NBER WORKING PAPER SERIES

SUBJECTIVE PARENTAL BELIEFS. 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

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. 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
JEL No. I21,I24,O1

ABSTRACT

We study the importance of maternal subjective beliefs about the technology of skill formation in determining parental investments on 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 of 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 were 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 the vast majority of 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 particular, mothers hold low subjective expectations, which means that mothers underestimate the returns to 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
37 Hillhouse Avenue
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
pjervis@dii.uchile.cl

1 Introduction

A large body of research, in developed and developing countries, show that differences in cognitive and non-cognitive skills across socio-economic groups appear early on in the lives of children and remain stable once these children start school. This evidence includes studies from developed countries, such as Cunha et al. (2006) in the United States, and developing countries such as Rubio-Codina et al. (2015) in Colombia. Lags accumulated in the first three years are important 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 interventions, based on stimulation and aiming at changing parenting practices, seem to be effective. 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). 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)).[1] Similar evidence, from different contexts, is accumulating.

To better understand the impacts of these interventions and possibly to design new ones, it is therefore key 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. (2018) then argue that the intervention's short run impacts are to a great extent explained by increased parental investments. 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, they show that a careful mediation analysis that takes into account the endogeneity of investment, can explain with these increases most of the impacts. The salient question, then, is, why do parents targeted by these interventions increase and improve parental investment?

[1] More recently, Michalopoulos et al. (2019) find more modest impacts of programs when implemented at a larger scale. In their study, they estimate that home visitation programs improve the quality of the home environment by 9% of a standard deviation but that these programs do not have an impact on child development as measured by expressive language and socio-emotional skills.

There are at least three channels through which home visitation programs and, more generally, stimulation interventions may influence parental investments. First, home visitors model parental behavior and thus present forms of interaction that are more conducive for positive development. According to this channel, home visitation programs improve child development by encouraging 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 provide information to parents about the importance of early parental investments for child development. 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, low-income 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 child development indicate that early stimulation is essential for subsequent development and that exposure to language and meaningful interactions drive subsequent developments. Under this hypothesis, low-income 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 estimate (dynamic) models where parents are assumed to solve an optimization problem where their welfare depends both on their consumption and on children outcomes and where they "know" the functional form of the technology of skill formation of child development or human capital (Del Boca et al. (2013)). Within such models, parental investment is driven by the nature of the technology of skill formation of human capital, by financial resources and the cost of investment in children 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 tough 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 investment and then relate these estimates to actual parental investment behavior. With our approach, mothers share beliefs about the set of possible technologies (i.e., they agree that a family of technology of skill formation adequately describe the process of child development), but we allow each mother to have individual beliefs about the parameters of the technology of skill formation. Therefore, in our framework, we allow mothers to have the right or the wrong expectation about the productivity of specific inputs. We can recover, for each mother in our sample, the expectation about the returns to parental investments, and we can identify the technology of skill formation that describes their expectations about the child development process (given inputs). Additionally, we can investigate if there is heterogeneity in expectations about the returns to early investments, what variables explain the heterogeneity in expected returns, if this heterogeneity in expectations about returns predict 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 make four important contributions. First, we propose a new tool to elicit information on beliefs by the respondent. Much work went into the design and implementation of the measurement tool and its validation. We wanted a measurement tool that could be implemented and used easily in a large scale survey and, at the same time, allows to derive measures of subjective beliefs about the process of child development that could be compared to actual data from the same population. From a measurement point of view, our approach consists in assuming that parents think that child development depends over a given period on the developmental status of a child and parental investment. We therefore ask mothers to relate different scenarios of children initial development and parental investment to certain developmental outcomes. In particular, we consider *high* and *low* levels of child development at the beginning of the period and *high* and *low* levels of parental investment. With these data we can derive very simply measures of investment returns under two different levels of child development at the beginning of the period. We then put more structure in the data and device an approach that allows to estimate, for each mother in the sample, the parameters of a *subjective technology of skill formation*.

Second, in our data, we can estimate the technology of skill formation objectively. Therefore, we can quantify whether maternal subjective beliefs are biased (or not) by comparing objective estimates with subjective beliefs. We find that maternal subjective beliefs are downward biased and that most mothers have expectations of returns to investments that are too low. Our findings, thus, mirror those of Boneva and Rauh (2018) who elicited maternal returns about early and late investments and find that mothers in the UK have low expectations about the returns to early investments in children.

Third, we relate our maternal subjective to actual parental investments in our data. We find that maternal subjective beliefs about the returns to parental investment correlate significantly with actual investment behavior. As mentioned above, we estimate returns to investments for two scenarios of child development at the beginning of the period. We find evidence that heterogeneity in both expectations predicts heterogeneity in parental investments, but that the correlation is stronger for the return under high level of child development at the beginning of the period.

Our final contribution is to assess whether the parenting stimulation program randomly targeted to half the sample 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 the fact that Andrew et al. (2018) find no differences in parental investments (or in child development) between control and treatment group one year after the end of the program. Our study suggests that the lack of permanent effects in parental investments is because the program did not permanently change maternal subjective beliefs.

The rest of the paper is organized as follows. Section 2 proposes a theoretical framework that highlights the contribution of our paper. Section 3 describes the context of our study. Section 4 presents the methodology to elicit maternal subjective beliefs. Section 5 develops econometric techniques to analyze our data. Section 6 presents our empirical results. Section 7 is the conclusion. Section A presents appendices that provide further details about our study.

2 The Process of Child Development: Subjective Beliefs and Objective Evidence

In this paper, we study a novel dataset that includes information on parental beliefs about the process of child development. The data we use was collected to evaluate a parenting stimulation program within a cluster Randomized Controlled Trial (RCT). We collected information on parental beliefs through a survey instrument, which we discuss extensively in Section 4. The module was informed by a model of parental behavior that incorporates a technology of skill formation whose output is child development measured by multiple skills. Such a model provides a useful conceptual framework that we use to interpret the collected data. We describe it in this section.

While the focus of this paper is not the estimation of parental preferences, it is useful to consider explicitly the problem faced by parents. We assume parents maximise an objective function and that parental preferences depend on household consumption (C_i), child development at the end of the period ($H_{i,1}$), and, possibly, parental investment (X_i). The direct dependence of preferences on parental investment is not standard but captures the potential psychic benefits of investing in children. We assume that parents' objective function is Cobb-Douglas in the three arguments:

$$U\left(C_i, H_{i,1}, X_i\right) = \ln C_i + \lambda_i \ln H_{i,1} + \kappa_i \ln X_i$$

The maximisation problem faced by parents is subject to a budget constraint. Let P_i and Y_i denote, respectively, the relative price of parental investment and household income. As we assume this is a one-period model without saving or borrowing, we write the budget constraint as:

$$C_i + P_i X_i = Y_i \tag{1}$$

In addition to the budget constraint, parents face the constraint imposed by the process of child development. Let $H_{i,0}$ denote the child development at the beginning of the period. Let ϵ_i and ν_i denote zero-mean variables that are, respectively, known and unknown by the parent at the time that parental investments are chosen. Both shocks are unobserved by the researcher. We assume that the technology of skill formation is a translog function:

$$\ln H_{i,1} = \delta_0 + \delta_1 \ln H_{i,0} + \delta_2 \ln X_i + \delta_3[\ln H_{i,0} \ln X_i] + \epsilon_i + \nu_i \tag{2}$$

The technology of skill formation in equation (2) reduces to a Cobb-Douglas when $\delta_3 = 0$. Cunha et al. (2013) assume that the technology of skill formation is a constant elasticity of substitution (CES), which can also yield, under certain assumptions, a Cobb-Douglas function. Attanasio et al. (2018) also use a CES technology of skill formation but consider different types of parental investments (time and material investment). They cannot reject the Cobb-Douglas specification and find that in the case of cognition, only material investment plays an important role. In our model, we consider only one dimension of child development (a combination language and cognition skills) and one dimension of parental investment (also a combination of materials and activities/time). These simplifications are driven by the nature of the beliefs data collected but, given the results obtained by Attanasio et al. (2018) on the same data, are not particularly stringent.

Equation (2) represents the actual process of child development. In standard models, parents are assumed to know such a process when making decisions about the allocation between consumption and investment goods. We relax this assumption. Instead, while we do assume that parents' beliefs about the technology of skill formation can be well approximated by the *right functional form*, we do not assume that they know the correct values for the parameters of such a technology. That is, we do not assume that the vector $\delta = (\delta_0, \delta_1, \delta_2, \delta_3)$ is in the information set of the parents. Therefore, parents do not necessarily observe or know the "true" technology of skill formation (the *objective technology of skill formation*), but have beliefs about it.

Let $\Omega_i = \{Y_i, P_i, \mu_{i,0}, \mu_{i,1}, \mu_{i,2}, \mu_{i,3}, \kappa_i, \lambda_i, \epsilon_i\}$. The set Ω_i is a subset of the parent's information set. We assume that parent i believes that the technology that produces child development in period $t = 1$ as a function of child development in period $t = 0$ and parental investment is:

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

Equations (2) and (3) play different roles in the model of child development and parental investment. The objective technology of skill formation in equation (2) describes the actual process of child development, given child development at the beginning of the period, $H_{i,0}$, and parental investment, X_i. Equation (3) and its parameters represent subjective beliefs about the process and are used by parents to determine parental investment choices.

The objective technology of skill formation, equation (2), can be estimated from actual data on child development at $t = 0$ and parental investment. The challenge one faces in obtaining consistent estimates of the parameters of the technology of skill formation is the fact that some of its arguments, and in particular parental investment, are chosen by parents and are therefore likely to be correlated to the residual terms in equation (2).

We assume parents maximise expected utility that depends on their consumption, child development and posssibly on investment:

$$\max_{X_i} E\left[\ln C_i + \lambda_i \ln H_{i,1} + \kappa_i \ln X_i \middle| \Omega_i, H_{i,0}, X_i\right]$$

subject to (1), (2) and (3).

From this problem it is possible to derive the following investment policy function which determines parental choices:

$$X_i = g(\Omega_i, H_{i,0}) \tag{4}$$

The investment equation (4) is a function of preference parameters (λ_i and κ_i), beliefs parameters ($\mu_{i,2}$ and $\mu_{i,3}$), child development at the beginning of period ($H_{i,0}$), income (Y_i) and prices (P_i). As our data comes from the evaluation of a stimulation intervention, among the determinants of investment we also have a dummy variable that indicates random assignment to control ($d_i = 0$) or treatment($d_i = 1$). Equation (4) makes clear that investment behavior is driven by a combination of parental preferences and subjective parental beliefs about the parameters of the technology of skill formation.

