Eliciting Maternal Subjective Expectations about the Technology of Cognitive Skill Formation*

Flávio Cunha†, Irma Elo‡, and Jennifer Culhane§

December 16, 2015

Abstract

We formulate a model of early childhood development that includes maternal subjective expectations (MSE) about the technology of skill formation. The model explores how MSE influences the amount and quality of maternal investments in the development of human capital of their child. The model cannot be estimated from data that are usually available to social scientists and thus requires primary data collection. We acquired such data by conducting studies where mothers were interviewed to elicit their MSE. The sample consists of socioeconomically disadvantaged African-American women. In addition, we used the data on child development and maternal investments from the Children of the National Longitudinal Survey of Youth 1979 (CNLSY/79) to objectively estimate the technology of skill formation. Our model suggests that a public policy or program that educated mothers about the objective technology of skill formation would lead to an increase in maternal investments by approximately 10% and the human capital at age 24 months would increase by 4.5%.

Key words: Cognitive skills, parental expectations, investments
JEL Code: A12

---

*We thank Dalton Banks, Michelle Gifford, Debbie Jaffe, Snejana Nihtienova, Ben Sapp, Shanae Smith, and Cheryl Tocci for excellent research assistance. We are grateful to Orazio Attanasio, Jere Behrman, Pietro Biroli, Hanming Fang, Limor Golan, Katja Kaufmann, James Heckman, Pedro Mira, Krishna Pendakur, Matt Wiswall, and Ken Wolpin for their comments in different stages of this research. We also thank participants at the Early Childhood Development and Human Capital Accumulation conference at University College London, the Institute for Research on Poverty's Summer Workshop at the University of Wisconsin–Madison, the Children's Human Capital Development workshop at Aarhus University in Denmark, the CES-ifo conference on the Economics of Education, the Conference on the Economics of the Family at the University of Chicago, and seminar participants at Columbia University, the University of Pennsylvania, UCLA, Stanford University, UC Berkeley, Insper São Paulo, CHPPP at the Harris School, the International Food and Policy Research Institute, Duke University, the EPGE/FGV Rio de Janeiro, the Pontifical Catholic University in Rio de Janeiro, UCI, and the Federal Reserve Bank in New York. This research was supported by grant INO12-00013 from the Institute for New Economic Thinking and grant 1R01HD073221-01A1 by the National Institute of Health. All remaining errors are ours.

† Department of Economics at Rice University and NBER.

‡ Population Studies Center and Department of Sociology at the University of Pennsylvania.

§ Children's Hospital of Philadelphia.

1 Introduction

In a pioneering study, Hart and Risley (1995) documented the variation in the language environment of children nine months to three years. Specifically, these authors measured the number of words spoken to the child during an hour, the number of conversation turns between adult and the child, and the quality and diversity of parental vocabulary. Results indicate that children living in poverty heard approximately half the words per hour compared to children of parents in professional occupations. Hart and Risley (1995) also showed that better early language environment at home predicted better language development, higher IQ, and superior school performance in the children.

In order to understand the factors associated with the quality of the language environment Rowe (2008) replicated Hart and Risley's original work and concluded that gaps in the language environment exist because poor, uneducated mothers do not understand the important role language environment plays in determining the cognitive development of their children. Rowe's finding suggested that heterogeneity in MSE about the technology of skill formation may explain variation in maternal investments.

Other research has demonstrated that expectations about returns from schooling influence educational attainment. For example, Attanasio and Kauffman (2009) showed that the higher the parental expectations of the returns from schooling, the more they invested in their children's education. Additionally, Jensen (2010) found that students' *subjective* expectations of returns from schooling extremely underestimated the documented *objective* returns. Relevant to this work, Jensen showed that subjective expectations are modifiable and changes in expectations correlate with behavior. Students in schools who were given information about the true returns completed on average 0.20–0.35 more years of school over the next four years compared to those who did not receive this educational intervention.

Recent literature provides evidence that public information campaigns can change prenatal investments. Aizer and Stroud (2010), for example, tracked the smoking habits of educated and uneducated pregnant women before and after the release of the 1964 Surgeon General's Report on Smoking and Health. Before the release of the report, educated and

uneducated pregnant women smoked at roughly the same rates. After the report, the smoking habits of educated women decreased immediately, creating a ten-percentage-point gap in pregnancy smoking rates between educated and uneducated women. These results suggest that beliefs and in investments in human capital can be influenced by information campaigns with differential effectiveness across socio-economic groups.

Motivated by this research, we formulate a model of early childhood development that includes MSE about the technology of skill formation. The model explores how MSE influences the amount and quality of maternal investments in the development of human capital of their child. The utility of this model is to predict the impact of public policies and programs designed to educate mothers about the importance of early investments in child development.

The model cannot be estimated from data that are usually available to social scientists and thus requires primary data collection. If we only observe investments and measures of human capital, it is impossible to decompose heterogeneity in expectations from heterogeneity in preferences (Manski, 2004).

To solve this identification problem, we developed and implemented a survey instrument to elicit MSE. Briefly, this survey presents pregnant women with hypothetical scenarios of "high" and "low" levels of investments over the course of the first two years of life together with varying levels of child's human capital at birth (e.g., birthweight, gestational length). For each scenario of investment and human capital at birth, we asked the mother to estimate the expected child developmental outcomes at age two. By comparing the expected developmental outcomes across the various scenarios, we are able to estimate MSE. In addition, we used the data on child development and maternal investments from the Children of the National Longitudinal Survey of Youth 1979 (CNLSY/79) to objectively estimate the technology of skill formation. Our model suggests that a public policy or program that educated mothers about the objective technology of skill formation would lead to an increase in maternal investments by approximately 10% and the human capital at age 24 months would increase by 4.5%.

