Using Randomized Information Shocks to Understand How Parents' Investments Depend on Their Children's Ability*

Rebecca Dizon-Ross
University of Chicago Booth School of Business

November 15, 2021

Abstract

Do parents invest more in children with higher or lower academic ability? This paper uses a new experimental method to uncover the relationship between investments and perceived ability. I randomly deliver information to parents in Malawi about their children's academic ability and analyze heterogeneity in the treatment effects based on both parents' baseline beliefs and the information delivered. I find that parents invest more years of schooling in children with higher academic ability. However, for expenditures and attendance, the results vary by parent socioeconomic status (SES), with high-SES parents investing more in their higher-ability children and low-SES parents investing more in their lower-ability children.

*I am very grateful to Ran Abramitzky, Abhijit Banerjee, Jim Berry, Marianne Bertrand, Pascaline Dupas, Esther Duflo, Caroline Hoxby, Seema Jayachandran, Asim Khwaja, Ben Olken, and Varun Satish for helpful comments and discussions. I thank Bridget Hoffmann, Rachel Levenson, and Michael Roscitt for help with the fieldwork, and Min Byung Chae and Lea Dousset 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. This study is registered in the AEA RCT Registry under the unique identifying number AEARCTR-0001808. All errors are my own. Contact: rdr@chicagobooth.edu.

1 Introduction

Parents' investments in their children have large and persistent impacts on their children's outcomes. Do parents invest more in children with higher ability, thereby reinforcing inequality in the distribution of child ability, or do they invest more in children with lower ability, thereby dampening inequality? The relationship between ability and investments is theoretically ambiguous, depending on the parents' perceived production function for human capital, financial constraints, and parents' preferences for equality. Understanding this relationship is important for predicting policy spillovers: if parents spend more on their high-ability children, then policies that improve children's ability will crowd-in other household spending. Although many papers have examined this relationship, causal identification is difficult, mainly because of the potential for reverse causality between investments and ability.

This paper uses a new experimental method to estimate how parental investments depend on their children's characteristics. I use child academic ability as the characteristic of interest, but the method could also be used for other characteristics. The first step is to conduct an experiment that delivers information to randomly-selected parents about their children's ability. The second step is to investigate how parental investments respond to the information as a function of both the ability information delivered and the parents' prior beliefs about their children's ability. The analysis essentially compares the treatment effects of information for parents who have similar beliefs about their children's ability but who receive different ability information, asking: do the investments of parents who receive high ability signals increase more or less than the investments of parents who receive low ability signals?

This paper first outlines the method and then implements it using data from an experiment that delivers information on children's school performance to 3,200 parents in Malawi. Malawian parents regard school performance as the most important measure of academic ability and it will serve as our ability measure throughout the paper. The experiment was previously analyzed in Dizon-Ross (2019); that paper examined whether information frictions impede parents' educational investments but did not aim to quantify how parents' investments depend on their children's ability, which is the goal of this paper.

I first quantify how investments depend on academic ability in the full sample and then estimate heterogeneity by a proxy for socioeconomic status (SES). The proxy I use is parental education, which, given the challenges of measuring income in developing countries, is the most cleanly-measured SES proxy in the data. Many theoretical models predict that the relationship between investments and ability will vary with wealth and SES because the relationship is mediated by financial constraints. For example, Becker (1993) argues that, relative to poorer parents, richer parents may invest more in their higher-ability children than their lower-ability children because, unlike poorer parents, they can afford to later offset inequalities in their children's labor market earnings by transferring non-human capital.

I present two main empirical findings. First, when I examine enrollment (the "extensive" margin), I find that parents invest more years of schooling in their higher-ability children. This suggests that parents believe that years of schooling and ability are complements. Second, when I examine how much parents invest in each child (the "intensive margin"), I find significant heterogeneity by parent SES: more-educated parents invest more in their higher-performers, whereas less-educated parents do the opposite. This pattern holds for both of the intensive margin investments examined, expenditures and attendance, with the pattern stronger for expenditures. This heterogeneity finding aligns with the classic Becker (1993) prediction. However, the reasons could differ from Becker (1993)'s, as even the more-educated households in my sample may not be able to afford to transfer non-human capital. Another potential explanation is that budget constraints vary with SES, and there are discontinuous benefits to hitting educational milestones, such as graduating from secondary school or learning to read. In particular, in Malawi, while primary school is freely available, secondary school is expensive and difficult to gain admission to. If only richer parents can afford secondary school, then there could be a higher perceived benefit for them to getting high achievers over the admission threshold. In contrast, for poorer parents, there may be higher perceived returns to helping low achievers acquire basic skills like learning to read.

This paper's method exploits random variation in who received information as well as non-random variation in children's true ability. To address the non-random variation, I perform robustness checks and supplementary analyses to show that correlates of true ability do not confound the estimates of the investment/ability relationship.

This paper contributes to a large literature examining how parents' investments depend on their children's ability and characteristics (see Almond and Currie (2011) and Almond and Mazumder (2013) for reviews). This literature finds mixed results. Identification is a key challenge: it is hard to find exogenous variation in children's characteristics. Many studies have exploited within-household variation in birthweight (e.g., Datar et al., 2010; Hsin, 2012; Restrepo, 2016); however, birthweight may not be exogenous to other time-varying household circumstances. Other studies use twin comparisons (e.g., Yi et al., 2015; Grätz and Torche, 2016), but the external validity of twin comparisons appears to be limited (Bharadwaj et al., 2013; Almond and Mazumder, 2013). More recent studies often use early-life biological shocks for identification, such as early-life grain yields (Leight, 2014), exposure to iodine (Adhvaryu and Nyshadham, 2016), Chernobyl (Halla and Zweimüller, 2014), and rainfall (Fan and Porter, 2020).

This paper's primary contribution is to introduce a new, experimental identification method that allows for three important differences from previous strategies. First, while previous strategies have only been able to understand how parents respond to early-life shocks to their children's traits, my method can be used to analyze how parents respond to shocks at any age. This is important because many policy interventions affecting children occur later in their lives (e.g., school interventions). Parents may respond very differently to early-life shocks than later-life shocks, as many parents believe that the production function varies by age (Boneva and Rauh, 2018).

Second, while the previous literature has relied on early-life biological shocks that affect children on many margins (health, cognitive ability, etc.), many policies are more limited in scope. A remedial education program, for example, may improve cognitive ability but is unlikely to affect health. My method allows us to identify how parents respond to more narrowly-defined characteristics, thus mapping more closely to many policy interventions.

Finally, the literature largely relies on household fixed effects for identification; in contrast, my strategy relies on experimental variation and does not require a household fixed effect. As Almond and Currie (2011) point out, the within-household investment relationship identified using a household fixed effect can differ importantly from the overall relationship, which is more relevant for many policy analyses. To give an example of how the two can diverge, if parents choose to invest equally in their children, as Berry et al. (2020) suggests is common, but parents' investments increase in their children's *average* ability, then within-household investments would not depend on child ability even though, overall, investments would increase in ability.

The next section motivates the empirical approach. The subsequent section describes the context and experimental design. I then presents the results and conclude.

2 Identification Approach

2.1 Parents' Optimization Problem

Child i's parent is choosing investments in child i's education. Denote one of the parent's investments as s_i. Denote child i's ability as α_i; the parent does not know α_i but rather has a distribution of beliefs about α_i, represented by the normal distribution $N(\hat{\alpha}_i, \sigma^2)$.[1] Our objective is to measure how investments depend on the mean of the beliefs distribution, $\hat{\alpha}_i$.

