provide motivating

whether parents are better assessors of children's future academic potential than the current test scores. I address these concerns in Section 5.1, e.g., by validating that current scores are much more predictive of future scores than current beliefs.

[17]The household average is used as the base specification since focus group discussions suggested that parents share information, but Section 5.1 shows robustness to a wide range of specifications.

[18]Parents could either be updating about underlying "ability" or about test-taking ability / the grading scale. Since test-taking ability and absolute performance matter for progression in this setting, any of these beliefs are relevant for behavior.

evidence of misallocations, and then turn to the experimental results.

Result 2: Control group parents attempt to tailor their investments to performance, but partly fail.

Data from the control group can be used to study how baseline parental investments depend on parents' beliefs about performance. This can give us insight into parents' preferred investment function and the likely production function that they have in mind. We can then compare this with how investments vary with *true* performance. A divergence between the two relationships would suggest that inaccurate beliefs may affect investments.

Figure 4 compares the preferred investment function (investments plotted against believed performance – the dashed lines) with the actual investment function (investments plotted against true performance – the solid lines). Note that the y-axes for both lines represent investments, but the x-axes differ. Both are locally linear regressions using control group data.[19] I first interpret the preferred functions. I then compare them to the actual.

The first immediate investment presented to parents is a choice among free workbooks with different difficulty levels. We gave parents four free books – an English and a math book for each of their two children. For each subject and child, parents were allowed to choose betwen three levels of difficulty: beginner, average, or advanced. Panel (a) of Figure 4 presents the choice results for math and English workbook difficulty choices graphically. The y-axis represents the chosen difficulty level. I focus on the dashed line, which represents parents' preferred investment given their beliefs about their child's math or English score, represented by the x-axis. The 3 choices are parametrized on the y-axis as 0/1/2 for simplicity, with the results robust to other parametrizations. The obvious prediction is that book difficulty choice should increase in perceived performance, and indeed, consistent with this prediction, the dashed lines for both English and math slope steeply upwards.

Panel (b) of Figure 4 presents similar results for a second investment, the willingness to pay (WTP) for subject-specific textbooks in math and English. WTP was evaluated using a Becker-DeGroot-Marschak (BDM) methodology, which gives respondents an incentive to report truthfully (see Online Appendix G for description). The textbooks are remedial (i.e., perceived by teachers as substitutes with performance), and so the prediction is that WTP will be higher for the subject in which parents think their child is doing worse. The use of the between-subject WTP (math − English) holds constant other factors such as the child's overall performance, which is advantageous for this test as it provides clean predictions.[20]

[19] For outcomes with multiple observations per household, clustered confidence intervals for the graphs are created using a block bootstrap, drawing bootstrap samples clustered at the household level, re-running the locally linear regressions, and displaying the 5th and 95th percentiles for each point.

[20] For example, consider a parent who received negative information about math. Because the math textbook is remedial, holding all else constant, the parent's WTP for the math textbook should increase. However, all else is not held constant: The negative shock to math performance is correlated with a negative

In Panel (b), as in Panel (a), the dashed lines are the preferred investment lines. The x-axis shows beliefs about performance in English relative to math. The y-axis shows the log of WTP for the math textbook minus the log of WTP for the English textbook:[21] For presentation purposes, I have flipped English relative to math, so that the prediction is that the line should have a positive slope. Indeed, the dashed line slopes steeply upwards, consistent with the prediction that WTP increases the more behind a child is in a given subject.

Both of these investment choices have clear predictions for parents' beliefs about the "right choice" (i.e., the perceived production function). An additional advantage is that both have clear predictions for the *actual* right choice and true production function. For example, the advanced workbook was designed specifically to be better for the higher performers in the sample. This enables a stronger argument that misallocations due to wrong beliefs about child performance lower actual returns.

That said, we may wish to consider larger investments where more is at stake. Secondary schooling is the first high-cost educational investment in Malawi. Few parents in the sample could afford these fees for all of their children; many cannot for even a single child. My next investment introduces a short-run, real-stakes proxy for secondary schooling. We conduct a lottery, in which the prize is four years of government secondary school fees for one child in every 100 households, worth roughly 120 - 240 USD at the time of the experiment. Parents were given nine tickets for the lottery and were asked to allocate the tickets across their two children. There are many "binary" choices in education where credit-constrained parents are forced to choose between a lumpy investment in one child or the other, for example if parents can only afford to send one child to secondary school or college. The lottery ticket allocation – and in particular, which child the parent allocates more tickets to – was designed to proxy for these types of decisions.[22]

shock to overall performance, which means that, say, the perceived probability that the child would drop out of school, rendering the textbook useless, might also increase. The net prediction is thus ambiguous for looking at the math textbook WTP on its own, but not ambiguous for looking at between-subject WTP, which holds the overall performance shock constant.

[21]Results are robust to using levels instead of logs. Only 6% of observations are 0's; I replace the 0's with the log of 10% of the lowest value of the price list, but the results are nearly the same if I drop these observations, replace them with 50% of the lowest value, etc.

[22]Although a single ticket could have also accomplished this goal, multiple tickets were employed for two reasons: first, to increase the power to detect small shifts that may have been inframarginal if there were just one ticket; and second, to allow me to make use of this lottery in a separate paper for studying inequality aversion. As expected, over 75% of parents split their nine lottery tickets as evenly as possible, consistent with an aversion to inequality between their children. The analysis thus reduces in most cases to analyzing which child the parents give their ninth ticket to, which proxies for the child they would choose for a binary choice. Some suggestive evidence that parents do make different discrete choices across children is that 25% of Malawian households with two children aged 14 or above have one child who has finished third grade and one who has not (Malawi DHS). Although this may partly reflect variation in children's choices rather than their parents', the evidence is suggestive. Presumably the differences amplify as children move through

There are two main channels through which academic performance would affect the expected return of a lottery ticket. First, most parents believe the earnings returns to secondary school is higher for higher-performing students. In the baseline survey, returns were perceived to be 70% higher for a hypothetical child in the top decile of performance versus for one in the bottom. This is consistent with studies in other countries (e.g., Aizer and Cunha (2012)). Second, the probability of admission to secondary school increases with performance, with admissions governed solely by performance on a standardized achievement test, and so the expected value of the fees paid and the probability of attending increase with performance.

Panel (c) of Figure 4 shows the lottery ticket allocation results. The dashed line plots the difference in tickets allocated to the older versus the younger child in the pair, with the x-axis the gap in perceived scores between the older and younger child.[23] Consistent with the prediction, the line slopes upwards: Parents give more tickets to the child they think is higher-performing.

I now compare the slope of the preferred investment functions just discussed with the slopes of the *actual* investment functions, depicted by the solid lines in the three panels of Figure 4. The solid lines have the same y-axis as the dashed lines, but a different x-axis: Their x-axis is true performance instead of believed performance. The prediction is that if parents base their investments on their inaccurate beliefs, then the slope of their investments on true performance will be attenuated relative to the slope on beliefs. The graphs show precisely this pattern: The slopes on true performance are only 20-37% as large as the slopes on beliefs. This suggests that parents try to tailor their investments to performance, but that their inaccurate beliefs prevent them from doing so. Since returns depend on *true* performance, if parents *knew* that, say, their child had a math score of 80, they would choose the highest difficulty book for him, but many parents do not know that and so fail to choose their preferred option. This evidence is suggestive, however, not causal; both beliefs and performance could be correlated with other factors that determine investments. An experiment, in contrast, can establish causality: I can test whether, in fact, information undoes the attenuation. It is this I turn to next.

Result 3: Information increases the slope of investments.

I now use the information experiment to test whether information increases the slope of investments on actual performance. Figure 5 shows locally linear regressions of investments against true performance for the treatment group (dashed line) and control group (solid

school, but I could not find data for older children.

[23]Since the lottery is a within-household allocation, to depict it graphically, we need to choose how to order the two children (i.e., to show child "1" relative to child "2"). Parents identified age in focus groups as the second most important factor for investment (behind performance), so I order using age, but the graphs look similar with any order.

line). Both lines have true performance as the x-axis. Note that Figure 4 examined only the control group, and thus the solid lines in Figures 4 and 5 are identical, as they depict the same data.

For all three investments considered, the figures show that the information treatment substantially increases the slope of the investment functions, thereby confirming that information frictions cause misallocations. I perform a formal test of whether information changes the slope of the investment function by running the following regression:

$$s_{ij} = c_0 + c_1 A_{ij} \times Treat_i + c_2 A_{ij} + c_3 Treat_i + c_4' X_{ij} + \varepsilon_{ij} \tag{2}$$

where i indexes households, j indexes siblings, s_{ij} is the investment decision, A_{ij} is the relevant academic performance metric (e.g., math for math workbooks), $Treat_i$ is an indicator for being assigned to the treatment group, and X_{ij} is a vector of control variables.[24] Standard errors are clustered at the household level.

The prediction is that the information treatment makes the slope steeper, so that $c_1 > 0$.[25] The key prediction regards c_1; c_3, the coefficient on $Treat_i$, is not particularly meaningful as it is just driven by the scaling of the A_{ij} variable, representing the treatment effect for those for whom $A_{ij} = 0$ for the particular A_{ij} measure used in that regression. For example, for the textbook regression, it is the treatment effect for those for whom math and English performance are the same (i.e., $math - English = 0$).

Table 3 presents the results using math and English workbook difficulty choices; the log of WTP for the math textbook minus the log of WTP for the English textbook; and the secondary school lottery tickets received, respectively. Since secondary school lottery tickets are inherently a within-household allocation (one child's allocation fully determines the other's), the lottery regression is estimated with a household fixed effect. Consistent with the graphical evidence, across all outcomes, c_1 is positive, statistically significant, and large in magnitude. Comparing the coefficient on $Score$ (slope in the control group) with the sum of the coefficients on $Score$ and $Treat \times Score$ (slope in the treatment group), information causes investments to become 2.7-5 times more steeply aligned with performance across the various investments, i.e., the slopes increase by 170-400%.

Section 5.2 examines the robustness of these patterns to different specifications, and addresses potential interpretation concerns.

[24] Results are robust to excluding the controls (see Section 5.2). Control variables include school fixed effects (FE's), the between-child score gap, and parents' education level. Note that this includes all variables underlying the stratification but not the stratum fixed effects themselves; I pre-specified that I would not control for stratum FE's because some of the stratum are very small, and so 20% of observations would be lost if stratum FE were included. The results are, however, robust to controlling for stratum FE's.

[25] The prediction would be $c_1 < 0$ for textbooks since they are a perceived substitute but, for presentation purposes, I have flipped the sign of the textbook dependent variable so that all coefficients should be positive.

Result 3A: The effects of information can be larger for less-educated parents.

As shown previously, less-educated parents appear to have less accurate beliefs. I now examine whether they respond more to information. Testing this is non-trivial, since it is difficult to define exactly what "respond more" means in the data. In particular, the size of parents' responses will depend on their preferred investment functions. These preferred investment functions may vary by parental education, since different parents face different constraints. For example, say that we were to study spending on college, and that only richer, more-educated parents could ever afford college. Because spending among less-educated parents would be constrained at 0, the effect of information on college spending would necessarily be larger for more-educated parents. However, this would not mean that inaccurate beliefs matter more for them in general, just for this particular input, for which their preferred investment function is different.

One clean way to examine the heterogeneity is to look at the effect of information on beliefs themselves and test whether treatment causes less-educated parents to update their beliefs more. We also wish to know, however, whether this translates into larger effects on their investment decisions. So as to avoid concerns regarding variation in preferred investment functions across parents, I wish to use an investment choice where the preferred investment function is as homogeneous as possible across parental education levels. The choice of difficulty level of free workbooks described above is most likely to meet this criterion, and was expressly included in the design to provide homogeneity across education levels. Since the workbooks are free, the choice should not be confounded by wealth. Moreover, we expect most parents to choose the workbook difficulty level that most closely matches their beliefs about their child's performance level, and there is no clear reason to expect heterogeneity in that behavior by parental education.

Table 4 shows the results of estimating equation (2) fully interacted with household-average years of parent education. I begin with the outcomes offering the cleanest test: beliefs, and the English and math workbooks. The results, presented in columns (1)-(3), suggest that information has a larger effect for less-educated parents. The baseline slopes in all cases are more attenuated for less-educated parents: The coefficient on *Score* $\times$ *Parent yrs of educ.* is positive, and is statistically significant in the regressions for beliefs and math workbooks. Moreover, the treatment effect on the slope is larger among the less-educated: The coefficient on *Treat* $\times$ *Score* $\times$ *Parent yrs of educ.* is negative and significant in all three cases. Extrapolating linearly, the magnitudes of the effects are economically significant. At baseline, the workbook choices of above-median-education parents are roughly 90% (40%) more steeply sloped for math (English) than the choices of below-median-education parents; information fully closes the gap.

Cols. (4) and (5) present similar estimations using the two additional investment choices described previously: WTP for textbooks and secondary schooling tuition lottery results. Here, in contrast to the workbook choice, the predictions are less clear. For both choices, there is greater potential heterogeneity in the preferred investment functions, for example due to credit constraints or different levels of aversion to inequality between children. Unsurprisingly, the estimates for these two choices are qualitatively consistent but weaker.

4.3 Longer-term outcomes

The above results demonstrate that inaccurate beliefs affect parents' investments in education. An open question, however, is the extent to which correcting these inaccuracies impacts investment decisions over the longer run. I next turn to longer-run data on retention in school, educational expenditures, and attendance to show that information frictions are also relevant for parents' larger, longer-term, investment decisions. The advantage of these data is that they allow us to gauge the persistence of the earlier results. However, the *ex ante* predictions for the preferred investment function are generally not as clear.[26]

Result 3B: Information affects the slope of longer-term investments.