This paper is organized as follows. In Section 2, we describe our survey instrument and the data used from the CNLSY/79. In Section 3, we introduce our model of investment in human capital of children. Section 4 describes the identification and estimation of the model, including the methodology for estimating maternal subjective expectations about the technology of skill formation. We present the estimation results in Section 5. In Section 5, we quantify the importance of beliefs in determining investments. We compare the results from our model simulation with policies that potentially affect maternal beliefs.

2 Data

The primary data come from two studies conducted in Philadelphia, PA. The first study is the Maternal Knowledge of Infant Development Study (MKIDS). MKIDS is a pilot study with the goal of developing a questionnaire to optimally elicit MSE and thus included various ways of presenting hypothetical scenarios of maternal investments and human capital at birth and asking about expected developmental outcomes. The study enrolled 323 African American pregnant women from low-income clinics during their second trimester of pregnancy. Data collected included socio-demographic information (e.g., age, education, marital status, household income) and items adapted from the Motor Social Development (MSD) Scale that were used to estimate MSE (see Table 1 for a description of the items). MKIDS tested two different forms to elicit MSE. In the first form, mothers were asked to report the likelihood (e.g., 42%) that a child would learn an item from the MSD Scale by age two for each hypothetical scenario varying human capital at birth and level of maternal investment. For example, participants were asked to estimate the percentage likelihood that a low birthweight baby exposed to high levels of investments would speak a partial sentence of three words or more by age two. In the second form, we ask mothers to report the age range (e.g., eighteen to twenty-four months) a child is likely to learn the tasks across the same scenarios. For example, participants were asked to report the upper and lower age by which a low birthweight baby exposed to high levels of investments would speak a partial sentence of three words or more. In addition, we also varied how the scenarios were described to study participants. For example, in the Baseline

presentation, study participants were told (BLA BLA BLA). [1] Two alternative descriptions of the scenarios were developed and tested. The first and second alternative descriptions collapse the difference between high and low human capital at birth and high and low investments, respectively. These scenarios were concretely defined in a five-minute video that the respondents watched before answering any questions. In the video for the baseline scenario, we designated "high" human capital as the one in which the baby's gestation lasted nine months, the baby weighed eight pounds and was 20 inches long at birth. In contrast, the "low" level of human capital at birth corresponds to a baby whose gestation was seven months, weighed five pounds and was 18 inches long at birth. The "high" human capital is around the 60th percentile in the distribution of human capital at birth in the CNLSY/79, while the "low" human capital is around the 1st percentile.

The video also showed examples of activities that mothers do with children. With the exception of breastfeeding, all activities depicted are part of the Home Observation for the Measurement of Environment–Short Form (HOME-SF) instrument (see Bradley and Caldwell 1980, 1984) The activities are the same for the "high" and "low" level of investments. The difference is in the amount of time. In the "high" level, the mothers spend six hours a day doing these types of activities, while in the "low" level they spend only two hours a day. These figures correspond, respectively, to roughly the 95th and 15th percentile of investments measured by the HOME-SF in the CNLSY/79.[2]

Finally, the MKIDS asked how much money a mother would spend on things likes books, experiences, etc. for their child given at various levels of household income and prices of these goods, experiences, etc. In our survey, we first told the respondent to assume that the baby's human capital at birth is "high." We then presented the respondent with nine hypothetical scenarios of monthly income and prices of investments. These nine hypothetical scenarios are the combination of three levels of monthly income ($1500, $2000, and $2500) and three levels for the price of investment goods ($30, $45, and $60).

[1] Appendix Table A2 provides a summary of the definitions used in the different scenarios presented to study participants.

[2] For a subset of the respondents, we provided alternative definitions of these hypothetical scenarios. As a result, we investigated the sensitivity of answers with respect to variations in how scenarios are framed.

These scenarios of income and prices are explained to MKIDS participants through a short video. In order to link investment to time, we prepared a three-minute video in which we explained to the respondent that the more time the mother interacts with her child, the more money she has to spend every month on educational goods, such as children's books and educational toys. The purpose of this exercise was to explain to the respondent that investments are costly.[3] We illustrated the concept by giving examples:

> If [the mother] spends two hours a day interacting with the child, she needs to buy two books and two educational toys per month… But if she spends three hours a day, she needs to buy three books and three educational toys per month… and so on.

We refer to this as stated-choice data. This information allows us to estimate the preference parameters that are necessary to quantitatively evaluate the impact of policies that affect MSE. Table 1 contains the distribution across these varying methods, scenarios, and stated-choice data.

The second study, the Philadelphia Human Development (PHD) Study, is longitudinal following primaparous mothers from the second trimester through the child's second birthday. Participants were recruited from the same low-income clinic as in MKIDS. Data for this paper come from the first wave in which MSE, using only the baseline scenario developed in MKIDS, and socio-demographic information is collected. Also, women in PHD were shown the same MKIDS explanatory video prior to data collection. Unlike MKIDS, all of the PHD Study participants were asked both the probability and the age range questions for a subset of the MSD items measuring expected developmental outcomes. The sample used in this analysis consists of 454 African-American participants. Appendix A contains a detailed explanation about procedures and features of each one of these studies.