Investments increase children's human capital. Denote the perceived production function for child i's human capital as:

[1]For expositional simplicity, I represent the uncertainty of the beliefs distribution σ^2 as constant across the population even though the mean $\hat{\alpha}_i$ can vary. I do this because the paper aims to identify the slope of investments with respect to $\hat{\alpha}_i$, not σ^2, and so my identification strategy exploits individual-level variation in $\hat{\alpha}_i$, not σ^2.

$$q_i = f(s_i, \alpha_i) \tag{1}$$

with f concave in s_i. Parent i chooses s_i to maximize expected household utility

$$s_i^* = \arg\max_s EU(q_i) = \arg\max_s EU(f(s_i, \alpha_i)) \tag{2}$$

subject to a budget constraint, with the expectation taken over both the beliefs distribution $N(\hat{\alpha}_i, \sigma^2)$ and any other uncertainty in the production of human capital.

The Investment Function

Define the *investment function*, or $s^*(\hat{\alpha}|\sigma^2)$, as the full set of solutions to equation 2 for all potential values of $\hat{\alpha}$ given a value of σ^2. Within this framework, our objective is to measure the slope of the investment function on $\hat{\alpha}$ for a given σ^2, $\frac{\partial s^*}{\partial \hat{\alpha}}$.

The slope $\frac{\partial s^*}{\partial \hat{\alpha}}$ depends on whether ability and investments are complements or substitutes in the perceived production function. If parents' utility functions maximize returns[2] and parents think that investments and ability are complements ($\frac{\partial^2 f}{\partial \alpha \partial s} > 0$), then parents' investments will increase in perceived ability, $\frac{\partial s^*}{\partial \hat{\alpha}} > 0$, thus "reinforcing" pre-existing gaps in ability. If instead investments and ability are perceived substitutes ($\frac{\partial^2 f}{\partial \alpha \partial s} < 0$), then investments will decrease in perceived ability, $\frac{\partial s^*}{\partial \hat{\alpha}} < 0$, "compensating" for gaps in ability.

The slope $\frac{\partial s^*}{\partial \hat{\alpha}}$ also depends on σ^2. While the exact relationship depends upon the specific utility function and beliefs distribution, in many models, a higher σ^2 decreases the magnitude of the slope (i.e., $|\frac{\partial s^*}{\partial \hat{\alpha}}|$ decreases in σ^2): the more uncertain the beliefs are, the less strongly parents would want to invest based on those beliefs. This dependence of the slope on σ^2 means that one cannot identify the slope of the investment function *in general* but rather for a specific level of beliefs uncertainty. I describe below for which level of σ^2 my method identifies the slope.

2.2 Identification Strategy

The goal of this paper is to characterize the slope of the investment function, $\frac{\partial s^*}{\partial \hat{\alpha}}$, for a given population of parents and given σ^2. For simplicity, parametrize the investment function as linear:

$$s^*(\hat{\alpha}|\sigma^2) = \gamma(\sigma^2) + \beta(\sigma^2)\hat{\alpha} \tag{3}$$

where the parameter of interest, $\beta(\sigma^2)$, varies with σ^2. Assume that parent i's actual investments represent $s^*(\hat{\alpha}_i|\sigma^2)$ plus an error term ε_i that reflects all other determinants of investment:

$$s_i(\hat{\alpha}_i|\sigma^2) = \gamma(\sigma^2) + \beta(\sigma^2)\hat{\alpha}_i + \varepsilon_i. \tag{4}$$

A simple regression of s on perceived ability $\hat{\alpha}$ would be biased due to omitted variable

[2] For example, if $U(q_i)$ is a linear function of q_i.

bias from the potential correlation between $\hat{\alpha}$ and ε. For example, a parent who cares more about education might consistently invest more in her child's education over time. Thus, at any given moment, the parent both invests more and has higher $\hat{\alpha}$ because of her higher investments in her child in the past. If caring about education is imperfectly measured, this generates a correlation between $\hat{\alpha}$ and ε.

This paper proposes an experimental approach to overcome this endogeneity problem. Consider an experiment that randomly divides parents into a "treatment group" that receives an information signal about α_i, and a "control group" that does not receive the signal. I denote the information signal as x_i. Because any measure of α_i will have some noise, I do not assume that x_i equals α_i itself, but rather that x_i is drawn from a normal distribution whose mean is α_i: $N(\alpha_i, \eta^2)$. I use "baseline" to denote the time period before the treatment group receives information and "endline" to denote the time period after, and I denote the baseline beliefs distribution for parent i as $N(\hat{\alpha_{0i}}, \sigma_0^2)$.

Beliefs Updating

Because the control group does not receive information, the mean and uncertainty of their beliefs distributions will remain the same at endline as at baseline:

$$\text{Control group endline beliefs} \sim N(\hat{\alpha_{0i}}, \sigma_0^2). \tag{5}$$

In contrast, for the treatment group, receiving information will move their mean beliefs towards the information signal and change their beliefs uncertainty. With Bayesian updating, the posterior mean can be expressed as a weighted average of the signal and the prior mean: $\lambda x_i + (1-\lambda)\hat{\alpha_{0i}}$. Hence,

$$\text{Treatment group endline beliefs} \sim N(\lambda x_i + (1-\lambda)\hat{\alpha_{0i}}, \sigma_1^2), \tag{6}$$

with λ denoting an updating parameter that shows us how far beliefs move towards the signal, and σ_1^2 denoting the *ex post* updated uncertainty.[3]

Treatment Effect on Investments

At endline, parents' investments will equal the investment equation (equation 4) evaluated at the mean and variance of their endline beliefs distributions. Control group investments will hence equal

$$s_i^C = \gamma(\sigma_0^2) + \beta(\sigma_0^2)\hat{\alpha_{0i}} + \varepsilon_i, \tag{7}$$

[3] In particular, the posterior density can be expressed as $N\left(\left(\frac{\hat{\alpha_{0i}}}{\sigma_0^2} + \frac{x_i}{\eta^2}\right)\left(\frac{1}{\sigma_0^2} + \frac{1}{\eta^2}\right)^{-1}, \left(\frac{1}{\sigma_0^2} + \frac{1}{\eta^2}\right)^{-1}\right)$, and so $\lambda = \frac{1}{\eta^2}\left(\frac{1}{\sigma_0^2} + \frac{1}{\eta^2}\right)^{-1}$ and $\sigma_1^2 = \left(\frac{1}{\sigma_0^2} + \frac{1}{\eta^2}\right)^{-1}$.

while treatment group investments will equal

$$s_i^T = \gamma(\sigma_1^2) + \beta(\sigma_1^2)\left(\lambda x_i + (1-\lambda)\hat{\alpha_{0i}}\right) + \varepsilon_i. \tag{8}$$

Thus treatment group investments differ from control group investments because both the mean and variance of the beliefs distribution have shifted.

The treatment effect, denoted τ_i, of delivering information to parent i, is equal to the parent's investment if assigned to the treatment group minus her investments if assigned to the control group:

$$\begin{aligned} \tau_i &= s_i^T - s_i^C \\ &= \left[\gamma(\sigma_1^2) + \beta(\sigma_1^2)(\lambda x_i + (1-\lambda)\hat{\alpha_{0i}}) + \varepsilon_i\right] - \left[\gamma(\sigma_0^2) + \beta(\sigma_0^2)\hat{\alpha_{0i}} + \varepsilon_i\right] && (9) \\ &= \left(\gamma(\sigma_1^2) - \gamma(\sigma_0^2)\right) + \lambda\beta(\sigma_1^2)(x_i - \hat{\alpha_{0i}}) + (\beta(\sigma_1^2) - \beta(\sigma_0^2))\hat{\alpha_{0i}} && (10) \\ &= \left(\gamma(\sigma_1^2) - \gamma(\sigma_0^2)\right) + \lambda\beta(\sigma_1^2)x_i + \left((1-\lambda)\beta(\sigma_1^2) - \beta(\sigma_0^2)\right)\hat{\alpha_{0i}} && (11) \end{aligned}$$

I make several observations about the treatment effect. First, because the treatment effect exploits the exogenous information shock, the ε term drops out. Hence, analyzing the treatment effect allows us to sidestep the omitted variable bias stemming from the correlation between ε and x or $\hat{\alpha_0}$.

