NBER WORKING PAPER SERIES

SUBJECTIVE PARENTALBELIEFS:
THEIR MEASUREMENT AND ROLE

Orazio Attanasio
Flávio Cunha
Pamela Jervis

Working Paper 26516
http://www.nber.org/papers/w26516

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2019, Revised August 2024

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 695300 - HKADeC - ERC-2015-AdG/ERC-2015-AdG). Cunha acknowledges funding from NIH grant 1R01HD073221-01A1. Jervis acknowledges funding from the Institute for Research in Market Imperfections and Public Policy MIPP (ICS13 002 ANID) and the Center for Research in Inclusive Education, Chile (SCIA ANID CIE160009). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Subjective Parental Beliefs: Their Measurement and Role
Orazio Attanasio, Flávio Cunha, and Pamela Jervis
NBER Working Paper No. 26516
November 2019, Revised August 2024
JEL No. I21,I24,O1

ABSTRACT

We study the importance of maternal subjective beliefs about the technology of skill formation in determining parental investments in child development. We describe our framework in three steps. First, we discuss the construction of the survey instrument we used to elicit maternal subjective beliefs. Second, we show how to convert the answers to the survey instrument into estimates of maternal subjective beliefs. Finally, we correlate maternal subjective beliefs with maternal investments in child development. We apply our framework to a unique dataset collected as part of an 18-month-long parenting stimulation program in Colombia, whose target population was low-income households with children aged 12 to 24 months at baseline and lasted 18 months. In this program, home visitors paid weekly visits to randomly chosen households to improve mother-child interactions and other maternal behaviors that foster the development of children's cognitive and non-cognitive skills. We show that most mothers believe that the technology of skill formation follows a Cobb- Douglas parameterization, but there is significant heterogeneity in coefficients of investments across mothers. In addition, mothers hold low subjective expectations, meaning they underestimate the returns on their investments. We also find that maternal subjective beliefs predict investments but that the program did not affect maternal subjective beliefs.

Orazio Attanasio
Department of Economics
Yale University
87 Trumbull Street
New Haven, CT 06511
and Institute for Fiscal Studies,
FAIR, BREAD and CEPR
and also NBER
orazio.attanasio@yale.edu

Flávio Cunha Department
of Economics Rice
University
P.O. Box 1892
Houston, TX 77251-1892
and NBER
flaviocunha@rice.edu

Pamela Jervis
Department of Industrial Engineering
University of Chile
Beauchef 851
Santiago, Chile
and Institute for Fiscal Studies
pjervisr@uchile.cl

1 Introduction

A large body of research shows that differences in cognitive and non-cognitive skills across socio-economic groups appear early in the lives of children and remain and even increase once these children start school. This evidence includes studies from developed countries, such as Cunha et al. (2006) in the US, and developing countries, such as Rubio-Codina et al. (2015) in Colombia. Lags accumulated in the first three years are substantial and have long-term consequences (see, for example, Currie and Thomas (1999), Behrman et al. (2009) and Heckman et al. (2010)).

There is also mounting evidence that development in the first 3 years of life is malleable and, therefore, salient for interventions. Many effective interventions have aimed at changing parenting practices to increase stimulation. The celebrated *Reach Up* program has been proven to be effective in the short and long run (see for instance Grantham-McGregor et al. (1991) for some of the short-run impacts and Gertler et al. (2014) for the long-run ones). Similar evidence, from different contexts, is accumulating.[1]

To better understand the impacts of these interventions and possibly design new ones, it is essential to understand what makes them work and, in particular, what drives parental behavior. Attanasio et al. (2014), for instance, show that a large-scale adaptation of the Jamaica program in Colombia increased cognitive and receptive language skills by 26% and 22% of a standard deviation, respectively. Attanasio et al. (2020) then argue that increased parental investments largely explain the intervention's short-run impacts. In particular, they report that both time spent interacting with children and the number of didactic materials at home increased by 30% and 23% of a standard deviation. Furthermore, through a careful mediation analysis that considers the endogeneity of investments, they show that these increases can explain most of the impacts. The salient question, then, is: why do parents targeted by these interventions increase and improve parental investment?

Home visitation programs and, more generally, stimulation interventions may influence parental investments through at least three channels. First, these inter-

[1] In the United States, the literature reports considerable impacts of home visitation programs when they were implemented in controlled, small-scale settings (Baker and Piotrkowski (1996), Caldera et al. (2007), Olds et al. (1998), Eckenrode et al. (2010), Drotar et al. (2009)). More recently, Michalopoulos et al. (2019) find more modest impacts of programs when implemented at a larger scale. Their study estimates that home visitation programs improve the quality of the home environment by 9% of a standard deviation but that these programs do not impact child development as measured by expressive language and socio-emotional skills.

ventions may shape parental behavior directly, presenting and encouraging parent-child interactions conducive to positive development. According to this channel, home visitation programs improve child development by inspiring parents to adopt technologies of skill formation that are more efficient in promoting child development. Second, some programs offer didactic materials that might increase the psychic benefits that parents experience when interacting with their children. Third, home visitors may provide, directly or indirectly, information to parents about the importance of early parental investments for child development.

While the first two interpretations of home visitations' successes explain them appealing to provision of additional *investment tools*, the third interpretation invokes a change in the *perceived effectiveness* of parental investment. According to this channel, some parents may choose low levels of parental investments because they expect returns to this investment to be low, which is consistent with anthropological and sociological studies in the US (Lareau (2003) and Putnam (2015)). In this case, while all parents might care equally about the development and well-being of their children, some parents might not be aware of the importance that some specific activities, such as talking and interacting in specific ways with a small child, might have for their development. And yet, the findings in developmental science indicate that early stimulation is essential for subsequent growth and that exposure to language and meaningful interactions drive later outcomes. Under this hypothesis, some parents, pursuing what Lareau (2003) defines "natural growth," could be making suboptimal parental investment choices.

The standard practice in economics to investigate what drives parental investment is to formulate and estimate (dynamic) optimization problems where parental welfare depends both on their consumption and on children's outcomes and where parents "know" the functional form and parameters of the technology of skill formation (Del Boca et al. (2013)). Within such models, parental investment is driven by the nature of the technology of skill formation of human capital, financial resources, the cost of investment in children, credit constraints, and by how much parents care for their children. Because these models assume parents "know" the technology of skill formation, these models are *ill-suited* to understand the importance of parental beliefs about the technology of skill formation in determining parental investments in children.

We can easily extend the theory to allow for misinformation about the technology of skill formation. Empirically, however, it is hard to provide credible estimates about the importance of misinformation because only under exceptional circumstances can one separately identify preferences from beliefs (Manski, 2004).

One possibility to study parental behavior, without assuming that parents *know* the nature of the process of child development or the technology of skill formation of human capital, is to elicit directly parental beliefs about the process of child development and, in particular, about the usefulness of parental stimulation and investment and how these inputs interact with child development before investing.

In this paper, we elicit maternal subjective beliefs in a sample of poor mothers in Colombia. We show how to convert the answers to a specific set of questions into estimates of expected rates of returns on specific investments and then relate these estimates to actual parental investment behavior.

We assume that mothers hold beliefs about the process of child development and how it relates to a set of inputs, which we call the *technology of skill formation*. While we consider likely determinants of child development, such as parental investment and an initial level of development, Still, we allow each mother to have individual beliefs about the parameters of the technology of skill formation. In particular, in our framework, we allow mothers to have the right or the wrong expectation about the productivity of specific inputs.

For each mother in our sample, we recover the expectation about the returns to parental investments. Furthermore we can identify and estimate the technology of skill formation that describes the mothers' beliefs expectations about the child development process (given inputs). This estimation allows us to compare the *perceived* productivity of the input considered to that estimate from objective data on child development, parental investment, and other controls. Additionally, we can investigate if there is heterogeneity in expectations about the returns to early investments, what variables explain this heterogeneity in expected returns, if this heterogeneity in expectations about returns predicts heterogeneity in parental investments, and if the parenting stimulation program influenced the heterogeneity in expectations about returns to investment.

Our work is closely related to Cunha et al. (2013), who elicit maternal beliefs about the technology of skill formation from disadvantaged mothers in Philadelphia, USA. As in that paper, we create scenarios of different inputs and ask mothers to report expectations about child development for each scenario of the inputs. However, we argue that our methods are more appropriate for researchers interested in eliciting such beliefs from populations with limited literacy and cognitive skills as our beliefs questions are far more straightforward than those used by Cunha et al. (2013). Our new beliefs elicitation survey instrument, in turn, requires us to develop new methods to analyze our data and to map maternal answers to expectations about returns to parental investment.

Our paper makes four innovative contributions. First, we propose a new belief elicitation tool. Much work went into designing, and validating a measurement tool that could be implemented and used easily in a large-scale survey to elicit subjective beliefs about the process of child development that could be compared to data on the actual process from the same population. From a measurement point of view, our approach assumes that parents believe that child development over a given period depends on a child's developmental status and parental investment, among other variables. We, therefore, ask mothers to relate different scenarios of initial development and parental investment to certain developmental outcomes for a hypothetical child. In particular, we consider *high* and *low* levels of initial development and *high* and *low* levels of parental investment. With these data, we can derive straightforward measures of subjective expectation of returns to parental investment under two different levels of initial development.

Second we find that maternal subjective beliefs correlate significantly with actual parental investments. As mentioned above, we estimate returns to investments for two initial development scenarios; the correlation of parental investments is stronger with the expected return under "high" initial development.

Third, putting more structure on the data, we estimate the parameters of a *subjective technology of skill formation* for each mother in the sample. As we can estimate the technology of skill formation objectively, with data on child development, we can compare objective and subjective processes and quantify whether maternal subjective beliefs are biased (or not). We find that maternal subjective beliefs are downward biased and that most mothers have expectations of returns to investments that are too low.

Fourth, we assess whether or not the randomly assigned parenting stimulation program affected maternal subjective beliefs or not. Our result is negative. The distribution of maternal subjective beliefs, elicited two years after the end of the intervention, whose impacts were measured in Attanasio et al. (2014), is the same for control and treatment groups. This result is consistent with Andrew et al. (2018) find ng no differences in parental investments (or in child development) between the control and treatment groups two years after the end of the program. Our study suggests that the lack of permanent effects on parental investments could be because the program did not change permanently maternal subjective beliefs.

The rest of the paper is organized as follows. Section 2 presents the methodology to elicit maternal subjective beliefs. Section 3 describes the context in which our study was developed and describes the data we use. Section 4 presents some evidence on the perceived returns. Section 5 discusses how to estimate subjective and objective *production functions* of human capital. In section 6, we present our empirical results of this estimation exercise. Section 7 conclusion closes the paper. Section A contains appendices that provide further details about our study.

2 Measuring child development, parental investment, beliefs.

When modelling the process of child development and its drivers, researchers typically work with latent variables, representing some abstract constructs, such as different dimensions of child development or what is often referred to as 'parental investment'. A useful approach has been to relate formally the latent factors of interest to available measures and obtain from the latter estimates of the former that one can use in empirical analysis. Such relationships, or *measurement systems*, under a set of assumptions and if enough measures satisfying such assumptions are available, can be estimated. Such an approach then efficiently summarizes the available measures tob obtain estimates of the relevant factors that can be used to analyse structural models linking them (e.g., see Cunha et al., 2010).

Based on this approach to measurement, we design questions to elicit parental beliefs about the process of child development. As we formalise below, we have in mind a process that relates the initial level of development, parental investment, and possibly other variables to subsequent development. In addition, we conjecture that parents have their own beliefs about such a process, which we elicit through a series of questions presenting parents with a number of hypothetical scenarios.

In this section, we describe in detail both the approach we use to estimate the latent factors of interest and the methods we use to construct hypothetical scenarios to elicit parental beliefs. We will start with an intuitive description of our general approach to eliciting parental beliefs and using these data. We then provide details of the measurement systems we use to estimate the latent factors representing child development and parental investment from the available data. These estimates are later used to estimate an *objective production function* of child development. Here, we also use them to construct hypothetical scenarios. We then describe how the hypothetical scenarios are used to elicit parental beliefs. That subsection is followed by one where we discuss the problem of establishing comparable metrics between the latent factors that enter our models of child development and those that determine parental beliefs.

2.1 An intuitive summary

To elicit individual beliefs about the process of child development, we present parents (typically mothers) with scenarios characterised by different levels of investment and baseline development and ask them how much they expect a *hypothetical* child to develop under these different scenarios. The answers to these questions will then enable us to compute the perceived expected rate of return of specific investment strategies under varying levels of baseline child development.

In the conceptual framework we use to model the process of child development, this construct is represented by a latent variable $H_{i,a}$, which evolves as the child ages. Such a variable, for child i aged a, $H_{i,a}$, depends on the initial development $a-1$, $H_{i,a-1}$, on parental investment, X_i, and other observed and unobserved environmental factors, represented by a vector Z_i and a variable ϵ_i, respectively. Parental investment X_i can also be seen as another unobserved latent factor that enters our model, which the following equation represents:

$$H_{i,a} = F_a(H_{i,a-1}, X_i, Z_i, \epsilon_i) \tag{1}$$

The variable ϵ_i is unobserved by the researchers and, possibly, by the parents. Henceforth in this paper, we refer to $H_{i,a}$ and $H_{i,a-1}$ as subsequent and initial development (or human capital), respectively.

In eliciting subjective parental beliefs about the process of child development, we posit that mothers believe that, as in equation (1), child development at age a depends on child development in period $a-1$, parental investments, and other variables. However, we do not assume that mothers know the "true" process of child development. Instead, if we denote with Ω_i the information available to mother i, we assume that the expected development by mother i, given a set of inputs, is given by:

$$E[H_{i,a}|\Omega_i] = E[\tilde{F}_a(H_{i,a-1}, X_i, Z_i, \epsilon_i)|\Omega_i] \tag{2}$$

In what follows, we assume that the functional forms of F_a and $\tilde{F}_a$ are the same but that their parameters might differ.

To elicit beliefs about the developmental process, we present mothers with several "scenarios" - that is, pairs of parental investment X_i and initial development $H_{i,a-1}$. We explicitly ask about a *hypothetical* child of a certain age and not their own child or any other specific child. We chose this approach partly to avoid that they would give answers too related to their own experience and, importantly, to allow for the possibility that variables other than parental investment and baseline development, that is Z_i and ϵ_i in equation (1), enter the process of development. The assumption is that the reference to a *hypothetical child* would induce respondents to *average* across these variables. Then, for each scenario, we ask mothers to report their expected child developmental outcomes. We therefore obtain the expected level of development $\hat{H}_a$ that is induced by a specific scenario on X and H_{a-1}, at average values of Z_i and ϵ_i, as perceived by respondent i.

As we discuss below, we design four different scenarios for H_{a-1} and X, corresponding to "low" and "high" values of H_{a-1} and X. Mothers are then asked about the age at which the hypothetical child is able to achieve certain tasks under each alternative scenario. We assume that respondents' answers consider an average of the Z and ϵ variables that enter equation (1). In analyzing these data, when we compute returns to parental investment under different baseline development scenarios, we also need to assume that the average values of Z and ϵ do not change as we move from one hypothetical scenario to another.

A first big challenge in the design of questions to elicit beliefs is how to create the hypothetical scenarios representing different levels of X_i and H_{a-1} and what tasks can represent H_a. In this respect, it is important to identify situations or variables that are *salient* for the parents and that are related to the relevant latent factors. We assume that parents relate a number of observable variables (such as language development for child development and toys for parental investment) to the latent factors of interest.

Therefore, our first step, described in detail in section 2.2, is to define the connections between the latent factors that populate our model and the available measures, that is, a set of *measurement systems*. We assume that parents use similar measurement systems, that is they relate some (and possibly) more observable variables to the latent factors of interest.

The measurement systems we estimate play two roles in our exercise. On the one hand, these systems aggregate the individual items of child development and parental investment measures into continuous scores that we later use to estimate the *objective* process of child development. On the other hand, we use them to identify measures of development and parental investments, among thise that are well understood by the parents, that relate well to the relevant latent factors. The implicit assumption is that mothers use the same mapping from the relevant latent variables (child development and investment) to observable variables that are salient indicators of child development and parental investment.

As we want to compare the estimates of the process of children's cognitive development with parental beliefs, it is crucial that the developmental metric used in the objective data is comparable to that used with the (subjective) beliefs data. We discuss how these issues are addressed in section 2.3 below.

In section 2.4, we describe how we use the estimated measurement systems to construct the scenarios. After designing the scenarios, which define the inputs in the production function, we need to use them to create specific questions to elicit respondents' beliefs. We discuss this step in section 2.5. In addition, we describe how to convert, using again estimates from the measurement systems, individual answers to the beliefs questions into values for the relevant factors. This information is key, as our estimates of parental beliefs derive from these variables.

2.2 Mapping child development and parental investment into observables

As children reach different degrees of development, they become capable of performing certain tasks, such as understanding and expressing certain words. Therefore, the ability to perform a certain task is a marker of child development. Similar considerations apply to the parental investment latent factor. We identify a number of observable variables, such the availability of toys and the amount of time spent by parents with the child in certain activities, that are assumed to be markers of parental investment. We use these variables to estimate measurement systems for child development at different ages and for parental investment.