First, I examine the effect of information on the slope of investments. Panel A of Table 5 presents estimations of equation 2, with each regression first estimated in the full sample, and then estimated fully-interacted with parent years of education. All regressions use overall scores as the performance measure. To have the statistical precision to test hypotheses, it is useful to use continuous variables that capture all the variation in the data as done in Panel A. To aid in interpretation, however, Panel B shows estimates using binary regressors for both performance and education, specifically: indicators for whether a student has above-median score and for whether a household has above-median parent education.

I consider three outcomes: primary school retention (dropouts), attendance, and expenditures. Of the three, primary school retention should provide the cleanest test: Most parents believe schooling is more valuable for higher-performing children, whereas parental beliefs about the complementarity of expenditures or attendance with performance, as elicited in interviews, varied widely across parents. The literature on attendance and expenditures is also limited, and there is little reason to expect the production function to be the same as for years of schooling. For example, conditional on having chosen to keep a child enrolled in school, parents may need to invest more in their lower-performing children to keep them on track. Moreover, parents in Malawi are extremely poor on average, and expenditures in

[26] It is also harder to use control group data to generate predictions for the likely production function parents have in mind; compared with the immediate investments, these investments have many more omitted determinants, and so the observational regressions are more difficult to interpret. However, we can still use the information treatment effects themselves to infer the perceived complementarity/substitutability of the investments with performance.

general are accordingly low.

Column (1) shows the primary school retention results in the full sample. Consistent with the fact that nearly all parents believe years of schooling are a complement with academic performance, information increases the slope of the investment function. High-performing students in the treatment group are more likely to be enrolled in school one year later, while low-performing students are less likely.[27] The change in the slope in Panel A is significant at the 1% level. Panel B suggests that the magnitudes are economically meaningful. Among children whose parents found out they had above-median performance, dropout falls by 2 percentage points to nearly 0% (from a control group mean of 2%), whereas it roughly doubles, increasing from 2% to about 4% for those with below-median performance. These results highlight that information does not improve educational outcomes for all: it leads to reallocations, which can decrease investments for some. Since the literature suggests that schooling and ability are complements, these results are consistent with an improvement in returns (Pitt et al., 1990; Aizer and Cunha, 2012). Column (2) shows that the retention point estimates are larger among less-educated households. However, precision is low, so the differences are not statistically significant, despite the fact that the treatment effects on the gap in retention between low- and high-performing students are roughly twice as large for below-median education parents as above-median. This may partly reflect the fact that baseline retention rates are higher for households with above-median parent education.[28]

In contrast to the results for primary school retention, but perhaps to be expected given parents' heterogeneous beliefs regarding complementarity with performance, I find no significant effects in the full sample (columns (3) and (5)) for either expenditures or attendance. This is not, however, because parents did not respond to the information. Rather, it is because less-educated and more-educated parents responded in opposite directions, thus obscuring the effects in the full sample (columns (4) and (6)). This can be seen most clearly in Panel B. For expenditures, information causes less-educated parents to increase their spending on their lower-performing children relative to their higher-performing children by 18% (*Treat* × *Above-median score*).[29] More-educated parents do the opposite, increasing spending on their *higher*-performing children relative to their lower-performing children by 10% (sum of the coefficients on *Treat* × *Above-median score* and *Treat* × *Above-med.score* × *Above-med.par.educ*). Here, the base effect for the less educated – the omitted category – is

[27]Many evaluations use self-reported enrollment as the outcome of interest (e.g., Bourguignon et al., 2003; Schultz, 2004), but Baird and Özler (2012) show that self-reported and school data do not always match. I have dropout data from 10% of the schools and, reassuringly, the coefficient on $Treat \times Score$ is the same regardless of the data source used, reflecting a high correlation between measures (0.5).

[28]Above-median households have a 99.5% retention rate relative to 96.7% for below-median.

[29]Logs are used for precision but only 1 percent of observations are 0; results are robust to other specifications (e.g., taking log of 1+expenditures or log of expenditures plus the lowest value ×10% or 50%).

negative (coefficient on *Treat* $\times$ *Above-median score*). Thus, a positive coefficient on *Treat* $\times$ *Above-med.score* $\times$ *Above-med.par.educ* that is lower in magnitude than the base effect would imply that the magnitude of the effect is larger for the less-educated; on the other hand, a positive coefficient that is larger in magnitude than the base effect, as we see here, "flips" the direction of the effect, and results in parents spending more on their *higher-performing* child rather than simply spending somewhat less on their lower-performing child.

A similar pattern holds for attendance. For the less-educated, information increases the attendance of low-performing children, whereas it does the opposite for the more-educated. For both expenditures and attendance, the results are not driven by selection into schooling, as the results are the same when one controls for or estimates bounds based on enrollment.

In Section 4.2, the heterogeneity by parent education in the treatment effects was caused by heterogeneity in belief accuracy by parent education. Here, that factor may also play some role, but cannot explain the change in the sign of the effects (positive for one group, negative for another). Rather, the results here suggest that there is heterogeneity in the *preferred investment function*. One potential explanation is that more-educated parents (who are likely to be richer) believe they can afford to send their children to secondary school, and so want to get their high achievers over the admission threshold, whereas less-educated parents do not see secondary as an option, and so have higher perceived returns to helping low achievers acquire basic skills like literacy. While this story is more about wealth than education, education is the best-measured proxy for wealth in the data. There are of course other possible explanations.

Importantly, uncovering different preferred investment functions does not mean some parents are "right" and some are "wrong" – as highlighted above, the constraints faced by the different types of households could differ substantially, causing the returns-maximizing action to differ as well. This finding also does not conflict with Result 3A (that information can have larger effects for less-educated parents). As discussed in that section, in order to evaluate whether less-educated parents respond "more," it is important to use investments where there is limited heterogeneity in the preferred investment function. Investments and expenditures obviously do not satisfy that condition. Thus, the results using beliefs and workbooks are the more instructive results for whether information matters "more" for less-educated parents. The expenditure and attendance results are more helpful for understanding how the perceived production functions for these particular investments vary by parent education.

4.4 Uncertainty

The previous sections indicate that information affects the slope of investments on true performance, thus suggesting that the slope was attenuated at baseline. As outlined in the conceptual framework, both inaccuracies in the mean of baseline beliefs and uncertainty of baseline beliefs could cause baseline attenuation in the slope. A reasonable question is thus whether the channel for the treatment effects is an effect on the mean or on the uncertainty of beliefs. I did not experimentally vary uncertainty separately from the mean, nor do I have an endline measure of beliefs certainty, so the analysis of the channels is suggestive in nature. Under an uncertainty channel, uncertainty could decrease the preferred slope of investments as a function of *beliefs*, since parents may not want to invest as steeply based on their beliefs if their beliefs are uncertain.[30] The attenuation of preferred investments on beliefs would then cause attenuation of actual investments on *true* performance – which is the attenuation that has been the focus of the analysis so far. In contrast, under the channel of inaccurate means, the slope of investments as a function of mean beliefs is not attenuated; rather, the attenuation of investments on true performance stems from the fact that, because beliefs are inaccurate, they themselves are attenuated functions of true performance. As a result, one empirical signature of the uncertainty channel is attenuation of investments on beliefs themselves; to assess uncertainty's role, I test whether information increases the slope of investments on beliefs.[31] Note that an implicit assumption is that information increases the certainty of beliefs. If information does not affect certainty, then changes to uncertainty cannot provide a channel for the treatment effects, and if it decreases it, we should see the opposite effect. I use two approaches, with results consistent for both.

Result 4: Decreasing the uncertainty of beliefs seems to affect parents' larger investments, but has limited effect on their smaller investments.

My first approach looks at the treatment effect on the slope for those who have relatively accurate beliefs at baseline. For this group, there is no belief accuracy effect of information (since beliefs were accurate to begin with). Any slope change therefore will likely represent an uncertainty effect. Panel A of Appendix Table A.2 shows the results of estimating equation 2 for parents whose beliefs regarding their children's performance were within 10 points of the true score. For the smaller investments, such as workbooks, the slope for these parents changes a little (i.e., there is a small uncertainty effect), but the effect is only 30% of the magnitude – and significantly different from – the change in slope in the full sample. This suggests that the effect presented earlier for the full sample effect is driven primarily by changes to belief accuracy. This is not surprising, since the preferred investment function was already steeply sloped in the control group. For the larger investments, on the other hand, the uncertainty effects are larger, with effects in the accurate beliefs sample representing 50% of the coefficient estimated in the full sample for the lottery, and 100% for retention. Of course, a caveat is that parents with accurate beliefs could be different from other parents.

[30]See Appendix C.1 for a framework yielding this prediction.

[31]I discuss other potential interpretations of a change in the slope of investments on beliefs in Section 5.2.

A second approach is to test whether the heterogeneity in the treatment effect by performance is equal and opposite to the heterogeneity by baseline beliefs. Suppose preferred investments as a function of baseline beliefs take the form $\beta_0 + \beta_1 \tilde{A}$. If information does not change the *preferred* slope, this means that all information does is move parents along the preferred function by the amount of the information shock $(A - \tilde{A})$. In that case, the treatment effect would be $\beta_1(A - \tilde{A})$, and the coefficients on $Treat \times A$ and $Treat \times \tilde{A}$ would be equal and opposite: β_1 and $-\beta_1$, respectively. If, instead, the magnitude of the coefficient on $Treat \times A$ is larger than that of $Treat \times \tilde{A}$, it suggests that beliefs about academic performance are more important to treatment parents' investments than to control parents', i.e., the slope of investments on beliefs has increased. To see this, denote the slope of the investment function in the control (treatment) group β_1^C (β_1). Parent i with baseline beliefs $\tilde{A}_i$ and true performance A_i would have investment of $s^C(\tilde{A}_i) = \beta_0^C + \beta_1^C \tilde{A}_i$ in the control group, and $s(A_i) = \beta_0 + \beta_1 A_i$ in the treatment group. Thus, the treatment effect as a function of A and $\tilde{A}$ is $\tau(A_i, \tilde{A}_i) = s(A_i) - s^C(\tilde{A}_i) = (\beta_0 - \beta_0^C) + \beta_1 A_i - \beta_1^C \tilde{A}_i$, and so heterogeneity in the treatment effect by A identifies β_1 and heterogeneity by $\tilde{A}$ identifies $-\beta_1^C$. Panel B of Appendix Table A.2 presents the results. The results of this test are consistent with the previous test, with the lottery and retention being the only investments where we can reject that the coefficients are equal and opposite at the 5% level.

Note that this section focused on a specific effect of uncertainty on investments, namely, whether changes to uncertainty contributed to the core treatment effects analyzed in this paper: the treatment effects of information on the alignment of investments with child performance. There are also several other ways that uncertainty can affect investments that are not the focus of this paper (see for example Bobba and Frisancho (2016)).

4.5 Welfare and average treatment effects

Decreasing information frictions about academic performance appears to have a substantial effect on parents' investments. A natural question to ask is: what are the welfare implications? If inaccurate beliefs about academic performance were the only market friction in the world, then we could unambiguously say that since parents respond to information, information increases welfare. In general, however, an intervention that corrects one market imperfection can decrease welfare if there are multiple interacting market failures (the "theory of the second best"). Evaluating the welfare effects of any intervention is thus difficult, since no single outcome can fully summarize the welfare impacts. One way to conceptualize the welfare issue is to think of the exercise in this paper as answering the following question: If all other market failures impacting education were fixed, would there still be large inefficiencies due to inaccurate beliefs? The results suggest that the answer is yes.

A second way to think about welfare is to think through the potentially interacting market failures and assess their likely impact. Although it is impossible to do this comprehensively, I discuss some key examples. In this context, my analysis suggests that information increases welfare in the face of the key interacting market failures. That said, these analyses are just suggestive, and there is an important caution about external validity: In some settings, the interactions may be different, and could lead to a decrease in welfare.

One example of a potentially interacting market failure is misinformation about the production function. Here, the concern is that parents are misinformed about the production function – and in particular, about the complementarity between investments and performance – and this causes them to in fact invest less efficiently as a result of receiving information about their children's performance. This is not a concern when analyzing some of the investment choices presented to parents in the experiment (e.g., the workbooks and remedial textbooks). These investments were designed to have clear predictions for increased returns, and, across the parental education spectrum, parents' reallocations align with the predictions. For the outcomes that proxy for years of schooling (i.e., primary school retention and the secondary school lottery), although there are no estimates of the production function in Malawi, estimates from other contexts suggest that years of schooling and ability are complements (Pitt et al., 1990; Aizer and Cunha, 2012), and there are reasons to believe that the complementarity may be greater in this setting.[32] My finding that parents allocate more years of schooling to their higher performers suggests they believe this complementarity exists, and is therefore consistent with parents being correct about the production function.

A second example is that any of several classic market failures (e.g., credit constraints, externalities) leads to underinvestment in education. The concern would then be if providing information about academic performance caused the average level of investments to fall due to, say, biased parental beliefs. In a world with no other market failures, even if information decreased the average level of educational investments, it would still represent an improvement to welfare, but because of the existence of other market failures, this could represent a potential concern. However, as we will see momentarily, reassuringly, information does not decrease the average level of investments.

Result 5: Information does not decrease the *average* level of investments.