Second, interpreting equation 10, the treatment effect has three components: first, an "intercept" term that is constant across the population and reflects the fact that information may affect uncertainty which could affect the investment function's intercept (i.e., investment chosen by a parent with $\hat{\alpha} = 0$); second, a term reflecting how investments shift from the shock to mean beliefs $\lambda(x_i - \hat{\alpha_{0i}})$; and third, a term reflecting how investments shift from the shock to uncertainty and corresponding change in slope.

Third, and most importantly for identification, equation 11 shows that the treatment effect varies with two individual-level parameters: $\hat{\alpha_{0i}}$ and x_i. This equation leads directly into our estimation procedure.

In particular, equation 11 shows that the heterogeneity in the treatment effect based on x_i (controlling for the heterogeneity by $\hat{\alpha_{0i}}$) identifies the parameter $\lambda\beta(\sigma_1^2)$. Importantly, since $0 \leq \lambda \leq 1$, $\lambda\beta(\sigma_1^2)$ tells us both the sign of the parameter of interest, $\beta(\sigma_1^2)$, as well as a lower bound estimate of its magnitude (since $|\lambda\beta(\sigma_1^2)| \leq |\beta(\sigma_1^2)|$).

In contrast, the heterogeneity in the treatment effect based on $\hat{\alpha_{0i}}$ (controlling for the heterogeneity in the treatment effect by x_i) identifies $\left((1-\lambda)\beta(\sigma_1^2) - \beta(\sigma_0^2)\right)$, which is less easy to interpret. Thus, unless λ is known, this method does not allow us to estimate $\beta(\sigma_0^2)$; rather we are only estimating the slope for uncertainty σ_1^2.[4]

[4] To further illustrate the intuition behind the three terms in equation 11, in the case where parents fully

Estimation Specification

To estimate the slope of the investment function, specifically $\beta(\sigma_1^2)$, equation 11 shows us that we can estimate the heterogeneity in the treatment effect by x, controlling for the heterogeneity by $\hat{\alpha_0}$. To do so, one can estimate the following regression:

$$s_i = b_0 + b_1 Treat_i \times x_i + b_2 Treat_i \times \hat{\alpha_{0i}} + b_3 Treat_i + b_4 x_i + b_5 \hat{\alpha_{0i}} + \mu_i \tag{13}$$

where s_i is a measure of parental investment; $Treat_i$ is an indicator for being in the treatment group; x_i is the information signal delivered through the experiment; and $\hat{\alpha_{0i}}$ is child i's parent's baseline (pre-intervention) mean belief about child i's ability. The coefficient of interest is b_1, which provides a lower bound estimate of the slope of the investment function, $\beta(\sigma_1^2)$. The approach essentially uses three instruments ($Treat_i, Treat_i \times x_i$, and $Treat_i \times \hat{\alpha_{0i}}$) to identify the three parameters in equation 11 ($(\gamma(\sigma_1^2) - \gamma(\sigma_0^2))$, $\lambda\beta(\sigma_1^2)$, and $((1-\lambda)\beta(\sigma_1^2) - \beta(\sigma_0^2))$, with $\lambda\beta(\sigma_1^2)$ the coefficient of interest.

Identification

As shown above, this specification uses experimental variation to overcome the classic endogeneity concern that perceived ability $\hat{\alpha_{0i}}$ is correlated with ε_i. However, it also relies on non-experimental variation in how the treatment effect of information varies with x_i and $\hat{\alpha_{0i}}$. In particular, if there were some correlate of x_i or $\hat{\alpha_{0i}}$ that also influenced the treatment effect of information, that would be problematic for identification.

There are two main robustness strategies to address this concern. The first is to control for other potential confounding factors interacted with treatment and test for stability in the coefficients of interest. The second is to measure treatment effects on investments for which we have clean *ex ante* predictions of what the investment function would be; one can then verify that the estimated investment function aligns with the *ex ante* predictions to provide reassurance that the estimation procedure works and that the identification assumptions hold.

I later present results from these strategies, both of which suggest that confounding factors do not drive the treatment effects here.

Differences from Dizon-Ross (2019)

Although both this paper's estimation strategy and that of Dizon-Ross (2019) rely on treatment effect heterogeneity, we learn different things from the specifications. Dizon-Ross

update their beliefs ($\lambda = 1$), equation 11 becomes

$$\tau_i = \left(\gamma(\sigma_1^2) - \gamma(\sigma_0^2)\right) + \beta(\sigma_1^2)x_i - \beta(\sigma_0^2)\hat{\alpha_{0i}} \tag{12}$$

and so the treatment effect heterogeneity by x_i and $\hat{\alpha_{0i}}$ would directly identify the slope of the investment function under different σ_2's: $\beta(\sigma_1^2)$ and $\beta(\sigma_1^2)$, respectively.

(2019) focuses on understanding whether information frictions affect the correlation between investments and true performance, not on quantifying the investment function, and so its primary specification examines whether information increased the slope of investments on x.[5] Unlike our specification, that specification does not provide a clean bound on the slope of the investment function and is thus not appropriate for accomplishing this paper's goal of estimating the slope of the investment function, except when there are no "uncertainty effects" (i.e., $b(\sigma_1^2) = b(\sigma_0^2)$), which does not appear to be the case for the investments analyzed here.[6]

A second distinction is that Dizon-Ross (2019) does not examine heterogeneity by SES in the investments presented here, which I consider a valuable contribution given the literature's longstanding theoretical interest in this relationship (Becker, 1993).[7]

3 Experimental Design

The experiment randomly provided some parents in rural Malawi with information on two of their children's "academic performance," which hereafter refers to average performance on achievement tests administered by schools during the term before the experiment took place. In Malawi, parents see academic performance as an important proxy for their children's underlying ability and the most important determinant of their own investments.[8] Dizon-Ross (2019) provides additional details on the experiment, reports the program evaluation results, and quantifies the extent to which information frictions distort investments. In contrast, this paper quantifies the slope of the investment function. This section describes the components of the experiment and data that are necessary to understand the analysis.

3.1 Context

Primary school in Malawi covers grades 1-8. Although it is technically free, it involves extra expenditures. Parents in the study sample spent an average of 1,750 Malawi Kwacha (MWK) annually per child, roughly 10.6 USD or 1.6% of annual household income. The main expenditures are uniforms (33%), 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 (World Bank, 2010).

[5] The primary specification was thus $s_i = b_0 + b_1 Treat_i \times x_i + b_3 Treat_i + b_4 a_i + \mu_i$.

[6] If there are no uncertainty effects, the coefficients on $Treat \times \hat{\alpha}$ and $Treat \times x$ should be equal and opposite, which does not appear to be the case in Table 2.

[7] Dizon-Ross (2019) additionally examines heterogeneity in other experimentally-generated investments that are not analyzed as primary outcomes in this paper; see the Robustness section for further discussion.