Most studies in the literature on child development do not have data on preferences or beliefs. Previous studies explored information on X_i, $H_{i,0}$, Y_i, P_i, or d_i to estimate an equation such as (4) while assuming that parents knew the parameters of the objective technology of skill formation. As we discuss below and as argued in Attanasio et al. (2018), it is possible to estimate the parameters of the objective production function in equation (2) without using data on parental beliefs, even when the latter differ from the former, that is when parents might have distorted views of the process of child development. However, without any other information, it is not possible to identify separately the μ_i's, which represent *subjective beliefs* about the process of child development, from the preference parameters κ_i and λ_i.

To get consistent estimates of the parameters that characterise the process of skill formation in equation (2) we use an IV approach. It is therefore essential that the policy function (4) for investment contains valid instruments. However, the approach we use does not necessarily require the correct specification and estimation of the investment function or the assumption that parents know the "true" technology of skill formation or even data on beliefs.

We describe the methodology we use in Section 5.3. The main intuition is simple and built on what was proposed by Williams et al. (2019). We first use a measurement system to obtain from available measures estimates of the latent factors (on child development and parental investment) that enter in equation (2). Then, we use the randomness of the allocation as an IV to obtain the parameter estimates from these factors. The randomness of the allocation, combined with the result found in Attanasio et al. (2018), who argue that the parenting stimulation program did not affect the technology of skill formation, but induced parents to change their parental investment behavior, provides us with a valid instrument to obtain consistent estimates of the objective technology of skill formation (2).

3 The Data: Origin and Content

In this section, we present the data we use in our study. First, we discuss its origin. Second, we describe the measures of child development and parental investment contained in the data. The measures are key to the design of the beliefs questions we discuss next.

3.1 The Evaluation of a Parenting Stimulation Program

As mentioned above, we use data that were collected to evaluate the impact of a parenting stimulation program aimed at fostering the development of young children living in poor families in Colombia. The basic structure of the program was guided by the Jamaica study of early years parenting stimulation by Sally Grantham-McGregor Grantham-McGregor et al. (1991). The 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 used the infrastructure of an existing welfare program and community women to deliver the stimulation component, in an attempt to test a scalable version of the program. The Jamaica curriculum (*Reach-Up*) was adapted to the Colombian context. The original curriculum aimed to promote child development in an integrated manner (language, cognitive, motor and socio-emotional skills) and to encourage caregivers to teach their children based on events surrounding daily routine activities. The curriculum was based on picture books, pictures to stimulate conversation, puzzles, cubes/blocks and patters, toys from recycled materials and 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 4 groups (Stimulation, Supplementation (micronutrients), Stimulation + Supplementation and Control) was done across towns, to avoid contamination of the control group.

The parenting stimulation program had significant impacts on a variety of 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, haemoglobin, 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 was collected using a general household survey. These included data on socioeconomic status, education, labour supply, time use, reproductive history, health conditions, depression, knowledge on parenting, parenting practices and the home environment, among others.

Attanasio et al. (2018) show that these early impacts on child development were, to a large extent, explained by an increase in parental investment. One possibility to explain 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. To investigate such a hypothesis, data on parental beliefs and on how they relate to parental investment can be very useful. 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 into the study were contacted again, to collect information on the medium-run impacts of the program, which are described in Andrew et al. (2018).[2] The follow-up two, that 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 Section 4.5. The survey instruments were administered to the primary caregivers, who were 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 that were collected in the baseline and first follow-up surveys of the evaluation study. Details on these data are important for this paper as we used them to design the beliefs elicitation survey instrument we describe in Section 4.5 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 (Bayley (2006)) and the short versions of the MacArthur-Bates Communicative Development Inventory (Jackson-Maldonado et al. (2012)). We use these instruments from the baseline and the first follow-up surveys to estimate, respectively, (the log of) child development at the beginning of the period ($\ln H_{i,0}$), and end of the period ($\ln H_{i,1}$).

The BSI-III is considered the gold-standard assessment of child development for children below age 42 months. It measures cognition, expressive language, receptive language, fine motor, 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. MLI-I is appropriate 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

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

of 100 words prompted by the interviewer, the parents report if the child "says the word" or "does not say the word" that is 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 100 words prompted by the interviewer, the parents report if the child "says the word" or "does not say the word" that is asked by the interviewer. The MLI-III was collected in follow-up one survey.

There are several important differences between the MLI and the BSID-III. First, the former is based on parental report while the latter is scored based on direct observation of the subject child. Second, the MLI was usually administered at the primary caregiver house by the interviewer that collected the household survey, while the BSID-III was administrated 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 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. We have a third source of data, directly reported by parents, that measure the amount of time parents interact with the children on a given day.

4 Elicitation of Maternal Subjective Beliefs

In this section, we describe the methodology we use to develop the survey instrument to elicit information on maternal beliefs about the technology of skill formation. In particular, we discuss how we used existing data on child development and parental investment to design the beliefs questions. We start with an intuitive explanation of our approach. We then provide the details of the techniques we use in the various steps to design the beliefs questions.

4.1 An intuitive summary

A first and important part of our approach consists in using the available data to relate the latent variables of our conceptual framework, child development and the parental investment, to observable variables. In formulating the beliefs questions, we assume that mothers use the same mapping from the relevant latent variables (child development and investment) to observable variables. Therefore, in our measurement tools, we want to include variables that, from the point of view of the mothers, are salient indicators of child development and parental investment. The relationship between child development and observable measures is important as we will explore it to estimate the process of child development, that is, the equation that describes the causal relationship between inputs and child developmental outcomes. As we want to compare the estimates of the process of children cognitive development with parental beliefs, it is crucial that the developmental metric used in the actual data is comparable to that used with the beliefs data.

In the conceptual framework we discussed in Section 2, child development 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 level of child development at the previous age $a-1$ and, among other environmental factors, on parental investment, X_i. Equation (2) represents such a process. In section 4.3.2, we explain the *measurement system* which relates the latent variables in our conceptual framework (and our representation of maternal beliefs) to observable variables and discuss how we estimate it.

In eliciting subjective maternal beliefs about the process of child development, we posit that mothers believe that child development at the end of the period depends on child development at the beginning of the period, parental investments, and other variables. However, we do not assume that mothers know the "true" process of child development. To elicit beliefs about the developmental process, we present mothers with several "scenarios" - that is, pairs of parental investment and child development at the beginning of the period. For each scenario, we ask mothers to report expected child developmental outcomes. Both the scenarios and the outcomes are formulated in terms of observable and, as we argue below, salient variables. In section 4.3.1, we explain the IRT model which aggregates the individual items of child development and parental investment measures into continuous scores. We note that the measurement systems we estimate plays two roles. On the one hand, we use them to design the beliefs questions in that they allow us to use well understood language and parental investments which is related to child development. On the other, we use to estimate from available data latent factors, which in turn are used to estimate the process of child development. To make the process estimates comparable to the estimates derived from the beliefs question we need to make sure we use the same metric in estimating the relevant latent factors.

Armed with estimates of the measurement systems that relate latent factors (child development and investment) to observable variables, we can convert both the scenarios and the outcomes into estimates of the relevant latent factor that appear in our framework. This approach allows us to map out the following function:

$$E[\ln H_{i,1}|\Omega_i, H_0, X] = f^i(H_0, X) = E[f^i(H_0, X) + \epsilon_i|\Omega_i, H_0, X] \tag{5}$$

where i indicates the mother. As i indexes the function f, we let each mother have a different belief about the process of child development that links $H_{i,0}$ and X_i to $H_{i,1}$. Equation (5) can be interpreted as a more general specification of an equation such as (3). The expression after the second equal sign in equation (5) stresses the fact that child development does not depend only on child development at the beginning of the period and parental investment, but also on a variety of other factors. As mothers are likely to be aware of the importance of these additional factors, we frame the beliefs questions referring to a hypothetical child, *explicitly not their child*. Our aim is to obtain information on their subjective beliefs about the role played by parental investment (X) and children's initial conditions (H_0) in the developmental process. ϵ_i represents heterogeneity in the process of child development with zero mean conditional on our exogenously chosen pairs, but it is not zero mean conditional on the actual realization of the pairs for mother i.[3]

Given the structure in equation (5) and the measurement system, the different pairs presented to the mothers describe different scenarios of values of $\ln H_0$ and $\ln X$; the possible answers represent measures of maternal expectations of child development at the end of the period, that is, they are error-ridden measures of $E[\ln H_{i,1}|\Omega_i, H_0, X]$, that correspond to each scenario.

Presenting mothers with different combinations of scenarios, for both child development at the beginning of the period and parental investment, allows us to get a better mapping of the function representing mothers' perception of the process of skill formation, by covering a larger subset of the domain of the relevant function. Moreover, asking several questions regarding different outcomes for each scenario allows us to reduce the influence of measurement error from the beliefs elicitation. However, the larger the number of pairs, the more costly it is for the mothers to answer our beliefs elicitation survey instrument. We design four different scenarios for H_0 and X, corresponding to "low" and "high" values of H_0 and X. Mothers are then asked about the age at which the hypothetical child is able to achieve certain tasks under each alternative scenario. The four scenarios are a compromise between the benefits

[3]This observation provides an intuition of why we use an IV approach to estimate the technology of skill formation objectively, but do not have to do so for the estimation of the subjective technology of skill formation.

and the costs associated with the elicitation of maternal beliefs.

In the rest of this section, we discuss how the scenarios are formulated, how they relate to $\ln H_0$ and $\ln X$ and how the answers given for each scenario relate to $E[\ln H_{i,1}|\Omega_i, H_0, X]$. Before that, however, we discuss how we ensure that we use the same and comparable metric to measure the various factors estimated from data and beliefs questions, and we sketch how the abstract constructs of $\ln H_1$, $\ln H_0$ and $\ln X$ are translated into variables and situations that can be easily understood by the mothers we interview.

4.2 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 that is going to be used 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. In this section, we discuss the child development metric we use.

When measuring development in children younger that 42 months, the Bayley Scales of Infant Development (BSID) are often considered some of the best available measures. Each BSID-III subscale is measured in terms of a raw score based on the number of items for which the examined child receives credit (i.e., 1 point). The number of items that children answer correctly increases, on average, with their age. We exploit this correlation to construct both the location and the scale of the child's latent stock of child development at both periods of our model, $\ln H_{i,0}$ and $\ln H_{i,1}$. In the education literature, these location and scale are known as "age-equivalent score."