Given the preceding discussion, if we were to see a negative average treatment effect (ATE), this would be reason to be concerned about potential welfare decreases due to interacting market failures. Panel A of Appendix Table A.3 presents the average treatment effects (ATEs) of information. Reassuringly, I do not find any statistically significant average treatment effect of information on the investments that proxy for the overall level of

[32]For example, Duflo et al. (2011) suggest that teachers in sub-Saharan Africa have incentives to target instruction to the high performers, potentially increasing complementarity relative to other contexts.

investments: retention, expenditures, and attendance.[33] It might also be concerning from an educational inequality perspective if information decreased the level of those investments among less-educated households relative to more-educated households, but Panel B of Appendix Table A.3 assuages that concern. This is not surprising since there was also no significant heterogeneity by parent education in overconfidence, just in belief accuracy.

One might be surprised by the absence of an average treatment effect for retention. Parents on average overestimate their children at baseline, and, for retention, invest more in their higher performers, suggesting that providing information might decrease retention. There are several non-mutually exclusive potential explanations for why I do not find an average effect. First, uncertainty in the control group may decrease investment, akin to uncertainty dampening investment in risky assets. Second, parents may already be spending as much as they can on education, and so the effect of information is primarily on the allocation of spending, not the level. Third, parents could respond more to positive than negative information. Finally, we may lack statistical precision.

I do not observe much evidence for the first channel: There is no positive average level effect for the parents who had more accurate beliefs at baseline, though the power of the test is low. See Panel C of Appendix Table A.3. The second channel is unfortunately difficult to test. I find evidence that the third channel may play a role, however, as I discuss next.

Result 6: Investments respond more to positive than negative shocks.

Appendix Table A.4 shows the results from estimating equation 2, fully interacted with a dummy for receiving a positive information shock ($A_{ij} > \tilde{A_{ij}}$). The model is estimated for all outcomes for which (a) one direction of shock is unambiguously positive; and (b) there is a treatment effect on the slope in the full sample.[34] I find that the change in slope (i.e., coefficient on $Treat \times Score$) is larger for parents who receive positive information shocks.[35] For retention, precision is lacking, but the magnitude of the coefficient is large, suggesting that this channel could help explain why there is no negative ATE. Of course, positive shocks are not randomly assigned, so the results are only suggestive.

[33]The immediate investments were designed specifically to look at reallocations, and their level does not proxy for overall spending; for example, a decrease in relative spending on math versus English textbooks or in the average difficulty level of the workbook chosen does not imply a decrease in overall educational investments. For completeness, however, these ATEs are also reported in Appendix Table A.3.

[34]As an example, the lottery outcome is not estimated, since it is a function of between-child performance and it is thus unclear which direction would be "positive."

[35]One potential concern is if the positive information shocks were larger, but that is not the case: The absolute gap between believed and true performance is roughly 40% smaller for the positive information shock sample. Another potential concern is that some actions are bounded (e.g., one cannot choose a less difficult workbook than beginner), but restricting the sample to parents whose predicted behavior based on baseline beliefs is in the middle of the range of potential outcomes yields similar results.

4.6 How much of the SES gap in outcomes can belief inaccuracy explain?

Children from higher SES households have better educational outcomes. Result 3A suggests that inaccurate beliefs may play a role. I now present a back-of-the-envelope calculation for the share of the SES gap in outcomes that information could close. This paper has examined a range of investments; I focus here on retention, since school enrollment is an outcome of standalone interest, and so fewer assumptions are needed to translate treatment effects into implications for the ultimate outcomes we care about, like completed schooling. Retention does have downsides, however, including the fact that baseline dropout rates are higher among less-educated households; Appendix D shows robustness to using other outcomes with more homogeneous preferred investment functions.[36]

It is important to note that this calculation is simply meant to be suggestive. First, it involves taking point estimates seriously: although the heterogeneity in the retention treatment effects between below-median education and above-median education households (hereafter: low-SES and high-SES) is large in magnitude, precision is low and I cannot reject equality. Second, the calculation relies on several assumptions; Appendix D details the assumptions and shows robustness. A third important caveat is that education is not welfare; we might care more about whether information closes the gap in welfare by SES than the gap in educational outcomes by SES. However, the assumptions needed to estimate the effects on welfare would be relatively heroic, so I perform the calculation with education.

The idea behind the calculation is to compare the projected SES gap in outcomes in the control group and in the treatment group. The information treatment reallocated dropouts from higher-performing to lower-performing students. This improves projected outcomes (e.g., primary school completion, earnings) since higher-performing students have higher expected school attainment (Hunt, 2008). Furthermore, if schooling and performance are complements, this implies an improvement in future earnings as well (Aizer and Cunha, 2012). The reason that information may narrow the projected gap in outcomes by SES is that the effects of information on retention are larger among low-SES households.

Primary school completion I first use the annual dropout rates in the control group among low-SES and high-SES households to project the baseline primary school completion rates by SES. In my data, baseline dropout rates are lower for high-SES (0.5% annual dropout rates for high-SES relative to 3.3% for low-SES), and thus projected primary school

[36] Besides heterogeneity in the preferred investment function, a second downside of retention is that precision is low due to a low base dropout rate. At the other end of the spectrum is workbooks, with the highest power and least heterogeneity in preferred investments, but farther removed from the ultimate outcomes of interest. Appendix D.4 presents a similar calculation for workbooks: Information closes 88% of the SES gap for math books and 100% for English.

completion rates higher. This is consistent with the literature from Malawi and elsewhere that dropouts are higher among low-SES than high-SES (World Bank, 2010; EPDC, 2007).

I then use the estimated treatment effects on retention at the (performance $\times$ SES) level to project primary school completion in the treatment group. The treatment effects are larger among low-SES households, with roughly twice as large an effect on the spread in dropouts between high-performing and low-performing students (4.6 percentage points for low-SES relative to only 2.3 percentage points among high-SES: see Panel B of Table 5).

Under the conservative assumption that dropout rates are twice as high among below-median students as above-median students,[37] the results of this exercise are as follows: At baseline, the projected primary completion rate is 0.96 among high-SES households and 0.76 among low-SES households, yielding a gap of 0.20. The projected rates in the treatment group are 0.89 and 0.79, yielding a gap of 0.10. Information thereby closes roughly 50% of the gap in projected primary school completion.[38]

Secondary school completion and earnings Transition rates to secondary school among those who have completed primary school are higher among high-SES households than low-SES households in Malawi, likely reflecting credit constraints. This means that if we extend the analysis to secondary school, the baseline SES gap is higher, but information narrows less of it: Low-SES households cannot capitalize as well as high-SES households on the secondary school option value of reallocating dropouts. Using assumptions outlined in Appendix D, I find that information decreases the projected gap in secondary completion rates by 14%, from a projected 0.37 in the control group to 0.32 in the treatment group. This is likely a lower bound, since one reason for the differential transition rates to secondary school by SES could be information frictions: If low-SES parents knew years in advance that their child was likely to get in to secondary school, they might be able to save enough.

Reallocating dropouts from low- to high-performing students can improve outcomes through two channels: higher projected attainment, and, with complementarity, higher earnings returns conditional on attainment. The school completion results only capture the first channel; to get at both, I look at projected earnings, and the results are similar, with information decreasing the projected SES earnings by 18%.

[37] Poor school performance is a widely recognized driver of dropout, with estimates suggesting that the dropout rate among below-median performance students is 5-10 times higher than above-median students (Liddell and Rae, 2001; Sabates et al., 2010). To be conservative, I assume the dropout rate is twice as high among below-median students. In my data, I cannot reject that dropouts are 4 times higher for that group. I do not use the point estimates from my sample for the base scenario since they are inconsistent with the literature but imprecise, but Panel B of the table in App. D.3 shows that the results are very similar if I do.

[38] Part of the effect is due to the projected primary school completion rate declining among high-SES households, but, as discussed in Appendix D.2, even if I adjust for that, the estimates still suggest that belief inaccuracies close upwards of 10% of the SES gap.

5 Robustness and additional analysis

I now examine the robustness and mechanisms behind the results presented in the previous section, and present additional results on secondary outcomes.

5.1 Beliefs: Robustness and mechanisms

One potential concern with Result 1 – that beliefs appear to be inaccurate – is whether the "inaccuracies" in beliefs simply reflect noise in the performance measure. The correlations in the data suggest otherwise: Recall that the correlation between tests taken during the term is 0.8 for overall performance, and 0.6-0.7 within subjects, which suggest high test reliability. Importantly, these correlations are much higher than the correlations between parents' beliefs and the term-average scores, which are 0.3 for overall and 0.2-0.3 for subject-level. I also have data on *future* test scores for a small subset of the sample that I can use to validate the use of the current test scores as a performance measure. Using control group data, Online Appendix Table E.1 shows that current test scores are nine times more predictive of future test scores than parents' baseline beliefs are, with coefficients of 0.74 relative to 0.08. This suggests that the inaccuracies in parents' beliefs do not simply reflect noise, and that current test scores are a better predictor of future test performance than parents are. Misunderstanding the difficulty of the grading scale also does not drive the results: The patterns are similar for within-class percentile ranks (Online App. Table E.2).

I now examine the robustness of Result 1A: that beliefs are less accurate among less-educated parents. The base specification uses the average years of education among parents as the parental education measure, but Online Appendix Table E.3 shows that the results are highly robust across a range of measures of parent education, as well as child performance. One may also wonder whether other correlates of parent education (besides SES, for which parental education is proxying) drive the result, such as school quality. Online Appendix Table E.4 shows robustness to controlling for other variables and their interactions with score, including school fixed effects interacted with score.[39]

One may also wonder about mechanisms. Why do less-educated parents have less accurate beliefs? I cannot answer definitively, but some suggestive evidence comes from tracing beliefs as children progress through school. If the primary cause is that more-educated parents can better judge their children's skills, then the gap in belief accuracy might grow as children use more advanced skills (e.g., multiplication instead of addition). If instead more-educated parents can better read report cards and talk to teachers, then the gap might

[39]Children of less-educated parents also have lower scores. Column (7) suggests this does not drive the belief accuracy gap by controlling for a quadratic and cubic in score. All regressions already control for score, so the concern would be if the relationship varied non-linearly with score and education picked this; the higher-order terms assuage this concern.

stay constant. Online Appendix Table E.5 shows that the gap grows as children progress.

5.2 Robustness of information treatment effects on the slope

This section investigates the robustness of the estimated treatment effects on the slope of investments, both the full-sample estimates and the heterogeneity by parent education. Online Appendix Tables E.6 and E.7 show that the results are invariant to excluding controls. For workbooks, the base linear specification makes a strong cardinality assumption; Online Appendix Tables E.8 and E.9 shows robustness to an ordered probit specification that relaxes that assumption. For retention, since the dependent variable mean is near 1, Online Appendix Table E.10 shows robustness to using probit. Online Appendix Table E.11 shows the "beliefs shock" specification.[40]

I now discuss the robustness of the *interpretation*. That is, to show that information increases the alignment of investments with performance, the analyses shown so far are sufficient, assuming the randomization was successful. This alone is important: If returns vary with performance, increasing alignment can increase returns. But, to interpret the channel as changes to beliefs, additional robustness checks are useful.

One potential concern is that performance is not randomly assigned. If there is heterogeneity in the effect of information based on an omitted correlate of performance, it could also change the slope. It is reassuring that, for the immediate investments where we have clear predictions for behavior from the conceptual framework and the control group analysis in Section 4.2, the results fit the predictions exactly. Columns (2) through (5) of Appendix Table A.5 provide further evidence that omitted factors do not drive the result by showing robustness to household fixed effects, and to controlling for child-level controls interacted with treatment. Online Appendix Table E.12 repeats the analysis for the longer-term investments. We lose statistical power quickly, but reassuringly the coefficient for the retention result stays stable in magnitude and the p-value remains ≤ 0.15 across all specifications.

One variant of this concern would be if the treatment effects were driven by information increasing the salience of education. If salience effects were uniform, it would affect the investment *level*, not slope, so the concern would be if salience effects varied and were correlated with performance. But, salience would likely be a household-level effect (or correlated with the child-level controls); the robustness checks above thus assuage the concern.

Another concern is priming: Perhaps the intervention caused treatment parents to invest based on perceived academic performance whereas they do not at baseline. The analysis in

[40]The specification looks at heterogeneity in the treatment effect based on the information received relative to baseline beliefs, $A-\tilde{A}$. This specification assumes that the coefficients on $Treat \times A$ and $Treat \times \tilde{A}$ are equal and opposite. That assumption was rejected for retention and secondary school due to beliefs uncertainty (see Section 4.4 and Appendix Table A.2), and so this specification is not as instructive for those investments; for the investments where the assumption was not rejected, the results are consistent.

Section 4.2 of the baseline data provide some reassurance against this. The control group did not receive information, but their investments are highly dependent on beliefs about performance. A related external validity concern is if having a baseline beliefs survey primed all parents to invest based on performance, thereby overstating information's effects. Such priming would likely fade over time, so the effects on longer-term outcomes like retention are reassuring. Moreover, before the study, parents almost universally identified academic performance as the primary determinant of investments, and baseline expenditures are highly correlated with believed performance.