Our third data source is the CNLSY/79. We obtained maternal sociodemographic characteristics (e.g., age, education, race, household income, marital status), infant birth

[3] We have implicitly assumed that the production function for investment goods is Leontief in maternal time and investment goods (such as children books). Obviously, this need not be the case.

outcomes (e.g., gestational age, birthweight and length), maternal investments (the Home Observation for the Measurement of the Environment Short Form, HOME-SF) and infant developmental outcomes from the MSD Scale at age two.

MSD scale plays an important role in our analysis because it is used to measure child development at age two which allows to estimate the technology of skill formation. The MSD is also used to create the instrument to measure MSE in MKIDS and PHD Studies discussed above. In the MSD instrument, mothers answer fifteen age-appropriate items regarding motor, language, and numeracy development. All items are dichotomous and the total raw score for children of a particular age is obtained by a simple summation (with a range 0 to 15) of the affirmative responses. The HOME is also important to estimate the technology of skill formation and to define the scenarios of investments shown in the video as described above. DESCRIBE AGE RANGE, INCOME, CHANGES IN TAXATION POLICY AS INSTRUMENTS.

3 Model

In this section, we first describe the parameterization of the technology of skill formation and explain how the CNLSY/79 data described above are used to objectively estimate the parameters of our model. Second, we use the MKIDS and PHD data to estimate MSE about the technology of skill formation. Third, we use the stated-choice data to recover the parameters that describe maternal preferences.

3.1 The Technology of Skill Formation

Let $q_{i,0}$ and $q_{i,1}$ denote, respectively, the stocks of human capital of child i at birth and at 24 months.[4] Let x_i denote maternal investments in the human capital of child i during the first two years of the child's life. Let ν_i denote shocks to the development process. We assume that the technology of skill formation is "approximately" Cobb-Douglas:

$$\ln q_{i,1} = \psi_0 + \psi_1 \ln q_{i,0} + \psi_2 \ln x_i + \psi_3 \ln q_{i,0} \ln x_i + \nu_i. \quad (1)$$

[4] In the empirical application below, we measure $q_{i,0}$ by birth weight, birth length, and gestation length. We measure $q_{1,i}$ by developmental tests around the time the child is 24 months old.

We are doing two things. First, we use the CNLSY/79 data to estimate the parameter vector ψ. Previous research showed that the technology of cognitive skill formation, seen in Equation (1), follows a Cobb-Douglas specification.[5]

Second, we use the MKIDS and PHD data to estimate MSE about the parameter vector ψ. The parameterization in Equation (1) is particularly convenient to make progress on these questions. To see why, let $\mathcal{H}_i$ denote the mother's information set. According to the technology function denoted in Equation (1), it follows that:

$$E\left(\ln q_{i,1} \middle| q_0, x, \mathcal{H}_i\right) = \mu_{i,\psi,0} + \mu_{i,\psi,1} \ln q_0 + \mu_{i,\psi,2} \ln x + \mu_{i,\psi,3} \ln q_0 \times \ln x \qquad (2)$$

where $\mu_{i,\psi,j} = E\left(\psi_j \middle| q_0, x, \mathcal{H}_i\right)$. Under specification (1), we can investigate, for example, if mother i believes that the technology is Cobb-Douglas by testing the hypothesis that $\mu_{i,\psi,3} = 0$.

Next, maternal utility function is presented. Let c_i denote household consumption. Preferences are described by the following parameterization:

$$U\left(c_i, q_{i,1}, x_i\right) = \ln c_i + \alpha_{i,1} \ln q_{i,1} + \alpha_{i,2} \ln x_i \qquad (3)$$

The utility function denoted in Equation (3) is Cobb-Douglas, and it has three arguments. Under our assumptions, parents care about consumption (c_i), about child development, and about investments in child development directly. Direct preference for investment is not usually present in most models of human capital formation, but here it is important to include it because we want to allow for investments to be determined by a component other than the one mediated by MSE. As we will demonstrate below, the Cobb-Douglas preferences are convenient because it allows us to focus on investigating the importance of mean beliefs, so it is not necessary to elicit beliefs about higher-order moments. More general formulations of the utility function require elicitation of higher order moments of maternal beliefs.

Let y_i denote household income, which we assume is exogenously determined in the model. The budget constraint reads:

[5] See Cunha, Heckman, and Schennach (2010).

$$c_i + px_i = y_i \tag{4}$$

Optimal parental investment is the one that maximizes utility—Equation (3)—subject to the perceived technology of skill formation—Equation (2)—and the budget constraint—Equation (4). It is easy to show that the policy function for investment is:

$$x_i^* = \left[\frac{\alpha_{i,1}(\mu_{i,\psi,2}+\mu_{i,\psi,3}\ln\ _0)+\alpha_{i,2}}{\alpha_{i,1}(\mu_{i,\psi,2}+\mu_{i,\psi,3}\ln q_0)+\alpha_{i,2}+1}\right]\left(\frac{y_i}{p}\right) \tag{5}$$

Typically, econometricians observe the vector $D_i = (q_{0,i}, q_{1,i}, x_i, y_i, p)$. As explained in Manski (2004), the major identification issue that arises when D_i is the only data available is that one cannot separately identify heterogeneity in preferences (captured in Equation [5] by $\alpha_{i,1}$ and $\alpha_{i,2}$) from heterogeneity in beliefs (represented in Equation [5] by $\mu_{i,\psi,2}$, and $\mu_{i,\psi,3}$). In the context of this simple model, one of the contributions of our research was to develop and implement a methodology to elicit $\mu_{i,\psi}$. Clearly, to the extent that investments are partly determined by these beliefs, these variables are interesting by themselves. More important, if we add $\mu_{i,\psi}$ to the data D_i, we are able to separately identify heterogeneity in preferences from heterogeneity in beliefs.