Let $BSID^{raw}_{i,j}$ denote the observed raw score for child i in scale j. Let $\ln a_i$ denote the logarithm of the child's age at the time that child development is measured. For each scale, we calculate the mean of the raw score for all the children in our data of a certain age a_i, $mean_{age_i}(BSID^{raw}_{i,j})$. We then regress log age ($\ln a$) on an intercept and these average raw scores. We denote the intercept and slope of this regression, which converts average scores into "age" as π^j_0 and π^j_1 and use them to define the developmental age-equivalent of each child corresponding to scale j.

$$\ln devage_{i,j} = \pi^j_0 + \pi^j_1 \times BSID^{raw}_{i,j} \tag{6}$$

Notice that equation (6) converts the BSID-III score of every child into an estimated log of age-equivalent score. The log of age-equivalent BSID-III score has a meaningful location and scale. We also note that the log of age-equivalent score is invariant to monotonic transformations to the $BSID_{i,j}^{raw}$ score. In our analysis, the log of age-equivalent expressive language BSID-III score is used as an anchor for the other measures of child development. As a result, our factors of child development $\ln H_{i,0}$ and $\ln H_{i,1}$ have both the location and the scale of the BSID-III log of age-equivalent expressive language score and, thus, have cardinality. This is a desirable property because the estimation of the technology of skill formation requires cardinal variables.

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 period and parental investments.

In what follows, the measures of child development we use are all located and scaled according to the log of age-equivalent scores of expressive language scale in the BSID-III. We now present the measurement system that provides a basic framework for our data analysis and for our construction of the beliefs elicitation survey instrument.

4.3 Mapping child development into observable variables

As children reach different degrees of development, they achieve the ability to perform certain tasks, such as understanding certain words and the ability of expressing them. The ability of performing certain task are, therefore, markers of child development. This insight is important both for estimating latent factors that can be used in the estimation of the production function in equation (2) from actual data and for the formulation of the beliefs questions, as we assume that mothers also use the same markers as indicators of child development. 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 child development, in reality or as perceived by mothers is related to a set of observable variables.

In order to relate observed variables to the abstract constructs that enter equations such as (2), 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 $\ln H_{i,1}$, $\ln H_{i,0}$ 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. Crucially, we assume that mothers use a similar model, relating the process of child development to some well-defined variables reflecting skills, such as the ability of a child to understand or express words. Analogously, we use a number of observable variables (purchase of toys and books, activities with the child and time spent in these activities) as markers of parental investment and assume that mothers relate *parental investment* to the same variables observable variables.

In this subsection, we describe the two systems that we estimate. The first, an IRT model, aggregates the individual discrete items of the MLIs, FCI and time use instruments into individual continuous scores which are scaled appropriately to guarantee consistency across the two system. The second is a standard factor model that use the aggregated MLI scores and the subscales of the Bayleys' tests, and the FCI instrument and time to estimate the relevant latent factors.

4.3.1 Item Response Theory Analysis of MLI and FCI

To design scenarios for the beliefs questions we focus on a set of items from the MLI. The MLI score, a measure of expressive language skills, is obtained through a simple summation of the performance in individual words. In the MLI-I, the answers can be assigned a score of 0, 1, or 2. Given that there are 104 words in MLI-I, a child's score can take any value between 0 and 208. In the MLI-II and MLI-III, the answers can be assigned a score of 0 or 1, and because there are 100 words, the score can range between 0 and 100.

We use the MLI as a foundation for our elicitation instrument. For this reason, it would be impossible to summarize scenarios of child development at the beginning of the period if we had to describe how a hypothetical child performs in each one of the MLI words. Additionally, it would be even more difficult to ask mothers to rate a hypothetical child's performance, for each one of the four scenarios, if she had to provide answers to each one of 100 MLI-III words.

For this reason, we need to reduce the number of words that we use to describe the scenarios, and the number of words we ask mothers to state whether the child will be able to say or not for each scenario. One way to reduce the dimensionality is to estimate an Item Response Theory model for MLI-I, MLI-II, and MLI-III.

The IRT analysis we perform serves two purposes. First, with the IRT model, we are able to identify sets of words that are informative and have different degrees of difficulty. We use this information to design our beliefs elicitation survey instrument. We return to this in Section 4.4. Second, from the estimates of the IRT model, we can obtain an MLI score for child i of age t which we denote by $m_{i,t}^{MLI}$. This score is an alternative and more efficient scoring algorithm than the simple sum of correct answers that is normally used. Such a score gives greater weights to words that are more informative and lower weights to words that do not have strong relationship with expressive language skills. The score $m_{i,t}^{MLI}$ is now a continuous measure, derived from the MLI data, that can be used to estimate the measurement system we describe below.

Let $w_{i,j,k}^{*}$ denote the latent performance of the child i with respect to word k in the MLI-j list. We assume that this latent variable is determined by a single index:

$$w_{i,t,k}^{*} = \alpha_{t,k,0} + \alpha_{t,k,}' z_{i,t} + \beta_{t,k} m_{i,t}^{MLI} + \epsilon_{i,t,k} \tag{7}$$

where $z_{i,t}$ is a vector of observable variables, which are allowed to shift the index $w_{i,t,k}^{*}$.

As we discuss below, to estimate the objective process of skill formation, we use data from both the MLI and the BSID-III. To guarantee comparability with our measures of maternal beliefs, we add the BSID-III to the IRT model. To estimate the parameters of the IRT model, however, we fix the intercept and the loading coefficient of the BSID-III, as we use the BSID-III only to provide the scale and location of the latent variable we estimate with the IRT models. When predicting the MLI scores, we discard the BSID-III data. Therefore the scoring algorithm we construct inherits the location and scale (in age equivalence) we fixed for the BSID expressive language subscale, as discussed in section 4.2. This property makes the latent factors that can be derived from the scenarios and from the answers to the beliefs question directly comparable to the estimates of the latent factor used in the estimation of the objective process of skill formation.

We use slightly different types of single index models for three different types of MLIs:

- Dichotomous variables (for MLI-II and MLI-III) $w_{i,j,k} \in \{0, 1\}$: $Prob\{w_{i,j,k} = 1\} = Pr\{w_{i,j,k}^{*} \geq 0\}$;

- Polytomous variables (for MLI-I) $w_{i,j,k} \in \{1, 2, ..., L\}$: $Prob\{w_{i,j,k} = l\} = Pr\{c_{l-1} \leq w_{i,j,k} \leq c_l\}$, where $c_0 = -\infty$;

- Continuous variables (for BSID-III): $w_{i,j,k} = w^*_{i,j,k}$; [4]

We apply the same procedures for the analysis of the FCI instrument and time use that we use as measures of parental investments. In particular, we classify FCI items as measures of didactic materials and measures of activities and time that involve parents and children. Appendix A.1 presents detailed description of the estimation procedure of the IRT models presented in this Section.

The estimation of the IRT model provides a framework to obtain a continuous score of MLI that we can use in the estimation of the objective technology of skill formation. However, the IRT analysis would not be necessary if the only goal of the study was to objectively estimate the technology of skill formation because one could use the MLI raw scores for this purpose. The estimation of the IRT model is useful for several reasons if the aim is to estimate subjective beliefs about the technology of skill formation. First, the model helps identify salient items for the elicitation instrument. Second, given the scenarios and maternal answers, the model provides a way to parsimoniously estimate the scores of the latent variables for the scenarios (log of human capital at the beginning of the period and parental investment) and the score of the latent variable for log of human capital at the end of the period. In what follows, we first show how to operationalize scenarios and then how to use the IRT parameters to construct the remainder part of the beliefs elicitation survey instrument.

4.3.2 A richer measurement system

Consider the latent factor $\theta_{i,j} \in \{\ln H_{i,0}, \ln X_i, \ln H_{i,1}\}$. The index i represents the child and the index j represents time-0 natural log of child development ($\ln H_{i,0}$), natural log of parental investment ($\ln X_i$), or time-1 natural log of child development ($\ln H_{i,1}$). We express the relationship between measure k, $m^k_{i,j}$, and latent factors, $\theta_{i,j}$, as follows:

$$m^k_{i,t} = g_{j,k}(\theta_{i,t}, z_{i,t}, \epsilon^k_{i,t}) \tag{8}$$

where $\epsilon^k_{i,j}$ represents measurement error and $z_{i,t}$ are observable variables that can potentially enter the relationship between child development and its markers represented by equation (8). The superscript k represents the different measures we have for each

[4] In the BSID-III case, the variable $w_{i,j,k}$ denotes the score in subscale k, $k = 1, 2, 3$.

one of the latent factors, including the scores derived from the MLI, the various subscales of the BSID-III, and the subscales of the FCI as well as the data on the amount of time of interaction between parents and children.

Having estimated a model like (8) with a very wide range of observed measurement variables, it is then possible to use the estimated parameters to obtain estimates of the unobservable latent factors, even when we explore only a subset of the measurement variables used to estimate the more general model. This observation will be important for the construction of the scenarios and for the use of the maternal subjective beliefs data.

It is useful to mention our procedure's basic steps.

1. We consider "dedicated" systems where a measurement loads on a single factor. It is worth it to remember that we consider explicitly three factors: natural log of child development at the beginning of the period ($\ln H_{i,0}$), natural log of child development at the end of the period ($\ln H_{i,1}$), and natural log of parental investment ($\ln X_i$);

2. Each latent factor we consider can affect the available measures on its own or in combination with some observable variables $z_{i,t}$ (such as gender);

3. In particular, we assume that the skill registered by measurement variable k, for child i or age t is determined by a single index:

$$m^k_{i,t} = \gamma^k_{t,0} + \gamma'^k_{t,1} z_i + \gamma^k_{t,2}\theta_{i,t} + \xi^k_{i,t}; \qquad \theta_{i,t} = \ln H_{i,0}, \ln H_{i,1}, \ln X_i. \tag{9}$$

   where $z_{i,t}$ is a vector of observable variables, which are allowed to shift the index $m^k_{i,t}$.

4. We assume that the joint distribution of the unobservable factors θ_i is a mixture of log normals, while the measurement error $\xi^k_{i,t}$ is a normal random variable, independent across different measures k.

We label this model as the *measurement system*. We estimate all of its parameters (subject to location and scaling parameters described in Section 4.2) and recover the distribution of the latent factors.

The measurement system we have described is used as an input in the estimation of both subjective and objective technologies of skill formation, but in different ways. For the objective estimation of the technology of skill formation, we use realizations of child development and investments and we divide the estimation in two stages. In the first stage, we estimate the parameters of the measurement system and predict factor scores. In the second stage, we regress the predicted factor scores for $\ln H_{i,1}$ against the predicted factor scores for $\ln H_{i,0}$ and $\ln X_i$ and apply bias-correction formulas. We present the estimation procedure in Section 5.3.