A final concern is that information causes parents to update about the production function, not academic performance. This would still be an information friction, but the interpretation would be different. For investments where the preferred investment function did not change much in Section 4.4, we can rule out this concern, since changing the perceived production function should change the preferred investment function. However, the preferred investment function did change for the lottery and retention. Above I interpret the reason as an increase in beliefs certainty. In theory, it could also reflect updating about the production function, but the robustness to household fixed effects assuages this concern: Changes to the perceived production function should affect both children in the household similarly.[41]

5.3 Secondary outcomes

In the endline survey, I also collected data on two outcomes which I considered secondary because I did not have *ex ante* hypotheses that there would be effects or expected power was low: transfers across schools and non-monetary investments, such as giving the child fewer chores, or homework assistance. For completeness, these results are presented in Online Appendix Table E.13. Parents indicated *ex ante* that non-monetary investments would respond to their children's performance, but expected power was low since it is difficult to measure these investments cleanly. I find positive average treatment effects but no significant impact on the slope. For transfers across schools, parents did not indicate *ex ante* that it was a margin which would respond. However, information increases transfers by 50%, from 6% to 9%. Although there is no change in the slope on performance, heterogeneity in the preferred slope by school type could explain this. At low-quality schools, finding out a child is doing well might make it worth the effort costs of changing him to a better school, so transfers would be positively sloped with performance. In contrast, at high-quality schools, finding out a child is doing poorly could indicate a poor match, and so transfers would have

[41] For retention, the concern would be that parents who found out their child had poor performance decided that schooling has low returns, and that caused retention to fall among low-performers. This should impact both the parents' children similarly; the fact that adding a household fixed effect does not diminish the point estimate assuages the concern. Parents may be inferring about their children's *individual-level* returns, but that is a semantic distinction: This paper uses academic performance as a proxy for individual-level returns.

the opposite slope. Indeed, if we look at the results separately by school quality (proxied by school-average achievement), there are slope effects, with the slope becoming more positive at low-quality schools and more negative at high-quality schools (Online Appendix Table E.14). Of course, this is just one of many potential explanations – and it implicitly assumes that parents know school quality which may not be the case – but the results are suggestive.

6 Conclusion

This paper tests whether inaccuracies in parents' beliefs about their children's academic performance impact their educational investments. I find that there are large discrepancies between believed and true performance. At baseline, parents try to tailor their investments to their children's performance, but partly fail as a result of inaccurate beliefs. Providing information has a large impact on parents' investments, roughly tripling their responsiveness to academic performance. The impacts are seen across a broad range of investments, from those with the cleanest predictions for efficiency to those which proxy most closely for overall attainment. Even within the fairly homogeneous, low-education context of Malawi, I find significant heterogeneity by parent education. Less-educated parents have less accurate beliefs, and update their beliefs and some investments more in response to information.

The heterogeneity in belief accuracy observed in this paper is also seen in other contexts, such as the U.S.. Taken together, the findings suggest that inaccurate beliefs may serve to perpetuate inequalities across generations, both within and across countries, with back-of-the-envelope calculations suggesting that the channel is quantitatively important. The findings thus relate to a large literature on inter-generational mobility, both in developing and developed countries (Hertz et al., 2007; Black and Devereux, 2011). They also advance our understanding of the role of misinformation in decision-making, relating to literature not just in education but other domains, like health (Dupas, 2011; Madajewicz et al., 2007).

It is perhaps surprising that baseline information is poor if the returns to knowledge are high and the information exists. However, parents may over-estimate their own knowledge, or (perceived) information acquisition costs may be high, as suggested in the U.S. by Bergman (2016). Interviews with parents also suggest that uneducated parents are intimidated to talk with their children's teachers. This is consistent with other studies showing that information constraints matter for education (e.g., Jensen, 2010; Dinkelman and Martínez A, 2014).

A second aspect of this paper is how parents' investments depend on their children's academic ability and endowments. This relationship is important for predicting policy spillovers. If parents spend more on their high-ability children, policies that increase ability will crowd-in household investments. The results here suggest that parents reinforce at the extensive margin, but that the results differ by parental education at the intensive margin of spending.

Lastly, this paper focused on identifying the causal chain between parents' beliefs and investments, not on designing a cost-effective information policy. There are still many open policy design issues, such as whether information delivery through schools can be improved, or how sustained the information delivery must be to obtain effects on test scores and longer-run outcomes, which would likely require a larger intervention than that evaluated here. These questions are left for future research.

References

Adhvaryu, A. and A. Nyshadham (2014). Endowments at birth and parents' investments in children. *The Economic Journal*.

Aizer, A. and F. Cunha (2012). The production of child human capital: Endowments, investments and fertility. *Unpublished paper, Brown University*.

Alexander, K. and D. Entwisle (2006). Baltimore Beginning School Study, 1982-2002. The Harvard-MIT Data Centers. Henry A. Murray Research Archive. Log# 01293.

Almond, D. and J. Currie (2011). Human capital development before age five. *Handbook of Labor Economics 4*, 1315–1486.

Andrabi, T., J. Das, and A. I. Khwaja (2012). What did you do all day? Maternal education and child outcomes. *Journal of Human Resources 47*(4), 873–912.

Andrabi, T., J. Das, and A. I. Khwaja (2016). Report cards: The impact of providing school and child test scores on educational markets. *Unpublished working paper*.

Arcidiacono, P. et al. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics 166*(1), 3–16.

Baird, S. and B. Özler (2012). Examining the reliability of self-reported data on school participation. *Journal of Development Economics 98*(1), 89–93.

Banerjee, A. V., R. Banerji, E. Duflo, R. Glennerster, and S. Khemani (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India. *American Economic Journal: Economic Policy*, 1–30.

Banerji, R., J. Berry, and M. Shotland (2013). The impact of mother literacy and participation programs on child learning: Evidence from a randomized evaluation in India. *Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL)*.

Behrman, J. R., M. R. Rosenzweig, and P. Taubman (1994). Endowments and the allocation of schooling in the family and in the marriage market: The twins experiment. *Journal of Political Economy*, 1131–1174.

Bergman, P. (2016). Parent-child information frictions and human capital investment: Evidence from a field experiment. *Unpublished working paper*.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012). The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *The Quarterly Journal of Economics 127*(3), 1205–1242.

Bharadwaj, P., J. Eberhard, and C. Neilson (2013). Health at birth, parental investments and academic outcomes. *UCSD Working Paper Series*, 1862–1891.

Black, S. E. and P. J. Devereux (2011). Recent developments in intergenerational mobility. *Handbook of Labor Economics 4*, 1487–1541.

Bobba, M. and V. Frisancho (2016). Learning about oneself: The effects of performance feedback on school choice. Technical report, TSE Working Paper.

Bourguignon, F., F. H. Ferreira, and P. G. Leite (2003). Conditional cash transfers, schooling, and child labor: micro-simulating brazil's bolsa escola program. *The World Bank Economic Review 17*(2), 229–254.

Chevalier, A., S. Gibbons, A. Thorpe, M. Snell, and S. Hoskins (2009). Students' academic self-perception. *Economics of Education Review 28*(6), 716–727.

Datar, A., M. R. Kilburn, and D. S. Loughran (2010). Endowments and parental investments in infancy and early childhood. *Demography 47*(1), 145–162.

Dinkelman, T. and C. Martínez A (2014). Investing in schooling in Chile: The role of information about financial aid for higher education. *Review of Economics and Statistics 96*(2), 244–257.

Duflo, E., P. Dupas, and M. Kremer (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review 101*(5), 1739–74.

Dupas, P. (2011). Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics*, 1–34.

EPDC (2007). Educational inequality within countries: Who are the out of school children? Policy report, Education Policy Data Center.

EPDC (2009). Wealth still matters: Wealth differentials in primary school attendance from 1990-2006 in developing countries. Working paper, Education Policy and Data Center.

Griliches, Z. (1979). Sibling model and data in econometrics: Beginnings of a survey. *Journal of Political Economy 87*, S37–S64.

Hastings, J. S., C. A. Neilson, and S. D. Zimmerman (2013). Are some degrees worth more than others? Evidence from college admission cutoffs in Chile. Technical report, NBER.

Hertz, T. et al. (2007). The inheritance of educational inequality: International comparisons and fifty-year trends. *The BE Journal of Economic Analysis & Policy 7*(2).

Hoxby, C. and S. Turner (2013). Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper*.

Hunt, F. (2008). Dropping out from school: A cross country review of the literature. Technical report, CREATE Pathways to Access, No. 16. ERIC.

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

Leight, J. (2014). Sibling rivalry: Endowment and intrahousehold allocation in Gansu Province, China. *Unpublished manuscript*.

Levhari, D. and Y. Weiss (1974). The effect of risk on the investment in human capital. *The American Economic Review*, 950–963.

Liddell, C. and G. Rae (2001). Predicting early grade retention: A longitudinal investigation of primary school progress in a sample of rural South African children. *British Journal of Educational Psychology 71*(3), 413–428.

Madajewicz, M. et al. (2007). Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh. *Journal of Development Economics 84*(2), 731–754.

Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from Madagascar. *Unpublished manuscript*.

Olson, L., H. White, and H. Shefrin (1979). Optimal investment in schooling when incomes are risky. *The Journal of Political Economy*, 522–539.

Paviot, L., N. Heinsohn, and J. Korkman (2008). Extra tuition in southern and eastern africa: Coverage, growth, and linkages with pupil achievement. *International Journal of Educational Development 28*(2), 149–160.

Pekkala Kerr, S., T. Pekkarinen, M. Sarvimäki, and R. Uusitalo (2015). Post-secondary education and information on labor market prospects: A randomized field experiment. *IZA Discussion Paper No. 9372*.

Pitt, M. M., M. R. Rosenzweig, and M. N. Hassan (1990). Productivity, health, and inequality in the intrahousehold distribution of food in low-income countries. *The American Economic Review*, 1139–1156.

Psacharopoulos, G. and H. A. Patrinos (2004). Returns to investment in education: A further update. *Education Economics 12*(2), 111–134.

Rosenzweig, M. R. and K. I. Wolpin (1994). Are there increasing returns to the intergenerational production of human capital? Maternal schooling and child intellectual achievement. *Journal of Human Resources*, 670–693.

Rosenzweig, M. R. and J. Zhang (2009). Do population control policies induce more human capital investment? Twins, birth weight and China's one-child policy. *The Review of Economic Studies 76*(3), 1149–1174.

Sabates, R., K. Akyeampomg, J. Westbrook, and F. Hunt (2010). School drop out: Patterns, causes, changes and policies. UNESCO education background paper.

Stinebrickner, R. and T. Stinebrickner (2014). A major in science? Initial beliefs and final outcomes for college major and dropout. *The Review of Economic Studies 81*(1), 426–472.

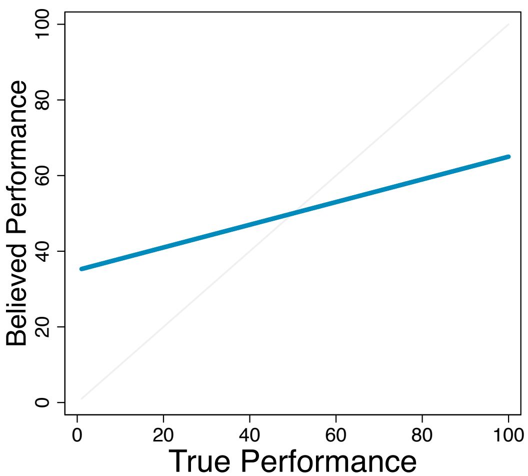
Stinebrickner, T. and R. Stinebrickner (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics 30*(4), 707–748.

Wiswall, M. and B. Zafar (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies 82*(2), 791–824.

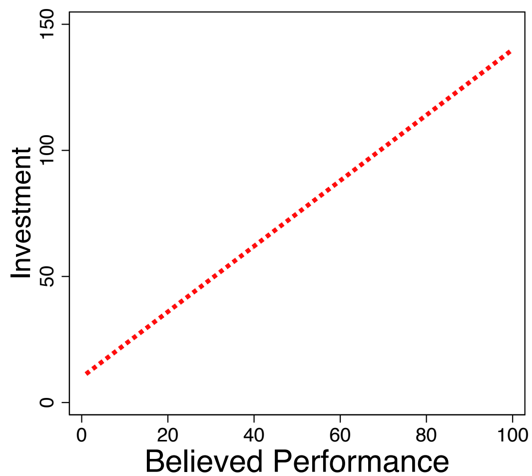
World Bank (2010). The education system in Malawi. Policy report, World Bank.

Figure 1: Conceptual framework: Inaccurate beliefs about performance can cause the slope of investments as a function of academic performance to differ from the slope as a function of beliefs

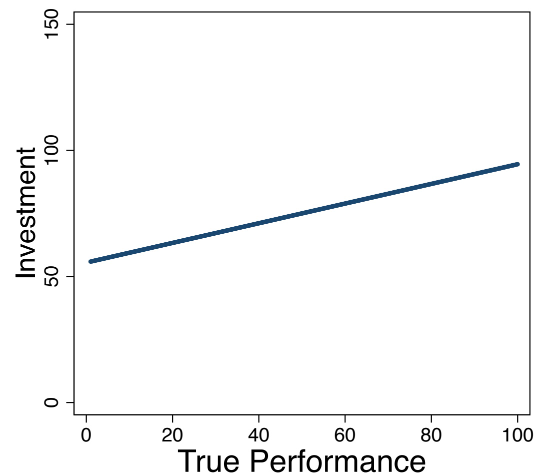
(a) Beliefs may be inaccurate, for example attenuated on true performance (slope < 1).



(b) Parents choose their investments based on their (inaccurate) beliefs.