3.2 Identification and Estimation of the Model

As described above child development is estimated in the CNLSY/79 by counting the number of affirmative answers in the MSD Scale. This summary score is problematic for our goals because of participant burden. Recall that we present participants with four scenarios with different levels of human capital at birth and investments. Thus, we would have these 15 MSD questions for times for each form of elicitation. Instead, we estimate an Item Response Theory (IRT) model so that we could reduce the number of items used in the survey instrument measuring MSE.

Let the variable a_i denote the child i's age at the time of the measurement of skills in the CNLSY/79 dataset. Let θ_i denote the child i's development relative to other children in the same age group. For example, $\theta_i = 0$ if child i's development is typical for his or her age; $\theta_i > 0$ if child i is advanced for the age; and $\theta_i < 0$ if child i has developmental delays

relative to children in his or her age group. The variable θ_i is a latent factor whose distribution we aim to estimate with the IRT model. For each child i and MSD item j, define the latent variable $d_{i,j}^*$ according to the following specification:

$$d_{i,j}^* = b_{j,0} + b_{j,1}\left(\ln a_i + \frac{b_{j,2}}{b_{j,1}}\theta_i\right) + \eta_{i,j} \tag{6}$$

We do not observe the variable $d_{i,j}^*$. We observe, however, that the binary variable $d_{i,j}$ is equal to one whenever $d_{i,j}^* \geq 0$ and equal to zero otherwise.

In the IRT model of Equation (6), the parameters $b_{j,0}$ are smaller for the harder items. The parameter $b_{j,1}$ describes how fast performance in task j improves as age increases. The parameter $b_{j,2}$ denotes the informational content of item j with respect to child development. The higher the value of $b_{j,2}$, the more information item j contains about child development θ_i.

Let Φ denote the cumulative distribution function (CDF) of a normal random variable with mean zero and variance one. If we assume that the idiosyncratic component $\eta_{i,j} \sim N(0,1)$, it follows that the probability that child i can perform MSD task j is equal to:

$$Pr\left(d_{i,j} = 1 \middle| a_i, \theta_i\right) = 1 - \Phi\left[-b_{j,0} - b_{j,1}\left(\ln a_i + \frac{b_{j,2}}{b_{j,1}}\theta_i\right)\right] \tag{7}$$

IRT models are, in fact, factor models in which the dependent variables take on discrete values. As in other factor models, we need to make two normalizations: one for the location and the other for the scale of θ_i. Thus, we restrict the mean of θ_i to be equal to zero, and we set $b_{2,j} = 1$ for one of the MSD items. In our empirical analysis, we assume that the distribution of the factor θ_i is equal to a mixture of two normal CDFs.

Equation (7) plays a major role in our analysis for two reasons. First, it allows us to estimate the actual natural logarithm of child development, which we then use to estimate the parameter vector ψ in the right-hand side of Equation (4) using the CNLSY/79 data.

Second, it allows us to translate answers from the questionnaire designed to elicit maternal expectations about the natural logarithm of child development. This information allows us

to recover MSE about the parameters of the technology of skill formation (the vector μ_ψ in the right-hand side of Equation [5]). By comparing the parameter vector ψ that we estimate with the maternal beliefs μ_ψ, we are able to estimate if mothers have biased or unbiased expectations about the technology of skill formation.

3.3 Eliciting Maternal Beliefs

In this section, we provide details about how the MSD instrument used in the CNLSY/79 was adapted to elicit MSE. As we now explain, although the questions used in the MKIDS and PHD are similar to the ones used in the CNLSY/79, they differ in one important detail. In the original MSD, mothers with a two-year-old child were asked if that child had reached a developmental milestone. We adapt several questions from the MSD to ask women to speculate the degree to which a child is likely to reach those same developmental milestones by age two for each scenario of human capital at birth and level of investment. For example, in the original MSD CNLSY/79 mothers or 24-month-old child were asked to responde yes or not to a statement such as "Does your child speak a partial sentence of three words or more?" In contrast, we asked study participants the following types of questions:

1. *How likely is it that a baby will learn how to say a partial sentence with three words or more by age two years if the child is small at birth and BLA BLA BLA?*

2. *What do you think is the youngest age and the oldest age a baby learns to speak a partial sentence of three words or more if the child is small at birth and BLA BLA BLA?"*

In the first type of question, the respondent uses a sliding scale to indicate the likelihood by age two years that the child will learn how to say a partial sentence. This type of question is more closely related to how the literature in economics elicits subjective expectation (see summary of this literature in Manski, 2004), so it is relatively straightforward to transform answers to measures of expectations about the natural log of child development.

In the second type of question, the respondent uses a sliding range scale to indicate the youngest and oldest ages by which the child will learn how to speak a partial sentence of

three words or more.[6] The second question is more in line with how the literature in child development measures parental knowledge (see Epstein, 1979; Ninio, 1988). This type of question requires additional steps in order to be able to transform answers into measures of expectations.