For the subjective technology, we use the same two stages, but instead of realizations of time-0 human capital and investments, we use the data from the hypothetical "scenarios" of time-0 human capital and investments. Additionally, instead of realizations of time-1 human capital, we use maternal reports of expectation of time-1 child development for each one of the "scenarios." We now start to describe the construction of the elicitation instrument.

4.4 Hypothetical Scenarios

The next step in our framework is to define the scenarios that were presented to the mothers. To design such scenarios, we use the data we collected on the families in the first two waves of our survey.[5] Our goal is to provide mothers with concrete verbal descriptions of scenarios for child development at the beginning of the period (time-0) and parental investments.

4.4.1 Hypothetical Scenarios for current Child Development

As we explained in 3.2, we measured child development with two distinct instruments: MLI and BSID-III. In spite of the BSID-III reliability, we build on the MLI instrument to construct the scenarios of child development at baseline. The main reason for such a strategy is because language is very salient to mothers. They associate language acquisition to child development and they seem to be acutely aware of language development (see Weigel et al. (2006), Masur and Gleason (1980), Furrow et al. (1979)). In addition, the MLI provide a stock of items among which we choose (on the basis of the available data) the more effective and salient.

The MLI-I has 104 words. From these we choose a small set of carefully selected words, using estimates of the parameters $\alpha_{j,k,0}$ and $\beta_{j,k}$ from the IRT model in equation (7), to construct a number of scenarios. These parameters are useful because they describe important properties of the elements of the MLI-I instrument. The parameter $\alpha_{j,k,0}$ in the IRT model for j =MLI-I captures the difficulty of word k in the MLI-I

[5] The beliefs questions were asked in the third wave.

instrument, while $\beta_{j,k}$ measures the information content of the same word in the MLI-I instrument. By selecting salient words (high $\beta_{j,k}$) and varying degrees of difficulty (different values of $\alpha_{j,k,0}$), we can use a small set of MLI-I items to construct informative groups of "easy" and "difficult" words. We then define the first scenario - verbal description of H_0^L - as one in which the child can only say the easy words at the beginning of the period (child aged 9 months). We define the second scenario - H_0^H - as one in which the child can understand and say both the easy and the difficult words at the beginning of the period.

4.4.2 Hypothetical Scenarios for Parental Investments

Similarly to the construction of scenarios for child development at the beginning of period, we use the estimated parameters of the IRT model for the FCI instrument and time to define scenarios for parental investments. We identify salient items of such a factor and choose these measures to define *low* and *high* scenarios for parental investment.

We describe the scenario that defines X^L as the one in which the mother spends little time with the child and spends little money on didactic materials or toys. In particular under this scenarios the toys available to the child are defined by simple toys that extremely common, such as a ball or a doll. In contrast, we define the scenario for X^H as the one in which the mother spends more time and money on the child's development. In this scenario, the toys made available to the child include puzzle, construction toys and material for drawing, painting and/or drawings.

The scenarios for parental investment were presented to the mothers in laminated illustrated cards so that the oral 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.

4.5 Beliefs Elicitation Survey Instrument

The instrument to elicit maternal beliefs provides mothers with a hypothetical scenario $s = (s_h, s_x)$, where s_h is a scenario for H_0 and s_x is a scenario for X. The scenario pairs s take value in the set $S = \{(H_0^L, X^L), (H_0^L, X^H), (H_0^H, X^L), (H_0^H, X^H)\}$. Then the beliefs elicitation survey instrument asks mothers to report the expected level of child development at the end of the period, $E\left[\ln H_{i,1} | \, \Omega_i, s\right]$. Specifically, for each of the four scenarios we construct, we asked the mother at what age a hypothetical child would start saying 3 sets of words in that particular scenario. The three set of words were chosen to be 'easy' (first set), 'medium' (second set) and hard (third set) and on the basis of the parameters of an IRT model. We used words from the MLI-II on baseline data for children between 19-24 months. All the chosen words, k, had high loading factors' $\beta_{j,k}$. 'Easy' words had low $\alpha_{j,k,0}$'s, 'medium' words had medium $\alpha_{j,k,0}$'s and 'hard' words had high values of $\alpha_{j,k,0}$'s in the equations of the IRT model for j =MLI-II. For ease of notation, we will label the 'easy' words as having $\{q = e_1, e_2, e_3\}$, the 'medium' words as having $\{q = m_1, m_2, m_3\}$ and the 'hard' words as having $\{q = h_1, h_2, h_3\}$.

For each scenario, we use different group of words to have multiple measures of the expected level of child development at the end of the period $E\left[\ln H_{i,1} | \, \Omega_i, s\right]$. The use of different groups of words with different levels of difficulty allows us to investigate and address measurement error in maternal responses about expected levels of child development.

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. Figure A.3 in Appendix A.2 shows the scenarios $s \in S$ used in the beliefs elicitation survey instrument tablets for which the mothers reported the age a hypothetical child would start saying each set of words. For each set of words and a scenario there was a string with a bead. The mother was asked to put the bead at the age at which the hypothetical child would be able to say a certain 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, before asking these questions the mothers were trained to use the wooden tablets with some practice questions. They were asked at what age a hypothetical baby (aged 6 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 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. We chose this procedure after extensive piloting in which alternative wordings of the questions were tried. We describe the procedure in more detail in Appendix A.2.

The next step in our framework is to describe a methodology that transforms maternal subjective beliefs answers to estimates of $E\left[\ln H_{i,1} \middle| \Omega_i, s\right]$.

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

In this section, we show how we used the answers to the beliefs elicitation survey questionnaire. In particular, we first show how to compute subjective expected rates of returns to parental investment under different levels of child development at the beginning of the period. We then move on to show how to use the subjective maternal beliefs data to estimate, under certain assumptions, the parameters of the *subjective technology of skill formation* for each respondent. Finally, we show how to estimate the parameters of the objective technology of skill formation.

5.1 Subjective Expected Returns to Parental Investment

The mothers in our sample answer the beliefs elicitation survey questionnaire by reporting the age at which a hypothetical child can say three sets of words under different scenarios of parental investments and child development at the beginning of the period. Given that, it is very simple to compute the subjective expected returns to parental investment under different levels of child development at the beginning of the period implied by these answers: one could compute the returns in terms of number of months and average across the gains in months caused by the parental investment for the three sets of words. The interpretation is that a child with high levels of parental investment will reach high performance in language development much faster than a child who receives a lower level of parental investments. The fact that mothers report ages in their answers is convenient because our measures of child development are located and scaled in a developmental age anchor, as explained in Section 4.2.

Slightly more complicated (and more useful) is to express the expected returns to parental investment in terms of an error-ridden measure of the latent factor that, in our conceptual framework, represents the subjective maternal conditional expectation about child development at the end of the period $E[\ln H_{i,1}|\Omega_i, H_0, X]$. To be clear, we represent this error-ridden measure as $\ln H_{i,1,s,q}$ where i indexes the mother, s indexes the hypothetical scenarios, and $q = e, m, h$ indexes the difficulty level of the words. To convert the answers from the beliefs elicitation survey questionnaire into the $E[\ln H_{i,1}|\Omega_i, H_0, X]$-metric, we follow a three-step approach.

In step one, we use the IRT model described in Section 4.3.1 to estimate the log of the age at which the median child learns how to say the set of words of difficulty q. Given these words in the beliefs elicitation survey questionnaire, let $M_{i,q}$ be equal to one 1 if the child can say all the three words considered in category q. The probability that $M_{i,q} = 1$ for child i with latent child development at the end of the period $\ln H_{i,1}$ is:

$$\Pr\left(M_{i,q} = 1 | \ln H_{i,1}\right) = \left[\prod_{w_q=1}^{3} \left[1 - \Phi\left(-\alpha_{j,w_q,0} - \alpha'_{j,w_q,1} z_i - \beta_{j,w_q} \ln H_{i,1}\right)\right]\right] \tag{10}$$

where the parameters $\alpha_{j,w_q,0}$, $\alpha_{j,w_q,1}$ and β_{j,w_q} were estimated in the IRT model.

Let $\ln \hat{H}_{q,1}$ denote the prediction of the factor $\ln H_{i,1}$ implied by our latent factor model when $\Pr\left(M_{i,q} = 1\right) = 0.5$. Note that $\ln \hat{H}_{q,1}$ is the median natural log of child development at the end of the period, measured in age-equivalent location and scale, as described in Section 4.2. Therefore it represents the *developmental age* at which half of the children have learned to say the three words within the difficulty category q. We can invert equation (10) to estimate $\ln \hat{H}_{q,1}$. We then exponentiate it to obtain $\hat{H}_{q,1}$ for the next step.

In the second step, we compare the log age for such a median child with the log of the age the mother reports for a given scenario Let $a_{i,s,q}$ denote the age reported by mother i, for scenario s, and word difficulty level q. We can then define maternal subjective beliefs about the *developmental delay*, $\Delta_{i,s,q}$ as follows:

$$\Delta_{i,s,q} = a_{i,s,q} - \hat{H}_{q,1} \tag{11}$$

As an example, assume that $\hat{H}_{e,1} = 21$, so that the median child learns the easy words at age 21 months. Suppose, additionally, that mother i states that for scenario $s = (H^H, X^L)$ the hypothetical child she refers to will learn the "easy" words at age 25 months, so that $a_{i,(H,L),e} = 25$. In this example, mother i's subjective beliefs about *developmental delay* implied by the (H,L) scenario is the difference between 25 and 21, so that $\Delta_{i,(H,L),e} = 4$. In other words, mother i believes that, under scenario $s = (H^H, X^L)$, this hypothetical child is 4 months behind in terms of child development with respect to the median child in our sample.

Finally, in step three, we obtain an error-ridden measure $\ln H_{i,1,s,q}$ of the latent factor $E[\ln H_{i,1} | \Omega_i, H_0, X]$. We want to make sure that we measure a mother's expected development in terms of a benchmark consistent with our data collection. In particular, as the data for $lnH_{i,1}$ were collected when the children were around 36 months old, we adopt this benchmark for the expectation data. Let τ denote the typical age of children at the time of the assessment of child development at the end of the period. Then,

$$\ln H_{i,1,s,q} = \ln\left(\tau - \Delta_{i,s,q}\right) \tag{12}$$

Given the mapping between different scenarios and developmental age given by equation (12), it is immediate to compute the subjective expected returns to parental investment under different levels of child development at the beginning of the period measured in this metric.