As mentioned above, the estimation of these measurement systems provides estimates of the latent factors of interest that can be used to analyse the structural model we consider. Furthermore, as we assume that mothers also use the same markers as indicators of child development and investment, we can use these results for formulating the beliefs questions. Although the specific system that we use to design the scenarios for the belief questions is slightly different from the richer one we use to estimate the latent factors used in estimating the *true* production function, the idea is substantially very similar: we assume that the child development process, in reality or as perceived by mothers, is related to a set of observable variables. To identify which variables to use to define scenarios, we choose the most informative and salient for parents. We note that to get estimates of the relevant latent factors, we do not need *all relevant measures*, as a consistent estimate (with some measurement error) can be obtained from a subset of them.

In order to relate observed variables to the abstract constructs that enter equations such as (1), we use relatively standard latent variable models, which, in some cases, we extend to accommodate the nature of the data available and use all the information we have efficiently. In particular, we assume that, corresponding to each of the three latent factors $\ln H_{i,a}$, $\ln H_{i,a-1}$ and $\ln X_i$, we have a number of observable indicators whose values are affected by one of the three latent variables and some measurement error. For child development, we use a wide range of variables, including tests of children's development and maternal reports. For parental investment, we use information on materials and activities.

To be specific, consider the latent factor $\theta_{i,j} \in \{H_{i,a-1}, X_i, H_{i,a}\}$. The index i represents the child and the index j represents the specific factor considered: child development at a $(H_{i,a})$ or parental investment (X_i). We express the relationship between measure k, m_i^{kj}, and latent factors, $\theta_{i,j}$, as follows:

$$m_i^{kj} = g_{j,k}(\theta_{i,j}, z_{i,j}, \epsilon_{i,j}^k) \tag{3}$$

where $\epsilon_{i,j}^k$ represents measurement error and $z_{i,t}$ are observable variables that can potentially enter the relationship between child development and its markers in equation (3). The superscript k represents the different measures we have for each one of the latent factors. We discuss the specific measures we use in section 3.

Having estimated a model like (3) with a vast range of observed measures, it is then possible to use the estimated parameters to estimate the unobservable latent factors. Note that estimates of the latent factors can be obtained even when considering only a subset of the measurement variables used to estimate the more general model. Moreover, and importantly, it is possible to identify available measures that are particularly informative about the various latent factors. These two observations are key for the construction of the scenarios and for the use of the subjective beliefs data.

In our application, we consider 'dedicated' systems, where a measurement loads on a single factor, that is child development at a given age $a-1$, $(H_{i,a-1})$, and at the subsequent age a, $(H_{i,a})$, or parental investment (X_i). Moreover, each latent factor $\theta_{i,j}$ can affect the available measures on its own or in combination with some observable variables $z_{i,t}$ (such as gender).

For the available continuous measures, we use a log-linear functional form for the measurement systems. In particular, we assume that the factor j registered by measurement variable k, for child i or age t is determined by a single index:

$$\begin{aligned} m_{i,a}^{kj} = \;& \gamma_{a,0}^{kj} + \gamma_{a,1}^{\prime kj} z_{i,a} + \gamma_{a,2}^{kj} \theta_{i,a}^{j} + \xi_{i,a}^{kj}; \quad j = 1, 2, 3; \\ & \{\theta_{i,a}^{1}, \theta_{i,a}^{2}, \theta_{i,a}^{3}\} = \{\ln H_{i,a-1}, \ln H_{i,a}, \ln X_i\}. \end{aligned} \tag{4}$$

where $z_{i,a}$ is a vector of observable variables, which are allowed to shift the index $m_{i,t}^{k}$. For binomial or discrete variables, we use the generalised IRT described in Appendix A.1. We assume that the joint distribution of the unobservable factors θ_i^j is a mixture of log-normals, while the measurement error $\xi_{i,t}^{kj}$ is a normal random variable, independent across different measures k. We estimate all the parameters (subject to location and scaling parameters discussed in Section 2.3) and recover the distribution of the latent factors.

2.3 Location and Scale of Child Development

As our conceptual framework deals with unobserved latent factors representing child development (and parental investment) for which we have a number of measures or markers, we need to establish a cardinal metric for such factors. Choosing a consistent and comparable metric is particularly important as we aim to compare objective data and relationships on the process of child development to those perceived by mothers. This section discusses the approach we use to develop a comparable child development metric.

In our sample, we have several measures of child development, capturing different dimensions of child development. We construct a cardinal metric by translating raw test scores into *developmental age* scores. We do so by running the following regression using the data on the first two waves of the survey, which were collected three years before the beliefs questions on the same respondents.

$$\ln a_i = \pi_0^j + \pi_1^j Y_i^j + u_i^j \tag{5}$$

where, as before, a_i is the age in months of child i, and Y_i^j is the score obtained by child i in dimension j. The specific scores are described in section 3.

With the estimates of the coefficients in equation (5), we can then construct estimates of the *developmental age* of a generic child i for developmental dimension j, who scores Y_i^j

$$\ln devage_i^j = \hat{\pi}_0^j + \hat{\pi}_1^j \times Y_i^j \tag{6}$$

Notice that equation (6) converts the raw test score for dimension j into an estimated log of the age-equivalent score. Such a measure has a location and a scale that we can compare in the objective and subjective data. In addition, it is invariant to monotonic transformations to the Y_i^j score.

In our analysis, we then use each sub-scale available, converted in terms of *developmental age*, as in equation (6), to estimate a measurement system for child development. We use the log of age-equivalent of an expressive language score as an anchor for the other measures of child development, meaning that we normalize the intercept and slope of the measurement system in equation (4) to zero and unity respectively for that variable. As a result, our factors of child development $\ln H_{i,a-1}$ and $\ln H_{i,a}$ have both the location and the scale of the (log) age-equivalent expressive language score and, thus, have cardinality. This property is required in the estimation of the technology of skill formation.

In contrast, the provision of cardinal metric for investments is easy because one of the measurement variables for investments is the amount of time of interaction between parents and children (as reported by the parent), which is cardinal.

This approach contrasts with that of Cunha et al. (2010), who obtain cardinality by anchoring raw scores onto adult outcomes. The advantage of our approach is that we can use the same cardinality to anchor not only child development scores but also the maternal answers to questions designed to elicit subjective beliefs about the process of child development without additional assumptions. Indeed, as we show below, we elicit maternal subjective beliefs by asking mothers to report the age they *believe* a hypothetical child will reach certain developmental milestones for each one of the scenarios of child development at the beginning of the period and parental investments.

2.4 Hypothetical Scenarios

To elicit subjective beliefs about the process of child development, we have to define the scenarios to be presented to parents to represent different points in the domain of the production function of child development in equation (1). Analogously, we also want to construct a description of potential outcomes of the same function. To design such scenarios and potential outcomes, we use the measurement systems in equation (3), as specified in equation (4) and estimated on the data from the first two waves of the survey.

Our goal is to provide parents with concrete verbal descriptions of scenarios for beginning-of-the-period child development $(a-1)$ and parental investments, which are the inputs of the technology of skill formation. Additionally, we aim to present more salient, verbal descriptions of child development in the subsequent period or age, a. Having estimated the measurement systems that map child development (at different ages) and parental investment to the relevant latent factors, we identify which items (of those available) are more *salient* for these latent factors.

We note that estimates of the measurement system not only provide a way to choose a set of items to describe the scenarios but can also be used to compute the values of initial level of development $H_{i,a-1}$ and investment X_i corresponding to that that scenario, as well as the values of the outcomes $H_{i,a}$ corresponding to the answers given by the respondents to the beliefs questions. As mentioned above, the metric (in terms of location and scale) used for the $H_{i,a-1}$ and $H_{i,a}$ is comparable to that used in the measurement system on the data on actual child development (developmental age) discussed in section 2.3.[2]

Hypothetical Scenarios for Current and Future Child Development. As we explain in Section 3, we measure child development with various measures, which we use to estimate the measurement systems for the child development factors at different ages. Given estimates of the parameters of such systems, we identify the most informative items among the available measures. In a system such as that in (4), those will be the items with relatively high loading factors $\gamma^{kj}_{a-1,2}$, as they are more effective in reflecting the influence of the unobservable latent factor.

Furthermore, we choose items with different levels of $\gamma^{kj}_{a-1,0}$ to identify different levels of development. For example, children with developmental delays can only answer easy items correctly, that is, the items in which $\gamma^{kj}_{a-1,0}$ is low. In contrast, children advanced in development can also answer correctly items in which $\gamma^{kj}_{a-1,0}$ is high.

We use this procedure to select the items to construct the scenarios that define different initial development levels. In particular, parents are faced with a hypo-

[2]We provide further details in Appendix A.3.1.

thetical nine-month-old child who can say a specific set of words. The alternative scenario is constructed by changing the group of words the hypothetical child can say. For instance, when initial development is low, the nine-month-old child can say (or understand) easy words (i.e., words with low $\gamma^{kj}_{a-1,0}$'s), but not difficult words (i.e., words with high $\gamma^{kj}_{a-1,0}$'s). In contrast, when initial development is high, the nine-month-old child can say both set of words. Given the set of words that describe a scenario and the estimates of the measurement system, it is possible to get an estimate of the level of development of a child described by a certain scenario. That is, a given scenario defines a certain level of $H_{i,a-1}$. Again, the implicit assumption is that the respondents use the same measurement system to relate the level of child development to the items we use to define the scenario.

We use an analogous approach to describe the level of child development development at age a, $H_{a,i}$, obtained under different scenarios for X_i and $H_{a-1,i}$. In particular, using the measurement system estimated for age a, we identify items that are particularly informative about the level of development at age a, $H_{a,i}$, that is words with high loading factors,$\gamma^{kj}_{a,2}$, and ask at what age the hypothetical child considered in a scenario will be able to say such words. We can then use the parameters of the measurement system and the appropriate metric to convert the answers given by respondents into an index of development which is comparable to that used to estimate the objective production function. As we discuss below, we repeat these questions for groups of words with different levels of difficulty $\gamma^{kj}_{a,0}$.

Hypothetical Scenarios for Parental Investments. In addition to using estimates of the measurement system for child development to define the scenarios for the level of initial development, we also use them to define the possible outcomes for future child development. Similarly to constructing scenarios for child development at age $a-1$, we use the estimated *measurement system* for parental investment to identify items that seem to be particularly informative for this unobserved latent factor. As before, given a set of informative items, we can construct estimates of the unobserved investment latent factor and then use a specific choice of items to build scenarios which correspond to *low* and *high* levels of parental investment. Given the observed distribution of parental investment in our data, we can infer to which percentiles of this distribution a particular scenario correspond. Again, the implicit assumption is that parents use the same measurement system to relate the items described in the scenario to the parental investment factor.

The scenarios for parental investment were presented to the mothers in laminated illustrated cards so that the verbal description could be reinforced with visual stimulation. Figure A.2 in Appendix A.2 shows the vignettes used during the beliefs elicitation survey instrument, which are assumed to represent the item chosen to define a scenario.

2.5 Beliefs Elicitation Survey Instrument

The instrument to elicit maternal beliefs provides mothers several hypothetical scenarios $s = (s_h, s_x)$, where s_h is a scenario for H_{a-1} for the level of initial development and s_x is a scenario for parental investment, X. The scenario pairs s, constructed following the protocol in section 2.4, take value in the set $S = \{(H^L_{a-1}, X^L), (H^L_{a-1}, X^H), (H^H_{a-1}, X^L), (H^H_{a-1}, X^H)\}$. Given these pairs, the beliefs elicitation survey instrument asks mothers to report the expected level of child development at the end of the period, $E\left[H_{i,a}|\Omega_i, s\right]$. Specifically, for each of the four scenarios we constructed, we asked the mother to report the age the hypothetical child would start saying three sets of words. These sets, in turn, correspond to different difficulty levels: "easy," "medium" and "hard" as determined by the value of $\gamma^{kj}_{a,0}$ in equation (4). This protocol will yield, for each scenario about the level of beginning-of-the-period development and parental investment, multiple measures of the expected level of end-of-the-period child development $E\left[H_{i,a}|\,\Omega_i, s\right]$. These multiple measures will allow us to investigate and address measurement errors in maternal responses.

To record their answers, mothers used wooden tablets that contained several strings with beads and markings of different ages (from 9 to 48 months) at the top. Figure A.3 in Appendix A.2 shows the scenarios $s \in S$ used in the tablets for which the mothers reported the age the hypothetical child would start saying each set of words. The tablet has a string with a bead for each set of words and a scenario. The mother was asked to put the bead at the age at which the hypothetical child could say a specific set of words under a given scenario. At the end of the exercise, each mother was presented with two wooden tablets (left and right diagrams from Figure A.3) with the 12 strings and beads and was asked whether she would want to revise any of the questions.

Consistency (in that easier words - or high investment - should correspond to earlier ages) was not forced. However, we trained the mothers to use the wooden tablets with some practice questions before asking the elicitation questions. They were asked at what age a hypothetical baby (aged six months) would start to crawl, walk and run using two different scenarios in terms of nutrition (low and high). During these practice questions, the interviewer would point out inconsistencies if, for instance, the mother would indicate that the hypothetical baby would start to run before starting to crawl or that a malnourished child (low nutrition) would run before a well-fed one. The point of this exercise was to familiarise the mothers with an instrument that is not standard in fieldwork, especially with a population with low levels of education. Appendix A.2 describes in detail this procedure, which we chose after extensive piloting.

2.6 Returns to parental investment

As mentioned above, the questionnaire that elicits individual beliefs about child development asks the respondents at which age a hypothetical child will be able to say three sets of words under different scenarios of parental investments and initial child development. Given that, a simple estimator of the subjective expected returns to parental investment, conditional on the level of the beginning-of-the-period child development, is the average gain in months between the high and low investment scenarios across the three sets of words considered as outcomes. A child exposed to high levels of parental investment will probably be able to say certain words earlier than a child exposed to a lower level. The subjective beliefs questions, therefore, can yield two sets of such measures of expected return, one for each level of initial development. The fact that the questions are posed about three possible three sets of words (hard, medium, and easy) provides several measures of such returns. The answers can then be scaled and combined using the developmental age discussed in Section 2.3 and express the return in terms of gains in developmental age.

The approach we propose is to express subjective expected returns to parental investment in terms of an (error-ridden) measure of the latent factor with which we represent child development and that, in what follows, we use to inform and influence parental investment in a structural model of parental behavior. To put these ideas more formally, in the model we discuss below, we represent the subjective maternal expectation about child development at age a, $H_{i,a}$, as the subjective expectation of such factor, conditional on individual information and its determinants, considered in equation (1), X_i and $H_{i,a-1}$, $E[H_{i,a}|\Omega_i, H_{i,a-1}, X_i]$.

If we denote with $\hat{H}^q_{i,a}(H_{i,a-1}, X_i)$ a measure of the subjective belief about the expected development held by mother i given inputs $H_{i,a-1}$ and X_i, the answers to our questions will provide the following quantities:[3]

$$\hat{H}^{q}_{i,a}(H^{k,j}_{i,a-1}, X^{k,j}_{i}) = E[H_{i,a}|\Omega_i, H^{k,j}_{i,a-1}, X^{k,j}_{i}] + \nu^{q,k,j}_{i}; \quad j,k = L,H; \quad q = e,m,h \tag{7}$$

where $\nu^{q,k,j}_{i}$ is a measurement error, $H^{k,j}_{i,a-1}, X^{k,j}_{i}$ represent the arguments of the production function in equation 1, and the superscript q refers to the different sets of words. The subscript i refers to a mother in our sample.

Each of the measures obtained from the answers to the subjective beliefs questions and the measurement system utilised to design the scenarios, provides an estimate, affected by measurement error, of the subjective expectations about factor $H_{i,a}$ for the hypothetical child in each hypothetical scenario. Given these estimates, we can obtain, for low and high level of initial development, three estimates of the subjective returns to parental investment, each affected by some measurement error:

$$r^{i,q}_{a,H_k} = \hat{H}^{iq}_{a}(H^{k}_{a-1}, X^{H}) - \hat{H}^{iq}_{a}(H^{k}_{a-1}, X^{L}); \quad k = H,L, \quad q = e,m,h \tag{8}$$

After computing the specific rates of returns implied by different scenarios, the next step is to add some additional structure to the our model and assume that mothers share the specification of the production function we specify, but not necessarily with the same parameters. We discuss these two final step in Section 5.

3 The Data: Origin and Content

This section presents the basic data we use in our study. We start discussing the data origin and then describe the measures of child development and parental investment contained in the data.

3.1 The Evaluation of a Parenting Stimulation Program

As mentioned above, the data were collected to evaluate the impact of a parenting stimulation program to foster the development of young children living in low-income families in Colombia. The program's basic structure was guided by the Jamaica study of early years parenting stimulation by Sally Grantham-McGregor

[3] As discussed above the answers to the beliefs questions are expressed in terms of developmental age, as mentioned in section 2.3 and detailed in Appendix A.3.1, using developmental age as a metric that makes it comparable to the measures used to estimate the objective production function.

(see Grantham-McGregor et al. (1991)). A Randomised Controlled Trial (RCT) in Colombia was designed to evaluate the effect of two different interventions and their combination, using a 2×2 design. The first treatment was a parenting stimulation program delivered through weekly home visits to stimulate the child and involve the caregiver and child in a number of structured visits, while the second was a nutritional supplementation.