[8] Academic performance depends both on "innate" ability and past inputs, and so we identify how investments depend upon this combined metric. While it could theoretically be interesting to isolate "innate" ability, the literature has consistently documented that it is difficult to measure "innate" ability, and that any measure combines innate ability with past inputs.

Although schools in the sample all send report cards home every term with average achievement test scores, the official report cards are often hard for parents to understand or do not reach them at all. Sixty percent of parents state that they do not know their child's performance from the last report, thus creating scope for an intervention providing academic performance information to affect parents' beliefs.

3.2 Experimental Design

The study worked with 39 schools in two districts in Malawi. We sampled 3,451 households enrolled at those schools during term 2 of the 2012 school year with at least two students enrolled in grades 2-6 and with academic performance data available.

Half of the households in the sample were randomly-assigned to a treatment group that received information about their children's academic performance and half to a control group that did not. The randomization was stratified on a test score measure and a proxy for parent education, since one *ex ante* goal was to look at heterogeneity by parent education. Among the 3,451 sampled households, 21% of households were found to be ineligible and, of the eligible households, 97% agreed to participate in the baseline survey, yielding a final experimental sample of 2,634 households.

Baseline survey visit: Surveyors visited all households and asked to speak with the parent who is the primary decision-maker about education. Surveyors then conducted a baseline survey measuring demographics, education spending, and beliefs. While eliciting baseline beliefs about test scores, surveyors explained the grading scale used by schools to parents and reviewed a sample report card with the same format as those later delivered to the treatment group. Immediately after administering the survey, surveyors conducted the information intervention for the treatment group.

Information intervention and report cards (Treatment group only): Surveyors walked parents through two report cards (one per child) describing their children's academic performance. The reports showed each child's performance on the 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. The statistics were term-level averages 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). Appendix A presents a sample report card.

3.3 Data and Outcomes

The analysis uses data from surveys with parents and administrative data from schools.

(1) School performance data: In March 2012, at the end of term 2 of the 2012 school year, surveyors gathered the term 2 achievement test data. I construct an overall score that

averages the math, English, and Chichewa scores as the measure of x. I use absolute scores instead of the relative (percentile) measures, since parents appear to have paid more attention to the absolute than relative information.[9]

(2) Baseline survey data: The baseline survey ran from April to June of 2012 and measured demographics, education spending, and beliefs. Mean beliefs about academic performance were measured by asking parents about the same performance metrics later delivered in the intervention report cards. I use parents' beliefs about the overall score x as the measure of $\hat{\alpha_0}$.

(3) First endline survey: This survey was conducted immediately after the baseline survey and information intervention. This survey measured "experimental outcomes" which were opportunities to take up educational resources offered to parents as part of the experiment. These investments were designed to produce clear *ex ante* predictions about the shape of the investment function; I hence do not use these as primary outcomes since backing out their investment functions is not of standalone interest. Instead, I use them in the robustness section to address potential identification concerns.

(3) Outcomes: Outcomes data come from two sources: (i) data on dropouts and expenditures from an endline survey of parents one year after the intervention (June-July 2013); (ii) administrative data on attendance gathered 1 month after the intervention (July 2012). There was sufficient budget to include roughly 900 households in the second endline survey sample. Of the households selected for the sample, 98% (893) were successfully surveyed, balanced by treatment. The attendance data were gathered by giving schools templates to record the data for the month following the intervention and were collected from 35% of the sample.[10]

3.4 Summary Statistics and Balance

Table 1 presents summary statistics and tests for balance by treatment. Identification relies on baseline beliefs differing from the information signal delivered, which seems true here: on average, parents' beliefs about their children's true test scores differ from their children's actual test scores by 20.4 percentage points (pp), over 1 standard deviation of the test score distribution. The differences between the treatment and control groups in the baseline variables are never large, with a joint test of equality failing to reject the null that all are 0.

[9] See Section 2.1 and Online Appendix F of Dizon-Ross (2019).

[10] See Online Appendix F2 of Dizon-Ross (2019) for more detail.

4 Empirical Results

I now use the experimental data to estimate the slope of the investment functions for three outcomes: enrollment, attendance, and expenditures. There is no reason to expect the perceived production function – and hence the investment function – to be the same for enrollment (the "extensive margin") as for the intensive margin; for example, years of schooling could be more helpful for higher-performing children but, conditional on enrollment, attendance and expenditures could be more critical in preventing lower-performing children from falling behind. I estimate the investment functions first in the full sample, and then estimate the heterogeneity by parents' education.

4.1 Full Sample Results

Panel A of Table 2 presents the results from estimating equation 13 in the full sample. Panel B shows estimates using binary regressors, specifically, indicators for whether a student has an above-median score and whether a parent has an above-median belief about her child's score. I include the Panel B estimates because they are easier to interpret, but they have lower statistical power since they do not leverage all the underlying variation in the data. All regressions include a standard vector of control variables to improve precision.[11] Standard errors are clustered at the household level.

Recall that the coefficient on $Treat \times Score$ gives a lower bound estimate of the slope of the investment function, $\beta(\sigma_1^2)$. Panel A shows that both enrollment (column 1) and attendance (column 3) are upward sloping in perceived child academic ability, although the latter is significant only at the 10% level. A 1 standard deviation increase in a parent's perception of her child's academic ability (18 score points) is associated with at least a 1.8 pp increase in the likelihood of enrollment one year later, and at least a 1.6 pp increase in the likelihood that the child is in school on a given day (these statistics represent the coefficient on $Treat \times Score$ in columns 1 and 3, respectively, multiplied by 18). These are meaningful increases given that the average dropout rate and absence rate in the control group were both below 10% (2.0% for dropouts and 8.9% for absence). In contrast, there is no significant effect on expenditures.

The binary versions in Panel B tell a similar story, with lower precision. The $Treat \times AboveMedianScore$ coefficients suggest that having a child with an above-median score instead of a below-median score increases the likelihood of enrollment by 3.5 pp and the attendance rate by 1.5 pp.

[11]The estimates are very similar without controls but precision decreases (Appendix Table A.1). See list of control variables in the Table 2 notes.

4.2 Heterogeneity by parent education

I now estimate the heterogeneity in the investment functions by parent education. The investment function could vary with SES if the perceived production functions, constraints, or preferences vary.

Panel A of Table 3 presents the results from estimating equation 13 fully interacted with household-average years of parental education. Panel B presents the "binary versions," using dummies for above-median scores and above-median beliefs as the measures of x and $\hat{\alpha_0}$ and a dummy for having above-median parental education as the parental education measure. Again, Panel B has lower precision and is included for ease of interpretation. For enrollment (column 1), the power to detect heterogeneity is limited.

Interestingly, for expenditures and attendance, the investment functions of less-educated and more-educated parents slope in opposite directions. Interpreting Panel A, column 2, the coefficient on $Treat \times Score$, which represents the slope of the investment function for parents with no education, is negative (significant at the 10% level): the least-educated parents spend more on their lower-performing children. However, the coefficient on $Treat \times Score \times ParentEduc$ is positive, suggesting that the more education parents have, the more they spend on their higher-performing children relative to their lower-performing children. Based on linear extrapolation of the coefficients, once parents have at least 5 years of schooling, the slope becomes positive, and substantially more positive as schooling increases. To more easily understand the magnitudes, Panel B shows that parents with below-median education (the omitted category) spend roughly 15% more on children who have lower scores (*Treat* × *Above-median score*).[12] More-educated parents do the opposite, increasing spending on *higher*-performing children relative to lower-performing children by 19% (sum of the coefficients on *Treat* × *Above-median score* and *Treat* × *Above-med.score* × *Above-med.ParentEduc*). Hence, the investment function has opposite signs for more and less-educated parents: a negative slope among less-educated parents and a positive slope among more-educated ones.