Notice that this procedure yields one measure of maternal expectations about child development at the end of the period, conditional on a scenario of child development at the beginning of the period and parental investment, for each set of words, $q \in \{e, m, h\}$. Therefore, these measures can be seen as different error-ridden measures of our target subjective moment, $E\left[\ln H_{i,1} \middle| \Omega_i, s\right]$, which we use to recover subjective expected returns and to estimate the subjective technology of skill formation. Specifically, we model:

$$\ln H_{i,1,s,q} = E\left[\ln H_{i,1} \middle| \Omega_i, s\right] + \eta_{i,1,s,q} \tag{13}$$

5.2 Estimation of the Subjective Technology of Skill Formation

As we mentioned in Section 4.4, the scenarios we presented the mothers were chosen with the help of the IRT analysis of the data of child development at the beginning of the period and parental investments. These IRT models allow us to compute a score for the log of child development at the beginning of the period ($\ln H_0^{s_h}$) and the log of parental investment ($\ln X^{s_x}$) corresponding to each scenario $s \in S$. Furthermore, we treat $\ln H_{i,1,s,q}$ as three measures of the same (unobserved) expected child development at the end of the period conditional on a given scenario s. In our conceptual framework, mothers know the "right" functional form for the technology of skill formation but not necessarily the "right" parameters. Therefore maternal reports are error-ridden measures of the term in the left-hand side of the subjective technology of skill formation, equation (3), which we reproduce here for convenience:

$$E[\ln H_{i,1}|\Omega_i, H_{i,0}, X_i] = \mu_{i,0} + \mu_{i,1} \ln H_{i,0} + \mu_{i,2} \ln X_i + \mu_{i,3}[\ln H_{i,0} \ln X_i] + E[\epsilon_i|\Omega_i, H_{i,0}, X_i] \tag{3}$$

As we mentioned above, in the elicitation of maternal beliefs, we ask mothers to think of a hypothetical child, not their own children. Furthermore, we describe this hypothetical 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 is one that refers to this hypothetical child. Therefore, we treat maternal reports as information about $E[\ln H_{i,1}|\Omega_i, H_0^{s_h}, X^{s_x}]$. If we combine equations (3) and (13), and condition on the information set that the mother had to answer the elicitation instrument, which simulates a specific scenario $s = (s_h, s_x)$, we obtain:

$$\ln H_{i,1,s,q} = \mu_{0,i} + \mu_{1,i} \ln H_0^{s_h} + \mu_{2,i} \ln X^{s_x} + \mu_{3,i} \left[\ln H_0^{s_h} \ln X^{s_x}\right] + E[\epsilon_i|\Omega_i, H_0^{s_h}, X^{s_x}] + \eta_{i,1,s,q} \tag{14}$$

We assume that $E[\epsilon_i|\Omega_i, H_0^{s_h}, X^{s_x}] = 0$, which implies that the scenarios $s \in S$ are uncorrelated with shocks ϵ_i. This is true for the hypothetical child. This assumption is fairly mild because the scenarios are constant across mothers. [6]

Equation (14) 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. Note that our procedure does not assume that mothers share the same expectations about the returns to investment or about the parameters of the subjective technology of skill formation. This feature is important for our goals because we aim to investigate if the parenting stimulation program affected

[6] However, in the estimation of the objective technology of skill formation, the relevant conditional expectation is $E[\epsilon_i|\Omega_i, H_{i,0}, X_i]$, which we do not assume is equal to zero. In essence, the estimation of the subjective technology of skill formation does not have endogeneity because the scenarios are exogenous and do not vary across mothers. In the objective technology of skill formation estimation, actual parental investments depend on variables that are known by the parents, but unobserved by the econometrician.

maternal beliefs or not. By estimating mother-specific beliefs parameters, we will be able to compare mean beliefs in the control group against those in the parenting stimulation group.

5.3 Estimation of the Objective Technology of Skill Formation

One of the main contributions of this paper is to compare individual subjective beliefs of the process of child development with actual evidence on a sample of children from the same population from which our sample of mothers is drawn. Having explained how we convert the answers mothers give to the beliefs elicitation survey questionnaire into parameters of a *subjective technology of skill formation* at the individual level, we need now to obtain the same set of parameters for a sample of actual children. A number of issues need to be tackled to make the comparison between the "subjective" and "objective" technology of skill formation sensible. First, as mentioned above, we make the relatively strong assumption that maternal subjective beliefs can be represented with a technology of skill formation which has a functional form that is similar to the one we fit to actual data, but does not have necessarily the same parameters. Second, we need to guarantee the metrics used for the latent factors that enter to the "objective" and "subjective" technology of skill formation are the same and have cardinality so we can meaningfully estimate the objective and subjective parameters of the technology of skill formation. Third, while the "subjective" technology of skill formation is obtained manipulating the scenarios exogenously (as we chose them) and asking mothers what outcomes they map into, in the case of actual data, we need to take into account that investment choices are made by parents, possibly in reaction to shocks and the process of child development. In other words, actual investment is endogenous. In this Section, we discuss briefly the last two issues.

To estimate the "objective" technology of skill formation we first estimate a measurement system, effectively using the same data to estimate the IRT model to define the scenarios and get estimates of $H_{i,1}$ in the formulation of the beliefs questions, but using a slightly different way to aggregate the information coming from discrete variables. We describe the procedure in detail in Appendix A.4. [7] Having obtained estimates for the relevant factors we scale them in a way which is consistent with the approach used to estimate the "subjective" technology of skill formation. This location and scaling makes the two sets of estimates comparable.

[7] We first estimate an IRT model to aggregate the items of the MLI measures. We then use a continuous factor model to aggregate this measures with continuous measures from the BSID-III.

As for the endogeneity of parental investment, we use a IV approach which requires the use of a valid instrument. We use the parenting stimulation program, which we know 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 would not be a valid instrument. However, Attanasio et al. (2018), who use a different set of instruments, argue that the observed impact of the parenting stimulation program is not caused by the program directly but is mediated by an induced increase in parental investment. Such a result makes the program, which was randomly assigned in our sample, a valid instrument.

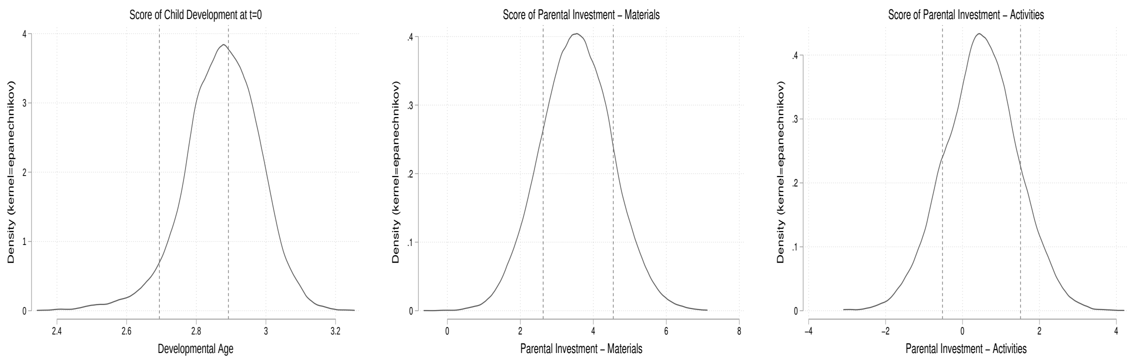
6 Data Analysis and Empirical Results

In this section, we first provide some information about the survey instrument used to elicit beliefs data and show how different scenarios span the domain of the production function of human development, both in terms of initial conditions and parental investment. We then use these data to characterise how maternal subjective beliefs vary in the population, how they are related to parental investment and how, translating them in beliefs about a technology of skill formation, they compare to the relationship between child development on one side and parental investment and child development at the beginning of the period as we estimate in the data.

6.1 Baseline and Follow-up one Data Analysis and the Construction of Scenarios

As discussed in Section 4.5, the scenarios used as a starting point for the beliefs elicitation model were constructed using the estimate of *IRT* models for child development at the beginning of the period and parental investment in materials and activities. In total we consider 4 scenarios that correspond to "low" and "high" values of parental investment and child development at the start of the period considered for the hypothetical child to whom the scenarios refer to. 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: Scenarios for Child Development at the beginning of the period and Parental Investment (Materials and Activities/Time)



6.2 Expected Child Development: Maternal Answers

We now start analysing the data on beliefs by computing average outcomes (the age at which a hypothetical child is able to say *'easy', 'medium' and 'hard'* words) corresponding to different scenarios $s \in S$ in terms of parental investment and child development at the beginning of the period. This evidence provides insights on whether the answers provided by mothers to the beliefs elicitation survey instrument are sensible.

Table 1: Expected Child Development at the end of the period: Maternal Answers

VARIABLES			Mean	St. Dv.	Min	Max
Low Child Development at $t = 0$	Low Investment	easy	18.3	6.3	9	48
		medium	23.4	7.3	10	48
		hard	29.4	8.8	11	48
	High Investment	easy	15.7	5.7	9	48
		medium	20.0	6.7	9	48
		hard	24.9	8.1	9	48
High Child Development at $t = 0$	Low Investment	easy	14.3	4.6	9	45
		medium	17.9	5.5	9	47
		hard	22.2	7.1	10	48
	High Investment	easy	13.5	5.1	9	48
		medium	16.6	5.7	9	48
		hard	20.3	7.1	9	48

Observations: 1200.

Table 1 report the mean, standard deviation and range of the answers to the 12 questions in the beliefs elicitation survey instrument presented in two wooden tablets (left and right diagrams from Figure A.3) with 12 strings each with one bead used to provided the answer. The age range presented to the mothers to choose from was between 9 and 48 months. For each scenario, $s \in S$, we report the relevant statistics for the easy, medium and hard words. The mean for easy words is always below that of medium words 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 the means for levels of high parental investment and the means for low levels of child development at the beginning of the period above the corresponding means for high levels of child development at the beginning of the period. As none of these consistencies were forced by the interviewers on the respondents, this evidence constitutes an indication that, on average, the questions worked, at least in some basic dimension, reasonably well. While at the individual level there are a few inconsistencies, they are few and do not affect the means.

There is a considerable amount of variability in the answers that the mothers provide. The coefficient of variations for all the 12 questions are between 0.3 and 0.4.

Using the answers to the basic questions, we can compute subjective maternal beliefs about the returns to parental investment. For each set of words, $\{q = e, m, h\}$, and for each mother we can compute the returns from moving from "low" to "high" parental investment, expressed in terms of months. That is, for each mother, we can simply compute the difference in the ages at which the hypothetical child start saying a certain set of words under the high and low parental investment scenario. Moreover, we can compute these returns both for children with low and high levels of child development at the beginning of the period. Figure 2 plots the distributions of returns to parental investment in the sample for words of medium difficulty, $\{q = m\}$, conditional on low (dashed line) and high (continuous line) levens of the child's initial development. The latter is clearly to the right of the former. A similar picture is obtained computing the returns using as outcomes words of different level of difficulty or their average.