The parenting stimulation program employed community women and used the infrastructure of an existing welfare program to deliver the stimulation component to test a scalable version of the program. The Jamaica curriculum (*Reach-Up*) was adapted to the Colombian context. The original curriculum promoted child development in an integrated manner (i.e., language, cognitive, motor and socio-emotional skills). It did so by encouraging caregivers to explore daily routine activities to teach their children. The curriculum was based on picture books to stimulate conversation, puzzles, cubes/blocks, toys from recycled materials, language games, and songs.

The evaluation sample included 1,429 children aged 12-24 months at baseline living in 96 semi-urban towns. The randomisation, over the four groups (Stimulation, Micronutrient Supplementation, Stimulation plus Supplementation, and Control) was done across towns to avoid contamination of the control group.

The parenting stimulation program significantly impacted various outcomes, which are discussed in Attanasio et al. (2014). This study used two surveys on the children in the evaluation sample and their primary caregiver: the baseline survey collected before the program started in 2009-2010 (children aged 12-24 months) and a first follow-up survey collected 18 months after the baseline, at the end of the program, in the last few months of 2011. In the baseline and first follow-up surveys, data were collected to measure children's cognitive, language, and socio-emotional skills as well as height, weight, hemoglobin, and morbidity. Parental investments were measured with data on food intakes, childcare arrangements, didactic materials, parent-child interactions, and time use. Finally, data on primary caregivers and other household members were collected using a general household survey. These included data on socioeconomic status, education, labour supply, time use, reproductive history, health conditions, depression, knowledge of parenting, parenting practices, and the home environment, among others.

Attanasio et al. (2020) show an increase in parental investment essentially explained these early impacts on child development. One possibility to justify such an increase, caused by a program that did not provide parents with any resources, is that the program changed parental beliefs about the process of child development. Data on parental beliefs and their relation to parental investment can be instrumental in investigating such a hypothesis. More generally, we can relate data on parental beliefs to parental investment, both as a way to validate our novel measures and to investigate the role that beliefs play in investment choices.

About two years after the end of the program, the children and their families that participated in the study were contacted to collect information on the medium-run impacts of the program, which are described in Andrew et al. (2018).[4] The second follow-up, which happened in the fall of 2013, included, among other things, a survey instrument to elicit parental beliefs about the process of child development, designed along the lines we describe in Sections 2.4 and 2.5. The survey instruments were administered to the primary caregivers, mostly mothers. In what follows, we refer to parental subjective beliefs as maternal subjective beliefs.

3.2 Measures of Child Development and Parental Investments

In this subsection, we summarize the measures of child development and parental investments collected in the evaluation study's baseline and first follow-up surveys. Details on these data are essential for this paper as we used them to design the beliefs elicitation survey instrument we describe in Section 2.5 and to estimate the parameters of the objective process of child development, which we then compare to the beliefs data.

To assess child development, we used Bayley Scales of Infant and Toddler Development, Edition III (BSI-III, Bayley, 2006) and the short versions of the MacArthur-Bates Communicative Development Inventory (MLI, Jackson-Maldonado et al., 2012). We use these instruments from the baseline and the first follow-up surveys to estimate (the log of) beginning- and end-of-the-period child development, or $\ln H_{i,0}$ and $\ln H_{i,1}$, respectively.

The BSI-III is considered the gold-standard assessment of child development for children under 42 months. It measures cognition, expressive and receptive language, and fine and gross motor skills. For our analysis, we use the subscales that relate to expressive and receptive language and cognition.

The MLI has three versions which depend on the child's age. The MLI-I is ap-

[4] Andrew et al. (2018) report a fade-out of the program's effects on measures of child development and parental investment.

propriate for children aged between 8 to 18 months old. For each of 104 words prompted by the interviewer, the parent reports if the child "understands and says the word," "understands, but does not say the word," or "neither understands nor says the word." The MLI-II is appropriate for children aged between 19 to 30 months old. For each of the 100 words prompted by the interviewer, the parents report if the child "says the word" or "does not say the word" asked by the interviewer. Both of these were collected at baseline. The MLI-III is appropriate for children aged between 31 and 48 months old. For each of the 100 words prompted by the interviewer, the parents report if the child "says the word" or "does not say the word" asked by the interviewer. The MLI-III was collected in the first follow-up survey.

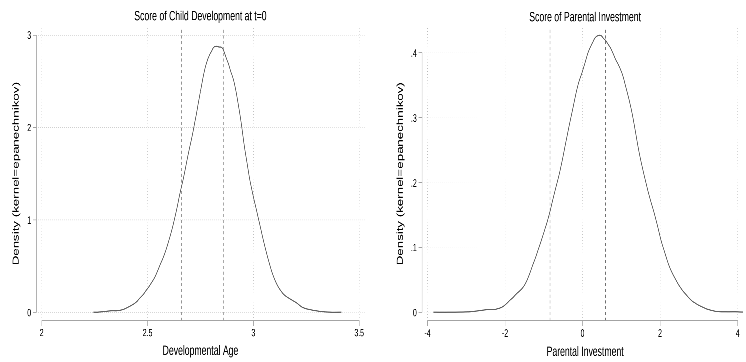
There are several important differences between the MLI and the BSID-III. First, the former is based on parental reports, while the latter is scored based on direct observation of the subject child. Second, the MLI was usually administered at the primary caregiver's house by the interviewer that collected the household survey, while the BSID-III was administered in community centers by a trained evaluator.

To assess parental investments, we used the UNICEF Family Care Indicator instrument (FCI, Frongillo et al. (2003)), which corresponds to X_i in our model, and we use the data collected in the follow-up one. This instrument contains questions about the types and number of play materials used by the child as well as the types and frequency of play activities. The data reflect the interactions between the primary caregiver (mostly the mother) and the child. A third source of data, directly reported by parents, measures the amount of time parents interact with children on a given day.

3.3 Scenarios and parental beliefs

As discussed in Section 2.5, the scenarios used as a starting point for the beliefs elicitation model were constructed using the estimates of the measurement systems for child development and parental investment estimated from the available measures.In total, we consider four scenarios for the hypothetical child. They correspond to "low" and "high" values of parental investment and beginning-of-period child development. In Figure 1, we plot the density function of the factor representing developmental age and parental investment. The dotted lines in the figure represents the 'low' and 'high' levels of the factor considered. The figure gives an idea of the part of the domain of the two inputs of the function in (2) that are covered by the scenarios.

Figure 1: Distribution of the initial conditions and parental investment factors and of positions of the scenarios



For each of the four scenarios, mothers were asked questions about the age at which the hypothetical child would be able to use the three different set of words (easy, medium, and hard). Therefore, for each scenario, we have three measures of the outcome of the perceived process of child development. Table 1 reports the mean, standard deviation, and range of the answers to the twelve questions in the beliefs elicitation survey instrument presented in two wooden tablets (left and right diagrams from Figure A.3). For each scenario, $s \in S$, we report the relevant statistics for the easy, medium, and hard words. For instance, in the first row, we see that on average, mothers think that a child with 'low development' at nine months and exposed to "low"' parental investment, start saying easy words at 18.4 months.

There is a considerable amount of variability in the answers the mothers provide. The coefficient of variations for all twelve questions is between 0.3 and 0.4. The mean for easy words is below that of medium ones for each scenario, which, in turn, is below that of hard words. Furthermore, the means for the low levels of parental investment are always above those for high. In addition, the means for low levels of beginning-of-the-period child development are above the corresponding means for high levels of the same variable. As our interviewers did not force these consistencies on respondents, this evidence suggests that, on average, parents believe that the technology of skill formation is an increasing function of inputs. While a few individuals may be inconsistent, they do not affect the means.

The next step in our analysis of the beliefs data consists in translating the out-

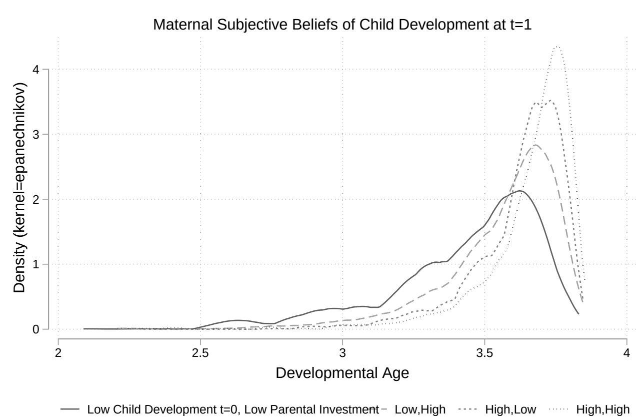
Table 1: Expected Endline Development by Scenarios of Baseline Development and Parental Investment

VARIABLES			Mean	St. Dv.	Min	Max
Low Baseline Development	Low Investment	easy	18.2	6.3	9	48
		medium	23.3	7.3	10	48
		hard	29.3	8.8	11	48
	High Investment	easy	15.6	5.5	9	48
		medium	19.9	6.5	9	48
		hard	24.8	8.0	10	48
High Baseline Development	Low Investment	easy	14.2	4.5	9	45
		medium	17.8	5.4	9	47
		hard	22.1	7.0	10	48
	High Investment	easy	13.4	5.0	9	46
		medium	16.6	5.6	9	48
		hard	20.2	7.0	9	48

Maternal Answers. Observations: 1112.

comes we elicit (the age at which a child can say different words) into an index of development, following the steps of the procedure we described in Section 2.3. Following that procedure allows us to convert the answers to the beliefs questions into the expected *developmental age* for each of the four scenarios (and for each of the three sets of words).

Figure 2: Maternal Beliefs of Child Development



As we have three different estimates of the perceived *developmental age* for each $s \in S$ corresponding to the set of words, $\{q = e, m, h\}$, we consider their average. In Figure 2, we plot the four density distributions of the (log of) average developmental ages corresponding to the four scenarios. Consistently with the evidence in Table 1, the distribution of developmental ages moves to the right as we move from the worst scenario (low child development at the beginning of the period and low investment) to the best one (high child development at the beginning of the period and high investment). Importantly, when moving from low to high investment keeping child development at the beginning of the period constant, the distribution also moves to the right.

4 Returns to parental investment

We start this section by presenting descriptive evidence on the perceived returns to investment we have elicited in our study. Next, we validate these measures by investigating whether or not they predict parental investment in two survey waves.

4.1 Subjective returns

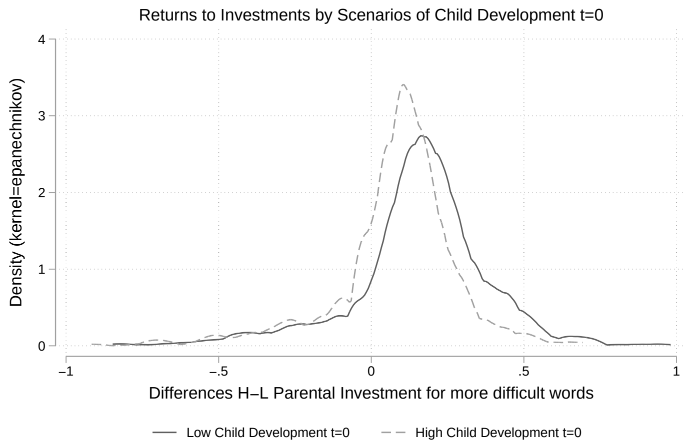
We can compute subjective maternal beliefs about the returns to parental investment by using the answers to the basic questions. For each respondent, we can compute the returns from moving from "low" to "high" parental investment, as expressed in equation (8), which we reproduce here.

$$r^{i,q}_{a,H_k} = \hat{H}^{iq}_a(H^k_{a-1}, X^H) - \hat{H}^{iq}_a(H^k_{a-1}, X^L); \quad k = H, L, \quad q = e, m, h, \qquad (8)$$

This measure of the subjective expectation about the productivity of parental investment can be computed for low and high scenarios of baseline development. Moreover, for each scenario of baseline development, we can calculate the return to parental investment for each set of words: easy, medium, and hard, $\{q = e, m, h\}$. Finally, we can express these returns in terms of the ages at which the hypothetical child starts saying a certain set of words under the high and low parental investment scenario or in terms of *developmental age*, as sketched in Section 2.3.

We start describing the returns obtained in terms of the gains in months, that is, before conversion to developmental age. Figure 3 plots the sample distributions of returns to parental investment for words of medium difficulty, $\{q = m\}$, conditional on low (dashed line) and high (solid line) levels of baseline development. The latter is clearly to the right of the former. A similar picture is obtained by computing the returns using other difficulty levels separately or exploring an aggregate measure such as their average. The means and standard deviations of the distribution of subjective returns are reported in Table 2, which shows that the mean

Figure 3: Subjective expectations about the returns to parental investment for High and Low initial development. *Months*



return is a decreasing function of baseline development for all difficulty levels of the words. The fact that mothers, on average, seem to think that the investment return is higher when baseline development is low is a robust result which we discuss further below.

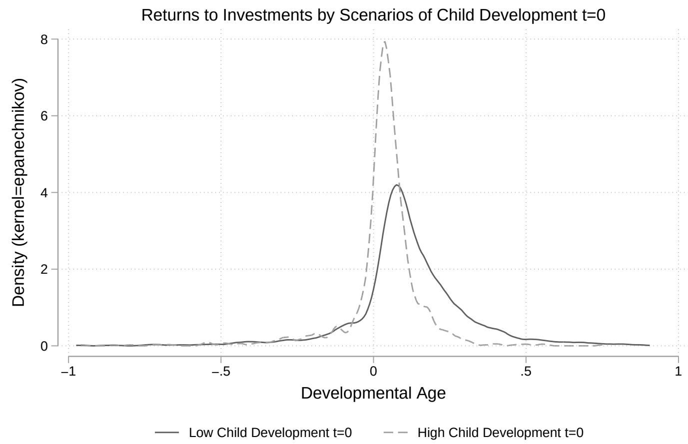
We note that there are a few mothers for whom the subjective return on investment is negative, signaling probably a problem in understanding the questions. As seen in Table 2, however, the return is negative for a relatively small fraction of the sample: for medium words, which register the highest fraction of negative returns, it is about 13% and 16.7% in the case of the low and high scenarios of baseline development, respectively.

Table 2: Returns to Parental Investment by set of words

VARIABLES		Mean	Std. Dev.	% of neg. values
Easy words	Low Initial Development	0.16	0.26	13.4
	High Initial Development	0.07	0.23	15.4
Medium words	Low Initial Development	0.16	0.24	13.4
	High Initial Development	0.08	0.21	17.5
Hard words	Low Initial Development	0.17	0.22	12.1
	High Initial Development	0.10	0.20	16.6
Observations: 1112.				

Having translated for each mother and scenario the expected outcome in terms

Figure 4: Returns to Parental Investments for Low and High Initial Development scenarios
Developmental age



of developmental age, we can, again, compute the returns to investment under low and high beginning-of-the-period child development. We plot the density for these returns in Figure 4. The graphs for the two returns are, not surprisingly, similar to those for the returns obtained directly from the beliefs data expressed in terms of months, which we plotted in Figure 3. The sample returns expressed in terms of developmental age are less dispersed. The returns when baseline development is high are, on average, higher than the returns when the baseline human capital is low. A possible interpretation of this finding is that the mothers in our sample think of parental investment as a *remedial* action which is particularly useful when a child has developmental problems. Furthermore, a small fraction of mothers appear to have negative expected returns. In the rest of the paper, we use this definition of returns, expressed in terms of developmental age.

As shown in Table 1, maternal subjective beliefs are substantially heterogeneous. Table 2 and Figures 3 and 4 show that this heterogeneity is also reflected in views about the returns to parental investment. As these measures are novel, we validate them showing they correlate with mothers' characteristics and parental investment.