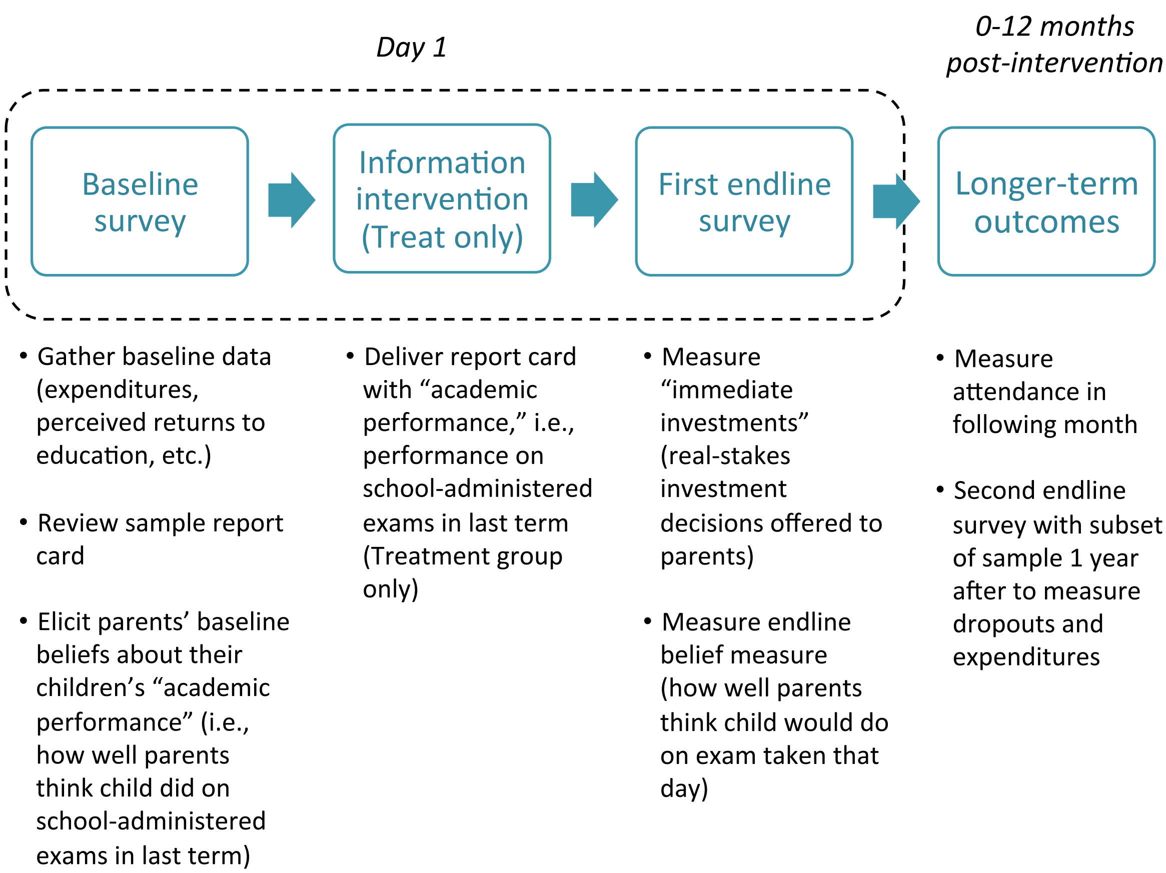


(c) The slope of investments on *true* performance may thus be attenuated relative to the slope on beliefs.



Notes: Graphs are illustrative. The conceptual framework illustrates a way to test whether parents' inaccurate beliefs affect their investments. A common type of belief inaccuracy is that beliefs will be "attenuated" on true performance, i.e., have a slope less than 1 on true performance [subfigure (a)]. Parents base their investments on their potentially inaccurate beliefs, and so plotting investments on beliefs shows us parents' "preferred" slope, i.e., the slope they would opt to choose if they knew their children's true performance [subfigure (b)]. However, because beliefs are inaccurate – and in particular, attenuated – the slope of investments as a function of children's *true* academic performance is flatter than the slope on beliefs [subfigure (c)]. The interpretation of the difference in slopes is that investments are not as well tailored to academic performance as parents would like.

Figure 2: Overview of data collection

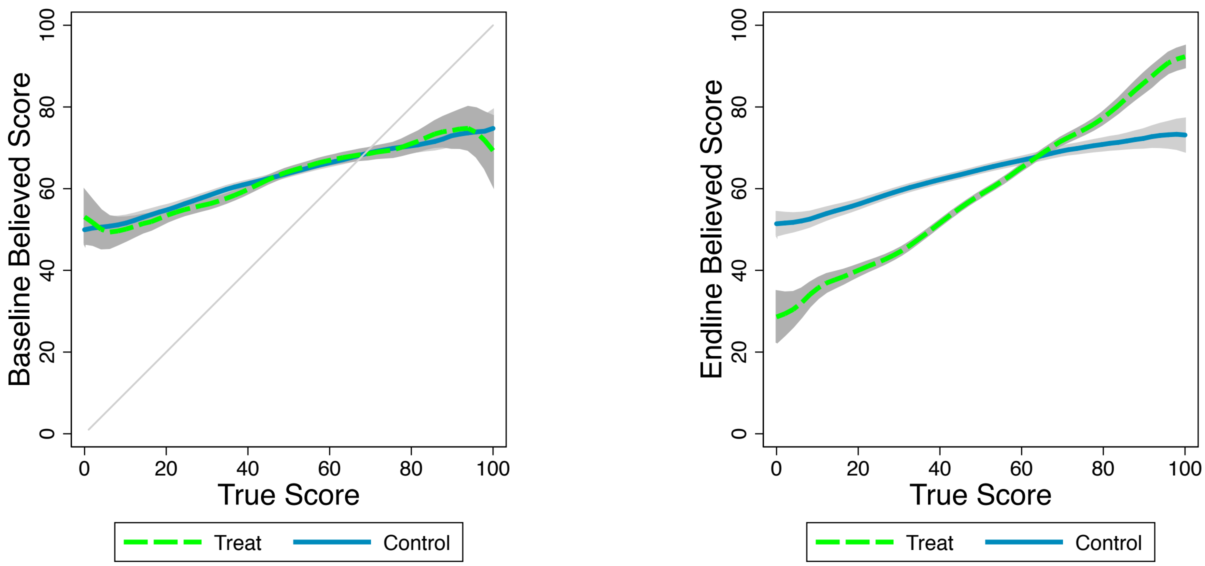


Notes: For any given household, all "Day 1" activities conducted on the same day as the baseline survey; across the sample, the baseline survey was rolled out over the course of two months.

Figure 3: Beliefs results

(a) Baseline beliefs are attenuated

(b) Information decreases the attenuation

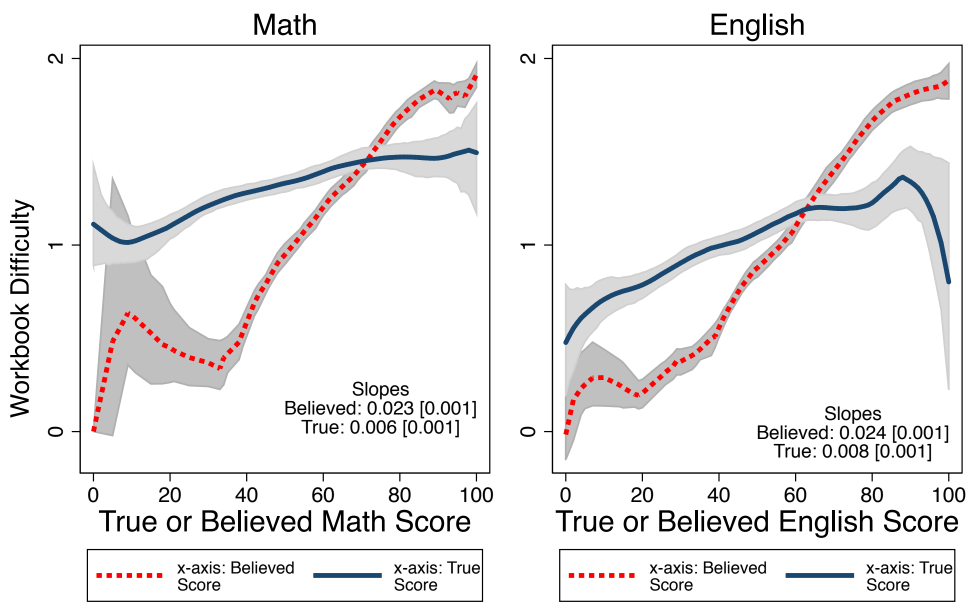


Notes: Data sources are survey data and administrative baseline test score data. Lines are locally linear regression lines with beliefs as the dependent variable and true baseline academic performance as the x-axis. Panel (a) shows *baseline* beliefs as the dependent variable and shows that beliefs are attenuated (i.e., that the slope is less than 1 and so they do not move 1-to-1 with true scores), and that this is balanced across the treatment and control groups. Panel (b) shows a belief measure measured during the first endline survey. This shows that information decreases the attenuation.

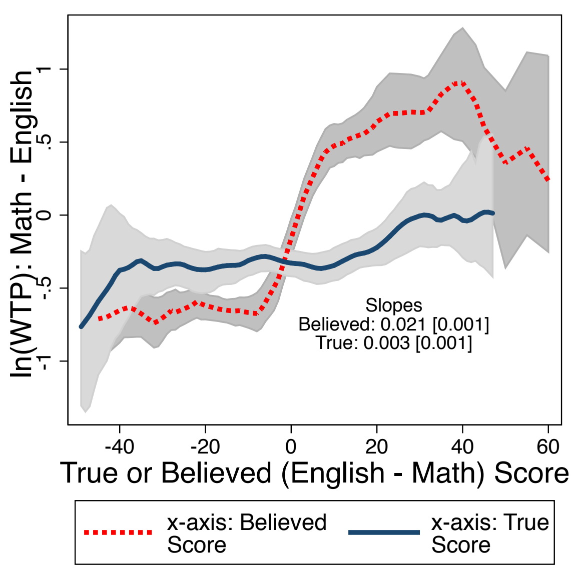
Figure 4: In the control group, the slope of investments on true academic performance is attenuated relative to the slope on believed performance

(Control group only)

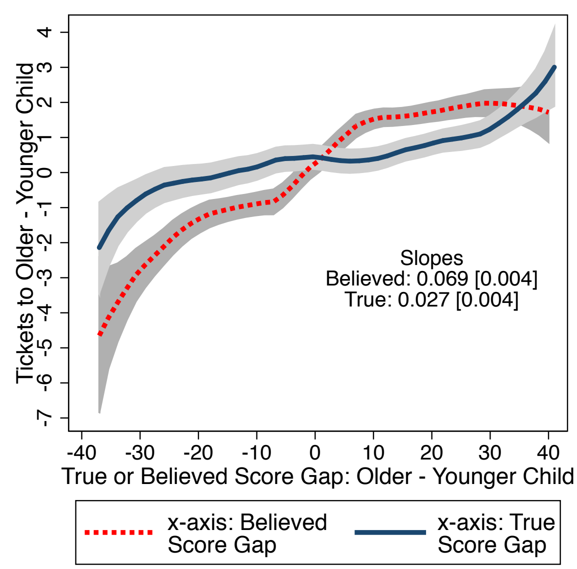
(a) **Difficulty level chosen for free workbooks**



(b) **WTP for remedial textbooks**



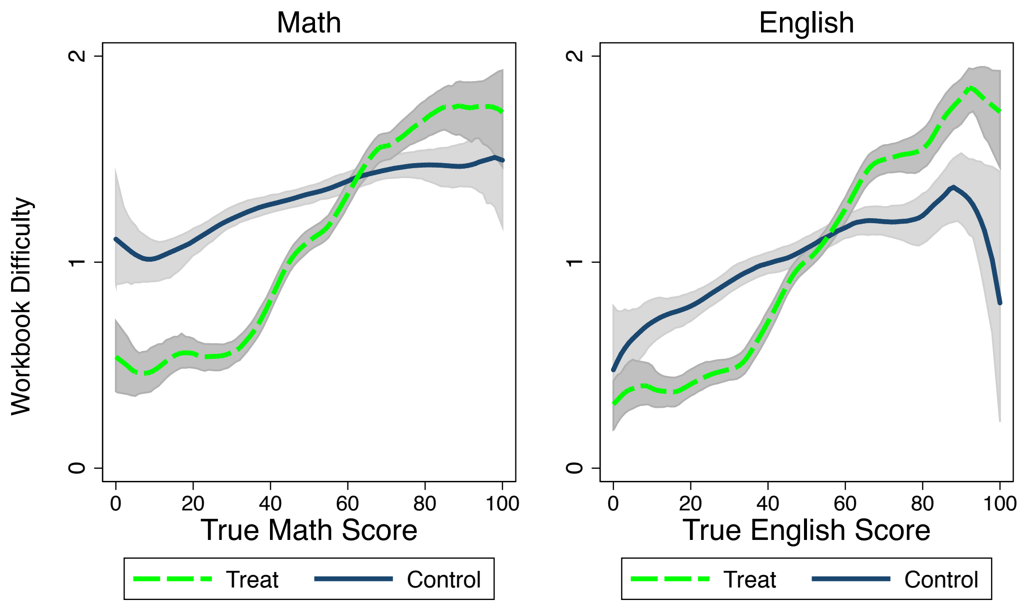
(c) **Secondary school lottery**



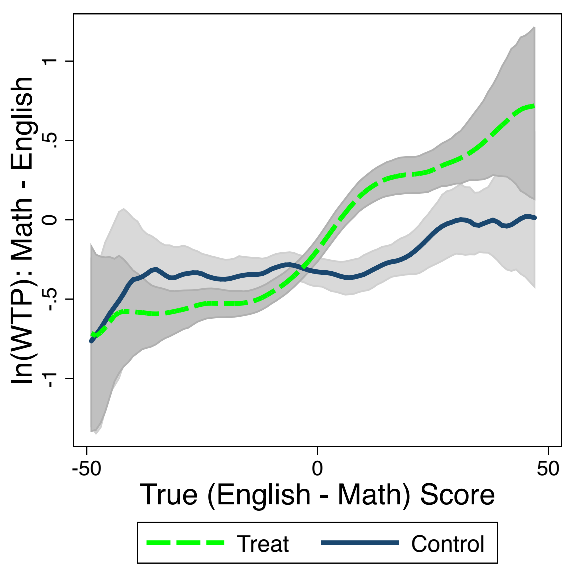
Notes: Control group data only. Data sources are survey data and administrative baseline test score data. Lines are locally linear regression lines with investments as the dependent variable and either true (solid line) or believed (dashed line) baseline academic performance as the x-axis. For the workbook graphs (panel (a)), the dependent variable is the parent's choice of difficulty for a free workbook, where 0 corresponds to the beginner workbook, 1 corresponds to the average, and 2 to the advanced. For textbook WTP (panel (b)), the dependent variable is the difference in the parent's log WTP for a remedial math textbook relative to a remedial English textbook. Because the textbooks are remedial, the prediction is that this should increase in the child's English relative to math performance. For the secondary school lottery, the dependent variable is the number of secondary school lottery tickets given to the older relative to younger child in the household. The grey areas are 95% confidence intervals.