Specifically, our survey instrument described to the expectant mother four different hypothetical scenarios of investments and the baby's human capital at birth. In the first scenario, the baby's human capital at birth is "high" $(\overline{q}_0)$, and the mother has a "high" level of investment $(\overline{x})$. In the second scenario, the mother also has a "high" level of investment $(\overline{x})$, but the baby's human capital at birth is "low" $(\underline{q}_0)$. In the third scenario, the baby's human capital at birth is "high," but the mother has a "low" level of investment $(\underline{x})$. Finally, in the fourth scenario, both the baby's human capital at birth and investments are low. To ensure that all participants received the same information, these scenarios were described in a five-minute video that the respondents watched before answering any questions.

3.3.1 Estimating MSE Using "How likely" Questions

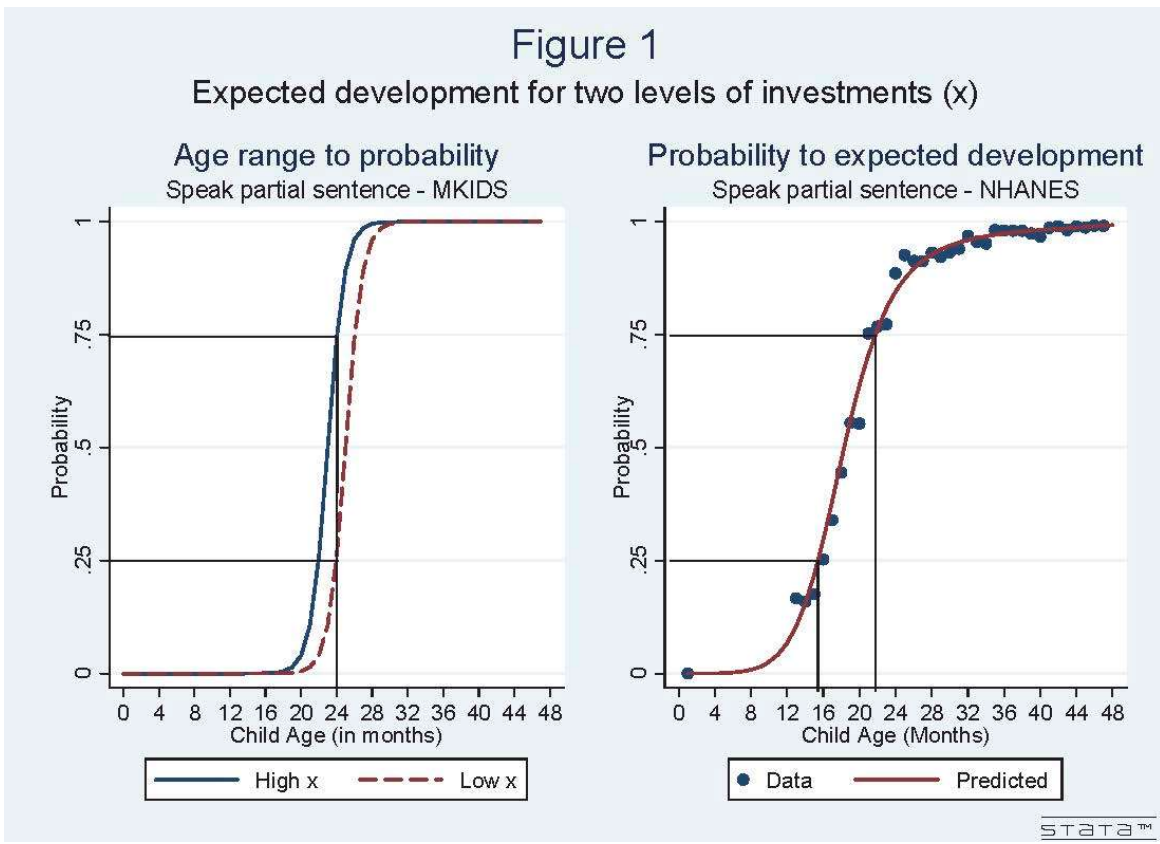
We now discuss how we transform the answer to the questions asked in our instrument into measurements of the MSE of child development at age 24 months. This expectation is conditional on three objects: the maternal information set $\mathcal{H}_i$, the level of human capital at birth and investment given to the respondent through the scenarios described above (see Equation [2]).

Let $p^L_{i,j,k}$ denote the likelihood reported by respondent i that a child will learn MSD item j by age 24 months if human capital at birth and investments are at the levels determined in scenario k. We explore the IRT model to derive an error-ridden measure of maternal expectation of the natural log of development at age 24 months, $\ln q^L_{i,j,k}$, from the reported probability $p^L_{i,j,k}$. To do so, we invert Equation (7) and solve for θ_i:

[6] The design of the survey instrument was influenced by Delavande, Giné, and McKenzie (2011), who showed that individuals report more accurately when their answers are represented with visual instruments.

$$\ln q_{i,j,k}^{L} \equiv \left(\ln 24 + \frac{b_{2,j}}{b_{1,j}}\theta_{i,j,k}^{L}\right) = -\left[\frac{b_{j,0}+\Phi^{-1}\left(1-p_{i,j,k}^{L}\right)}{b_{j,1}}\right] \tag{8}$$

The algorithm described above can be easily explained graphically. For illustration purposes, Figure 1 (right panel) shows the data and the resulting prediction from the IRT model for the MSD item "speak a partial sentence of three words or more." The algorithm above simply transforms the probability into the equivalent age. Building on the example shown in Figure 1, suppose that the mother believes that there is a 75% chance that the child will learn how to speak a partial sentence by age two years when investment is "high." According to the IRT model, this statement means that the mother believes that at age 24 months, $\ln \hat{q}_{i,j,k}^{L} = \ln 22$.