In Table 3, we relate subjective expected returns for low and high scenarios of initial development to the respondent mothers' socioeconomic characteristics regressing subjective expected returns (under low and high scenarios of baseline development) on age, two education dummies (indicators for primary or secondary

Table 3: Returns to Child Development on Investment and SE characteristics

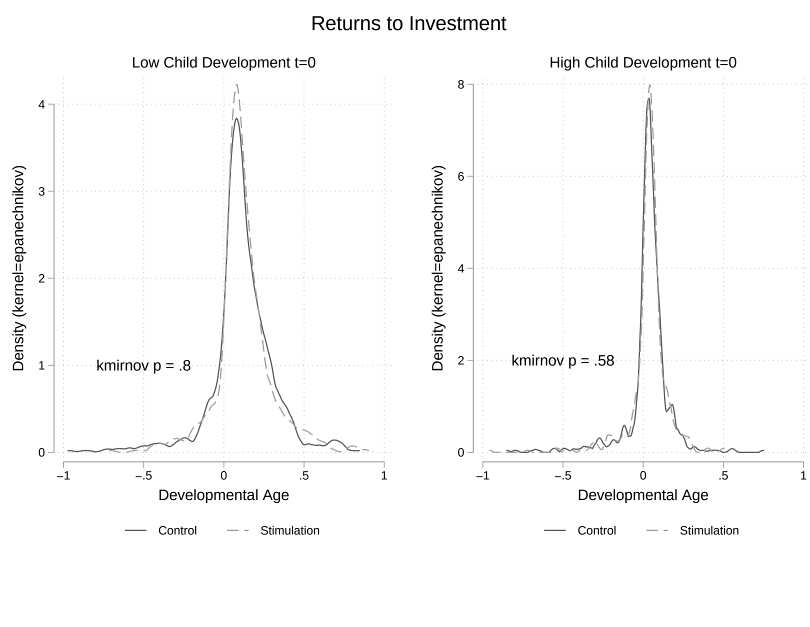
VARIABLES	Return to Low Initial Development		Return to High Initial Development	
Mother's age	0.001	0.001	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)
Mother's education (primary)	-0.004	-0.004	0.002	0.003
	(0.015)	(0.015)	(0.013)	(0.013)
Mother's education (secondary and more)	-0.009	-0.008	0.015*	0.016*
	(0.018)	(0.017)	(0.009)	(0.009)
Mother's depression (CES-D)	-0.006	-0.006	0.003	0.003
	(0.012)	(0.012)	(0.008)	(0.008)
Mother's IQ (standardized Raven's score)	0.015**	0.015**	0.018***	0.018***
	(0.007)	(0.007)	(0.004)	(0.004)
Dummy for Male (child)	-0.014	-0.014	-0.007	-0.006
	(0.012)	(0.012)	(0.007)	(0.007)
Child's age	-0.020	-0.020	0.008	0.008
	(0.024)	(0.023)	(0.018)	(0.018)
Wealth index (standardized)	-0.003	-0.003	0.011**	0.011**
	(0.007)	(0.007)	(0.004)	(0.004)
Treatment Assignment (dummy)		0.002		0.006
		(0.013)		(0.009)
R^2	0.008	0.008	0.043	0.043
F	1.574	1.398	3.874	3.546
Observations	1112	1112	1112	1112

Standard errors (in parentheses) are clustered at municipality level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

education), the CES-D depression index, the Raven progressive matrices test score and children gender. In the second and fourth columns, we also add a treatment dummy referring to the stimulation intervention described in Section 3.1. Of the variables considered, the only one that appears to be significantly related to the expected returns on parental investment is the Raven score, indicating that women with higher Raven tests have higher expected returns to maternal investment, both for low and high initial development.

Attanasio et al. (2014) have shown that the parenting stimulation program evaluated with the first two waves of the survey we are using had an impact on several measures of child development, including cognition and language. Attanasio et al. (2020), analyzing the mechanisms that generated such impacts, argued that a significant increase in parental investment triggered by the parenting stimulation program explained these effects. A possible explanation of these results could be that the increase in parental investment was driven by a shift in maternal beliefs about the nature of the developmental process. The results in Table 3 show that this is not the case, at least a few years after the parenting stimulation program was finished. The intervention does not seem to affect the mean of expected returns. Confirming this evidence, we plot the densities of subjective expected returns for

Figure 5: Subjective Expected Returns to Investment and Parenting Stimulation Program



low and high levels of initial child development for the randomly assigned control or parenting stimulation groups. Figure 6 shows the densities of subjective expected returns in the two groups are virtually identical regardless of the scenario of initial development.

We stress that the beliefs data were collected only in the second follow-up survey. Andrew et al. (2018) report that, at that point, the effects of the parenting stimulation program on parental investment (and on child development measures) had faded out. Therefore, any impact on subjective beliefs had possibly faded out, too. To explore this possibility and validate our belief measures, we now analyze associations between parental investment and parental beliefs.

4.2 Does Parental Investment Vary with Maternal Subjective Beliefs?

Having obtained measures of parental beliefs from questions referring to a hypothetical child, we relate them to actual parental investment. We start by considering investment as a function of its possible determinants, *including its perceived returns*. We consider returns conditional on "low" and "high" baseline development scenarios separately.

As mentioned above, we have three measures of the subjective returns to investment, one corresponding to each set of words $\{q = e, m, h\}$. The availability of these multiple measures of the same variable can help account for measurement error.

Table 4: Investment and Returns on Investment

VARIABLES	OLS	OLS[a]	OLS[a]	IV	IV[a]	IV[a]
Inv. Return: Low	-0.500	-0.141	-0.140	-0.221	-0.170	-0.172
Baseline Development	(0.305)	(0.150)	(0.146)	(0.155)	(0.150)	(0.145)
Inv. Return: High	1.457***	0.565**	0.560**	1.109***	0.536**	0.523 **
Baseline Development	(0.407)	(0.220)	(0.214)	(0.104)	(0.208)	(0.203)
Treatment Assignment			0.165*			01642*
			(0.087)			(0.088)
Controls		yes	yes		yes	yes
R^2	0.008	0.135	0.146	0.005	0.132	0.138
F p-value	0.002	0.000	0.000	0.000	0.000	0.000
Observations	1112	1112	1112	1112	1112	1112

[a] Regression controls by mother's age, education, depression (CES-D) and IQ (standardised Raven's score) as well as gender and age of the child and standardized wealth index. [b] Regression controls also by a dummy for the treatment assignment.

Standard errors (in parentheses) are clustered at municipality level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We run two different specifications. In the first, we use the average of the three measures of returns (both for a low and high baseline development scenarios), with the idea that possible measurement error in the three measures could average out. In the second specification, we use returns as measured by the ability of the hypothetical child to use hard words and use the other two measures of returns (i.e., the ability to say easy and medium words) as instruments. For both specifications, we report a simple regression that does not include any other variable and another that contains controls for the possible determinants of parental investment, including the mother's age, education, depression, and Raven scales.

We compute the latent investment factor factor by first constructing two latent factors from individual items, representing time and material investment. We then aggregate them into a single investment factor using an additional factor analysis. We stress that the investment data come from the second wave of the survey, which was collected at the end of the stimulation intervention. As mentioned above, we also note that the parenting intervention that motivated the data collection had just finished before the second wave and was shown to have a significant effect on parental investment, while that effect petered out in the third wave. The data on beliefs, instead, were collected on the same households, in the third wave of the survey, collected more than 2 years after the second.

In Table 4, we report the results obtained using both the approach of averaging across the three measures of returns we have for low and high initial conditions and the one of using an IV strategy, where two of the measures are used as an instrument for the third. For each of these approaches we first include only the two measures of return and then add a number of control variables, reflecting parental background and wealth, as well as, in column 3 and 6, a dummy for the treatment.

We notice that the using OLS on the mean of the three measures does not yield results dramatically different from the IV approach. While the coefficient on the return to investment under low initial condition is never statistically different from zero (with the point estimate being negative) the coefficient on the return under high initial conditions is positive (as one would expect) and significantly different from zero. These results indicate that parent with higher expected returns to investment invest more in their children.

The size of this coefficient is reduced greatly by the consideration of additional variables, but it is still significantly different from zero. Furthermore this result is not affected by the consideration of a dummy variable indicating the presence in the village of the stimulation intervention Such a variable is mildly significant.

The coefficient on the return to low baseline development scenario, instead never becomes significantly from zero and its point estimate remain negative. Neither the size nor the precision with which these coefficients are estimated is affected by the introduction of control variables or the use of IV.

These results constitute a crucial validation exercise for our belief measures. We find that some of our measures of subjective beliefs predict parental investment. To an extent, this is true for parental investment measured at the same time as the parental beliefs and for parental investment (performed by the same respondents) at a different point in time.

5 Expected Returns to Parental Investment and Technologies of Skill Formation: subjective and objective views

The associations between parental investment and data derived from the beliefs questions we have estimated in the previous section are very simple. We regress parental investment on a measure of expected returns (expressed in months) under different initial conditions and take into account the possible presence of measurement error in perceived returns either by averaging across multiple measures associated to the same latent factor or using some of the measures as instruments for the others, under the assumption that the measurement error is uncorrelated across different measures.

In this section, we put more structure on the beliefs data we elicited. In particular, we show how to use the subjective maternal beliefs data to estimate, under certain assumptions, the parameters of a *subjective technology of skill formation* for each respondent. We also briefly discuss how we estimate the parameters of an *objective technology of skill formation* , that we can then compare to the parameters of the subjective production functions.

5.1 Estimation of the Subjective Technology of Skill Formation

As we mentioned in Section 2.4, the scenarios and the potential outcomes we presented to the mothers in our sample were determined with the help of measurement systems estimated on data on child development and parental investments. The estimates of these measurement systems allow us to compute the scores of baseline development ($H_{a-1}^{s_h}$) and parental investment (X^{s_x}) corresponding to each scenario $s \in S$. Furthermore, the same applies to $H_{i,a,s,q}, \quad q \in \{e, m, h\}$, which we treat as three measures of the same (unobserved) expected child development at age a conditional on a given scenario s, *as perceived by respondent* i.

We assume mothers know the "right" functional form for the technology of skill formation but not necessarily the "right" parameters. Therefore maternal reports of $H_{i,a,s,q}, \quad q \in \{e, m, h\}$ are error-ridden measures of the left-hand side of the equation (2), implied by the subjective technology of skill formation, which we reproduce here with a specific functional form, for a given scenario s:

$$\begin{aligned} E[\ln H_{i,a}|\Omega_i, H_{i,a-1}, X_i] =& \mu_{i,0} + \mu_{i,1} \ln H_{i,a-1} + \mu_{i,2} \ln X_i + \\ & \mu_{i,3}[\ln H_{i,a-1} \ln X_i] + E[\epsilon_i|\Omega_i, H_{i,a-1}, X_i] \end{aligned} \tag{9}$$

In eliciting maternal beliefs, we ask mothers to think of a hypothetical child, not their own children. Furthermore, we associate this child with hypothetical scenarios of child development at a certain age and parental investment. Therefore, from the mother's point of view, the conditional expectation in equation (9) refers to a 'typical' child in their environment The reference to this hypothetical child implies that the intercept $\mu_{i,0}$ captures, in addition to beliefs about the intercept of the technology of skill formation, the average level of determinants of child development, $\bar{Z}_i$. Similar considerations can be made for the shock ϵ_i, so that it is not strictly necessary to assume that $E[\epsilon_i|\Omega_i, H^{s_h}_{a-1}, X^{s_x}] = 0$, which implies that the scenarios $s \in S$ are uncorrelated with shocks ϵ_i. We can think of such an average as a fixed effect that is absorbed in $\mu_{i,0}$, like the average of the variables Z that are not considered in the scenarios.

In equation (9), we make explicit the dependence of development at age a on parental investment and development at age $a - 1$ because they vary counterfactually across the scenarios used to elicit beliefs.

When we consider the answers to the beliefs questions for a given scenario s, we can re-write equation (9) as:

$$\begin{aligned} \ln H_{i,a,s,q} =& \mu_{0,i} + \mu_{1,i} \ln H^{s_h}_{a-1} + \mu_{2,i} \ln X^{s_x} + \mu_{3,i} \left[\ln H^{s_h}_{a-1} \ln X^{s_x}\right] + \\ & E[\epsilon_i|\Omega_i, H^{s_h}_{a-1}, X^{s_x}] + \eta_{i,a,s_x,s_h,q}; \qquad s_h, s_x = H, L \end{aligned} \tag{10}$$

The residual $\eta_{i,a,s_x,s_h,q}$ reflects measurement error in eliciting beliefs. It is natural to assume that the shocks to child development ϵ_i are uncorrelated with the variables considered in the scenarios, conditional on the information available to parents.

$$E[\epsilon_i|\Omega_i, H_{i,a-1}, X_i] = E[\epsilon_i|\Omega_i] \quad \forall \;\; H_{i,a-1} \;\; \text{and} \;\; X_i.$$

This assumption is made for the hypothetical child, and it is reasonably mild because the scenarios are constant across mothers. The identification and estimation of the parameters of the subjective production function are conceptually very different from the same tasks relating to its counterpart in the objective technology of skill formation. The analysis of the subjective technology of skill formation using answers to the questions we designed does not require researchers to address the endogeneity of parental investment because it uses exogenous counterfactual variation across investment scenarios that do not vary across mothers. In contrast, recovering the parameters of the objective technology of skill formation requires using data on actual parental investments, which correlate with variables that parents know, but the econometrician does not observe.

Equation (10) can be seen as a factor model where the μ_i's are the factors and where the factor loadings are known. For each mother in our sample, we have twelve data points that we can use to estimate mother i's subjective expectations about the parameters in the technology of skill formation. We can then estimate, for each individual, the vector $\{\widehat{\mu}_{i,0}, \widehat{\mu}_{i,1}, \widehat{\mu}_{i,2}, \widehat{\mu}_{i.3}\}$ by running an Ordinary Least Squares (OLS) regression separately for each mother in our sample. Alternatively, we can improve the precision of the individual-level estimates by using the Swamy (1970) estimator, described in the detail in Appendix A.3. Regardless of the method used, we stress that the subjective production functions vary in the cross- section.

5.2 Estimation of the Objective Technology of Skill Formation

One of the goals of this paper is to compare subjective beliefs about the parameters of the technology of skill formation with their objective counterparts. Having explained how we convert mothers' answers into parameters of a *subjective technology of skill formation* at the individual level, we now need to obtain the objective estimates for the same parameters. However, we must tackle several issues to make the comparison between the "subjective" and "objective" technology of skill formation sensible.

First, we need to ensure the latent factors that enter the "objective" and "subjective" technology of skill formation have cardinality and are measured with the same metric. This will allow us to have meaningful comparisons of the estimates of the “objective” and “subjective” models of skill formation. Second, while the “subjective” technology of skill formation is estimated by manipulating exogenous scenarios, estimating the “objective” technology requires accounting for the endogeneity of actual investments. Third, as mentioned above, we make the relatively strong assumption that one can represent maternal subjective beliefs about the child development process with a technology of skill formation that has a functional form similar to the one we fit to the objective data, albeit the parameter values are allowed to differ. Here we briefly discuss the first two issues.

To estimate the “objective” technology of skill formation, we use variation in the child development and parental investment factor scores derived from the same measurement system we used to design the hypothetical scenarios. Using the same measurement system (see Appendix B) and the same data, we can scale the estimated latent factors following the procedure discussed in Section 2.3. Therefore, the objective factors' scales are comparable to those used to estimate the “subjective” technology of skill formation, making the estimated parameters comparable. We can therefore ask questions such as whether on average parents over or underestimate the productivity or usefulness of parental investment.

As for the endogeneity of parental investment, we use an IV approach which requires the use of a valid instrument. As an instrument, we use the parenting stimulation program, which was randomly allocated to half the towns in our sample. We know the intervention affected parental investment over the period considered. It could be argued that the parenting stimulation program might have affected the technology of skill formation directly and therefore the random assignment would not be a valid instrument. However, Attanasio et al. (2020), who use a different set of instruments, rule out this potential mechanism. Thus, random assignment to the program is a valid instrument.

A final caveat should be mentioned. Attanasio et al. (2020) divide parental investment into *time* and *material* investments and estimate the effect of both components onto child development. Although such a decomposition is conceptually easy for the formulation of questions on beliefs, in practice it is hard because the number of scenarios grows exponentially with the number of inputs in the technology of skill formation. Therefore, unlike Attanasio et al. (2020), we restrict ourselves to an aggregate measure of investment, so to maintain comparability between the estimates of the “objective” production function and those of the “subjective” one.

6 Estimating Production Functions

In this section, we present estimates of the parameters of the subjective technology of skill formation. Next, to set a benchmark against which we can compare the results obtained, we present estimates of the same functional form, which we estimate on data on actual child development and parental investment. The latter constitutes the objective technology of skill formation.

6.1 Subjective Technology of Skill Formation

Section 2 modeled parental investment as a function of individual preferences, resources, and the *subjective perception* of the technology of skill formation. Section 5.1 discussed how, starting from an assumption about the subjective technology of skill formation, we can estimate its parameters for each respondents, the vector $(\widehat{\mu}_{i,0}, \widehat{\mu}_{i,1}, \widehat{\mu}_{i,2}, \widehat{\mu}_{i.3})$ in equation (9) or equation (10). In this section, we report the results of this exercise.

The mother's subjective technology of skill formation is defined by equation (9) or, equivalently, by equation (10), which we reproduce for convenience:

$$\begin{aligned} \ln H_{i,a,s,q} =& \mu_{0,i} + \mu_{1,i} \ln H^{s_h}_{a-1} + \mu_{2,i} \ln X^{s_x} + \mu_{3,i} \left[\ln H^{s_h}_{a-1} \ln X^{s_x}\right] + \\ & E[\epsilon_i | \Omega_i, H^{s_h}_{a-1}, X^{s_x}] + \eta_{i,a,s_x,s_h,q}; \qquad s_h, s_x = H, L; \quad q = e, m, h \end{aligned} \tag{10}$$

In equation (10), η represents measurement error. At the same time, the subscript q refers to the different outcomes considered in the scenario, in particular, whether the outcome is represented by easy (e), medium (m), or hard (h) words. We recall that we assume $E[\epsilon_i|\Omega_i, H^{s_h}_{a-1}, X^{s_x}] = 0$. As we have twelve observations, we can estimate via OLS the coefficients $\mu_{k,i}, k = 0, 1, 2, 3$, for each each mother i.

Our procedure, obtained with the Swamy (1970) estimator mentioned in Section 5.1 and described in detail in Appendix A.3, yields more precise estimates of the coefficients of these equations. Similar results can be obtained with a factor model.[5] Given the scaling of the relevant factors, these results can be compared to estimates of the *objective* technology of skill formation obtained from data on actual child development.

[5]In equation (10), the μ_i's are factors to be estimated whose loading factors are the variables that define the scenarios and which, in our data, are observed.