We note that there are a few mothers for whom the subjective return on investment is negative, signalling probably a problem in understanding the questions. In Table 2, we report that the return is negative for about 14% in the case of low initial condition and in 17.5% in the case of high initial condition. However, as mentioned above, on average, the returns to parental investment are positive. Importantly, the returns to investment are perceived to be higher under low initial child development than high initial child development.

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

Figure 2: Returns to Parental Investment by Scenarios of Child Development at $t = 0$

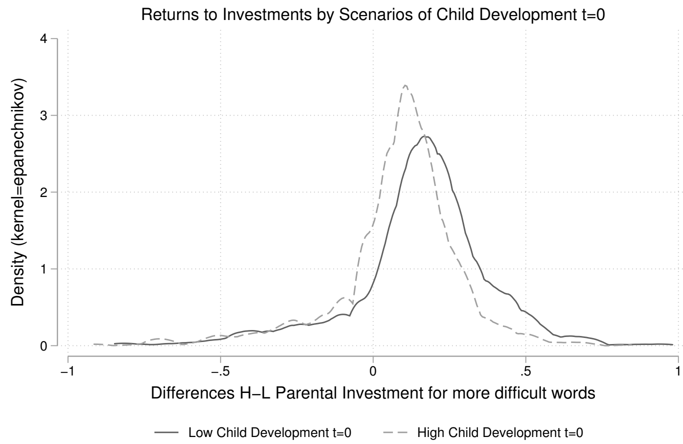


Table 2: Returns to Parental Investment for more difficult words

VARIABLES	Mean	Std. Dev.	% of negative values
Return to Low Child Development t=0	0.16	0.24	13.8
Return to High Child Development t=0	0.08	0.21	17.5

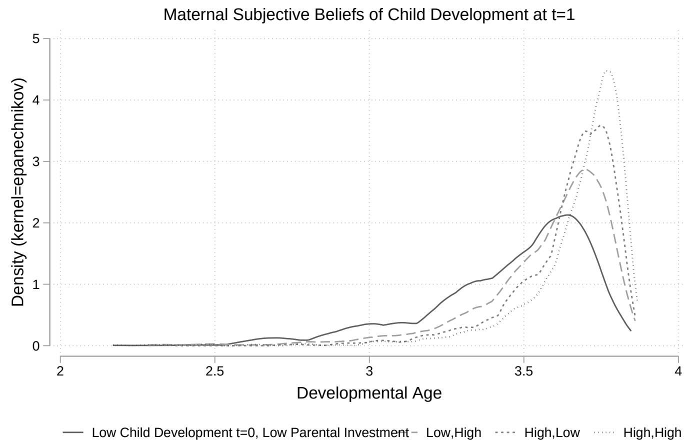
Observations: 1200.

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 5. 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 3 sets of words). As we have three different estimates of the perceived *developmental age* for each $s \in S$ corresponding to set of words,$\{q = e, m, h\}$, we consider their average.

Figure 3 plots 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). Analogously, when moving from low to high child development at the beginning of the period keeping investment constant or moving from low to high investment keeping child development at the beginning of the period constant, the distribution also moves to the right.

We use the expression defined in equation (12) to estimate subjective beliefs about the return to investment across different scenarios s. In particular, we show how child

Figure 3: Maternal Beliefs of Child Development



development at the beginning of the period influences maternal subjective expectations about the returns to investment. For high levels of initial development the return is:

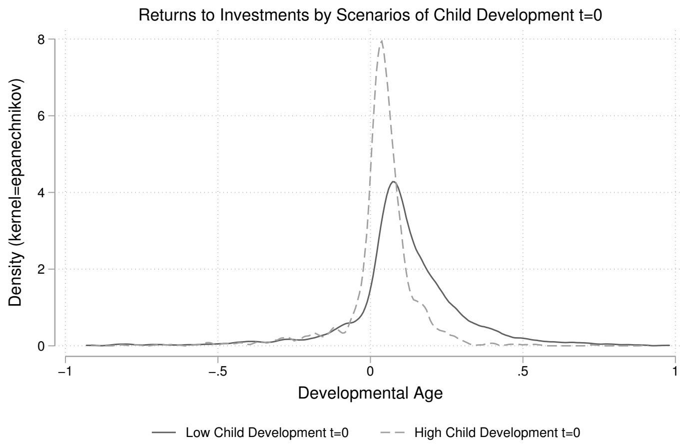
$$r_{i,q}\left(H_0^H\right) = \ln H_{i,1,(H^H,X^H),q} - \ln H_{i,1,(H^H,X^L),q}, \tag{15}$$

while for low initial development is:

$$r_{i,q}\left(H_0^L\right) = \ln H_{i,1,(H^L,X^H),q} - \ln H_{i,1,(H^L,X^L),q}. \tag{16}$$

Having translated for each mother and for each scenario the expected outcome in terms of developmental age, we can, again, compute the returns to investment (under low and high child development at the beginning of the period) using equations (15) and (16). We plot the density distribution 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, and expressed in terms of months, which we plotted in Figure 2. Once again, we notice that the return with low child development at the beginning of the period are, on average, higher than the returns with low child development at the beginning of the period. 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.

Figure 4: Subjective Expected Returns to Investments by Scenarios of Child Development $t = 0$



6.3 How Maternal Subjective Beliefs Vary

As shown in Table 1, maternal subjective beliefs about the process of child development are substantially heterogeneous in our sample. We now study whether they co-vary with mothers' observable characteristics. Table 3 relates the subjective expected returns to parental investment for low and high levels of child development at the beginning of the period to socioeconomic characteristics of the mothers who expressed those subjective beliefs. In our data, in addition to standard variables, such as age and education of the mothers, we have a wealth of other variables. In the table, we report the results obtained regressing subjective expected returns to parental investment (under low and high child development at the beginning of the period) on age, two education dummies (indicators for primary or secondary education, with the no-education being the excluded group), the CES-D index of depression, the score in the Raven progressive matrices test taken by the mothers and an indicator of whether the child in the original evaluation sample of the intervention is male. In the second and fourth columns, we also add a dummy that identifies mothers living in villages targeted by the intervention.

Of the variables considered, the only one that appears to be significantly related to the expected returns on parental investment is the score in the Raven tests, indicating that women with higher Raven tests have higher expected returns to maternal investment, both for low and high initial child development.

Are Maternal Subjective Beliefs about Returns of Parental Investment af-

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

VARIABLES	Return to Low Child Development at $t = 0$		Return to High Child Development at $t = 0$	
Mother's age	0.002*	0.002*	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Mother's education (primary)	0.156*	0.156*	-0.006	-0.006
	(0.088)	(0.088)	(0.052)	(0.052)
Mother's education (secondary and more)	0.151*	0.151*	0.015	0.015
	(0.088)	(0.088)	(0.052)	(0.051)
Mother's depression (CES-D)	0.001	0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Mother's IQ (standardized Raven's Score)	0.015**	0.014**	0.018***	0.018***
	(0.007)	(0.007)	(0.004)	(0.004)
Dummy for Male (child)	-0.018	-0.018	-0.006	-0.006
	(0.012)	(0.012)	(0.008)	(0.008)
Dummy for Treatment		0.008		0.001
		(0.013)		(0.009)
Adjusted R^2	0.014	0.014	0.026	0.026
F	2.582	2.224	4.367	3.751
Observations	1200	1200	1200	1200

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

fected by the Parenting Stimulation Program? Attanasio et al. (2014) have shown that the parenting stimulation program evaluated by the data we are using had an impact on several measures of child development, including cognition and language. Attanasio et al. (2018) analysing the data from the same parenting stimulation program show that the program triggered a large increase in parental investment and argue that such an increase explains a large part if not all the increase in children cognition caused by the parenting stimulation program.

A possible hypothesis, therefore 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. On average, expected returns do not seem to be affected by the parenting stimulation program at all. To look beyond means, using the beliefs data, we plot the distribution of subjective expected returns to parental investment for low and high levels of child development at the beginning of the period for the randomly assignment to control or parenting stimulation program. Figure 5 shows the distribution of subjective expected returns in the two sample is virtually identical.

While this result might be disappointing, Andrew et al. (2018) report that, in the follow-up two, in which the maternal subjective beliefs data were collected, the effect of the parenting stimulation program on parental investment had faded out.[8] It is therefore possible that both the effects of the parenting stimulation program on maternal subjective beliefs and parental investment have faded out, 2-years after the end of the parenting stimulation program. We now turn the analysis of correlations between parental investment and maternal subjective beliefs about its return.

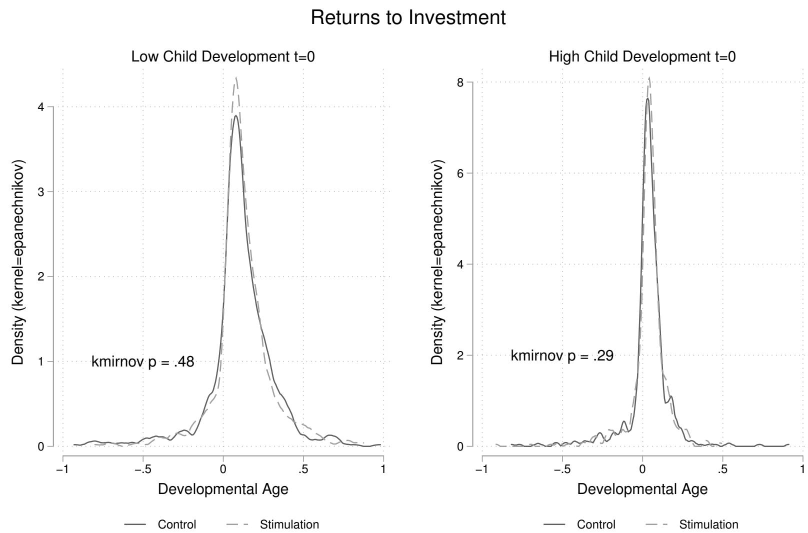
6.4 Does Parental Investment Vary with Maternal Subjective Beliefs?

Having estimated parental beliefs on returns to investment we relate them to investment. We start by considering a reduced form equation where investment is regressed on its determinants, *including its perceived returns*. We consider separately returns conditional on "low" child development at the beginning of the period and "high" child development at the beginning of the period.

As mentioned above, we have three measures of these returns, one corresponding to each set of words $\{q = e, m, h\}$. We run three separate regressions: first, we consider

[8] Andrew et al. (2018) report that also the impact of the parenting stimulation program on child development, as measured 2-years after the end of the parenting stimulation program, had faded out.