A similar, but less statistically strong, pattern holds for attendance. For the less-educated, information if anything increases the attendance of low-performing children ($Treat \times$ *Score*), whereas it appears to do the opposite for the more-educated ($Treat \times$ *Score* $\times$ $ParentEduc$).

One potential explanation for the heterogeneity is that budget constraints vary by education and that the production function exhibits discontinuous benefits for hitting educational milestones, such as graduating from secondary school or learning to read. Thus, more-educated parents (who are likely to be richer) may believe they can afford to send

[12]Logs are used for precision but only 1 percent of observations are 0; Appendix Table A.3 shows the results are robust to other specifications.

their children to secondary school and want to get their high achievers over the admission threshold. In contrast, less-educated parents may not see secondary school as an option and so could have higher perceived returns to helping low achievers acquire basic skills like reading. There are of course other possible explanations.

These heterogeneity results are not driven by selection into schooling. The expenditure results are robust to controlling for or conditioning on enrollment (Appendix Table A.2), and attendance was measured immediately after the intervention before dropouts had occurred.

Although heterogeneity in updating λ by SES could cause heterogeneity in the estimated lower bound of the investment function (i.e., the coefficient on $Treat \times Score$), it cannot be responsible for heterogeneity in the *sign* of the estimate (positive for one group, negative for another). Rather, the results here suggest that, for the intensive margin investments of attendance and expenditures, there is heterogeneity in the slope of the investment function itself ($\beta(\sigma_1^2)$).

4.3 Robustness

This section presents two pieces of evidence that confounding factors do not drive the results.

First, to address the concern that the treatment effect heterogeneity by score might be confounded by other treatment effect heterogeneity, Table 4 shows that the results are robust to controlling for other observables interacted with treatment, such as gender and baseline educational expenditures. Panels A through C show robustness of the full-sample results. Reassuringly, the coefficients of interest remain stable as I control for different variables across columns. Panels D through F repeat the exercise for the parental heterogeneity results, with columns successively adding controls interacted with treatment variables (i.e., in parallel to score); the coefficients are again stable.[13]

Second, I summarize the analysis from Dizon-Ross (2019) of the "experimental outcomes" which suggests that the primary factor driving heterogeneity in the response to information was in fact variation in scores and baseline beliefs, not other correlated factors. Recall that the experimental outcomes (described in the notes to Appendix Table A.4) were designed to have clear *ex ante* predictions about the shape of the investment function. For example, one outcome is willingness to pay for a remedial textbook, which should decrease in perceived ability since the textbooks are remedial.

Because we already know the shape of the perceived investment function for these investments and do not need to back it out from the treatment effects, we can instead use

[13]Parent education is also not randomized and so the results could be picking up heterogeneity by other dimensions of SES correlated with parent education. For our purposes this is not a concern: our goal is to estimate the heterogeneity by parent SES and all its correlated dimensions.

the treatment effects and our knowledge of the investment function to back out whether the identification assumptions hold.

Appendix Table A.4 shows that the heterogeneity in the treatment effects by x_i and $\hat{\alpha_{0i}}$ aligns with what theory predicts, thus providing evidence that the identification assumptions hold. All of the treatment effects have the expected sign. Moreover, for the smaller investments for which one would expect limited uncertainty effects, the signs on $Treat \times Score$ and $Treat \times Beliefs$ are nearly equal and opposite, which exactly fits our model but would be unlikely if instead confounding factors were driving the effects.[14] Finally, the larger the investment, the larger the magnitude of the $Treat \times Score$ coefficient relative to the $Treat \times Beliefs$ coefficient, which matches a model where uncertainty plays a greater role for larger investment decisions.

5 Conclusion

This paper presents a new experimental method for identifying how parents' investments depend on their perceptions of their children's ability. Using data from an information experiment in Malawi, I find that at the extensive margin, parents invest more years of schooling in children who are higher performing. However, at the intensive margin, the results vary across parents. For example, for expenditures, while no pattern emerges in the full sample, this masks heterogeneity: less-educated parents spend more on their lower performers and more-educated parents spend more on their higher performers.

Understanding the dependence of investments on perceived ability is important for predicting policy spillovers. If parents "reinforce," spending more on their high-ability children, policies that increase ability will crowd-in household investments. My intensive margin results suggest that policies that increase ability may increase socioeconomic inequality among enrolled students by crowding in investment among high-SES households but crowding it out among low-SES households.

Two areas for future work are estimating the *actual* production function to determine whether parents appear to accurately understand it, and applying this paper's method elsewhere to examine how the investment function varies across settings.

[14] If there are no uncertainty effects ($\beta(\sigma_1^2) = \beta(\sigma_0^2)$), then the coefficient on $Treat \times Belief$s becomes $-\lambda\beta(\sigma_1^2)$, the opposite of the coefficient on $Treat \times Score$.

References

Adhvaryu, A. and A. Nyshadham (2016). Endowments at birth and parents' investments in children. *The Economic Journal 126*(2593), 786–820.

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

Almond, D. and B. Mazumder (2013). Fetal Origins and Parental Responses. *Annual Review of Economics 5*(1), 37–56.

Becker, G. (1993). *A Treatise on the Family* (Enlarged ed.). Cambridge, Massachusetts: Harvard University Press.

Berry, J., R. Dizon-Ross, and M. Jagnani (2020). Not playing favorites: An experiment on parental fairness preferences. Technical report, National Bureau of Economic Research.

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

Boneva, T. and C. Rauh (2018). Parental beliefs about returns to educational investments—the later the better? *Journal of the European Economic Association 16*(6), 1669–1711.

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

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

Fan, W. and C. Porter (2020). Reinforcement or compensation? Parental responses to children's revealed human capital levels. *Journal of Population Economics 33*(1), 233–270.

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

Halla, M. and M. Zweimüller (2014). Parental response to early human capital shocks: Evidence from the chernobyl accident.

Hsin, A. (2012). Is biology destiny? birth weight and differential parental treatment. *Demography 49*(4), 1385–1405.

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

Restrepo, B. J. (2016). Parental investment responses to a low birth weight outcome: who compensates and who reinforces? *Journal of Population Economics 29*(4), 969–989.