Table 5: Estimation of Subjective Technology of Skill Formation

Dependent Variable: Expected Log of Follow Up Development				
	Cobb-Douglas	Fraction $\|t\| > 2$	Trans-log	Fraction $\|t\| > 2$
Intercept	2.519 (0.042)	93.80%	2.175 (0.056)	80.76%
Initial Development	0.350 (0.013)	73.29%	0.468 (0.018)	66.37%
Investment at Follow Up	0.077 (0.005)	44.42%	0.692 (0.058)	23.47%
Investment at Follow Up x initial Development			-0.212 (0.019)	20.50%

Observations: 1112. Numbers in parentheses are *not* standard errors of the estimated coefficients, but the standard deviation of the estimated coefficients in the sample.

The individual coefficients of the subjective technology of skill formation are summarised in Table 5. Our approach delivers a set of coefficients for the perceived production function for each mother in our sample. As with our estimates of subjective returns to parental investment, reported in Table 2, the estimated perceived technologies of skill formation, which imposes some structure on these returns, exhibit a considerable amount of heterogeneity.

In the first and third columns of Table 5, we report the average of the individual level coefficients of two specifications for the technology of skill formation. The specification in the first column assumes a Cobb Douglas form, forcing the $\mu_{i,3}$ coefficients to zero, while the specification in the third column assumes a translog specification as in equation (2). Below the average of each coefficient, we report the standard deviation for that coefficient *in the sample*. These standard deviations, therefore, represent the sample heterogeneity in perceived production functions, rather than the precision of our estimates. In the second and fourth columns, we report the fraction of coefficients in the sample which are estimated to be statistically significant (that is with a t-value greater than 2). As the individual coefficients are estimated with 12 observations, the lack of precision of the individual estimates is not too surprising.

For the Cobb-Douglas specification, we note that we do not impose constant returns to scale and that the average effect of investment, that is, its marginal product, is 0.077. They are (on average) much smaller than that on initial level of development. The fraction of coefficients for parental investment with a t-value greater than 2 is 44%, while it is 73% for the coefficient on H_0.

Moving to the translog specification, we find that about 20% of the interaction terms are significantly different from zero. This finding indicates that the Cobb-Douglas case represents a good approximation (relative to the translog) for many mothers. However, this result could be driven by the attempt to estimate four coefficients with twelve observations.

6.2 Objective Production Function

As mentioned above, comparing the subjective and objective production functions is straightforward because we assume they have the same functional form and, importantly, the latent factors used in the estimation have the same metric. We can then compare the objective and subjective functions by comparing the coefficients of these functions. Therefore, in both cases, we estimate two specifications: a Cobb-Douglas form and a translog form.

$$\ln H_{i,a} = \delta_0 + \delta_1 \ln H_{i,a-1} + \delta_2 \ln X_i + \delta_3 \left[\ln H_{i,a-1} \ln X_i\right] + \beta' Z_i + \epsilon_i \tag{11}$$

where Z_i is a set of other observable variables that affect child development in addition to initial conditions and parental investment and ϵ_i capture unobservable factors. The Cobb-Douglas specification imposes $\delta_3 = 0$, (which corresponds to $\mu_{i,3} = 0$).

Among the Z_i variables which enter the technology of child development in addition to the initial conditions and parental investment, we consider mother's education attainment and IQ tests. Furthermore, we also control for children's age and gender. We utilize these variables to standardize the measurements used to construct the developmental factors in our analyses of the subjective beliefs data. Therefore, we must consider their effect on *measured* development and its relation to the variables of interest. As discussed above, we interpret the fact that the questions about the subjective perceptions refer to a *hypothetical child* as implying an averaging of the Z_i, so that the effects of these variables would be reflected in the intercept of subjective production function, as represented in equation (10).

As we want to compare the parameters of the subjective and objective production functions, we are particularly interested in identifying the *structural parameters* that reflect the effect of the initial level of child development (at age $a-1$) and of parental investment on children subsequent development. To recover these parameters, we need to account for the potential endogeneity of parental investment. To deal with this issue, for all specifications considered, we use a control function approach where the variable excluded from the production function, but included in the first stage for investment, is the random assignment to the stimulation intervention whose impact was studied by the RCT. Attanasio et al. (2020) show that the intervention increased parental investment substantially and, in their mediation analysis, argue that it did not have a direct effect on child development. This evidence, therefore, makes it a valid instrument. The control function approach allows us to easily consider the interaction between initial conditions and parental investment with a single excluded instrument.

For the objective production function, we use data from the FU-I, when the children's ages correspond roughly to the ages used in describing the scenarios in the beliefs questions. We report the estimates of the objective production function in Table 6.

The first-stage results indicate that the treatment substantially increases parental investment, as reported in Attanasio et al. (2020). These results also imply that parents invest more in more developed children.

As for the second stage results, for the Cobb-Douglas specification, we remark that the investment coefficient on children's cognitive development is 0.154. The coefficients on initial development is 0.391 Both coefficients are strongly significant.

For the translog specification, while the coefficient on initial development is significantly different from zero, neither the coefficient on parental investment nor that on the interaction term is significant, although the latter, is nearly significant. Interestingly, the point estimate of the coefficient on the interaction term is estimated to be positive, providing suggestive evidence that parental investment is more productive on more developed children.

There are two reasons why the translog specification yields noisier estimates. When we use the Cobb-Douglas formulation, for each parent i, we have six moments to identify $\mu_{i,1}$ and an additional six moments to identify $\mu_{i,2}$. Thus, we have twelve moments to identify two parameters. With the translog specification, we have three moments to identify $\mu_{i,1}$, three to identify $\mu_{i,2}$, and three to identify $\mu_{i,3}$. Thus, we have nine moments to identify three parameters. This smaller number of moments is one factor that explains why our translog estimates are noisier.

Additionally, the nine moments in the translog case are very similar, leading to a higher correlation of these identifying moments than those in the Cobb-Douglas

Table 6: Objective Estimation of the Production Function

	First Stage	Second Stage	
		Cobb-Douglas	Translog
Intercept	-0.443 (0.289)	2.390 (0.066)	2.480 (0.081)
Initial Development	0.314 (0.099)	0.391 (0.027)	0.359 (0.032)
Log of Parental		0.154 (0.054)	-0.025 (0.107)
Log of Parental Investment x Initial Development			0.064 (0.033)
Treatment Assignment (dummy)	0.161 (0.039)		
Control function		-0.123 (0.055)	-0.131 (0.055)

Note: Dependent Variable for the Fist Stage is the Log of Investments at Follow Up and while for the Second Stage is the Log of Follow Up Development. The specifications control by mother's age, education, depression and IQ (standardised Raven's score) as well as gender and age of the child and wealth index (standardised). Observations: 1112.

case.[6] This higher correlation is another factor contributing to the noisiness of the translog estimates.

Finally, we note that the control function is strongly significant in all specifications, indicating that investment is, indeed, endogenous. The negative sign of the control function coefficient suggests that ignoring the endogeneity of investment reduces the estimated coefficient, a result consistent with Attanasio et al. (2020), which can be interpreted as indicating compensating behavior by the parents.

6.3 Comparing Subjective and Objective Production Functions

A significant advantage of imposing the structure of the production function on the subjective beliefs data is that we can now compare the productivity of parental investment as estimated from actual data to that perceived by parents, an exercise that was not possible by looking at the simple returns data reported in Table 1 or in Figure 1. As we mentioned before, the beliefs variables we use to estimate the subjective technology of skill formation are scaled so to make them comparable

[6] see Appendix C for a derivation of the identifying moments for both specifications.

Table 7: Subjective and Objective Marginal Products of Parental Investment

	Cobb-Douglas	Translog	
		H^L_{a-1}	H^H_{a-1}
Subjective Production Function	0.077	0.122	0.039
	[0.005]	[0.007]	[0.005]
Objective Production Function	0.154	0.148	0.173
	(0.054)	(0.054)	(0.055)

Note: Number in parentheses are standard errors. The number in square brackets are the standard deviation of the subjective expectations coefficients in the estimation sample. Observations: 1112.

with the estimates obtained on actual data: the average coefficients in Table 5 are therefore directly comparable with the estimates in Table 6.

For the Cobb-Douglas specification, the comparison is easy, as the marginal product of investment is constant and does not depend on the initial level of development. Here the evidence is clear, on average, the coefficient on the subjective Cobb-Douglas function is substantially lower than that in the objective functions whose coefficients are reported in Table 6: 0.077, compared to 0.154. For ease of comparison, we report these coefficients in Table 7.

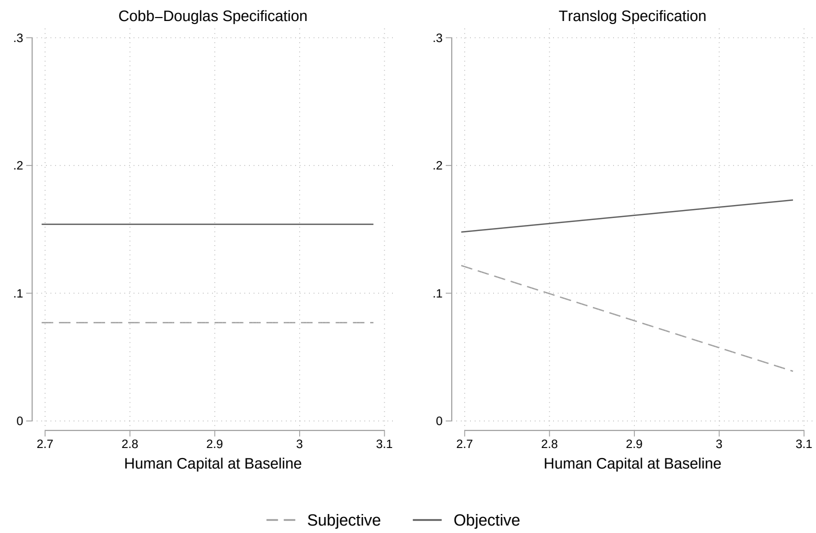
For the translog specification, the comparison is slightly more complicated as the marginal product of investment changes with baseline development. In particular, given the functional form, the expressions for the marginal product of parental investments as a function of baseline development for the subjective and objective production functions are given by:

$$\begin{aligned} ret^i(H^{s_h}_{a-1}) &= \mu_{2,i} + \mu_{3,i} H^{s_h}; \quad s_h = L, H \\ ret^{obj}(H^{s_h}_{a-1}) &= \delta_2 + \delta_3 H^{s_h}; \quad s_h = L, H \end{aligned} \tag{12}$$

where the δ's and the μ's are taken from the estimates in Tables 5 and 6. For H^{s_h}, we use the two values corresponding to the low and high initial development used in constructing the scenarios for the belief elicitation. We, therefore, report the returns for the objective and (average) subjective translog production functions computed at low and high baseline development in columns 2 and 3 of Table 7.

As with the Cobb-Douglas production function, we observe that the perceived return to parental investment is considerably lower than the estimated objective return. Furthermore, we notice that the difference between objective and subjective estimates is larger at high levels of initial development than at low levels.

Figure 6: Subjective and Objective Marginal Products of Parental Investment



While the results in Table 7 are interesting, they only report the returns for the objective and subjective production functions for two values of the initial level of development, albeit those that represent the scenarios in the beliefs questions. To better compare the objective and subjective production functions, Figure 6 plots the functions in equation (12) for both of them. On the left panel, we plot the returns implied by the Cobb-Douglas specification, which do not vary with the initial level of development and can be used as a reference point. On the right panel, we plot the (average) subjective return with a dotted line along with the objective return (solid line). We see that the return on the subjective production function is always below that of the objective function. Moreover, the former is decreasing in H_{a-1} while the other is increasing, indicating that in the objective production function investment and initial conditions are complements, while they are perceived as substitute in the subjective functions.

These results are consistent with anthropological evidence, such as that discussed in Lareau (2003), claiming that low-income parents might not be investing much in their children because they do not see the usefulness of such interventions. Other evidence (Putnam, 2015), consistent with a model where the underestimation of the productivity of investment is less severe at low level of initial development, indicates parents in disadvantaged environments might think parental investment to be beneficial only to deal with emergencies or adverse shocks.

Table 8: Subjective Estimates and SE characteristics: Cobb Douglas

	μ_1	μ_2
Mother's age	0.006	0.007*
	(0.004)	(0.004)
Mother's education (primary)	-0.015	-0.034
	(0.095)	(0.086)
Mother's education (secondary and more)	-0.003	0.022
	(0.092)	(0.081)
Mother's depression (CES-D)	-0.008	-0.041
	(0.055)	(0.061)
Mother's IQ (standardized Raven's score)	0.013	0.132***
	(0.038)	(0.037)
Dummy for Male	-0.085	-0.107*
	(0.064)	(0.061)
Child's age	0.031	-0.062
	(0.136)	(0.124)
Standardized Log of Baseline Development	0.024	0.033
	(0.033)	(0.032)
Wealth index (standardized)	-0.041	0.004
	(0.033)	(0.035)
Treatment Assignment (dummy)	-0.077	0.001
	(0.064)	(0.070)
Adjusted R^2	-0.001	0.015
F	0.924	2.264
Observations	1112	1112

Note: Standard errors (in parentheses) are clustered at municipality level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.4 How do the subjective coefficient vary?

Having estimated the parameters of the subjective technology of skill formation for each mother in the sample, we now analyse whether these coefficients relate to other variables. This exercise is similar to those reported in Table 3, except that rather than considering the two simple returns that can be computed directly from the beliefs data, we use the parameters of the subjective technology of skill formation and, therefore, introduce more structure on the data, which allows us to exploit efficiently all the information we have about individual beliefs.

In Table 8, we regress the estimated parameters of the subjective technology of skill formation for the Cobb=Douglas specification ($\mu_{i,1}$ and $\mu_{i,2}$) on the same set of observables used in Table 3: mother's age, two indicators of education, a depression index, the results of the Raven tests, and an indicator of the gender of the target child and a treatment dummy.

As in Table 3, most variables are insignificantly different from zero. In that table, we found that the expected return on children with *high baseline development* positively correlated with the maternal IQ, as measured by the results of the Raven test. Here, similarly, we find that the only significant variable in the regression for $\mu_{i,2}$, which measures the productivity of investment, is indeed the maternal performance in the Ravn test.[7].

Finally, we check whether the parameters of the subjective technology of skill formation predict investment. To do so, we assume parents solve the following maximization problem. Parent's utility function is:

$$U(c_i, h_{i,1}, x_i) = \ln c_i + (\beta + \zeta_i) \ln h_{i,1} + Z_i \pi \ln x_i$$

The parental utility function depends on household consumption (c), the child's human capital, and psychic benefits (or costs) of parental investments. Note that the psychic benefits vary according to household characteristics Z. Parents face two constraints: a budget constraint and a *perceived* production function of children development. The budget constraint is:

$$c_i + p x_i = y_i$$

where p is the price of investment, and y is household income. As the metric we use for parental investment is time, in our empirical analysis, p is the wage rate. [8]

[7]We do not report the results for the Translog specification, as they are much noisier and harder to interpret, as discussed above. Effectively, however, no variable seems to be related to the $\mu_{i,2}$ and $\mu_{i,3}$ coefficients

[8]We estimate a Heckman selection model to address selection of caregivers' participation in the labor market.

We assume that both the *perceived* and objective production function of child development are of the Cobb-Douglas form. Investment is determined by the *perceived* production function as well as by parental tastes and the budget constraint. Given these assumptions, the investment function that solves this optimisation problem satisfies the following equation:

$$\frac{px_i}{y_i - px_i} = \beta\mu_{i,2} + Z_i\pi + \zeta_i \tag{13}$$

where Z_i are several household controls.

Given this structure, to establish a relationship between investment and individual beliefs, we estimate equation (13), where investment is measured by a factor model using measurements collected at follow-up 1. As we are interested in the relationship between subjective beliefs and investment, we focus. on the parameter β. To estimate equation (13), we must address the fact that $\mu_{i,2}$ is estimated and, thus, it has errors. To do so, we derive a maximum likelihood estimator, which we describe in detail in Appendix C.

Our estimate for β is 0.085 with a standard error of 0.008, thus indicating that parental beliefs predict investments in our data. This result confirms that our measures of subjective beliefs are predictive of parental investment. The estimation of a structural model of behavior allows to make more precise statements about the role that beliefs play. It also validates the comparison between the objective and subjective production function and the result that, on average, in our sample, parents seem to underestimate the effectiveness of parental investment.

7 Conclusions

In this paper, we have proposed a new method to elicit maternal beliefs about the technology of skill formation. Our elicitation methodology allows us to directly estimate subjective beliefs about the returns to investment under mild assumptions. However, we go beyond and develop an analytical framework to generate and compare subjective and objective estimates of the technology of skill formation. Our paper is the first to use data from the same population and to explore random variation to a major parenting education program. In our context, we show that mothers underestimate parental investment's productivity severely.

We also show that our estimates of subjective beliefs about the productivity of parental investment co-vary with actual investment. Parents who report higher expected returns also tend to invest more in their children. This finding holds when we use estimates of the beliefs about the rates of return or the parameters of the subjective technology of skill formation.

Finally, our subjective expected returns are not higher in the group of mothers that, a few years before the belief elicitation module was implemented, received a stimulation intervention.

In the future, we think it is crucial to combine measures of subjective beliefs with data on expenditure allocation within families to estimate parental preferences and how parental beliefs change over time and with experience. Finally, it would be interesting to determine how parental beliefs might affect the allocation of resources among different children.