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



the average of the three measures of returns; second, we use rgw ewturn estimnated only hard words as the outcome and finally, to take into account that each of the three avaiable measures could be affected by measurement error, we use one measure of expected returns and instrument it with the other two. Table 4 shows that maternal beliefs about high initial child development correlate with maternal investments. The coefficients on the first and third column are much larger than those in the first and thrid, indicating that the measures of returns might be affected by measurement error. The returns on investment under low initial child development do not seem to matter for investment. In column 2, this variable takes a negative and significant coefficient, which is difficult to rationalise.

6.5 Subjective Technology of Skill Formation

As we discussed in Section 2, one can frame parental investment as depending on individual preferences, resources and the *subjective* technology of skill formation. In Section 5.2, we have 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 (3) or equation (14). In this section, we perform such an exercise.

The mother's subjective technology of skill formation is defined by equation (3) or,

Table 4: Investment and Returns on Investment

Dependent Variable: Maternal Investment				
VARIABLES	OLS	OLS[a]	IV	IV[a]
Return to Low Child Development at $t = 0$	-0.285*	-0.289*	-0.214	-0.246
	(0.164)	(0.154)	(0.162)	(0.149)
Return to High Child Development at $t = 0$	0.806***	0.332	0.846***	0.384*
	(0.231)	(0.217)	(0.229)	(0.210)
Controls		yes		yes
Adjusted R^2	0.007	0.117	.	0.115
F	6.670	19.604	6.865	19.432
Observations	1200	1200	1200	1200

[a] Regression controls by mother's age, education, depression and IQ (standardised Raven's score) as well as gender of the child.

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

equivalently, by equation (14), which we reproduce for convenience:

$$\begin{aligned}\ln H_{i,1,s,q} =& \mu_{0,i} + \mu_{1,i} \ln H_0^{s_h} + \mu_{2,i} \ln X^{s_x} + \\ & \mu_{3,i} \left[\ln H_0^{s_h} \ln X^{s_x}\right] + E[\epsilon_i | \Omega_i, H_0^{s_h}, X^{s_x}] + \eta_{i,1,s,q} \qquad (14)\end{aligned}$$

Our procedure, described in Section 5.2 yields estimates of the coefficients of these equations for each mother. We report the results obtained with the Swamy (1970) estimator, although similar results can be obtained with a factor model.[9] These results can then be compared to estimates of the *objective* technology of skill formation obtained from data on actual child development.

Our approach delivers a set of coefficient for the perceived production function for each mother in our sample on which we have beliefs data. As with the data on subjective returns to parental investment, the estimated technologies of skill formation exhibit a considerable amount of heterogeneity.

The individual coefficients of the subjective technology of skill formation are summarised in Table 5. In particular, in the first two columns of the Table we report the average coefficients of two specifications for the technology of skill formation: one which assumes a Cobb Douglas form, forcing the $\mu_{i,3}$ coefficients to zero and one that assumes a translog specification as in equation (3). Below the average of each coefficient, we report the standard deviation for that coefficient *in the sample*. These standard deviation, therefore, represent the sample heterogeneity in perceived production function,

[9] In equation (14), the μ_i's can be considered factors to be estimated whose loading factors are the variables that define the scenarios and which, in our data, are observed.

rather than the precision of our estimates. In the third column we report the fraction of coefficients in the sample which, in column 2, are estimated to be statistically significant (that is with a t-value greater than 2).

Table 5: Estimation of Subjective Technology of Skill Formation

Dependent Variable: Expected Child Development at $t = 1$			
	Cobb Douglas (1)	Translog (2)	Fraction $\|t\| > 2$
Intercept	2.8025 (0.0302)	2.5297 (0.0425)	93.25%
Child Development at $t = 0$	0.2582 (0.0093)	0.3523 (0.0136)	63.33%
Parental Investment	0.0472 (0.0030)	0.3206 (0.0274)	22.17%
P. Investment x Child Dev. at $t = 0$		-0.0946 (0.0091)	18.25%

The estimates for the coefficient of child development at the beginning of the period is considerably less volatile and more concentrate than the estimate for maternal investment. We find that very few of the interaction terms are significantly different from zero. This indicates that the Cobb Douglas case represents a good approximation (relative to the translog) for many mothers. We note that we do impose constant return to scale and that the average effect of investment is quite low, at 0.047. Moreover, for only 22% of mothers this coefficient is statistically significant. The coefficient of child development at the beginning of the period, instead, is statistically significant for 63% of mothers and is, on average, much larger, at 0.258.

To have an idea of how these coefficients compare with the estimates one would obtain fitting a similar technology of skill formation to actual data about child development, we fit such a specification to the first follow up survey, where the ages of the children correspond roughly to the ages used in describing the scenarios in the beliefs questions. To estimate such a specification, we need to take into account that, in actual data on child development and parental investment, the latter is endogenous, as it is decided by parents, possibly in reaction to individual shocks in the process of development.

Attanasio et al. (2018) estimate a technology of skill formation and an investment function to interpret the effect of the parental stimulation program on child development. They argue that investment is indeed endogenous and, taking that into account, changes substantially the size of the coefficient. While Attanasio et al. (2018) do not use the program as their primary exclusion restriction to identify the effect of parental investment on child development, they find that the parenting stimulation program does not enter the technology of skill formation directly. Therefore, here we use the randomness of the program as an exclusion restriction to obtain a consistent estimate of the effect of parental investment on child development.

Table 6: Estimation of Objective Technology of Skill Formation

	First Stage Log of Parental Investment	Second Stage Log of Child Development at $t = 0$ at First Follow Up
Intercept	-0.7033 (0.2626)	2.5834 (0.0912)
Log of Child Development at $t = 0$	0.5568 (0.0901)	0.2870 (0.0564)
Log of Parental Investment		0.1892 (0.0915)
Dummy for Treatment	0.1432 (0.0483)	

We report our results in Table 6. The first stage, which can be read as a model of parental investment, indicates that parents invest more in more developed children. It also indicates that the the treatment, as reported in Attanasio et al. (2018) increases substantially parental investment. As for the second stage result, reported in the second column, we notice that the coefficient of investment on children's cognitive development is 0.189, substantially higher than the average coefficient (0.047) of the subjective technology of skill formation. [10]

The beliefs variables we use to estimate the subjective technology of skill formation are scaled so to make them comparable with the estimates obtained on actual data:

[10] In addition to the results we report, we also estimated a version of the 'objective' technology of skill formation assuming a translog function. The coefficient on the interaction term is not statistically different from zero and small in magnitude.

the average coefficients in Table 5 are therefore directly comparable with the estimates in Table 6. The main difference between the point estimates in the latter table and the averages in the former, is on the coefficient on parental investment. While such a difference is not statistically significant at the 5% level, the evidence is suggestive of the fact that many parents tend to underestimate the productivity of investment in the developmental process. This result is consistent with anthropological evidence, such as that discussed in Lareau (2003), that claims that poor parents might not be investing much in their young children because they do not see the usefulness of such interventions.

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. In a way this exercise is similar to that whose results are 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 7 we regress the estimated parameters of the subjective technology of skill formation (the μ's) on the same set of observables used in Table 3. In particular, we regress each of the μ's on age, two indicators of education, a depression index, the results of the Raven tests, and an indicator of the gender of the target child. In addition, for each parameter we present a regression which includes the treatment dummy.

Table 7: Subjective Estimates and SE characteristics

VARIABLES	μ_0	μ_0	μ_1	μ_1	μ_2	μ_2
Mother's age	-0.007	-0.008	0.009*	0.010*	0.009	0.009*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Mother's education (primary)	-0.160	-0.172	0.195	0.208	0.101	0.102
	(0.403)	(0.413)	(0.415)	(0.425)	(0.377)	(0.378)
Mother's education (secondary and more)	-0.165	-0.175	0.219	0.229	0.156	0.157
	(0.391)	(0.401)	(0.402)	(0.413)	(0.376)	(0.377)
Mother's depression (CES-D)	0.006	0.006	-0.006	-0.007	0.001	0.001
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Mother's IQ (standardized Raven's Score)	0.015	0.016	-0.002	-0.003	0.117***	0.117***
	(0.035)	(0.035)	(0.036)	(0.035)	(0.037)	(0.037)
Dummy for Male (child)	0.069	0.072	-0.050	-0.052	-0.099	-0.099
	(0.068)	(0.068)	(0.067)	(0.067)	(0.067)	(0.068)
Standardized ln H_0	-0.034	-0.033	0.031	0.030	0.034	0.034
	(0.033)	(0.033)	(0.034)	(0.034)	(0.037)	(0.037)
Ratio ln H_0/ln H_1	0.002*	0.002	-0.002	-0.002	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Standardized Household Wealth	0.022	0.021	-0.017	-0.017	0.022	0.022
	(0.033)	(0.033)	(0.033)	(0.033)	(0.036)	(0.036)
Standardized ln X	0.060*	0.056*	-0.053	-0.049	-0.019	-0.019
	(0.032)	(0.033)	(0.032)	(0.033)	(0.035)	(0.035)
Dummy for Treatment		0.090		-0.092		-0.008
		(0.067)		(0.066)		(0.075)
Adjusted R^2	0.001	0.002	-0.000	0.001	0.013	0.012
F	1.599	1.783	1.245	1.416	1.669	1.556
Observations	1017	1017	1017	1017	1017	1017

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

As in Table 3, most variables are not significant. There are two important differences with the results in Table 3, however. First, unlike in that Table, the coefficient on the mother's IQ is not statistically different from zero. Second, and importantly, both the intercept and the coefficient on child development at the beginning of the period seem to be affected by the parenting stimulation program. In particular, the former is significantly increased while the latter is decreased. The program parents seem to perceive the process to be more effective and give less weight to the child development at the beginning of the period.

Finally, we check whether the parameters of the subjective technology of skill formation affect investment. In Table 8 we regress different measures of investment on various determinants of investment (such as household wealth), a dummy for treatment and the subjective productivity of investment, as measured by μ_2. The table indicates that investment measured by activities is positively related to the subjective productivity of investment, regardless of whether we control for other variables. Instead, when we consider total investment, and investment in materials or activities), μ_2 is important only when we do not control for additional variables. When we introduce child development at the beginning of the period, household wealth and mothers Raven tests results, the coefficient on μ_2 becomes much smaller and not statistically different from zero.