Figure 5: The information treatment increases the slope of investments on true academic performance

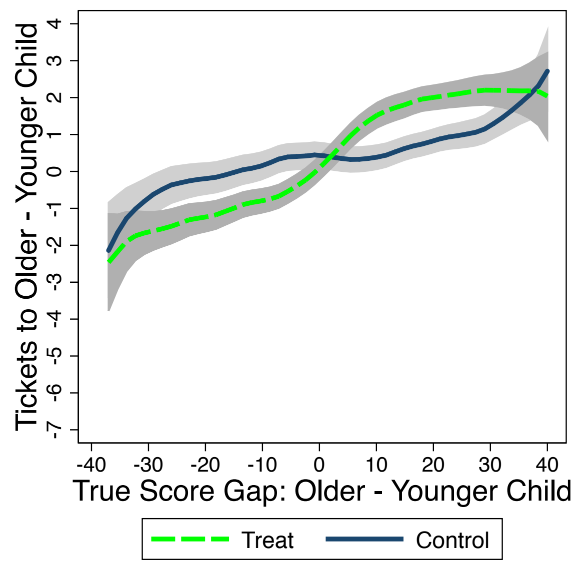
(a) **Difficulty level chosen for free workbooks**



(b) **WTP for remedial textbooks**



(c) **Secondary school lottery**



Notes: Data sources are survey data and administrative baseline test score data. Lines are locally linear regression lines with investments as the dependent variable and either true (solid line) or believed (dashed line) baseline academic performance as the x-axis. For the workbook graphs (panel (a)), the dependent variable is the parent's choice of difficulty for a free workbook, where 0 corresponds to the beginner workbook, 1 corresponds to the average, and 2 to the advanced.For textbook WTP (panel (b)), the dependent variable is the difference in the parent's log WTP for a remedial math textbook relative to a remedial English textbook. Because the textbooks are remedial, the prediction is that this should increase in the child's English relative to math performance. For the secondary school lottery, the dependent variable is the number of secondary school lottery tickets given to the older relative to younger child in the household. The grey areas are 95% confidence intervals.

Table 1: Baseline summary statistics

	Full sample		Control	Treat	Treat − Control		
	Mean	SD	Mean	Mean	Mean	Std. error	p-val T=C
A. Respondent Background							
Female	0.77	0.42	0.77	0.76	-0.01	0.02	0.37
Primary education decision maker	0.92	0.27	0.91	0.92	0.01	0.01	0.31
Age	40.8	11.0	40.6	41.0	0.32	0.44	0.47
Education (years)	4.44	3.57	4.42	4.45	0.04	0.13	0.78
Respondent has secondary education +	0.11	0.31	0.11	0.11	0.01	0.01	0.62
Parent can read or write Chichewa	0.67	0.47	0.67	0.68	0.01	0.02	0.67
Respondent is farmer	0.46	0.5	0.47	0.46	-0.01	0.02	0.7
Respondent's weekly income	2,126	4,744	2,051	2,203	197	194	0.31
B. Household Background							
Family size (Number of children[a])	5.13	1.74	5.16	5.1	-0.05	0.07	0.47
One-parent household	0.19	0.39	0.19	0.2	0.01	0.02	0.47
Parents' average education (years)	4.66	3.25	4.68	4.64	-0.04	0.12	0.74
Any parent has secondary education +	0.18	0.38	0.17	0.19	0.02	0.01	0.24
C. Student Information							
Child's grade level	3.72	1.37	3.72	3.72	0	0.04	0.94
Child's age	11.6	2.68	11.7	11.6	-0.1	0.08	0.21
Child is female	0.51	0.5	0.52	0.5	-0.02	0.01	0.25
Baseline attendance	0.91	0.13	0.92	0.91	0	0	0.72
Annual per-child education expenditures	1,742	2,791	1,712	1,772	58.0	83.0	0.48
Fees paid to schools	381	1,128	384	378	-6.84	23.9	0.78
Uniform expense	576	1,019	548	603	49.9	36.1	0.17
School supplies, books, tutoring, etc.[b]	785	1,819	780	790	14.3	62.3	0.82
Any supplementary expenditures on child	0.9	0.3	0.9	0.89	-0.01	0.01	0.49
D. Academic Performance (Average Achievement Scores)							
Overall score	46.8	17.5	47.1	46.4	-0.74	0.46	0.11
Math score	44.9	20.2	45.4	44.4	-1.08	0.54	0.04
English score	44.2	20.1	44.5	43.9	-0.56	0.53	0.29
Chichewa score	51.3	22.6	51.5	51.0	-0.57	0.59	0.34
(Math − English) Score	0.71	19.5	0.93	0.5	-0.53	0.51	0.3
E. Respondent's Beliefs about Child's Academic Performance							
Believed Overall Score	62.4	16.5	62.7	62.0	-0.78	0.48	0.11
Believed Math Score	64.7	19.0	65.2	64.3	-0.94	0.55	0.09
Believed English Score	55.3	20.9	55.6	54.9	-0.71	0.62	0.25
Believed Chichewa Score	66.8	19.4	66.8	66.7	-0.1	0.6	0.87
Beliefs about (Math − English) Score	9.48	21.5	9.59	9.37	-0.23	0.63	0.71
SD of Individual Beliefs about Score	7.69	10.1	8.08	7.28	-0.8	0.38	0.03
F. Gaps Between Believed and True Academic Performance							
Abs Val [Believed − True Overall Score]	20.4	14.5	20.4	20.3	-0.12	0.43	0.77
Abs Val [Believed − True Math Score]	25.8	18.0	25.8	25.7	-0.1	0.52	0.85
Abs Val [Believed − True English Score]	21.4	16.4	21.6	21.1	-0.57	0.48	0.23
Abs Val [Believed − True Chichewa Score]	23.8	17.5	23.7	23.9	0.18	0.51	0.73
Abs Val [Believed − True (Math-English) Score]	22.1	17.4	22.3	21.9	-0.44	0.51	0.39
Abs Val [Believed − True Overall Score (Child1-2)]	18.7	15.1	18.9	18.5	-0.35	0.59	0.55
Believed - True Overall Score	15.6	19.5	15.6	15.6	-0.07	0.58	0.9
Believed Score Higher than True Score	0.79	0.41	0.79	0.79	0.01	0.01	0.65
G. Beliefs about Complementarity							
Believes educ. and achievement complementary[c]	0.91	0.29	0.9	0.91	0	0.01	0.68
Sample Sizes							
Sample Size–HHs	2,634		1,327	1,307			
Sample Size–Kids	5,268		2,654	2,614			

Notes: Data source is baseline survey. Standard errors for the test of equality across treatment and control clustered at the household level.
a. Counted as a child if either of the primary caregivers for the sampled children is a parent of the child.
b. Includes exercise books and pencils, textbooks and supplementary reading books, backpacks, and tutoring expenses.
c. Respondent said that they thought the earnings of a more able child would increase "more" or "much more" than the earnings of a less able child from getting a secondary education.

Table 2: Heterogeneity in the attenuation of beliefs by parent education

Dep. Var.	Parent beliefs about child's score in:					
	Overall	Math	English	Chichewa	Math-Engl	Child 2 - 1
Score × Parents' yrs educ.	0.014***	0.018***	0.014***	0.0098***	0.013***	0.017***
	[0.0038]	[0.0038]	[0.0043]	[0.0036]	[0.0047]	[0.0049]
Score	0.25***	0.13***	0.22***	0.20***	0.091***	0.32***
	[0.023]	[0.023]	[0.026]	[0.021]	[0.029]	[0.028]
Parents' years education	-0.53***	-0.98***	-0.065	-0.32	-0.78***	0.044
	[0.20]	[0.20]	[0.21]	[0.23]	[0.094]	[0.12]
Observations	5,220	5,222	5,222	5,222	5,222	5,218

Notes: Data sources are baseline survey and baseline test score data. Each observation is a child. Standard errors are clustered at the household level. "Parents' years education" measures average years of education among the child's parents. Table displays regressions of parents' beliefs on their child's true score, the parents' education, and the interaction. The prediction is that true scores will be more highly correlated with the beliefs of more-educated parents, which means that the coefficient on "Score × Parents' yrs educ." will be positive. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Immediate outcomes: Information treatment effects on the slope of investments on academic performance

Dep. Var.	Math workbook difficulty level	English workbook difficulty level	ln(math textbook WTP) - ln(English textbook WTP)	Secondary school lottery tickets
	(1)	(2)	(3)	(4)
Treat × Score	1.34***	1.26***	0.013***	0.035***
	[0.093]	[0.096]	[0.0022]	[0.0053]
Score	0.65***	0.76***	0.0030*	0.017***
	[0.065]	[0.073]	[0.0016]	[0.0052]
Treat	-91.1***	-68.5***	0.15***	
	[4.91]	[4.83]	[0.041]	
Observations	5,239	5,239	5,183	5,258
R-squared	0.217	0.204	0.033	0.119
Score Used	Math	English	English – Math	Overall
Household FE	No	No	No	Yes

Notes: Data sources are baseline survey, baseline test score data, both endline surveys, and endline administrative data. Each observation is a child. Standard errors are clustered at the household level. Workbook difficulty choices are coded as 0 for beginner, 100 for average, and 200 for advanced.

The regressions test for whether information changes the slope of investments on children's academic performance (where academic performance is measured as children's average scores on school-administered achievement exams). One way to interpret the results is to compare the baseline slope in the control group (coefficient on Score) with the increase in the slope in the treatment group (coefficient on Treat × Score) to see how much the slope has increased as a result of information. Take for example column (1). The ratio of the coefficient on Treat × Score (1.34) to the coefficient on Score (0.65) shows us that the slope has increased by roughly 200% (1.34/0.65), so that the treatment slope is roughly 3 times as large as the control slope. The rough interpretation of the slope in the control group for that column is that, if the child's math score increases by one point, the chance that her parent chooses the next higher difficulty level of the free book increases by .65%.

Regressions control for school FE, parents' education, the between-child score gap, child baseline performance, and grade FE; column (4) also has a household FE.

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Immediate outcomes: Heterogeneity in the treatment effect on the slope, by parent education

Dep. Var.	Endline beliefs	Math workbook difficulty level	English workbook difficulty level	ln(math textbook WTP) - ln(English textbook WTP)	Secondary school lottery tickets
	(1)	(2)	(3)	(4)	(5)
Treat × Score × Parent yrs of educ.	-0.025***	-0.12***	-0.066**	-0.00073	-0.0011
	[0.0071]	[0.027]	[0.029]	[0.00059]	[0.0023]
Treat × Score	0.53***	1.92***	1.57***	0.017***	0.040***
	[0.044]	[0.16]	[0.17]	[0.0037]	[0.014]
Score × Parent yrs of educ.	0.022***	0.078***	0.032	0.00058	0.00048
	[0.0051]	[0.020]	[0.022]	[0.00038]	[0.0016]
Score	0.21***	0.29**	0.61***	-0.000055	0.027***
	[0.031]	[0.11]	[0.13]	[0.0026]	[0.0092]
Treat × Parent yrs of educ.	1.22***	6.48***	2.29	-0.032***	
	[0.39]	[1.46]	[1.53]	[0.012]	
Treat	-31.9***	-121.5***	-79.1***	0.30***	
	[2.31]	[8.58]	[8.59]	[0.071]	
Parent yrs of educ.	-0.79***	-3.86***	-0.29	0.024***	
	[0.27]	[1.08]	[1.18]	[0.0084]	
Observations	5,208	5,203	5,203	5,183	5,222
R-squared	0.342	0.220	0.207	0.035	0.117
p-val: Treat × Score × Yrs.Educ.=0	0.000	5.0e-06	0.022	0.222	0.652
Score Used	Overall	Math	English	English − Math	Overall

Notes: Data sources are baseline survey, baseline test score data, both endline surveys, and endline administrative data. Standard errors clustered at household level. Table shows the heterogeneity by parent education in the information treatment effect on the gradient of the investment function. Each observation is a child. Parents' years of education (Parent yrs of educ.) is the average across parents in the household. Regressions control for school FE, the main effect of parental years of education, the between-child score gap, and child achievement. Column (5) also controls for a household FE.
** p<0.01, ** p<0.05, * p<0.1

Table 5: Treatment effects on the slope for longer-term outcomes: Full-sample estimates, and heterogeneity by parent education