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

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

Table 1: Baseline summary statistics

	Full sample		Treatment group			Parents' education		
	Mean	SD	Control	Treat	Std Diff	Below Median	Above Median	Std Diff
A. Respondent Background								
Female	0.77	0.42	0.77	0.76	-0.04	0.74	0.8	0.13
Primary education decision maker	0.92	0.27	0.91	0.92	0.04	0.93	0.91	-0.05
Age	40.8	11.0	40.6	41.0	0.03	43.2	38.3	-0.47
Education (years)	4.44	3.57	4.42	4.45	0.01	2.08	7.22	2.04
Respondent has secondary education	0.11	0.31	0.11	0.11	0.01	0	0.24	0.79
Parent can read or write Chichewa	0.67	0.47	0.67	0.68	0.02	0.45	0.93	1.24
Respondent is farmer	0.46	0.5	0.47	0.46	-0.01	0.52	0.39	-0.26
Respondent's weekly income	2,126	4,744	2,051	2,203	0.03	1,528	2,839	0.27
B. Household Background								
Family size (Number of children[a])	5.13	1.74	5.16	5.1	-0.04	5.21	5.04	-0.1
One-parent household	0.19	0.39	0.19	0.2	0.03	0.21	0.18	-0.08
Parents' average education (years)	4.66	3.25	4.68	4.64	-0.01	2.22	7.54	2.79
Any parent has secondary education	0.18	0.38	0.17	0.19	0.04	0	0.38	1.1
C. Student Information								
Child's grade level	3.72	1.37	3.72	3.72	0	3.7	3.74	0.03
Child's age	11.6	2.68	11.7	11.6	-0.04	12.1	11.1	-0.38
Child is female	0.51	0.5	0.52	0.5	-0.04	0.52	0.5	-0.04
Baseline attendance	0.91	0.13	0.92	0.91	-0.01	0.91	0.92	0.15
Annual per-child education expenditures	1,742	2,791	1,712	1,772	0.02	1,276	2,284	0.36
Fees paid to schools	381	1,128	384	378	0	276	503	0.19
Uniform expense	576	1,019	548	603	0.05	480	687	0.2
School supplies, books, tutoring, etc.[b]	785	1,819	780	790	0.01	519	1,095	0.31
Any supplementary expenditures on child	0.9	0.3	0.9	0.89	-0.02	0.89	0.91	0.07
D. Perceived and True Academic Performance								
True Overall Score	46.8	17.5	47.1	46.4	-0.04	46.1	47.5	0.08
Believed Overall Score	62.4	16.5	62.7	62.0	-0.04	61.6	63.3	0.1
SD of Individual Beliefs about Score	7.69	10.1	8.08	7.28	-0.08	9.18	5.94	-0.33
Abs Val [Believed − True Overall Score]	20.4	14.5	20.4	20.3	-0.01	20.6	20.1	-0.04
Believed score higher than true score	0.79	0.41	0.79	0.79	0.02	0.78	0.8	0.07
E. Beliefs about Complementarity								
Believes educ. and performance complementary[c]	0.91	0.29	0.9	0.91	0.02	0.92	0.89	-0.12
Sample Sizes								
Sample Size–HHs	2,634		1,327	1,307		1,417	1,217	
Sample Size–Kids	5,268		2,654	2,614		2,834	2,434	

Notes: Data source is baseline survey. "Std Diff" refers to the standardized difference between groups (i.e., the difference in means between groups divided by the square root of half of the sum of variances of the groups.)
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, pencils, textbooks, supplementary reading books, backpacks, and tutoring expenses.
c. Respondent said that they thought the earnings of a higher-performing child would increase "more" or "much more" than the earnings of a lower-performing child from getting a secondary education.

Table 2: Across the sample, parents invest more years of schooling and attendance in higher-performing children

	Enrollment	ln(Total educ. expenditures)	Attendance rate
	(1)	(2)	(3)
PANEL A. CONTINUOUS VERSIONS			
Treat × Score	0.10**	-0.00051	0.089*
	[0.046]	[0.0025]	[0.053]
Treat × Beliefs	0.016	-0.0034	-0.18***
	[0.061]	[0.0028]	[0.057]
Score	-0.013	0.0035**	0.042
	[0.030]	[0.0017]	[0.036]
Beliefs	-0.014	0.00041	0.11**
	[0.037]	[0.0020]	[0.044]
Treat	-6.08*	0.24	6.73*
	[3.15]	[0.16]	[3.76]
p-val: (Treat × Score) +(Treat × Beliefs)= 0	0.020	0.133	0.129
PANEL B. BINARY VERSIONS			
Treat × Above-median score	3.45**	0.0098	1.46
	[1.47]	[0.079]	[1.58]
Treat × Above-median beliefs	0.77	-0.10	-4.50***
	[1.44]	[0.081]	[1.60]
Above-median score	-1.56	0.086	0.093
	[1.32]	[0.070]	[1.35]
Above-median beliefs	-0.42	0.027	2.81***
	[1.03]	[0.057]	[1.05]
Treat	-2.42*	0.047	1.04
	[1.24]	[0.067]	[1.41]
p-val: (Treat× Above-med.sc.) +(Treat× Above-med.bel)= 0	0.023	0.298	0.116
Control group mean	97.938	7.386	91.142
Observations	1,780	1,703	1,834

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. The regressions control for *Score*, *Beliefs*, *Treat*, grade FE, school FE, between-child score gap, household-average years of parental education, child and parent gender, a parental education proxy used for stratification and the baseline value of the dependent variable, if available. Enrollment is defined as being enrolled in school 1 year after the intervention. Both enrollment and attendance scaled to be out of 100, so enrollment, for example, is equal to 100 if the child is still enrolled and 0 otherwise. In Panel B, *Above-median score* (resp. beliefs) means the child had an above-median baseline overall score (resp. beliefs about child's overall score). *** p<0.01, ** p<0.05, * p<0.1

Table 3: Heterogeneity by parent education: At the intensive margin, relative to less-educated parents, more-educated parents invest more in their higher-performers

	Enrollment	ln(Total educ. expenditures)	Attendance rate
	(1)	(2)	(3)
Panel A. Continuous versions			
Treat × Score × Parent yrs of educ.	-0.0048	0.0017**	0.030*
	[0.0093]	[0.00082]	[0.016]
Treat × Score	0.12	-0.0084*	-0.058
	[0.077]	[0.0044]	[0.089]
Score × Parent yrs of educ.	-0.0010	-0.00053	-0.031***
	[0.0065]	[0.00058]	[0.0099]
Treat × Beliefs × Parent yrs of educ	-0.0090	-0.00060	0.015
	[0.013]	[0.00090]	[0.020]
Treat × Beliefs	0.046	-0.00035	-0.24**
	[0.11]	[0.0049]	[0.095]
Beliefs × Parent yrs of educ.	0.0036	0.00054	0.0069
	[0.0078]	[0.00062]	[0.013]
Panel B. Binary versions			
Treat × Above-med.score × Above-med parent educ	-3.56	0.34**	2.40
	[2.58]	[0.16]	[3.00]
Treat × Above-median score	4.92**	-0.15	-0.036
	[2.29]	[0.11]	[2.16]
Above-median.score × Above-median parent educ	0.55	-0.13	-1.52
	[1.68]	[0.098]	[1.90]
Treat × Above-med.beliefs × Above-med parent educ	-0.46	-0.068	4.69
	[2.43]	[0.16]	[3.01]
Treat × Above-median beliefs	1.03	-0.075	-6.67***
	[2.19]	[0.11]	[2.22]
Above-median.beliefs × Above-median parent educ	0.36	-0.0022	-1.19
	[1.59]	[0.10]	[1.65]
Control group mean:			
Below-median parent educ.	96.684	7.293	89.466
Above-median parent educ.	99.468	7.487	92.884
Observations	1,764	1,688	1,819