Table 8: Subjective Estimates and Investment

VARIABLES	Log Investment		Time		Activities		Materials	
Standardized μ_2	0.024	-0.014	0.101*	0.011	0.020	-0.022	0.033	0.003
	(0.030)	(0.031)	(0.052)	(0.026)	(0.023)	(0.020)	(0.026)	(0.030)
Dummy for Treatment	0.180*	0.182**	0.067	0.026	0.063	0.169***	0.049	0.015
	(0.095)	(0.090)	(0.130)	(0.068)	(0.067)	(0.057)	(0.086)	(0.081)
Dummy for Male	0.007	-0.011	0.019	0.006	0.036	0.024	0.003	-0.029
(child)	(0.062)	(0.063)	(0.120)	(0.056)	(0.057)	(0.039)	(0.054)	(0.056)
Standardized ln H_0		0.149***		0.067***		0.064**		0.120***
		(0.035)		(0.025)		(0.025)		(0.034)
Standardized Household Wealth		0.122***		0.025		0.049**		0.145***
		(0.035)		(0.029)		(0.019)		(0.035)
Mother's IQ		0.196***		0.074***		0.098***		0.144***
(standardized Raven's Score)		(0.034)		(0.029)		(0.021)		(0.032)
Adjusted R^2	0.005	0.100			0.001	0.073	-0.001	0.072
F	1.045	15.500	1.324	4.506	1.626	7.027	0.643	10.756
Observations	1017	1017	1200	1017	1200	1017	1200	1017

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

7 Conclusions

In this paper, we have proposed a new method to elicit maternal beliefs about the technology of skill formation. While our basic measures allows us to compute very simply the rate of return to parental investment under different scenarios, it also allows us to compare our estimates of parental beliefs to the estimates of the actual process of child development that we can obtain from child development and parental investment data from the same population. We show that, in our context, mothers tend to underestimate the productivity of parental investment.

We also show that our estimates of subjective beliefs about the productivity of parental investment covary with actual investment. Parents who report higher expected returns also tend to invest more in their children. This is true both when we use simple estimates of rates of return and we use the esimated parameters of the subjective technology of skill formation.

Finally, our expected return 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 important to combine measures of subjective beliefs with further measures about the way expenditure is allocated within families to estimate parental preferences and how parental beliefs about the process of child development 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 key to design of effective interventions to foster the development of children living in adverse environments. Parental beliefs about the process of child development is obviously a key driver of parental bbehaviour. Characterising parental beliefs and their potential biases is therefore important, both from a research and a policy point of view.

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:

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

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:

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

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:

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

The contribution to the likelihood can be written as:

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

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

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:

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

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 $m_{i,j,1}$:

$$L_i = \int \left[\prod_{k=1}^{K} G_{i,j,k}^{MLI}\right]\left[\prod_{l=1}^{L} G_{i,j,l}^{BSID-III}\right] f(m)\, dm \tag{17}$$

And the likelihood function is

$$L = \prod_{i=1}^{N} \int L_i \tag{18}$$

We can estimate the parameters of the IRT model by maximising the likelihood (18).

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:

$$L_i = \int \left[\prod_{k=1}^{K} G_{i,j,k}^{MLI}\right] f(m)\, dm \tag{19}$$

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.

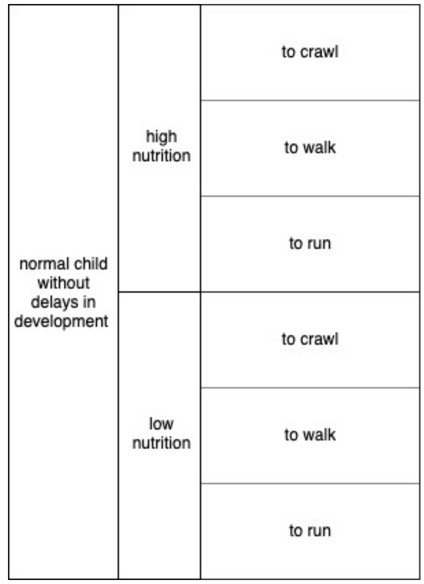


Figure A.1: Beliefs Elicitation Survey Instrument: Practice module

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.

<table>
<tr><td rowspan="6">child can only say the easy words (H_0^L)</td><td rowspan="3">mother spends little time with the child and spends little money on didactic materials or toys (X^L)</td><td>Easy words</td></tr>
<tr><td>More difficult words</td></tr>
<tr><td>Even more difficult words</td></tr>
<tr><td rowspan="3">mother spends little time with the child and spends little money on didactic materials or toys (X^H)</td><td>Easy words</td></tr>
<tr><td>More difficult words</td></tr>
<tr><td>Even more difficult words</td></tr>
</table>

<table>
<tr><td rowspan="6">child can say both the easy and difficult words (H_0^H)</td><td rowspan="3">mother spends little time with the child and spends little money on didactic materials or toys (X^L)</td><td>Easy words</td></tr>
<tr><td>More difficult words</td></tr>
<tr><td>Even more difficult words</td></tr>
<tr><td rowspan="3">mother spends little time with the child and spends little money on didactic materials or toys (X^H)</td><td>Easy words</td></tr>
<tr><td>More difficult words</td></tr>
<tr><td>Even more difficult words</td></tr>
</table>

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

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

A.3.1 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.3.2 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 η_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 $A_i = \left(\hat{\Upsilon}^{-1} + \tilde{V}_i^{-1}\right)^{-1} \hat{\Upsilon}^{-1}$:

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

A.4 Estimation of the Objective Technology of Skill Formation

In this section we describe how to objectively estimate the technology of skill formation so that we can compare maternal subjective beliefs against a model that describes the process of child development as represented by equation (2) in Section 2. In this methodology, we explore the dimensionality reduction provided by the measurement system and only use the factor scores for each measure of child development (both in baseline and follow-up one) and parental investments (in follow-up one).

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 Section 4.3.1 produces an additional continuos scale of expressive language development. Therefore, we have $K = 4$ continuous measures of child development at the beginning of the period. Furthermore, all of these measures have location and scale in age-equivalent scores:

$$\mathbf{m}_{i,\ln H_0} = \Gamma_{\ln H_0,0} + \Gamma'_{\ln H_0,1} z_i + \Gamma_{\ln H_0,2} \ln H_{i,0} + \Xi_{i,\ln H_0} \tag{20}$$

Note that the equation is written 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:

$$\mathbf{m}_{i,\ln H_1} = \Gamma_{\ln H_1,0} + \Gamma'_{\ln H_1,1} z_i + \Gamma_{\ln H_1,2} \ln H_{i,1} + \Xi_{i,\ln H_1} \tag{21}$$

Because our measures of child development have location and scale according to age-equivalent score, we could - in principle - fix the intercepts to zero and factor loadings to one for all of the measures of child development at the beginning of the period and at the end of the period. However, we adopt a less aggressive approach and, consistent with our methodology, we only do so for 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.

Finally, the IRT analysis of the FCI instrument and time spend with children in activities, explained in Section 4.3.1, produces two continuous measures of parental investments. The first measure summarizes parental investments as captured by activities and the second by materials. Therefore:

$$\mathbf{m}_{i,\ln X} = \Gamma_{\ln X,0} + \Gamma'_{\ln X,1} z_i + \Gamma_{\ln X,2} \ln X_i + \Xi_{i,\ln X} \tag{22}$$

The objective estimation of the technology of skill formation has three steps. In the first step, we factor analyze equations (20, 21, and 22) separately. Then, we use estimated intercepts, factor loadings, and variances of the factors, to generate the factor scores:

$$\hat{\theta}_i = \left(\Gamma'_{\theta,2}\mathbf{\Sigma}_\theta^{-1}\Gamma_{\theta,2}\right)^{-1}\left(\Gamma'_{\theta,2}\mathbf{\Sigma}_\theta^{-1}\mathbf{m}_{\mathbf{i},\theta}\right) \tag{23}$$

where $\mathbf{\Sigma}_\theta^{-1}$ is the variance of factor $\hat{\theta}_i$ and $\hat{\theta}_i = \ln\hat{H}_{i,0}, \ln\hat{H}_{i,1}, \ln\hat{X}_i$. Next, let $\delta = (\delta_0, \delta_1, \delta_2, \delta_3)'$ and $\Psi_i = \left(1, \ln\hat{H}_{i,0}, \ln\hat{X}_i, \ln\hat{H}_{i,0}\ln\hat{X}_i\right)$, so that we can write the technology of skill formation (2) in the following form:

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

Now, define $\tilde{R}_i = \left(1, \ln\hat{H}_{i,0}, d_i, \ln\hat{H}_{i,0}d_i\right)$. Note that we are using the random assignment to control or parenting stimulation program as a valid instrument for parental investments in children. Now, consider the following inconsistent estimator of the vector δ, $\tilde{\delta}$:

$$\tilde{\delta} = (\tilde{R}'\Psi)^{-1}(\tilde{R}'\ln\hat{H}_1)$$

where $\tilde{R}$ and Ψ are matrix of dimensions $(N \times 4)$ and $\ln\hat{H}_1$ is a row vector of dimension N, where N is the number of children in the study. The inconsistency of the estimator arises in every cell of $(\tilde{R}'\Psi)^{-1}$ that has a quadratic term of a predicted factor score. Because we instrument parental investments with the dummy for random allocation to control or parenting stimulation program, the only cells in $(\tilde{R}'\Psi)^{-1}$ that have biased estimates of moments are the ones that include quadratic terms of $\ln\hat{H}_{i,0}$. The bias terms depends on the factor loadings and variances of the measurement errors in (20).

The reason why this happens is because any predicted factor score contains not only the contribution from the unobserved latent variable, but also inherits parts of the measurement errors. We can use the estimated values of factor loadings and variances of measurement errors to adjust the specific cells in $(\tilde{R}'\Psi)^{-1}$ that are affected by this contamination. This is the central idea in Williams et al. (2019). We denote by R the bias-corrected matrix of $\tilde{R}$. Let $\hat{\delta}$ denote the following consistent estimator of δ:

$$\hat{\delta} = (R'\Psi)^{-1}(R'\ln\hat{H}_1)$$

The standard errors of $\hat{\delta}$ are given by:

$$\hat{\delta} = (R'\Psi)^{-1}(R'R)^{-1}[(R'\Psi)^{-1}]'$$

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. (2018). Estimating the production function for human capital: Results from a randomized control trial in colombia. NBER Working Paper No. 20965.

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 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.

Furrow, D., Nelson, K., and Benedict, H. (1979). Mothers' speech to children and syntactic development: Some simple relationships. *Journal of child language* 6:423–442.

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.

Masur, E. F. and Gleason, J. B. (1980). Parent-child interaction and the acquisition of lexical information during play. *Developmental Psychology* 16:404.

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.

Weigel, D. J., Martin, S. S., and Bennett, K. K. (2006). Contributions of the home literacy environment to preschool-aged children's emerging literacy and language skills. *Early Child Development and Care* 176(3-4):357–378.

Williams, B., Heckman, J., and Schennach, S. (2019). Nonparametric factor score regression with an application to the technology of skill formation preliminary -not for circulation .