Understanding parental behaviour and its determinants is critical to designing effective interventions to foster the development of children living in adverse environments. This paper shows that parental beliefs are vital drivers of parental behaviour. Therefore, characterising parental beliefs and their potential biases is essential, both from a research and a policy point of view.

References

Andrew, A., Attanasio, O., Fitzsimons, E., Grantham McGregor, S., Meghir, C., and Rubio-Codina, M. (2018). Impacts 2 years after a scalable early childhood development intervention to increase psychosocial stimulation in the home: A follow-up of a cluster randomised controlled trial in colombia. *PLOS Medicine* Https://doi.org/10.1371/journal.pmed.1002556.

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

Attanasio, O., Fernández, C., Fitzsimons, E., Grantham-McGregor, S., Meghir, C., and Rubio-Codina, M. (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in colombia: cluster randomized controlled trial. *BMJ* 349:g5785.

Baker, A. J. L. and Piotrkowski, C. S. (1996). Parents and children through the school years: The effects of the home instruction program for preschool youngsters. *New York: National Council of Jewish Women, Center for the Child* .

Bayley, N. (2006). Bayley scales of infant and toddler development. harcourt assessment. *San Antonio, TX* 3rd edition.

Behrman, J. R., Calderon, M. C., Preston, S. H., Hoddinott, J., Martorell, R., and Stein, A. D. (2009). Nutritional supplementation in girls influences the growth of their children: prospective study in guatemala. *The American journal of clinical nutrition* 90(5):1372–1379.

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

Caldera, B., D., L., K., Rodriguez, Crowne, S. S., Rohde, C., and Duggan, A. (2007). Impact of a statewide home visiting program on parenting and on child health and development. *Child Abuse and Neglect* 31(8):829–852.

Cunha, F., Elo, I., and Culhane, J. (2013). Eliciting maternal expectations about the technology of cognitive skill formation. NBER Working Paper No. 19144.

Cunha, F., Heckman, J., and Navarro, S. (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers* 57(2):191–261.

Cunha, F., Heckman, J., and Schennach, S. (2010). Estimating the technology of cognitive and non-cognitive skill formation. *Econometrica* 78(3):883–931.

Cunha, F., Heckman, J. J., Lochner, J. L., and Masterov, D. (2006). *Interpreting the Evidence on Life Cycle Skill Formation in E. Hanushek and F. Welch, eds., The Handbook of Economics of Education*. Amsterdam: North Holland.

Currie, J. and Thomas, D. (1999). Early test scores, socioeconomic status and future outcomes. Technical report, National bureau of economic research.

Del Boca, D., Flinn, C., and Wiswall, M. (2013). Household choices and child development. *Review of Economic Studies* 81(1):137–185.

Drotar, D., Robinson, J., Jeavons, L., and Lester Kirchner, H. (2009). A randomized, controlled evaluation of early intervention: The born to learn curriculum. *Child: Care, Health and Development* 35(5):643?649.

Eckenrode, J., Campa, M., Luckey, D. W., Henderson, C. R., Cole, R., Kitzman, H., Anson, E., Sidora-Arcoleo, K., Powers, J., and Olds, D. (2010). Long-term effects of prenatal and infancy nurse home visitation on the life course of youths: 19-year follow-up of a randomized trial. *Archives of Pediatrics and Adolescent Medicine* 164(1):9–15.

Frongillo, E., Sywulka, S., and Kariger, P. (2003). Unicef psychosocial care indicators project. final report to unicef. Mimeo, Cornell University.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M., and Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in jamaica. *Science* 344(6187):998–1001.

Grantham-McGregor, S., Powell, C., Walker, S., and Himes, J. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the jamaican study. *The Lancet* 338:1–5.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). The rate of return to the highscope perry preschool program. *Journal of public Economics* 94(1-2):114–128.

Jackson-Maldonado, D., Marchman, V., and Fernald, L. (2012). Short form versions of the spanish macarthur-bates communicative development inventories. *In Applied Psycholinguistics. Cambridge University Press, Cambridge* .

Judge, G. G., Hill, R. C., Griffiths, W. E., Lutkepohl, H., and Lee., T. C. (1985). *The Theory and Practice of Econometrics*. 2d ed. New York: John Wiley and Sons.

Kotlarski, I. (1967). On characterizing the gamma and normal distribution. *Pacific Journal of Mathematics* 20:729–38.

Lareau, A. (2003). *Unequal Childhoods: Race, Class, and Family Life*. University of California Press.

Manski, C. F. (2004). Measuring expectations. *Econometrica* 72(5):1329–1376.

Michalopoulos, C., Faucetta, K., Hill, C. J., Portilla, L. B., Lee, H., Duggan, A., Knox, V., Margie, N. G., and Nerenberg, L. (2019). Impacts on family outcomes of evidence-based early childhood home visiting: Results from the mother and infant home visiting program evaluation. Technical report, Mathematica Policy Research.

Olds, D., Henderson Jr, C. R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettitt, L., Sidora, K., Morris, P., and Powers, J. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior: 15-year follow-up of a randomized controlled trial. *Jama* 280(14):1238–1244.

Putnam, R. (2015). *Our Kids: The American Dream in Crisis*. Simon Schuster.

Rubio-Codina, M., Attanasio, O., Meghir, C., Varela, N., and Grantham-McGregor, S. (2015). The socioeconomic gradient of child development: Cross-sectional evidence from children 6–42 months in bogota. *Journal of Human Resources* 50(2):464–483.

Swamy, P. A. (1970). Efficient inference in a random coefficient regression model. *Econometrica: Journal of the Econometric Society* 311–323.

A Appendix

A.1 Appendix 1. Item Response Theory

In what follows, we use the indexes i and k to denote, respectively, a child in our sample and a word in the MLI-I or MLI-II instrument. We remind the reader that the index j is used to denote one of the three latent factors relating to child development at the beginning of the period, child development at the end of the period, and parental investment. We define the single index $w^*_{i,j,k}$ in the following way:

$$w^*_{i,j,k} = \alpha_{j,k,0} + \alpha'_{j,k,1} z_i + \beta_{j,k} m_{i,j,1} + \epsilon_{i,j,k}$$

where $\epsilon_{i,j,k} \sim N(0,1)$, z_i is a vector of observable variables (gender and age), which are allowed to shift the index $w^*_{i,j,k}$. The variable age is adjusted for the age at which the observation i is done but centering it around ages 18 months, so then, a_i, which denote the logarithm of the child's age, is equal to zero if the observation of the MLI-I is done at age 18 months. The parameters $\alpha_{j,k,0}$ and $\beta_{j,k}$ capture the difficulty level and information content of word k. The variable $\epsilon_{i,j,k}$ is measurement error. With some mild assumptions about the distribution in the cross section of $m_{i,j,1}$, and assumptions about the distribution of error terms, it is possible to identify the parameters $\alpha_{j,k,0}$, $\alpha_{j,k,1}$, and $\beta_{j,k}$ as well as the parameters of the distribution of the factor (see Kotlarski (1967)).

Let $w_{i,j,k} \in \{0,1,2\}$ denote the observed score for child i in word k from the MLI-I. The relationship between the score $w_{i,j,k}$ and the index $w^*_{i,j,k}$ is determined by the following rule:

$$w_{i,j,k} = \begin{cases} 0, & \text{if } w^*_{i,j,k} \leqslant 0, \\ 1, & \text{if } 0 < w^*_{i,j,k} \leqslant c_j, \\ 2, & \text{if } c_j < w^*_{i,j,k} \end{cases}$$

where c_j is the cut-off constant in the ordered discrete variable model. Assume, for now, that we observe $m_{i,j,1}$. Let Φ denote the CDF of a standard normal random variable. Let $G_{i,j,k}$ denote the contribution to the likelihood of observing score $w_{i,j,k}$ for child i in word k is:

$$\begin{aligned} G^{MLI-I}_{i,j,k} = & \left[\Phi\left(-\alpha_{j,k,0} - \alpha'_{j,k,1} z_i - \beta_{j,k} m_{i,j,1}\right)\right]^{\mathbf{1}(w_{i,j,k}=0)} \times \\ & \left[\Phi\left(c_j - \alpha_{j,k,0} - \alpha'_{j,k,1} z_i - \beta_{j,k} m_{i,j,1}\right) - \Phi\left(-\alpha_{j,k,0} - \alpha'_{j,k,1} z_i - \beta_{j,k} m_{i,j,1}\right)\right]^{\mathbf{1}(w_{i,j,k}=1)} \times \\ & \left[1 - \Phi\left(c_j - \alpha_{j,k,0} - \alpha'_{j,k,1} z_i - \beta_{j,k} m_{i,j,1}\right)\right]^{\mathbf{1}(w_{i,j,k}=2)} \end{aligned}$$

Second, note that MLI- II and MLI-III are dichotomous variable. Therefore:

$$w^*_{i,j,k} = \alpha_{j,k,0} + \alpha'_{j,k,1} z_i + \beta_{j,k} m_{i,j,1} + \epsilon_{i,j,k}$$

where $\epsilon_{i,j,k} \sim N(0,1)$ .Let $w_{i,j,k} \in \{0,1\}$ denote the observed score for child i in word k from the MLI-II or MLI-III. It follows that:

$$w_{i,j,k} = \begin{cases} 0, & \text{if } w^*_{i,j,k} \leqslant 0, \\ 1, & \text{if } w^*_{i,j,k} > 0. \end{cases}$$

The contribution to the likelihood can be written as:

$$\begin{aligned} G^{MLI-II}_{i,j,k} &= \left[\Phi\left(-\alpha_{j,k,0} - \alpha'_{j,k,1} z_i - \beta_{j,k} m_{i,j,1}\right)\right]^{\mathbf{1}\left(w_{i,j,k}=0\right)} \times \\ & \quad \left[1 - \Phi\left(-\alpha_{j,k,0} - \alpha'_{j,k,1} z_i - \beta_{j,k} m_{i,j,1}\right)\right]^{\mathbf{1}\left(w_{i,j,k}=1\right)} \end{aligned}$$

and we can define a similar equation for the contribution to the likelihood for $G^{MLI-III}_{i,j,k}$.

The data for MLI-I and MLI-II were collected at baseline. Therefore, depending on the age of the child, we have data for MLI-I or for MLI-II, but never for both instruments for the same child. Let the variable $\chi_{i,j} = 0$ if the observation for child i is from MLI-I and $\chi_{i,j} = 1$ if the observation for child i is from MLI-II. The contribution to the likelihood from child i and word k is:

$$G^{MLI}_{i,j,k} = \mathbf{1}\left(\chi_{i,j} = 0\right) \times G^{MLI-I}_{i,j,1} + \mathbf{1}\left(\chi_{i,j} = 1\right) \times G^{MLI-II}_{i,j,1}$$

In our procedure, it is crucial to select words from the MLI instrument that are informative and have different degrees of difficulty. One approach would be to estimate the MLI IRT by itself and let the maximization algorithm find the optimal values of the parameters of interest. We take a different approach. Because the BSID-III is the "gold standard" in measuring child development, we add the BSID-III to the IRT model for the MLI. In this sense, the parameters of the IRT model are now related to the BSID-III. Therefore, informative items are items that also correlate with the BSID-III. Difficult items are items who only the children with high scores in the BSID-III (once we control for age and gender) can understand and say. Therefore, although we only use the MLI in forming the beliefs elicitation survey questionnaire, the items we choose have some information from the BSID-III. In what follows, we describe how we use the BSID-III in our IRT model.

Unlike the MLI, the BSID-III is a continuous variable. Thus, let $BSID_{i,j,l}$ denote the observed score for child i in BSID-III subscale l. The relationship with the variable $m_{i,j,1}$ is captured by:

$$BSID - III_{i,j,l} = \alpha_{j,l,0} + \alpha'_{j,l,1} z_i + \beta_{j,l} m_{i,j,1} + u_{i,j,l}$$

where $u_{i,j,l} \sim N\left(0, \sigma_l^2\right)$. Therefore, the contribution to the likelihood is:

$$G_{i,j,l}^{BSID-III} = \left(\frac{1}{\sigma_l}\right)\left(\frac{1}{\sqrt{2\pi}}\right)\exp\left\{\frac{\left(BSID-III_{i,j,l} - \alpha_{j,l,0} - \alpha_{j,l,1}' z_i - \beta_{j,l} m_{i,j,1}\right)^2}{2\sigma_l^2}\right\}$$

Finally, the likelihood takes into account the fact that $m_{i,j,1}$ is not observed for any child. Therefore, we must integrate out the distribution of $m_{i,j,1}$:

$$L_i = \int \left[\prod_{k=1}^{K} G_{i,j,k}^{MLI}\right]\left[\prod_{l=1}^{L} G_{i,j,l}^{BSID-III}\right] f(m)\, dm \tag{14}$$

And the likelihood function is

$$L = \prod_{i=1}^{N} \int L_i \tag{15}$$

We can estimate the parameters of the IRT model by maximising the likelihood (15).

Finally, we use the estimates for $\alpha_{j,k,0}$, $\alpha_{j,k,1}$ and $\beta_{j,k}$, to estimate a child-specific $m_{i,j,1}$ by maximising the following function for each child i:

$$L_i = \int \left[\prod_{k=1}^{K} G_{i,j,k}^{MLI}\right] f(m)\, dm \tag{16}$$

Therefore, while we use the BSID-III to carefully select the words we use in the beliefs elicitation survey instrument, we do not use the BSID-III to predict the child's Bartlett scores in the MLI instrument. We follow a similar procedure for the IRT model for the MLI-III.

We also estimate an IRT model for parental investments (materials and activities). However, in the estimation of parental investments, we do not use the BSID-III in any way. Similar to the analysis of the MLI, we use the output of the IRT model to describe the scenarios of investments and to produce Bartlett scores of investments in time and material. These scores are then used as error-ridden measures of parental investments in the estimation of the objective production function. We have a third measure, directly reported by parents, that measure the amount of time parents interact with the children on a given day.

A.2 Appendix 2. Beliefs Elicitation Survey Instrument

The final beliefs elicitation survey instrument was designed after extensive piloting in which alternative wordings of the beliefs elicitation questions were tried. In particular, we pilot multiple wordings to define a child with low or high levels of child development at the beginning of the period, low and high levels of parental investments and levels of child development at the end of the period. During the pilot, we also try different ways to elicit expectations of child development: "What is the probability of a hypothetical child can say...?" "What is the minimum and maximum age a hypothetical child can say...?", "What is the minimum, the medium and maximum age a hypothetical child can say...?" among others.

Before asking to mothers of each child of our sample the expectation about child development for different scenarios for a hypothetical child, mothers were trained in the equipment (wooden tablets) used for the elicitation questions. For do so, we designed a practice module. Figure A.1 shows the scenarios presented to mothers for which the aim is to elicit maternal subjective beliefs about the effect of nutrition conditions (high and low: top and bottom in the diagram) on what age a hypothetical baby (aged 4 months) would start to crawl, walk and run. The use of different groups of physical activities with different levels of difficulty allows us to investigate and address measurement error in maternal responses about expected levels of physical activities. To indicate their answers, mothers used wooden tablets that had been marked with different ages (from 6 to 48 months) at the top and that contained a number of strings with a bead. Figure A.1 shows the two-scenarios used in the practice module tablets for which the mothers reported the age a hypothetical child would start to crawl, walk and run. For each set of physical activities and scenario there was a string with a bead (in total 6 strings). The mother was asked to put the bead at the age at which the hypothetical child would be able to crawl, walk and run under a given scenario. During these practice questions, the interviewer would point out to inconsistencies, if, for instance, the mother would indicate that the hypothetical baby would start to run before starting to crawl or she would indicate that a malnourished child (low nutrition) would run before a well fed one. The point of this exercise was to familiarise the mothers with an instrument that is not standard in field work and especially with a population with low levels of education.

normal child without delays in development	high nutrition	to crawl
		to walk
		to run
	low nutrition	to crawl
		to walk
		to run

Figure A.1: Beliefs Elicitation Survey Instrument: Practice module

The main module of the instrument to elicit maternal beliefs asks mothers to report the expected level of child development at the end of the period. For each of the four scenarios presented in Figure A.3, we asked the mother at what age a hypothetical child would start saying 3 sets of words in that particular scenario. To explain the scenarios to each mother, Figure A.2 was presented when explaining the combinations of scenario pairs s in the set $S = \{(H_0^L, X^L), (H_0^L, X^H), (H_0^H, X^L), (H_0^H, X^H)\}$. The scenario (H_0^L, X^L) is represented by the low amount of time doing activities like the ones shown at the left in Figure A.2 and low didactic materials like the ones shown at the right bottom in Figure A.2, on the opposite, the scenario (H_0^H, X^H) is represented by the high amount of time doing activities like the ones shown at the left in Figure A.2 and high didactic materials like the ones shown at the right top in Figure A.2.

Figure A.2: Parental Investment: $S = \{(H_0^L, X^L), (H_0^L, X^H), (H_0^H, X^L), (H_0^H, X^H)\}$

The three set of words were chosen to be easy, more difficult and even more difficult. To indicate their answers, mothers used wooden tablets that had been marked with different ages (from 9 to 48 months) at the top and that contained a number of strings with a bead. For each set of words and a scenario there was a string with a bead. At the end of the exercise each mother was presented with two wooden tablets (left and right diagrams from Figure A.3) with the 12 strings and beads and was asked whether she would want to revise any of the questions.