Importantly, the lower the subjective probability that the mother reports for a given item j, the lower the corresponding expectation about the natural log of child development at age 24 months. Again, we refer to Figure 1 for a visual explanation of the mechanics of the

algorithm. Suppose that for the "low" investment scenario, the mother believes that there is a 25% chance that the child will learn how to speak a partial sentence with three words or more by age 24 months. As shown in Figure 1, this statement means that the mother believes that at age 24 months the natural log of child development is such that $\ln \hat{q}^{L}_{i,j,k} = \ln 16$.

3.3.2 Estimating MSE Using Age-Range Questions

For the MSD item j and scenario k, suppose that the survey respondent i states that the youngest and oldest age at which a child will learn how to speak partial sentences of three words or more is $\underline{a}_{i,j,k}$ and $\overline{a}_{i,j,k}$ months, respectively. Our interpretation of the answer is that the respondent believes that the probability that the child will be able to speak a partial sentence of three words or more before age $\underline{a}_{i,j,k}$ is a number Δ_0 (arbitrarily) close to zero, and the probability after age $\overline{a}_{i,j,k}$ months is a number Δ_1 (arbitrarily) close to one. To infer the respondent's subjective probability that the child will learn how to speak partial sentences by age 24 months, we need to somehow construct how the probability varies with age. To do so, we show how the age-range data can be used to estimate a respondent i specific IRT model along with the parameterization used in Equation (6). To do so, let $\tilde{d}^{*}_{i,j,k}$ denote the latent variable that is determined according to:

$$\tilde{d}^{*}_{i,j,k} = \tilde{b}_{i,j,k,0} + \tilde{b}_{i,j,k,1} \ln a_{i,j,k} + \tilde{\eta}_{i,j,k} \tag{9}$$

where the shock $\tilde{\eta}_{i,j,k}$ is normally distributed with mean zero and variance one. Similar to the model described in Section 2.2, let $\tilde{d}_{i,j,k}$ denote the binary variable that takes the value one if $\tilde{d}^{*}_{i,j,k} \geq 0$ and zero, otherwise.

Note the parallelism between the IRT model described by Equation (6) and its counterpart represented in Equation (9). The parameters $\tilde{b}_{i,j,k,0}$ and $\tilde{b}_{i,j,k,1}$ in Equation (9) have the same interpretations to the parameters $b_{j,0}$ and $b_{j,1}$ in Equation (6).

However, there are two important differences between the models in Equations (6) and (9). First, the IRT model in Equation (9) describes maternal beliefs about typical development if investment and human capital at birth are defined according to scenario k. Because it

reflects typical from the point of view of the mother, the factor θ_i in Equation (6) is set to zero in Equation (9). In addition, because the IRT model is specific for scenario k, the parameters in (9) are indexed by k.

Second, the model represented in Equation (6) is fitted using actual developmental data from the CNLSY/79 study, while the one in (9) uses respondent i age-range data collected in MKIDS and PHD studies.

If we combine the model in Equation (9) with the age ranges provided by the respondent, we conclude that, according to the respondent i, the probability that the child will learn how to do MSD task j in scenario k by age $\underline{a}_{i,j,k}$ is:

$$\Delta_0 = 1 - \Phi\left[-\tilde{b}_{i,j,k,0} - \tilde{b}_{i,j,k,1} \ln \underline{a}_{i,j,k}\right]. \tag{10a}$$

Analogously, the probability that the child will learn how to do MSD task j in scenario k by age $\overline{a}_{i,j,k}$ is

$$\Delta_1 = 1 - \Phi\left[-\tilde{b}_{i,j,k,0} - \tilde{b}_{i,j,k,1} \ln \overline{a}_{i,j,k}\right]. \tag{10b}$$

If we manipulate the system of Equations (10a) and (10b), we conclude that for arbitrary j and k, the following equalities hold:

$$\tilde{b}_{i,j,k,1} = \frac{\Phi^{-1}(1-\Delta_0) - \Phi^{-1}(1-\Delta_1)}{\ln \overline{a}_{i,j,k} - \ln \underline{a}_{i,j,k}} \qquad \tilde{b}_{i,j,k,0} = \frac{\Phi^{-1}(1-\Delta_1) \ln \underline{a}_{i,j,k} - \Phi^{-1}(1-\Delta_0) \ln \overline{a}_{i,j,k}}{\ln \overline{a}_{i,j,k} - \ln \underline{a}_{i,j,k}}$$

Having estimated the parameters in the IRT model as perceived by respondent i, the next step in the algorithm is to estimate the probability that the child will learn how to say a partial sentence with three words or more by age 24 months. According to individual-specific IRT model, this probability is $p^A_{i,j,k}$ and defined according to:

$$p^A_{i,j,k} = 1 - \Phi\left[-\tilde{b}_{i,j,k,0} - \tilde{b}_{i,j,k,1} \ln 24\right].$$

We use this estimate of the probability, together with the IRT model defined in Equation (6), to derive an error-ridden measure of maternal expectation of the natural log of development at age 24 months, $\ln q^A_{i,j,k}$, from the implied probability $p^A_{i,j,k}$. To do so, we invert Equation (7) and solve for $\theta^A_{i,j,k}$:

$$\ln \hat{q}^A_{i,j,k} = \ln 24 + \frac{\tilde{b}_{i,j,2}}{\tilde{b}_{i,j,1}} \theta^A_{i,j,k} = -\left[\frac{b_{j,0} + \Phi^{-1}\left(1 - p^A_{i,j,k}\right)}{\tilde{b}_{i,j,1}}\right]. \quad (11)$$