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. In Panel A, the regressions control for *Treat*, *Score*, *Beliefs Treat* × *Parent yrs of educ* and *Parent yrs of educ*. In Panel B, the regressions control for *Treat*, *Above-median score*, *Above-median beliefs*, *Treat* × *Above-med.par.educ* and *Above-med.par.educ*. All regressions also control for grade FE, school FE, the between-child score gap, household-average years of parental education, child gender, parent gender, a parental education proxy used for stratification and the baseline value of the dependent variable, if available. Enrollment is defined as being enrolled in school 1 year after the intervention. Both enrollment and attendance scaled to be out of 100, so enrollment, 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 child's parents). *Above-median score* means the child had an above-median baseline overall score. *Above-median beliefs* means the parent had an above-median baseline belief about child's overall score. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Results robust to adding control variables interacted with treatment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Enrollment: Full Sample Results								
Treat × Score	0.10**	0.100**	0.099**	0.10**	0.099**	0.10**	0.10**	0.091**
	[0.046]	[0.046]	[0.046]	[0.046]	[0.046]	[0.046]	[0.047]	[0.044]
Observations	1,780	1,764	1,764	1,764	1,764	1,745	1,745	1,745
Panel B. ln(Educ. Expenditures): Full Sample Results								
Treat × Score	-0.00051	-0.00095	-0.00090	-0.00090	-0.00089	-0.00099	-0.00080	-0.00062
	[0.0025]	[0.0025]	[0.0025]	[0.0025]	[0.0025]	[0.0025]	[0.0025]	[0.0025]
Observations	1,703	1,688	1,688	1,688	1,688	1,669	1,669	1,669
Panel C. Attendance: Full Sample Results								
Treat × Score	0.089*	0.083	0.083	0.083	0.083	0.087	0.087*	0.096*
	[0.053]	[0.052]	[0.052]	[0.053]	[0.053]	[0.053]	[0.053]	[0.052]
Observations	1,834	1,819	1,819	1,819	1,819	1,799	1,799	1,799
Panel D. Enrollment: Parent Education Heterogeneity Results								
Treat × Score × Parent yrs of educ.		-0.0048	-0.0052	-0.0061	-0.0054	-0.0038	-0.0038	-0.0082
		[0.0093]	[0.0093]	[0.0092]	[0.0092]	[0.0093]	[0.0094]	[0.0093]
Observations		1,764	1,764	1,764	1,764	1,745	1,745	1,745
Panel E. ln(Educ. Expenditures): Parent Education Heterogeneity Results								
Treat × Score × Parent yrs of educ.		0.0017**	0.0017**	0.0017**	0.0017**	0.0016**	0.0016**	0.0016*
		[0.00082]	[0.00082]	[0.00082]	[0.00082]	[0.00081]	[0.00082]	[0.00082]
Observations		1,688	1,688	1,688	1,688	1,669	1,669	1,669
Panel F. Attendance: Parent Education Heterogeneity Results								
Treat × Score × Parent yrs of educ.		0.030*	0.030*	0.031*	0.032**	0.031*	0.032**	0.032**
		[0.016]	[0.016]	[0.016]	[0.016]	[0.016]	[0.016]	[0.016]
Observations		1,819	1,819	1,819	1,819	1,799	1,799	1,799
Includes controls interacted with relevant treatment variables[a]								
Treat × Parent yrs of educ		✓	✓	✓	✓	✓	✓	✓
Treat × Parent Female			✓	✓	✓	✓	✓	✓
Treat × Female				✓	✓	✓	✓	✓
Treat × Grade Level					✓	✓	✓	✓
Treat × Beliefs Uncertainty						✓	✓	✓
Treat × ln(Baseline Educ. Expenditures)							✓	✓
Treat × Baseline Attendance								✓

Notes: The dependent variables vary by panel and are listed in the panel labels. 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 control for all variables in Table 2. Enrollment is defined as being enrolled in school 1 year after the intervention. Both enrollment and attendance scaled to be out of 100 (so enrollment, for example, is equal to 100 if the child is still enrolled and 0 otherwise). *** p<0.01, ** p<0.05, * p<0.1

a. For panels A-C (full-sample results), the columns progressively add in the listed control variable interacted with $Treat$. For panels D-F (parent heterogeneity results), the columns progressively add in the listed control variable interacted with $Treat$, $Treat \times ParentYrsOfEduc$ and $ParentYrsOfEduc$. In panels D-F, there is no column 1 because all regressions already control for Parent yrs of educ (the control variable added in column 2) and its interactions with treatment.

APPENDIX (for online publication)

A 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.

Appendix Table A.1: The estimates are robust to excluding the control variables, but precision goes down

	Enrollment	ln(Total educ. expenditures)	Attendance rate
	(1)	(2)	(3)
PANEL A. FULL SAMPLE RESULTS			
Treat × Score	0.10**	-0.00018	0.081
	[0.047]	[0.0026]	[0.054]
Treat × Beliefs	0.022	-0.0037	-0.17***
	[0.062]	[0.0029]	[0.057]
Score	-0.0059	0.0036**	0.059
	[0.030]	[0.0018]	[0.037]
Beliefs	-0.033	0.0029	0.13***
	[0.037]	[0.0020]	[0.043]
Treat	-6.60**	0.26	6.75*
	[3.20]	[0.17]	[3.80]
Observations	1,780	1,703	1,834
PANEL B. PARENT HETEROGENEITY RESULTS			
Treat × Score × Parent yrs of educ.	-0.0047	0.0016*	0.029*
	[0.0097]	[0.00083]	[0.016]
Treat × Score	0.13	-0.0080*	-0.069
	[0.079]	[0.0046]	[0.091]
Score × Parent yrs of educ.	-0.0047	-0.00024	-0.033***
	[0.0064]	[0.00058]	[0.010]
Treat × Beliefs × Parent yrs of educ	-0.011	-0.00062	0.016
	[0.014]	[0.00091]	[0.020]
Treat × Beliefs	0.062	-0.00050	-0.24**
	[0.11]	[0.0051]	[0.095]
Beliefs × Parent yrs of educ.	0.0092	0.00042	0.0037
	[0.0075]	[0.00062]	[0.014]
Observations	1,764	1,688	1,819

Notes: This table shows the main results replicated without including the standard vector of control variables. 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. Enrollment is defined as being enrolled in school 1 year after the intervention. Both enrollment and attendance scaled to be out of 100 (so enrollment, for example, is equal to 100 if the child is still enrolled and 0 otherwise). The only regression controls included are: *Treat*, *Score*, *Beliefs*, *Treat × Parent yrs of educ* and *Parent yrs of educ*. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A.2: The expenditure results are robust to conditioning on enrollment

Dep. var.: ln(Total educ. exp.)	Sample: Enrolled Only (1)	With Enrollment Control (2)
PANEL A. CONTINUOUS VERSIONS		
Treat × Score × Parent Yrs of Educ.	0.0017**	0.0017**
	[0.00082]	[0.00082]
Treat × Score	-0.0086*	-0.0085*
	[0.0045]	[0.0044]
Score × Parent Yrs of Educ.	-0.00057	-0.00052
	[0.00059]	[0.00058]
Treat × Beliefs × Parent Yrs of Educ.	-0.00063	-0.00059
	[0.00091]	[0.00090]
Treat × Beliefs	0.00000011	-0.00030
	[0.0050]	[0.0049]
Beliefs × Parent Yrs of Educ.	0.00061	0.00053
	[0.00063]	[0.00063]
PANEL B. BINARY VERSIONS		
Treat × Above-med Score × Above-med Educ.	0.33**	0.34**
	[0.16]	[0.16]
Treat × Above-med Score	-0.15	-0.16
	[0.11]	[0.11]
Above-med Score × Above-med Educ.	-0.13	-0.13
	[0.098]	[0.098]
Treat × Above-med Beliefs × Above-med Educ.	-0.067	-0.066
	[0.16]	[0.16]
Treat × Above-med Beliefs	-0.075	-0.075
	[0.11]	[0.11]
Above-med Beliefs × Above-med Educ.	0.0029	-0.0012
	[0.10]	[0.10]
Control group mean:		
Below-median Parent Educ.	7.289	7.293
Above-median Parent Educ.	7.487	7.487
Observations	1,673	1,688