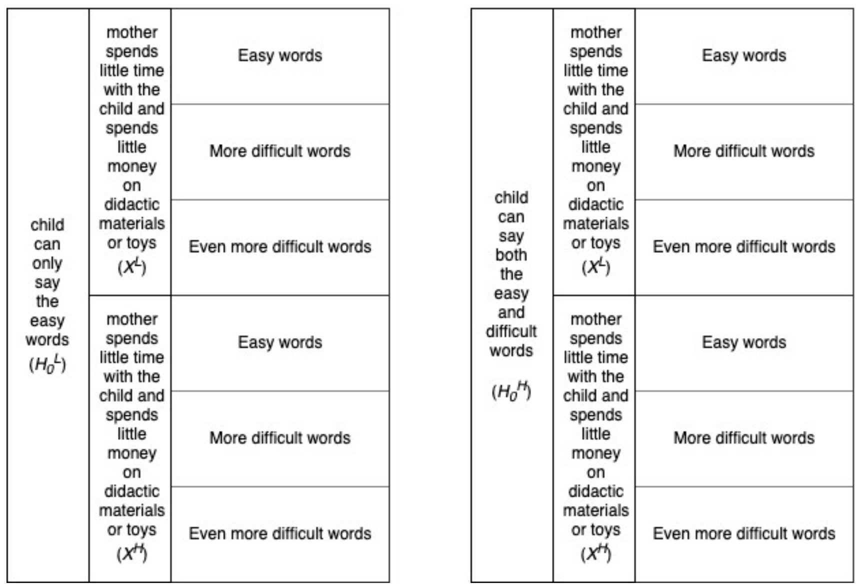


Figure A.3: Beliefs Elicitation Survey Instrument: child development elicitation instrument

A.3 Appendix 3. Estimation of the Subjective Technology of Skill Formation

Here we describe how we ensure that the metric used to construct the hypothetical scenario and their outcomes is comparable to that used to scale the data used to estimate the objective production function.

A.3.1 Using the measurement system to design scenarios

The beliefs questions are framed to elicit mothers' beliefs about the age at which a hypothetical child can understand or say three groups of three words, chosen on the basis of the difficulty and salience parameters of an estimated latent factor model under different scenarios. To convert the answers to such a question into a developmental index we use the factor model estimated on the baseline data, as explained in section 2.2. As we mentioned above, such a model includes all the MLI words and the BSI-III.

The relationship between child i 's development and their ability to say word j_k is described by a probit (or an order probit) model, where this ability is affected the factor θ_i through an index $m_i^{j_k *} = \alpha^{j_k} + \beta^{j_k}\theta_i + \epsilon_i^{j_k}$, where $k = e, m, h$ refers to words that are easy, medium and hard and j_k refers to the specific words in category k.

Given the words used in the beliefs questions, we define a variable d_i^k for child i to be equal to 1 if the child can say all the three words considered in category k. The probability that such a variable is 1 for child i with development θ_i will be given by:

$$\Pr\left(d_i^k = 1\right) = \left[\prod_{j=1}^{3}\left[1 - \Phi\left(-Z_i^{j_k}\right)\right]\right] \tag{17}$$

Let $\hat{\theta}_k$ denote the prediction of the factor θ implied by our latent factor model when $\Pr\left(d^q = 1\right) = 0.5$. Notice that, given the scaling we have described in section 3.1.1, this prediction is measured in terms of the developmental age at which the median child already know the words with difficulty level q :

$$\ln \tau_q = \hat{\theta}_q \tag{18}$$

Equation 18 is key for converting the elicited ages under different scenarios to developmental outcomes. We focus on the estimates obtained using equation (17), which uses the words in the latent factor model that are included in the beliefs elicitation questions.

A.3.2 Developmental delays and perceived rates of return to investment

Up to now we have not used the data on elicited beliefs. We now transform maternal answers to the beliefs questions into measurements of maternal beliefs about child development given a specific scenario, expressed in terms of the developmental age, as defined in the previous section.

As discussed above, we have four possible scenarios, $s = 1, 2, 3, 4$, and, for each scenario, we have questions about the age at which the child will be able to use three different sets of words, which differ in terms of their difficulty, $q = e, m, h$. Let $a_{i,s,q}$ denote the age reported by mother i for the set in which the word difficulty-level q and the scenario is s. Define maternal beliefs about the developmental delay for scenario s and word difficulty-level q, $d_{i,s,q}$, as follows:

$$d_{i,s,q} = a_{i,s,q} - \tau_q \tag{19}$$

where τ_q is defined by equation (18). For instance, assume that $\tau_e = 21$, so that the median child in our baseline sample has already learned the easy words by age 21 months. Suppose, additionally, that mother i states that for scenario $s = 4$ the hypothetical child she refers to will learn the "easy" words at age 25 months, so that $a_{i,4,e} = 25$. In this example, mother i 's beliefs about developmental delay implied by the 4 th scenario is 4 months, so that $d_{i,4,e} = 4$. In other words, mother

i 's beliefs that this hypothetical child is 4 months behind in terms of the median child.

Finally, given that we have three categories of words (easy, medium and hard $q = e, m, h$), and we define a benchmark age τ_q for each group, we want to re-scale developmental delays in terms of a unique benchmark. Let T denote an arbitrary age, which could be appropriate for our sample. In particular, in what follows we will set $T = 36$. We can then derive the maternal expectations about end-of-period child development for each scenario as follows:

$$\ln H_{i,1,s,q} = \ln\left(\tau - d_{i,s,q}\right) \tag{20}$$

Notice that this procedure yields one measure of developmental age for each set of words, $q \in \{e, m, h\}$ (easy, medium and difficult words). These can be seen as different measures of the same theoretical concepts and we use them as such in the last step of our approach.

The expressions in equation (20) are used to construct subjective beliefs about the return to investment across different scenarios s. In what follows, we discuss individual perceptions of returns to parental investment and their complementarity with initial conditions. We therefore characterise beliefs about the returns to investment when the initial conditions are high and when they are low. The former can be obtained from equation (20) as:

$$r_{i,q}\left(H_0^H\right) = \ln H_{i,1,1,q} - \ln H_{i,1,2,q}, \tag{21}$$

while the latter is computed as:

$$r_{i,q}\left(H_0^L\right) = \ln H_{i,1,3,q} - \ln H_{i,1,4,q}. \tag{22}$$

A.3.3 Consistency of Scenarios and Data

Before we describe the estimator we use to recover individual belief parameters, we briefly provide additional details that explain an important feature of our analysis to ensure comparability in location and scale between inputs in the estimation of the objective technology of skill formation and the hypothetical scenarios we use in the estimation of subjective parameters.

The scenario $\ln H^{s_h}$ is derived from words that we selected from the MLI instrument. However, the MLI is only one of the four measurements of child development at the beginning of the period. Thus, to impose consistency in location and scale, we proceed in the following way.

First, we compute the values for the scenarios from the MLI using the IRT analysis.

Second, we conduct an IV regression in which the BSID scores are the dependent variable and the MLI is the predictor variable. The IV regression is necessary because, under the assumptions of our measurement system, the OLS estimator is biased and inconsistent. We then use the results from the IV regression to predict the values for the scenarios for BSID.

Third, we predict $\ln H^{s_h}$ using the same four variables, and the same prediction rule (i.e., Bartlett), that we used to generate factors scores of $\ln H_{i,0}$, which we use in the estimation of the objective technology.

Fourth, we ensure that the location and scale of $\ln H^{s_h}$ are exactly the same as in the factor scores for $\ln H_{i,0}$.

We execute the same steps for the scenarios of $\ln X^{s_x}$. Next, we provide the formulas for the Swamy (1970) estimator.

A.3.4 Swamy estimator

We can use Swamy (1970)'s estimator to obtain, for each household, an efficient estimator for the vector $\mu_i = (\mu_{i,0}, \mu_{i,1}, \mu_{i,2}, \mu_{i,3})$. This approach is efficiently equivalent to a GLS estimator. Let $\widehat{\mu}_i$ denote the Swamy's estimator.

Before we proceed, we introduce some notation. Let $\ln \mathbf{H}_{i,1}$ denote the vector of maternal subjective beliefs reports. Note that, for each one of the three sets of words q, mothers provide answers to four scenarios s. Therefore, $\ln \mathbf{H}_{i,1}$ is a vector with twelve rows for each mother. Similarly, let $\mathbf{\jmath}_i$ denote the vector of measurement error in maternal reports. Let $Z_s = (1, \ln H^{s_h}, \ln X^{s_x}, \ln H^{s_h} \ln X^{s_x})$ denote the values for the scenarios. Thus, Z_s is a vector with four columns. We can arrange the information in a matrix Z which has twelve rows and four columns. This matrix Z has two properties. First, it does not vary across mothers. Second, the rows 1 to 4 are identical to rows 5 to 8 as well as rows 9 to 12 as the description of the scenarios are identical across word difficulty level q.

Let T denote the number of measures for each mother. Let L denote the dimension of the vector μ_i. In our case, $T = 12$ and $L = 4$. The first step is to run an OLS regression for each $i = 1, ..., N$:

$$\begin{aligned}
\tilde{\mu}_i &= (Z'Z)^{-1}(Z' \ln \mathbf{H}_{i,1}) \\
\tilde{\sigma}_i^2 &= \frac{\tilde{\eta}_i' \tilde{\eta}_i}{T - L} \\
\tilde{V}_i &= \tilde{\sigma_i^2}(Z'Z)^{-1}
\end{aligned}$$

The estimators $\tilde{\mu}_i$, $\tilde{\sigma}_i^2$, and $\tilde{V}_i$ are used as inputs in the second step which computes $\bar{\mu} = \frac{1}{N}\sum_{i=1}^{N} \tilde{\mu}_i$ and estimates:

$$\begin{aligned}
\hat{\Upsilon} &= \frac{1}{N-1}[\sum_{i=1}^{N} \tilde{\mu}_i \tilde{\mu}_i' - N\bar{\mu}\bar{\mu}'] - \frac{1}{N}\sum_{i=1}^{N} \tilde{V}_i \\
\hat{\Pi}_i &= \tilde{\sigma}_i^2 I + Z\hat{\Upsilon}Z'
\end{aligned}$$

Then, we can efficiently estimate the first two moments of the distribution of the vector μ_i: the mean, $E(\mu_i)$, and the variance, $Var(\mu_i)$:

$$\begin{aligned} E(\mu_i) &= (\sum_{i=1}^{N} Z'\hat{\Pi}_i Z)^{-1}(\sum_{i=1}^{N} Z'\hat{\Pi}_i \ln \mathbf{H}_{i,1}) \\ Var(\mu_i) &= \frac{1}{N}\sum_{i=1}^{N}(\hat{\Upsilon}+\tilde{V}_i)^{-1} \end{aligned}$$

The third step is to derive an efficient linear estimator of $\hat{\mu}_i$. To do so, we follow Judge et al. (1985). Let $A_i = \left(\hat{\Upsilon}^{-1}+\tilde{V}_i^{-1}\right)^{-1}\hat{\Upsilon}^{-1}$:

$$\begin{aligned} \hat{\mu}_i &= E(\mu_i) + \hat{\Upsilon}Z'\left(\tilde{\sigma}_i^2 I + Z\hat{\Upsilon}Z'\right)^{-1}(\ln \mathbf{H}_{i,1} - ZE(\mu_i)) \\ Var(\hat{\mu}_i) &= Var(\mu_i) + (I - A_i)\left[\tilde{V}_i^{-1} - Var(\mu_i)\right](I - A_i)' \end{aligned}$$

B Estimation of the Objective Technology of Skill Formation

In this appendix, we describe the procedure to estimate the technology of skill formation objectively. Our goal is to compare subjective beliefs against objective estimates of the equation (11) in Section 6.2. There are two problems that we must address: measurement error (in factor scores) and the endogeneity of investments.

B.1 Prediction of Factor Scores

Let $m_{i,\ln H_0,k}$ denote continuous error-ridden measures of the natural log of child development at the beginning of the period produced by child assessment instrument k, for $k = 1, ..., K$. As described in Section 3.2, we have three continuous subscales of the BSID-III (cognition, expressive language, and receptive language), and the estimation of the MLI data with the IRT model in Appendix A.1 produces an additional continuous scale of expressive language development. Therefore, we have $K = 4$ continuous measures of child development at the beginning of the period. Furthermore, we have fixed the measures' location and scale in age-equivalent scores:

$$\mathbf{m}_{i,\ln H_0} = \Gamma_{\ln H_0,0} + \Gamma'_{\ln H_0,1} z_i + \Gamma_{\ln H_0,2} \ln H_{i,0} + \Xi_{i,\ln H_0} \tag{23}$$

We write the equation in vector form, so $\mathbf{m}_{i,\ln H_0}$ is a vector of dimension $K = 4$ and so are $\Gamma_{\ln H_0,0}$, $\Gamma_{\ln H_0,2}$. $\Xi_{i,\ln H_0}$. Also, $\Gamma_{\ln H_0,1}$ is a (4x2) matrix. The same data are available for child development at the end of the period. Thus:

$$\mathbf{m}_{i,\ln H_1} = \Gamma_{\ln H_1,0} + \Gamma'_{\ln H_1,1} z_i + \Gamma_{\ln H_1,2} \ln H_{i,1} + \Xi_{i,\ln H_1} \tag{24}$$

We could fix the intercepts to zero and factor loadings to one in equations (23) and (24) because our measures have well-defined location and scale. However, we adopt a less aggressive approach. Consistent with our methodology, we only do so for the BSID-III expressive language scale, which is also highly correlated with the MLI. Thus, we do not restrict the values of any of the remaining measures of child development at any period.

The IRT analysis of the FCI instrument and time spent with children in activities, explained in Appendix A.1, produces two continuous measures of parental investments. The time use data from the time diaries produces a third measure of parental investments. Therefore:

$$\mathbf{m}_{i,\ln X} = \Gamma_{\ln X,0} + \Gamma'_{\ln X,1} z_i + \Gamma_{\ln X,2} \ln X_i + \Xi_{i,\ln X} \tag{25}$$

The objective estimation of the technology of skill formation has three steps. In the first step, we factor analyze equations (23, 24, and 25) separately. Then, we use estimated intercepts, factor loadings, and variances of the factors, to generate the factor scores:

$$\hat{\theta}_i = \left(\Gamma'_\theta \boldsymbol{\Sigma}_{\grave{}}^{-1} \Gamma_\theta\right)^{-1} \left(\Gamma'_\theta \boldsymbol{\Sigma}_{\grave{}}^{-1} \mathbf{m}_{\mathbf{i},\grave{}}\right) \tag{26}$$

where $\boldsymbol{\Sigma}_{\grave{}}^{-1}$ is the variance of factor $\hat{\theta}_i$ and $\hat{\theta}_i = \ln \hat{H}_{i,0}, \ln \hat{H}_{i,1}, \ln \hat{X}_i$. Note that:

$$\hat{\theta}_i = \theta_i + \left(\Gamma'_\theta \boldsymbol{\Sigma}_{\grave{}}^{-1} \Gamma_\theta\right)^{-1} \left(\Gamma'_\theta \boldsymbol{\Sigma}_{\grave{}}^{-1} \Xi_{i,\theta}\right) \tag{27}$$

Because the measurement error has a mean equal to zero, it follows that the first moment of the Bartlett factor scores is an unbiased estimator of the first moment of the actual factor:

$$\mathbf{E}\left(\hat{\theta}_i\right) = \mathbf{E}\left(\theta_i\right). \tag{28}$$

However, the second moment of $\hat{\theta}_i$ is biased because the variance of the measurement error is not zero. In fact, the bias is a function of the factor loadings and variances of the measurement error, which we estimate as part of the factor analysis:

$$\mathbf{E}\left(\hat{\theta}_i^2\right) = \mathbf{E}\left(\theta_i^2\right) + \left(\Gamma'_\theta \boldsymbol{\Sigma}_{\grave{}}^{-1} \Gamma_\theta\right)^{-1}. \tag{29}$$

B.2 Endogeneity of Investments

Next, consider the policy function for investments:

$$\ln \hat{X}_i = \rho_0 + \rho_1 \ln \hat{H}_{i,0} + \rho_2 d_i + \zeta_i \tag{30}$$

where d_i is the random assignment to treatment and ζ_i is the error term potentially correlated with ν_i. Suppose that we estimate the first-stage regression (30) using the factor scores for investments and human capital at baseline via OLS. Let $P = (\rho_0, \rho_1, \rho_2)$ and $\tilde{R}_i = \left(1, \ln \hat{H}_{i,0}, d_i\right)$. Then:

$$\tilde{P} = (\tilde{R}'\tilde{R})^{-1}(\tilde{R}' \ln \hat{X})$$

Unfortunately, $\tilde{P}$ is an inconsistent estimator of P because $\ln \hat{H}_{i,0}$ is an error-ridden measure of $\ln H_{i,0}$. To fix on the important ideas, note that:

$$\tilde{R}'\tilde{R} = \begin{bmatrix} 1 & \frac{1}{N}\sum_{i=1}^{N} \ln \hat{H}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} d_i \\ \frac{1}{N}\sum_{i=1}^{N} \ln \hat{H}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \ln \hat{H}_{i,0}^2 & \frac{1}{N}\sum_{i=1}^{N} d_i \ln \hat{H}_{i,0} \\ \frac{1}{N}\sum_{i=1}^{N} d_i & \frac{1}{N}\sum_{i=1}^{N} d_i \ln \hat{H}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} d_i^2 \end{bmatrix} \tag{31}$$

The inconsistency in the OLS estimator is due to the term $\frac{1}{N}\sum_{i=1}^{N} \ln \hat{H}_{i,0}^2$ because, as illustrated in (29), the quadratic term includes variation due to the true factor and the measurement error. To account for this bias, let $\Delta_{H_0} = \left(\Gamma_{H_0}' \boldsymbol{\Sigma}_{\mathbf{H_0}}^{-1} \Gamma_{H_0}\right)^{-1}$, and define matrix B_1 as:

$$B_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Delta_{H_0} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{32}$$

Next, define $R'R = \tilde{R}'\tilde{R} - B_1$. Then, let $\hat{P}$ denote the consistent estimator:

$$\hat{P} = (R'R)^{-1}(\tilde{R}' \ln \hat{X}) \tag{33}$$

We use $\hat{P}$ to estimate the parameters of the first-stage regression (30). We use the estimated parameters and the observations on investments, child development at baseline, and the random assignment to treatment to estimate the error term ζ_i:

$$\hat{\zeta}_i = \ln \hat{X}_i - \hat{\rho}_0 - \hat{\rho}_1 \ln \hat{H}_{i,0} - \hat{\rho}_2 d_i \tag{34}$$

Note that $\hat{\zeta}_i$ is a function of $\ln \hat{X}_i$ and $\ln \hat{H}_{i,0}$. Thus, it inherits measurement errors from these variables.