3.3.3 Accounting for Measurement Error

Note that $\ln q^L_{i,j,k}$ and $\ln q^A_{i,j,k}$ are two error-ridden measures of maternal expectations about the natural log of child development. We assume that:

$$\ln q^L_{i,j,k} = R_{i,j}\delta_L + E\left(\ln q_{i,1} \left| q_0, x, \mathcal{H}_i\right.\right) + \epsilon^L_{i,j,k},$$

$$\ln q^A_{i,j,k} = R_{i,j}\delta_A + E\left(\ln q_{i,1} \left| q_0, x, \mathcal{H}_i\right.\right) + \epsilon^A_{i,j,k}.$$

The vector $R_{i,j}$ captures heterogeneity across individuals, MSD items, or the type of question asked. In this sense, the measurement error model that we use in this paper allows for the error-ridden measures $\ln q^L_{i,j,k}$ and $\ln q^A_{i,j,k}$ to be biased indicators of the latent variable of interest, $E\left(\ln q_{i,1} \left| q_0, x, \mathcal{H}_i\right.\right)$.

Let $\ln z_k = \ln q_{0,k} \times \ln x_k$ and use Equation (5) to arrive at the following model:

$$\ln q^L_{i,j,k} = R_{i,j}\delta_L + \mu_{\psi,i,0} + \mu_{\psi,i,1} \ln q_{0,k} + \mu_{\psi,i,2} \ln x_k + \mu_{\psi,i,3} \ln z_k + \epsilon^L_{i,j,k} \quad (12a)$$

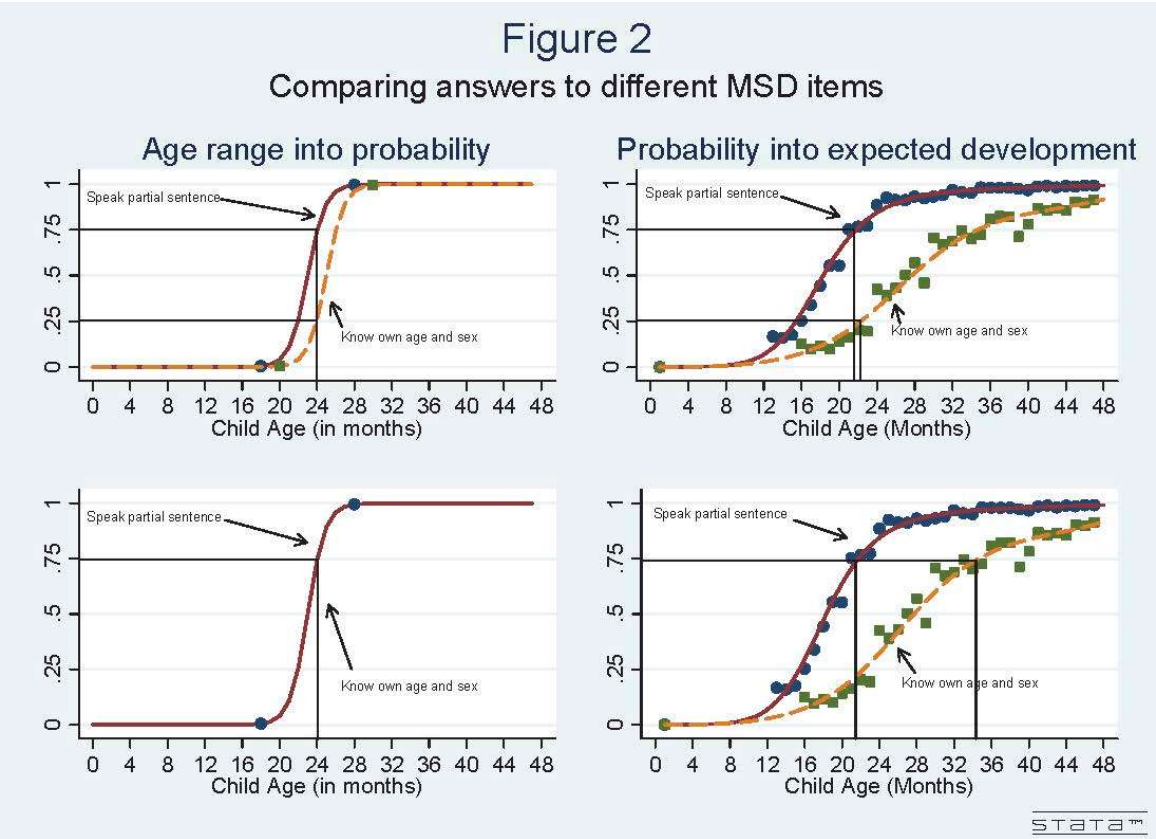
$$\ln q^A_{i,j,k} = R_{i,j}\delta_A + \mu_{\psi,i,0} + \mu_{\psi,i,1} \ln q_{0,k} + \mu_{\psi,i,2} \ln x_k + \mu_{\psi,i,3} \ln z_k + \epsilon^A_{i,j,k} \quad (12b)$$

Equations (12a) and (12b) constitute a linear factor model in which $\mu_{\psi,i} = \left(\mu_{\psi,i,0}, \mu_{\psi,i,1}, \mu_{\psi,i,2}, \mu_{\psi,i,3}\right)$ are the factors, $\epsilon_{i,j,k} = \left(\epsilon^L_{i,j,k}, \epsilon^A_{i,j,k}\right)$ are the measurement errors, and $\lambda_k = \left(1, \ln q_{0,k}, \ln x_k, \ln z_k\right)$ are the factor loadings. Interestingly, note that λ_k is known and fully determined by the description of the scenarios of investments and human capital at birth. The fact that the factor loadings are known reduces the estimation of the model in Equations (12a) and (12b) to the estimation of the distribution of $\mu_{\psi,i}$ and the distribution of $\epsilon_{i,j,k}$.