Notes: Column 1 is restricted to only children enrolled at the endline survey, and column 2 controls for endline enrollment. Both columns have ln(Total Educational Expenditures) as the dependent variable. 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. In Panel A, the regressions control for *Treat*, *Score*, *Beliefs*, *Treat* × *Parent yrs of educ* and *Parent yrs of educ*. In Panel B, the regressions control for *Treat*, *Above-median score*, *Above-median beliefs*, *Treat* × *Above-med.par.educ* and *Above-med.par.educ*. All regressions also control for grade FE, school FE, the between-child score gap, household-average years of parental education, child gender, parent gender and a parental education proxy used for stratification. In Panel B, *Above-med.par.educ.* means the household was above-median for parent years of education (average years of education across the child's parents). *Above-median score* means the child had an above-median baseline overall score. *Above-median beliefs* means the parent had an above-median baseline belief about child's overall score *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Table A.3: Expenditure results robust to different ways of handling 0's

	ln(exp)	ln(1+exp)	ln(.1min(exp)+exp)	ln(.5min(exp)+exp)	IHS(exp)
	(1)	(2)	(3)	(4)	(5)
Panel A. Continuous Versions					
Treat × Score × Parent yrs of educ.	0.0017**	0.0037***	0.0034***	0.0029***	0.0039***
	[0.00082]	[0.0014]	[0.0012]	[0.0010]	[0.0015]
Treat × Score	-0.0084*	-0.018**	-0.017***	-0.015***	-0.019**
	[0.0044]	[0.0070]	[0.0064]	[0.0055]	[0.0074]
Score × Parent yrs of educ.	-0.00053	-0.0010	-0.00092	-0.00078	-0.0011
	[0.00058]	[0.00075]	[0.00070]	[0.00064]	[0.00078]
Treat × Beliefs × Parent yrs of educ	-0.00060	-0.0029**	-0.0026*	-0.0021*	-0.0031**
	[0.00090]	[0.0015]	[0.0013]	[0.0011]	[0.0016]
Treat × Beliefs	-0.00035	0.017*	0.014*	0.011	0.018*
	[0.0049]	[0.0088]	[0.0079]	[0.0067]	[0.0094]
Beliefs × Parent yrs of educ.	0.00054	0.0010	0.00094	0.00082	0.0011
	[0.00062]	[0.00074]	[0.00071]	[0.00066]	[0.00076]
Panel B. Binary Versions					
Treat × Above-med.score × Above-med.par.educ	0.34**	0.56**	0.52***	0.48***	0.57**
	[0.16]	[0.22]	[0.20]	[0.18]	[0.23]
Treat × Above-median score	-0.15	-0.28	-0.26*	-0.24*	-0.29
	[0.11]	[0.17]	[0.15]	[0.13]	[0.18]
Above-median.score × Above-median.par.educ	-0.13	-0.22*	-0.21*	-0.20*	-0.23*
	[0.098]	[0.13]	[0.12]	[0.11]	[0.13]
Treat × Above-med.beliefs × Above-med.par.educ	-0.068	-0.31	-0.27	-0.21	-0.33
	[0.16]	[0.21]	[0.20]	[0.18]	[0.22]
Treat × Above-median beliefs	-0.075	0.25	0.20	0.13	0.28
	[0.11]	[0.17]	[0.16]	[0.14]	[0.18]
Above-median.beliefs × Above-median.par.educ	-0.0022	0.037	0.027	0.014	0.042
	[0.10]	[0.13]	[0.12]	[0.11]	[0.13]
Control group mean:					
Below-median parent educ.	7.293	7.256	7.264	7.286	7.944
Above-median parent educ.	7.487	7.426	7.437	7.459	8.113
Observations	1,688	1,708	1,708	1,708	1,708

Notes: The columns all have different dependent variables, as indicated by the column labels. In the labels, "exp" represents total educational expenditures, "min(exp)" represents the minimum of total educational expenditures across the sample, and IHS represents the inverse hyperbolic sine transform. 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. The regressions control for *Score*, *Beliefs Treat*, grade FE, school FE, between-child score gap, household-average years of parental education, child and parent gender, a parental education proxy used for stratification and the baseline value of the dependent variable. In Panel B, *Above-median score* (resp. beliefs) means the child had an above-median baseline overall score (resp. beliefs about child's overall score). *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A.4: Treatment effects using "experimental outcomes" suggest that correlated factors do not drive the results

	Experimental Outcomes			
	Math workbook difficulty level	English workbook difficulty level	ln(English textbook WTP) - ln(math textbook WTP)	Lottery tickets
	(1)	(2)	(3)	(4)
Treat × Score	1.64***	1.66***	0.015***	0.048***
	[0.090]	[0.088]	[0.0021]	[0.0056]
Treat × Beliefs	-1.52***	-1.54***	-0.011***	-0.034***
	[0.10]	[0.087]	[0.0021]	[0.0063]
Score	0.14**	0.14**	-0.00035	0.0020
	[0.056]	[0.058]	[0.0015]	[0.0039]
Beliefs	2.24***	2.34***	0.021***	0.069***
	[0.072]	[0.059]	[0.0014]	[0.0045]
Treat	-4.81	-0.21	-0.041	
	[6.79]	[5.18]	[0.044]	
p-val: (Treat × Score) +(Treat × Beliefs)= 0	0.291	0.209	0.147	0.030
Observations	5,233	5,233	5,213	5,250
Score and Beliefs Measure Used	Math	English	Math - English	Overall
Household FE	N	N	N	Y

Notes: Replicated from Table C20 of the Online Appendix of Dizon-Ross (2019).
In all regressions, each observation is a child. The dependent variables are listed at the top of each column, and the specific score and beliefs measures used listed at the bottom.
Columns 1 and 2 have as dependent variables the parents' choice of difficulty for a free subject-specific workbook they were offered for their children. Workbooks came in three difficulty levels – Beginner, Average, and Advanced – and parents could choose a level for each of their children in math (column 1) and English (column 2). Workbook difficulty choices are coded as 0 for beginner, 100 for average, and 200 for advanced. The higher difficulty workbooks were designed for children with higher performance, so the prediction is that the investment function will be upwards sloping (highly positive coefficient on $Treat \times Score$). In addition, since the workbooks are small and free, we expect limited uncertainty effects and so predict that the coefficients on $Treat \times Score$ and $Treat \times Beliefs$ will be nearly equal and opposite, which they nearly are.
The dependent variable for Column 3 is the difference in the parent's log willingness to pay for a remedial English textbook relative to a remedial math textbook. Since the textbooks are remedial, we expect willingness to pay to be decreasing in subject-specific performance. Since the score and beliefs measures are flipped for presentation purposes relative to the dependent variable (math - English instead of English - math), the prediction is again that the coefficient on $Treat \times Score$ will be strongly positive.
The dependent variable for Column 4 is the number of secondary school lottery tickets a parent gave a given child. Each parent was given 9 tickets and allowed to split them between children; the winning child was entered in a lottery for their secondary school fees to be paid. Since this is a within-household decision, the regression includes a household fixed effect. Because higher-performing children have a higher chance of admission to secondary school and since most parents perceive the returns to secondary school conditional on admissions to be higher for higher-performing children, the prediction is that the $Treat \times Score$ coefficient will be positive.
The magnitude of the investment grows across columns, from columns 1 and 2 that are free, to column 3 that involves a monetary investment, and then to column 4 that involves the largest investment. Correspndingly, we might expect that uncertainty effects would also increase across the columns. We find that the magnitude of the $Treat \times Score$ relative to the $Treat \times Beliefs$ coefficient increases across columns, and we can only reject equality in column 4.
Regressions control for school FE, parents' education, the between-child score gap, a parental education proxy used for stratification, child baseline performance, grade fixed effects, treatment, and the main effects of any variable interacted with treatment. Standard errors clustered at the household level. *** p<0.01, ** p<0.05, * p<0.1