Dep. Var.	Retention		ln(Total educ. expenditures)		Attendance rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Continuous versions						
Treat × Score	0.11***	0.15**	-0.0019	-0.0089**	0.013	-0.15*
	[0.038]	[0.069]	[0.0022]	[0.0039]	[0.048]	[0.080]
Treat × Score × Parent yrs of educ.		-0.0091		0.0014**		0.033**
		[0.0092]		[0.00069]		[0.014]
Treat	-5.51***	-6.83*	0.093	0.37*	-0.80	7.97*
	[2.11]	[3.87]	[0.11]	[0.20]	[2.61]	[4.43]
Treat × Parent yrs of educ.		0.33		-0.056		-1.83**
		[0.53]		[0.037]		[0.82]
Panel B. Binary versions						
Treat × Above-median score	3.71**	4.57*	-0.040	-0.18*	-0.17	-3.13
	[1.45]	[2.37]	[0.074]	[0.11]	[1.53]	[2.15]
Treat × Above-med.score × Above-med.par.educ		-2.31		0.28*		5.84*
		[2.78]		[0.14]		[3.02]
Treat	-2.20*	-2.06	0.020	0.082	-0.15	1.94
	[1.15]	[1.93]	[0.061]	[0.090]	[1.24]	[1.73]
Treat × Above-median parent educ.		-0.13		-0.14		-3.98
		[2.19]		[0.13]		[2.47]
Observations	1,786	1,768	1,709	1,692	1,827	1,812
Control group mean	97.9		7.4		91.1	
Score Used	Overall	Overall	Overall	Overall	Overall	Overall

Notes: Data sources are baseline survey, baseline test score data, endline survey and endline data collected from schools. Each observation is a child. Standard errors clustered at the household level. In the interest of brevity, not all regression coefficients are shown, but the regressions showing heterogeneity by parent education (shown in the even-numbered columns) are fully-interacted and so control for all interactions and main effects of all variables shown (e.g., in Panel A, the regressions control for Score, Score × Parent yrs of educ., and Parent yrs of educ.). All regressions also control for grade FE, school FE, the between-sibling achievement gap, and the baseline value of the dependent variable, if available (not available for retention). Retention is defined as being enrolled in school 1 year after the intervention. Both retention and attendance scaled to be out of 100 (so retention, for example, is equal to 100 if the child is still enrolled and 0 otherwise). In Panel B, Above-med.par.educ. means the household was above-median for parent years of education (average years of education across the parents). Above-median score means the child had an above-median baseline overall score. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix

Table of Contents

Note: Appendices E through I (including all "Online Appendix Tables") can be found in the "Online Appendix" document found at the below link:
http://faculty.chicagobooth.edu/rebecca.dizon-ross/research/papers/perceptionsOnlineApp.pdf

Appendix Table A.1: Heterogeneity by parent education in belief accuracy, uncertainty, overconfidence, and children's academic performance

	Belief inaccuracy		**Uncertainty**		**Overconfidence**		**Performance**	
Dependent Variable:	Abs.val.[believed - true score]		Std. dev. of beliefs		Believed - true score		Score	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parents' years education	-0.202***	-0.200***	-0.614***	-0.611***	-0.079	-0.078	0.348***	0.354***
	[0.064]	[0.066]	[0.055]	[0.056]	[0.088]	[0.090]	[0.076]	[0.077]
Child and parent controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	5,220	5,019	5,171	4,974	5,220	5,019	5,230	5,029
Dep. Var. Mean	20.385		7.658		15.626		46.718	

Notes: Data sources are baseline survey and baseline test score data. Each observation is a child. Standard errors are clustered at the household level. Child and parent controls include a control for child gender, grade FE, parent gender, and whether the parent is the primary education decisionmaker. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A.2: Uncertainty tests: Effect of information on the slope of the *preferred* investment function

	Immediate outcomes				Longer-term outcomes		
	Math workbook difficulty level	English workbook difficulty level	ln(English textbook WTP) - ln(math textbook WTP)	Lottery tickets	Retention	ln(Total educ. expenditures)	Attendance rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Treatment effect on the slope for those with beliefs within 10 pts of truth							
Treat × Score	0.36*	0.38**	0.0014	0.018*	0.15**	-0.00072	-0.12
	[0.22]	[0.16]	[0.0049]	[0.010]	[0.069]	[0.0049]	[0.095]
Score Measure	Math	English	Math − English	Score	Score	Score	Score
Treat × Score (full sample)	1.348	1.248	0.013	0.036	0.109	-0.002	0.013
p-val: Treat × Score equal in full sample	0.000	0.000	0.002	0.051	0.970	0.852	0.354
Observations	1,106	1,450	1,416	1,786	534	508	489
Panel B. Heterogeneity in treatment effects by score vs. beliefs (equal and opposite indicates no change in slope)							
Treat × Score	1.65***	1.66***	0.015***	0.049***	0.10**	-0.00059	0.077
	[0.090]	[0.087]	[0.0021]	[0.0056]	[0.047]	[0.0025]	[0.053]
Treat × Beliefs	-1.52***	-1.55***	-0.011***	-0.035***	0.017	-0.0035	-0.17***
	[0.10]	[0.086]	[0.0021]	[0.0063]	[0.061]	[0.0028]	[0.057]
p-val: (Treat × Score) +(Treat × Beliefs)= 0	0.240	0.230	0.150	0.028	0.016	0.116	0.121
p-val: Treat × Score = 0	0.000	0.000	0.000	0.000	0.026	0.815	0.144
Observations	5,233	5,233	5,213	5,250	1,780	1,703	1,822

Notes: Data sources are baseline survey, baseline test score data, the endline surveys, and endline administrative data. Panel A takes parents whose baseline beliefs were within 10 points of their children's true academic performance as the sample, and examines the treatment effect on the slope of investments on children's true score. Panel B uses the entire experimental sample and looks at the heterogeneity in the treatment effect based on both the true score and parents' beliefs, where the prediction for no change in slope (i.e., no uncertainty effects) is that the coefficients are equal and opposite. Standard errors clustered at the household level. Regressions control for school FE, parents' education, the between-child score gap, child baseline performance, grade fixed effects, the baseline value of the dependent variable (baseline value not available for retention or immediate outcomes), treatment, and the main effects of any variable interacted with treatment. Workbook difficulty choices are coded as 0 for beginner, 100 for average, and 200 for advanced. Retention defined as being enrolled in school 1 year after the intervention; retention and attendance scaled to be out of 100 (so, for example, retention is equal to 100 if the child is still enrolled and 0 otherwise).
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Table A.3: Average treatment effects

	Immediate outcomes				Longer-term outcomes		
	Endline beliefs	Math workbook difficulty level	English workbook difficulty level	ln(math textbook WTP) - ln(English textbook WTP)	Retention	ln(Total educ. expenditures)	Attendance rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Average treatment effects							
Treat	-7.13***	-31.3***	-12.6***	0.14***	-0.40	0.0013	-0.21
	[0.48]	[2.08]	[2.15]	[0.041]	[0.71]	[0.049]	[0.78]
Observations	5,244	5,239	5,239	5,219	1,786	1,709	1,827
Panel B. Average treatment effect: Heterogeneity by parent education							
Treat × Parent Yrs. of Educ.	0.11	0.027**	0.93	-0.13	-0.070	0.012	-0.26
	[0.14]	[0.012]	[0.61]	[0.65]	[0.16]	[0.018]	[0.26]
Treat	-7.63***	-0.27***	-35.3***	-12.3***	-0.21	-0.049	1.12
	[0.89]	[0.071]	[3.64]	[3.82]	[1.27]	[0.097]	[1.44]
Observations	5,208	5,183	5,203	5,203	1,768	1,692	1,812
Panel C. Uncertainty Level effects: Beliefs within 10 pts of truth							
Treat	0.42	-6.56*	2.33	0.071	-0.42	0.070	1.02
	[0.66]	[3.64]	[3.20]	[0.072]	[0.88]	[0.086]	[1.32]
Observations	1,571	1,299	1,657	1,589	579	550	541

Notes: Data sources are baseline survey, baseline test score data, immediate endline survey, endline survey, and endline administrative data. Each observation is a child. Standard errors are clustered at the household level. Regressions control for school FE, parents' education, the between-child score gap, child baseline performance, grade fixed effects, and the baseline value of the dependent variable (baseline value not available for retention or immediate outcomes). Workbook difficulty choices are coded as 0 for beginner, 100 for average, 200 for advanced. Retention defined as being enrolled in school 1 year after the intervention; retention and attendance scaled to be out of 100 (so, for example, retention is equal to 100 if the child is still enrolled and 0 otherwise). "Parent Yrs. of Educ." measures average years of education among the child's parents. Panel C uses the relevant measure of beliefs (e.g., overall for beliefs, math - English for textbooks; see Table 4 for details.) *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A.4: Asymmetric responses to positive vs. negative information shocks

Dep. Var.	Endline beliefs (1)	Math workbook difficulty level (2)	English workbook difficulty level (3)	Retention (4)
Treat × Score × Pos. Shock	0.423***	1.223***	1.618***	0.087
	[0.040]	[0.180]	[0.149]	[0.121]
Treat × Score	0.208***	0.639***	0.253**	0.127***
	[0.029]	[0.107]	[0.111]	[0.045]
Observations	5,244	5,239	5,239	1,786
R-squared	0.407	0.264	0.279	0.054
Score Used	Overall	Math	English	Overall

Notes: Data sources are baseline survey, baseline test score data, and the endline surveys. The table shows the results of estimating equation 1 (i.e., the equation estimated in Table 3, which shows how information affected the slope of the investment function), fully interacted with an indicator for whether a household was a "positive shock" household, where "positive shock" means that the child's true performance was higher than the parent's baseline beliefs. In the interest of brevity, not all coefficients are shown. Each observation is a child. Standard errors are clustered at the household level. Regressions control for school FE, parents' education, the between-child score gap, child baseline performance, grade fixed effects, and all of the main effects and interaction terms (i.e., Treat, Score, Pos. Shock, and all of their double and triple interactions). *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Table A.5: Robustness of information treatment effects: Immediate outcomes

	Treatment effect on slope *(Columns vary the control variables)*				
	(1)	(2)	(3)	(4)	(5)
Panel A. Dependent Var: Endline Beliefs					
Treat × Score	0.41***	0.36***	0.36***	0.37***	0.37***
	[0.025]	[0.046]	[0.046]	[0.046]	[0.046]
Treat	-26.3***				
	[1.32]				
Observations	5,244	5,244	5,244	5,244	5,244
R-squared	0.337	0.760	0.760	0.763	0.764
Panel B. Dependent Var: ln(Math Textbook WTP) - ln(English Textbook WTP)					
Treat × (English − Math Score)	0.013***	0.013***	0.013***	0.014***	0.014***
	[0.0022]	[0.0038]	[0.0038]	[0.0039]	[0.0039]
Treat	0.15***				
	[0.041]				
Observations	5,183	5,183	5,183	5,183	5,183
R-squared	0.033	0.601	0.601	0.602	0.602
Panel C. Dependent Var: Math Workbook Choice					
Treat × Math Score	1.34***	1.20***	1.20***	1.13***	1.13***
	[0.093]	[0.17]	[0.17]	[0.17]	[0.17]
Treat	-91.1***				
	[4.91]				
Observations	5,239	5,239	5,239	5,239	5,239
R-squared	0.217	0.695	0.695	0.696	0.696
Panel D. Dependent Var: English Workbook Choice					
Treat × English Score	1.26***	1.27***	1.26***	1.33***	1.33***
	[0.096]	[0.17]	[0.17]	[0.17]	[0.17]
Treat	-68.5***				
	[4.83]				
Observations	5,239	5,239	5,239	5,239	5,239
R-squared	0.204	0.710	0.710	0.714	0.715
Panel E. Dependent Var: Lottery tickets received					
Treat × (Overall Score)		0.035***	0.035***	0.037***	0.036***
		[0.0053]	[0.0053]	[0.0052]	[0.0054]
Observations		5,258	5,258	5,258	5,080
R-squared		0.119	0.122	0.155	0.170
Includes controls for (all panels):					
Household FE	No	Yes	Yes	Yes	Yes
Treat × Female	No	No	Yes	Yes	Yes
Treat × Grade Level	No	No	No	Yes	Yes
Treat × Educ. Expenditures	No	No	No	No	Yes

Notes: Data sources are baseline survey, baseline test score data, and the endline survey data. Each observation is a child. Standard errors are clustered at the household level. Regressions control for school FE, parents' education, the between-child score gap, child baseline performance, grade fixed effects, and the main effect of any variable interacted with Treat. Workbook difficulty choices are coded as 0 for beginner, 100 for average, 200 for advanced. The regressions test for a change in the slope, with the prediction being that information will increase the slope (positive coefficient on Treat × Score).
*** p<0.01, ** p<0.05, * p<0.1

B Sample information intervention report card

Report Card

Name: **NDEMA LONGWE** **Standard:** **2**

	Score	**Grade**	**Position**
Maths:	75/100	3	10/100
English:	33/100	1	71/100
Chichewa:	67/100	3	38/100
Overall:	58/100	2	52/100

Number of Exams Administered in Class: 3

Grades
1 = Needs support
2 = Average
3 = Good
4 = Excellent

Note: "Positions" are a measure of children's relative performance within their classes, equal to 100 minus the percentile. For ease of interpretation, the measure is converted to percentiles for the analysis. See Section I.1 for details.