Next, we use $\hat{\zeta}_i$ to account for the endogeneity of investments in the technology of skill formation (11). Let **ffi** $= (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4)'$ and $\Psi_i = \left(1, \ln \hat{H}_{i,0}, \ln \hat{X}_i, \ln \hat{H}_{i,0} \ln \hat{X}_i, \hat{\zeta}_i\right)$, so that we can write the technology of skill formation (11) in the following form:

$$\ln \hat{H}_{i,1} = \Psi_i \delta + \nu_i$$

To save on notation, let $\hat{y}_i = \ln \hat{Y}_i$. Consider the matrix $A = \tilde{\Psi}'\tilde{\Psi}$:

$$A = \begin{bmatrix} 1 & \frac{1}{N}\sum_{i=1}^{N} \hat{h}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i & \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i\hat{h}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \hat{\zeta}_i \\ \frac{1}{N}\sum_{i=1}^{N} \hat{h}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \hat{h}_{i,0}^2 & \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i\hat{h}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i\hat{h}_{i,0}^2 & \frac{1}{N}\sum_{i=1}^{N} \hat{\zeta}_i\hat{h}_{i,0} \\ \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i & \frac{1}{N}\sum_{i=1}^{N} \hat{h}_{i,0}\hat{x}_i & \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i^2 & \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i^2\hat{h}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \hat{\zeta}_i\hat{x}_i \\ \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i\hat{h}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i\hat{h}_{i,0}^2 & \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i^2\hat{h}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i^2\hat{h}_{i,0}^2 & \frac{1}{N}\sum_{i=1}^{N} \hat{\zeta}_i\hat{x}_i\hat{h}_{i,0} \\ \frac{1}{N}\sum_{i=1}^{N} \hat{\zeta}_i & \frac{1}{N}\sum_{i=1}^{N} \hat{\zeta}_i\hat{h}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \hat{\zeta}_i\hat{x}_i & \frac{1}{N}\sum_{i=1}^{N} \hat{\zeta}_i\hat{x}_i\hat{h}_{i,0} & \frac{1}{N}\sum_{i=1}^{N} \hat{\zeta}_i^2 \end{bmatrix}$$

Define:

$$\Delta_{H_0} = \left(\Gamma'_{H_0} \mathbf{\Sigma}_{\mathbf{H_0}}^{-1} \Gamma_{H_0}\right)^{-1}$$
$$\Delta_X = \left(\Gamma'_X \mathbf{\Sigma}_{\mathbf{X}}^{-1} \Gamma_X\right)^{-1}$$

Following our discussion above, each matrix A's cell with a quadratic term carries terms that represent the variation due to measurement error. Therefore, we need to derive these terms and propose a matrix that accounts for this excess variation. In addition, note that $\hat{\zeta}_i$ inherits measurement errors from $\ln \hat{X}_i$ and $\ln \hat{H}_{i,0}$. Therefore, expressions in which $\hat{\zeta}_i$ multiplies $\ln \hat{X}_i$ or $\ln \hat{H}_{i,0}$ also are biased because of the excess variance from these errors. Let $a_{j,k}$ denote the element in row j and column k in the matrix $A = \tilde{\Psi}'\tilde{\Psi}$. All of the following elements are corrupted by excess variation: $a_{2,2}$, $a_{2,4}$, $a_{2,5}$, $a_{3,3}$, $a_{3,4}$, $a_{3,5}$, $a_{4,2}$, $a_{4,3}$, $a_{4,4}$, $a_{4,5}$, $a_{5,2}$, $a_{5,3}$, $a_{5,4}$, $a_{5,5}$. Next, we derive the probability limit of each one of these elements, and use this derivation to construct a matrix B_2 that accounts for the excess variation due to measurement error.

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \hat{h}_{i,0}^2 = \mathbf{E}\left(h_0^2\right) + \Delta_{H_0}$$

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \hat{x}_i \hat{h}_{i,0}^2 = \mathbf{E}\left(x h_0^2\right) + \mathbf{E}\left(x\right) \Delta_{H_0}$$

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \hat{\zeta}_i \hat{h}_{i,0} = -\rho_1 \Delta_{H_0}$$

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \hat{x}_i^2 = \mathbf{E}\left(x^2\right) + \Delta_X$$

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \hat{x}_i^2 \hat{h}_{0,i} = \mathbf{E}\left(x^2 h_0\right) + \mathbf{E}\left(h_0\right) \Delta_X$$

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \hat{\zeta}_i \hat{x}_i = \mathbf{E}\left(\zeta^2\right) + \Delta_X$$

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \hat{x}_i^2 \hat{h}_{0,i}^2 = \mathbf{E}\left(x^2 h_0^2\right) + \mathbf{E}\left(x^2\right) \Delta_{H_0} + \mathbf{E}\left(h_0^2\right) \Delta_X + \Delta_{H_0} \Delta_X$$

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \hat{\zeta}_i \hat{x}_i \hat{h}_{i,0} = \mathbf{E}\left(\zeta x h_0\right) + \mathbf{E}\left(h_0\right) \Delta_X - \rho_1 \mathbf{E}\left(x\right) \Delta_{H_0}$$

$$\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\zeta_i^2=\mathbf{E}\left(\zeta^2\right)+\Delta_X+\rho_1^2\Delta_{H_0}$$

Now, define the matrix B_2 as:

$$B_2=\begin{bmatrix}0 & 0 & 0 & 0 & 0\\ 0 & \Delta_{H_0} & 0 & \mathbf{E}(x)\,\Delta_{H_0} & -\rho_1\Delta_{H_0}\\ 0 & 0 & \Delta_X & \mathbf{E}(h_0)\,\Delta_X & \Delta_X\\ 0 & \mathbf{E}(x)\,\Delta_{H_0} & \mathbf{E}(h_0)\,\Delta_X & \mathbf{E}\left(x^2\right)\Delta_{H_0}+\mathbf{E}\left(h_0^2\right)\Delta_X+\Delta_{H_0}\Delta_X & \mathbf{E}(h_0)\,\Delta_X-\rho_1\mathbf{E}(x)\,\Delta_{H_0}\\ 0 & -\rho_1\Delta_{H_0} & \Delta_X & \mathbf{E}(h_0)\,\Delta_X-\rho_1\mathbf{E}(x)\,\Delta_{H_0} & \Delta_X+\rho_1^2\Delta_{H_0}\end{bmatrix}$$

For the Cobb-Douglas case:

$$B_2=\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & \Delta_{H_0} & 0 & -\rho_1\Delta_{H_0}\\ 0 & 0 & \Delta_X & \Delta_X\\ 0 & -\rho_1\Delta_{H_0} & \Delta_X & \Delta_X+\rho_1^2\Delta_{H_0}\end{bmatrix}$$

Let $\Psi'\Psi$ denote the excess-variation adjusted matrix:

$$\Psi'\Psi=\tilde{\Psi}'\tilde{\Psi}-B_2.$$

Consider the consistent estimator $\hat{\delta}$:

$$\hat{\delta}=\left(\Psi'\Psi\right)^{-1}\left(\tilde{\Psi}'\ln H_1\right)\tag{35}$$

B.3 Monte Carlo

We assessed the quality of our estimation methodology with a Monte Carlo simulation. We set N = 12,000 and we conducted S = 500 simulations of the model. Appendix Table B.1 compares three different estimators. The first is the infeasible estimator that uses the actual factors in the estimation of the model's parameters. The second is the feasible inconsistent estimator that does not account for the excess variance due to measurement error. The third is the feasible consistent estimator that uses the formulas (33) and (35) for the first and second stages, respectively. Note that the infeasible and feasible consistent estimators recover the true parameter values, but the feasible inconsistent estimator does not.

C Accounting for Measurement Error in Beliefs

In this section, we describe how we account for measurement error in beliefs in our analysis of the association between investments and beliefs. First, we describe the Cobb-Douglas case. Second, we describe the translog case.

Table B.1: Monte Carlo Simulation

First Stage							
Parameters		Unfeasible		Feasible Inconsistent		Feasible Consistent	
	True	Mean	SD	Mean	SD	Mean	SD
ρ_0	0.500	0.499	0.030	0.550	0.029	0.498	0.032
ρ_1	0.250	0.250	0.016	0.216	0.015	0.250	0.018
ρ_2	1.000	1.000	0.026	1.002	0.027	1.002	0.027
Second Stage							
Parameters		Unfeasible		Feasible Inconsistent		Feasible Consistent	
	True	Mean	SD	Mean	SD	Mean	SD
δ_0	1.000	0.999	0.041	1.108	0.043	1.003	0.046
δ_1	0.500	0.501	0.022	0.428	0.021	0.500	0.025
δ_2	0.700	0.700	0.029	0.675	0.030	0.698	0.031
δ_3	-0.100	-0.101	0.008	-0.085	0.008	-0.100	0.009
δ_4	N/A	0.693	0.028	0.661	0.029	0.693	0.029

C.1 Cobb-Douglas Case

Let x_i denote parental investments. Suppose that the perceived production function is Cobb-Douglas. Then, consider the following regression equation:

$$x_i = \beta \hat{\mu}_{i,2} + Z_i \pi + \zeta_i \tag{36}$$

Let Ω_i denote the mother's information set. Let s and q denote the scenario (of investments and human capital at baseline) and elicitation item (difficulty level of the words). Note that:

$$\ln H_{i,1,s,q} = \mu_{i,0} + \mu_{i,1} \ln H_{i,0} + \mu_{i,2} \ln X_i + \epsilon_{i,1,s,q}$$

For scenario (1) and word q, this equation reads:

$$\ln H_{i,1,1,q} = \mu_{i,0} + \mu_{i,1} \ln \underline{H}_{i,0} + \mu_{i,2} \ln \underline{X}_i + \epsilon_{i,1,1,q}$$

For scenario (2), it is:

$$\ln H_{i,1,2,q} = \mu_{i,0} + \mu_{i,1} \ln \underline{H}_0 + \mu_{i,2} \ln \overline{X} + \epsilon_{i,1,2,q}$$

Next, consider the difference $\Delta_{i,2,1,q} = \ln H_{i,1,2,q} - \ln H_{i,1,1,q}$. Note that:

$$\frac{\Delta_{i,2,1,q}}{(\ln \overline{X} - \ln \underline{X})} = \mu_{i,2} + \frac{\epsilon_{i,1,2,q} - \epsilon_{i,1,1,q}}{(\ln \overline{X} - \ln \underline{X})}$$

Note that we can define $\Delta_{i,4,3,q} = \ln H_{i,1,4,q} - \ln H_{i,1,3,q}$. Then:

$$\frac{\Delta_{i,4,3,q}}{(\ln \overline{X} - \ln \underline{X})} = \mu_{i,2} + \frac{\epsilon_{i,1,4,q} - \epsilon_{i,1,3,q}}{(\ln \overline{X} - \ln \underline{X})}$$

This derivation shows that $\Delta_{i,2,1,q}$ and $\Delta_{i,4,3,q}$ are measures of $\mu_{i,2}$. The first measure conditions on the low value of H_0, while the second conditions on the high value. In our empirical analyses, we consider three measurement-error models. The first model only uses the first set of measures for the three groups of words. The second model only uses the second set of measures for the three groups of words. The third model uses both sets of moments. However, we also include an error term that is specific to each type of measurement equation. Indeed, we arrive at this model by proposing:

$$\epsilon_{i,1,2,q} - \epsilon_{i,1,1,q} = \left(\ln \overline{X} - \ln \underline{X}\right) \left(\alpha_{2,q}\tau_{i,2} + \nu_{i,2,1,q}\right) \tag{37}$$

Note that $\tau_{i,2}$ is an error term that is specific to the first set of measures. In contrast, we define the error term $\tau_{i,3}$ by using the second set of measures:

$$\epsilon_{i,1,4,q} - \epsilon_{i,1,3,q} = \left(\ln \overline{X} - \ln \underline{X}\right) \left(\alpha_{3,q}\tau_{i,3} + \nu_{i,4,3,q}\right) \tag{38}$$

Therefore, the third model has three factors: $\mu_{i,2}$, $\tau_{i,2}$, and $\tau_{i,3}$. For completeness, we write the full model here:

$$\frac{\Delta_{i,2,1,q}}{(\ln \overline{X} - \ln \underline{X})} = \mu_{i,2} + \alpha_{2,q}\tau_{i,2} + \nu_{i,2,1,q} \tag{39}$$

$$\frac{\Delta_{i,4,3,q}}{(\ln \overline{X} - \ln \underline{X})} = \mu_{i,2} + \alpha_{3,q}\tau_{i,3} + \nu_{i,4,3,q} \tag{40}$$

In addition, each model includes equation (36). Cunha et al. (2005) discuss identification of these models. We estimate their parameters via maximum likelihood.

C.2 Translog Case

Next, suppose that the perceived production function is translog. Then, consider the regression equation:

$$x_i = \beta \left(\hat{\mu}_{i,2} + \hat{\mu}_{i,3} \ln H_{i,0}\right) + Z_i \pi + \zeta_i \tag{41}$$

Again, we aim to estimate the parameter β. The inclusion of $\mu_{i,3}$ requires us to use both sets of differences. Indeed, the model is:

$$\frac{\Delta_{i,2,1,q}}{\left(\ln \overline{X} - \ln \underline{X}\right)} = \mu_{i,2} + \mu_{i,3} \ln \underline{H}_0 + \frac{\epsilon_{i,1,2,q} - \epsilon_{i,1,1,q}}{\left(\ln \overline{X} - \ln \underline{X}\right)} \tag{42}$$

$$\frac{\Delta_{i,4,3,q}}{\left(\ln \overline{X} - \ln \underline{X}\right)} = \mu_{i,2} + \mu_{i,3} \ln \overline{H}_0 + \frac{\epsilon_{i,1,4,q} - \epsilon_{i,1,3,q}}{\left(\ln \overline{X} - \ln \underline{X}\right)} \tag{43}$$

A sufficient condition for identification is to assume that the error term $\epsilon_{i,1,s,q}$ is independent across scenarios s and words q. In this case, the system of seven equations identify the joint distribution of $\mu_{i,2}$ and $\mu_{i,3}$. To understand why the model is identified, manipulate the system above to obtain the following measurement equations:

$$y_{i,2,q} \equiv \frac{\ln \overline{H}_0 \Delta_{i,2,1,q} - \ln \underline{H}_0 \Delta_{i,4,3,q}}{\left(\ln \overline{X} - \ln \underline{X}\right)\left(\ln \overline{H}_0 - \ln \underline{H}_0\right)} = \mu_{i,2} + \frac{\ln \overline{H}_0 \Delta\epsilon_{i,2,1,q} - \ln \underline{H}_0 \Delta\epsilon_{i,4,3,q}}{\left(\ln \overline{X} - \ln \underline{X}\right)\left(\ln \overline{H}_0 - \ln \underline{H}_0\right)} \tag{44}$$

$$y_{i,3,q} \equiv \frac{\Delta_{i,4,3,q} - \Delta_{i,2,1,q}}{\left(\ln \overline{X} - \ln \underline{X}\right)\left(\ln \overline{H}_0 - \ln \underline{H}_0\right)} = \mu_{i,3} + \frac{\Delta\epsilon_{i,4,3,q} - \Delta\epsilon_{i,2,1,q}}{\left(\ln \overline{X} - \ln \underline{X}\right)\left(\ln \overline{H}_0 - \ln \underline{H}_0\right)}, \tag{45}$$

where $\Delta\epsilon_{i,k',k,q} = \epsilon_{i,1,k',q} - \epsilon_{i,1,k,q}$. Unfortunately, $y_{i,2,q}$ and $y_{i,3,q}$ can be highly correlated. For example, if $\ln \overline{H}_0 = 1$ and $\ln \underline{H}_0 = -1$. In this case, the correlation between $y_{i,2,q}$ and $y_{i,3,q}$ is minus one.