C Appendix to the conceptual framework

Discussion of Prediction 1: Attenuation in the slope of investments

Assume the preferred investment function is: $s^*(\tilde{A}) = \beta_0 + \beta_1 \tilde{A}$.[42] The slope of the preferred investment function (i.e., the slope of investments on beliefs) is thus β_1, and, using the standard OLS formula, the slope of the actual investment function (i.e., the slope of investments on true performance) is $\frac{cov(\beta_0+\beta_1\tilde{A},A)}{var(A)} = \beta_1 \frac{cov(\tilde{A},A)}{var(A)}$. Thus, whenever $\frac{cov(\tilde{A},A)}{var(A)} \neq 1$, inaccurate beliefs will cause the actual slope to differ from the preferred slope, and whenever $\frac{cov(\tilde{A},A)}{var(A)} < 1$, there is attenuation. Since $\frac{cov(\tilde{A},A)}{var(A)}$ is the slope from regressing believed performance on true performance, this means that the condition for attenuation in the slope of investments on true performance is that beliefs are an attenuated function of true performance, i.e., have a slope less than 1.

There are several ways to express the assumptions that lead to attenuation in the slope of beliefs on true performance. One way is that (i) A and $\tilde{A}$ are positively correlated, and (ii) the variance of $\tilde{A}$ is not "too much larger" than the variance of A, which, more rigorously means that $\frac{SD(\tilde{A})}{SD(A)} < \frac{1}{corr(\tilde{A},A)}$, where *corr* is correlation and SD is the standard deviation. Note that, since correlations are bounded above by 1, a sufficient condition is that the variance of $\tilde{A}$ is smaller than that of A. As shown in footnote 5, one can see this by re-expressing the slope of believed on true performance as $corr(\tilde{A}, A)\frac{SD(\tilde{A})}{SD(A)}$. The level of attenuation is thus driven by the correlation between believed and true performance: the lower the correlation, the more attenuated the slope of believed on true performance, and thus the more attenuated the slope of investments on true performance.

Discussion of Prediction 2: If there is baseline attenuation, information increases the slope of investments

Assume that parents make investments according to the above model plus an error term, u_i, representing all omitted determinants of investments: $s = s^*(\tilde{A}) + u_i = \beta_0 + \beta_1\tilde{A} + u_i$. I first outline the bias in an observational data approach, and then outline how an experiment addresses this bias.

The observational approach would be to compare the slopes estimated from regressing baseline (or control group) s on $\tilde{A}$ with the slope from regressing baseline s on A. The slope from regressing on $\tilde{A}$ will be the true causal slope, β_1, plus an omitted variable bias (OVB) term: $\beta_1 + \frac{cov(\tilde{A},u_i)}{var(\tilde{A})}$. The slope from regressing on A will be the true causal slope, $\beta_1\frac{cov(\tilde{A},A)}{var(A)}$, plus an OVB term: $\beta_1\frac{cov(\tilde{A},A)}{var(A)} + \frac{cov(A,u_i)}{var(A)}$. Thus, the difference in slopes will be $\left(\beta_1 - \beta_1\frac{cov(\tilde{A},A)}{var(A)}\right) + \left(\frac{cov(\tilde{A},u_i)}{var(\tilde{A})} - \frac{cov(A,u_i)}{var(A)}\right)$ and so will only give us an unbiased estimate of the

[42]Note that, for expositional simplicity, I focus on the linear case, but one can interpret this as the best linear predictor function in the case where investments are non-linear in $\tilde{A}$.

true difference in slopes, $\beta_1 - \beta_1 \frac{cov(\tilde{A},A)}{var(A)}$, if the second term (i.e., the difference between the OVB terms $\left(\frac{cov(\tilde{A},u_i)}{var(\tilde{A})} - \frac{cov(A,u_i)}{var(A)}\right)$) is equal to 0.

An experiment can solve this problem. Consider comparing the slopes of the actual investment functions (s regressed on A) for parents who have received information about their children's true academic performance, A, (treatment group) vs. those who have not (control group). Parents in the treatment group will now base investments on true performance A, so their investments will be $s^*(A) + u_i = \beta_0 + \beta_1 A + u_i$.[43] The slope in the treatment group will thus be $\beta_1 + \frac{cov(A,u_i)}{var(A)}$, whereas in the control group it will be the same as above: $\beta_1 \frac{cov(\tilde{A},A)}{var(A)} + \frac{cov(A,u_i)}{var(A)}$. Since, unlike for the observational approach, the omitted variable terms are now identical, comparing the slope between treatment and control groups will allow us to estimate the true difference in slopes $|\beta_1 - \beta_1 \frac{cov(\tilde{A},A)}{var(A)}|$. If investments were attenuated at baseline, that difference will be positive, meaning that information will increase the magnitude of the slope.

C.1 Uncertainty predictions

There are many ways to model uncertainty in beliefs. Here, I show one potential framework which yields the prediction that uncertainty in parents' beliefs about academic performance, $\tilde{A}$, leads to attenuation in the slope of investments on beliefs. The framework captures the intuition described in the main text: that uncertainty in parents' beliefs may make them not want to make their investments depend as strongly on their beliefs. This is a richer model than the one used in Section 2.

Assume there is some true unobserved underlying academic skill. Call this $\mathbf{A}$ and call parents' beliefs about it $\tilde{\mathbf{A}}$. Assume this underlying academic skill is what determines returns and thus what parents truly want to base decisions on. Assume further that academic skill is distinct from academic performance, A (where A is what we measured baseline beliefs on, and what we delivered information about in the intervention). Rather, academic performance A is taken by parents as a signal of $\mathbf{A}$.

In this context, we can model beliefs about academic skill $\tilde{\mathbf{A}}$ as being a convex combination of beliefs about school performance, $\tilde{A}$, and beliefs about all other aspects or signals of academic skills, $\tilde{\mathbf{A}}_{-\tilde{A}}$, given by:

$$\tilde{\mathbf{A}} = \lambda \tilde{A} + (1 - \lambda)\tilde{\mathbf{A}}_{-\tilde{A}}$$

where λ is the weight on the academic performance.

[43] Note that this assumes that parents fully update their beliefs in response to the intervention. If they only partially update their beliefs, then the difference in slope between treatment and control groups would be weighted downwards by the updating parameter (i.e., if updated beliefs were a weighted combination of A and $\tilde{A}$ with γ the weight on A, then the difference in slopes would uncover $\gamma(\beta_1 - \beta_1 \frac{cov(\tilde{A},A)}{var(A)})$).

Since preferred investments would be a function of $\tilde{\mathbf{A}}$, not $\tilde{A}$, we could write the preferred investment function as $\tilde{s}^*(\tilde{\mathbf{A}})$. For expositional simplicity, let's look at the linear case where $\tilde{s}^*(\tilde{\mathbf{A}}) = \tilde{\beta}_0 + \tilde{\beta}_1\tilde{\mathbf{A}}$ (where the $\tilde{\beta}_1$ notation distinguishes this from the preferred investment function in the simpler model from Section 2.) Preferred investments could then be written as:

$$\begin{aligned} s^*(\tilde{\mathbf{A}}) &= \tilde{\beta}_0 + \tilde{\beta}_1\tilde{\mathbf{A}} \\ &= \tilde{\beta}_0 + \tilde{\beta}_1\lambda\tilde{A} + \tilde{\beta}_1(1-\lambda)\tilde{\mathbf{A}}_{-\tilde{A}} \end{aligned}$$

In this context, information about academic performance, A, should increase the certainty of beliefs about academic performance, $\tilde{A}$. This could increase the weight that parents place on beliefs about academic performance when forming their beliefs about underlying academic skill, that is, increase λ. As a result, under most assumptions for the form that $\tilde{\mathbf{A}}_{-\tilde{A}}$ would take,[44] the slope of investments on beliefs about school performance $\tilde{A}$ should increase, since λ has increased.

Note that this is a channel for uncertainty to change the slope of investments on beliefs about academic *performance*, $\tilde{A}$, even if the underlying slope of the true preferred investment function on beliefs about academic *skill*, $\tilde{\mathbf{A}}$, does not change.[45]

D Appendix to the back-of-the-envelope calculation

D.1 Assumptions

When I use secondary school completion as the outcome, the base scenario makes the following assumptions based on World Bank (2010) data. It assumes that the transition rate to secondary school for primary graduates is (i) 50% overall; (ii) 90% higher among high-SES than low-SES households; and (iii) has the same gap between high- and low-performing students as primary school retention rates do. For simplicity, I assume that there are no dropouts during secondary.

When I use earnings as an outcome, the base scenario assumes that the average earnings return to schooling is 10% and that the return is 19% higher among students with above-

[44]Specifically, the regression of investments on $\tilde{A}$ would have slope $\lambda\beta_1 + (1-\lambda)\beta_1\frac{cov(\tilde{A},\tilde{\mathbf{A}}_{-\tilde{A}})}{Var(\tilde{A})} = \lambda\beta_1 + (1-\lambda)\beta_1 corr(\tilde{A},\tilde{\mathbf{A}}_{-\tilde{A}})\frac{sd(\tilde{\mathbf{A}}_{-\tilde{A}})}{sd(\tilde{A})}$. Thus, since $corr(\tilde{A},\tilde{\mathbf{A}}_{-\tilde{A}}) \leq 1$ increasing λ increases the slope as long as the variance of $\tilde{\mathbf{A}}_{-\tilde{A}}$ is not too much larger than the variance of $\tilde{A}$.

[45]For example, with a linear preferred investment function and a quadratic loss function, the slope of the true preferred investment function should not change. It is useful to note that, in this richer model, although providing information about A should unambiguously increase the certainty of $\tilde{A}$, it is ambiguous whether it will decrease or increase the uncertainty of beliefs about $\mathbf{A}$. For example, if the information were very different from parents' prior beliefs, it could increase the uncertainty of $\tilde{\mathbf{A}}$.

median performance than below-median performance (Psacharopoulos and Patrinos, 2004; Aizer and Cunha, 2012).

Appendix Section D.3 below shows robustness to both of these sets of assumptions.

D.2 Discussion of primary schooling result

As discussed in the text, information is projected to close roughly 50% of the gap by SES in projected primary school completion, but part of the effect is due to the PCR declining among high-SES households. This is because information has two impacts on dropouts. The primary effect is reallocation of dropouts from high- to low-performing students, but there is an additional (not statistically significant) effect among high-SES households: a positive point estimate for the effect of information on *average* dropout rates (.01 among high-SES relative to .0005 among low-SES). Since only the first channel (reallocations from high-performing to low-performing students) is statistically significant and we do not have a theoretical reason to expect the second, we may also be interested in the effect if we do not allow the projected PCR to worsen for high-SES households. In that case, the projected SES gap in the treatment group is 0.96-0.79=0.17, and information would close 15% of the gap. Alternatively, one can hold the average dropout rates constant for both SES groups and just allow for the difference in dropouts between high- and low-performing students to change. Information then closes 11% of the SES gap. Thus, although magnitudes vary, all suggest that belief inaccuracy plays a non-trivial role in explaining SES gaps.

D.3 Robustness of calculation to changing the assumptions

	Percent of SES gap in [...] that is closed by information		
	Primary school completion rate (1)	Secondary completion rate (2)	Earnings (3)
A. Base			
Base	.48	.14	.18
B. Change assumptions affecting primary completion, sec. completion, and earnings			
Use actual data (not assumption) on baseline dropout gap between high and low performing students	.51	.15	.2
Decrease baseline dropout gap between low SES and high SES by 50%	.71	.13	.15
Increase baseline dropout gap between low SES and high SES as much as possible (30%)	.43	.15	.19
C. Change assumptions affecting secondary completion and earnings			
Decrease secondary transition rate gap between high and low SES by 50%		.2	.25
Increase secondary transition rate gap between high and low SES by 50%		.11	.15
D. Change assumptions affecting earnings only			
Decrease returns to education by 50%			.16
Increase returns to education by 50%			.19
Decrease complementarity by 50%			.17
Increase complementarity by 50%			.18

Notes: This table shows to what extent information would close the projected gap between low-SES and high-SES households (where low-SES is defined as below-median parent education households, and high-SES is above-median parent education households) in terms of projected (1) primary school completion rate (conditional on starting, since nearly 100% of students start), (2) secondary school completion rate (not conditional on starting secondary), and (3) projected earnings based on primary and secondary school. Base assumptions described in Section 4.6 and Appendix D.1. In panel B, for the third scenario, the reason that the baseline dropout gap between low SES and high SES cannot be increased more than 30% while maintaining the same average level of dropouts is that the dropout rate for high-SES would have to be negative.

D.4 Back-of-the-envelope calculation using workbooks instead of dropouts

The idea for this calculation is the same as for the retention calculation: Compare the gap between low- and high-SES households in the control and treatment groups and see how much of the gap is closed by information.

For workbooks, a parent's goal is to match the difficulty of the book to the performance level of their child. I assume that the highest-possible-returns mapping is the (average) choice made by the treatment group (i.e., that the best book for a child is the average difficulty chosen by the treatment group for children of her performance level), and that the returns to the workbooks are linearly decreasing in difficulty away from the optimum (i.e., that for a student who should have the remedial workbook, the decrease in returns from the advanced is twice as big as for the average, and that for an average child the returns are equally bad for a remedial or advanced). Under these assumptions, households with above-median education parents had higher returns at baseline than those with below-median education due to their more accurate beliefs, but information closes 88% of the gap for math books, and 100% of the gap for English books. Note that this calculation differs slightly from the magnitude sizing in Section 4.2 because here we are not just using averages but taking the averages of the absolute value of the deviations. That said, the conclusion are very similar to the Section 4.2 results: there also information nearly fully closes the gap.