Thus, it would be possible to estimate the model using only one item of the MSD scale. However, because we have multiple items, we can investigate how the respondents' answers vary across MSD items for a fixed scenario of human capital at birth and investments. For example, the top right panel in Figure 2 shows, for each age, the fraction

of children who can "speak a partial sentence of three words or more" (solid curve). Also shown in the same top right panel in Figure 2 is the fraction of children who "know own sex and age" (dashed curve). Clearly, at each age, there are children who can speak a partial sentence of three words or more but who do not know their own sex and age. This fact indicates that the latter is a more difficult item than the former.

If the respondents understand the survey instrument, we would expect them to assign a lower probability or higher age ranges to items that are more difficult. This is the case depicted in the top left panel of Figure 2. Fixing the scenario in which the baby's health at birth is "good" and investments are "high," this hypothetical respondent provided answers that imply a high probability of "speak[ing] a partial sentence" but a low probability of "know[ing] own age and sex." As a result, once we transform the probability into measures of expected development, the two different measures are quite close in a quantitative sense (top right panel).



It is also possible that respondents report similar probabilities or age ranges for the same scenario across different items. Such a possibility is depicted at the bottom half of Figure 2. In that case, we would see measures of expected development that vary widely from easier to more difficult items. If the results indicate such constancy of age ranges, we would be worried about the possibility that respondents do not understand the instrument very well.

3.3 Eliciting Preferences using Stated-Choice Data

Typically, preference parameters are estimated by using revealed preference data. Unfortunately, this is not possible in our case because we do not observe investments. To estimate $\alpha_{i,1}$ and $\alpha_{i,2}$, we follow a different route. Our approach is to elicit the preference parameter by stated-choice data obtained from the MKIDS and described in Section 2.

For each combination of prices and income, we presented the respondents with the following instructions: "Suppose that your household income is $\$y$ per month and that for each hour per day that the mother spends interacting with the child she has to spend $\$p$ per month on educational goods. Consider the following four options."

The four options represent different levels of investments: two, three, four, or five hours per day interacting with a child. For example, if the mother i chooses $x_{i,m,n}$ hours per day when the price is $\$p_m$ and income is $\$y_n$ then her monthly expenditure is $\$p_m x_{i,m,n}$, and the share of income allocated to investment is $s_{i,m,n} = \frac{\$p_m x_{i,m,n}}{\$y_n}$. Note that variability in the share $s_{i,m,n}$ across respondents i arises strictly because of variability in choices $x_{i,m,n}$ (all respondents face the same set of prices and incomes).

The manipulation of Equation (5), together with the definition of $s_{i,m,n}$, allows us to conclude that:

$$\frac{s_{i,m,n}}{1-s_{i,m,n}} = \alpha_{i,1}\left(\mu_{i,\psi,2} + \mu_{i,\psi,3} \ln q_0\right) + \alpha_{i,2} + \epsilon_{i,m,n}. \tag{13}$$

Equation (13) suggests a factor model in which $\frac{s_{i,m,n}}{1-s_{i,m,n}}$ is the measurement associated with two latent factors ($\alpha_{i,1}$ and $\alpha_{i,2}$) with corresponding factor loadings $\left(\mu_{i,\psi,2} + \mu_{i,\psi,3} \ln q_0\right)$ and one, respectively. The $\epsilon_{i,m,n}$ is a mean-zero error term.

OBJECTIVE ESTIMATION OF THE TECHNOLOGY OF SKILL FORMATION USING THE CNLSY/79 AND INSTRUMENTAL VARIABLE.

4 Results

In this section, we describe the empirical results from our analysis of MKIDS, the PHD Study, and the CNLSY/79.

4.1 Sample Characteristics from MKIDS, PHD Study, CNLSY/79

The analysis in this paper focuses on a very homogenous group of 777 participants. All of the respondents were black. The participants in both studies tended to be young (about 80% of them were at most 25 years old) and had little schooling (18% of the respondents were high school dropouts or have received a GED, 43% had a high school diploma, and 39% had some post-secondary schooling, but only about 5% of them have completed a college diploma). The sample was economically disadvantaged. The median income was below $20,000 per year, which is about the second decile in the US distribution of household income.[7] Finally, the vast majority of the respondents were single. Appendix Table A1 presents additional information on demographic characteristics of respondents.

ADD CNLSY/79

4.2 Subjective Expectations About the Technology of Skill Formation

Before we report our findings about maternal expectations, Table 2 presents the rank ordering of the MSD item by difficulty as estimated from the CNLSY/79 as well as the MKIDS and PHD participants' answers about expectations of child developmental outcomes according to our two forms; stated probability at age two and probability inferred from the age range data (see Sections 3.3.1 and 3.3.2). These probabilities are presented for

[7] For comparison, in 2010 the black median household income was $33,460. The median household income in our survey was roughly half that amount.

the four different scenarios where human capital at birth and investments are varied. Several important results are evident from this table. The probabilities reported by respondents vary in ways that are consistent with basic assumptions of the technology of skill formation. Ceteris paribus, the higher the stock of human capital at birth, or the level of investment, the higher the probability that a baby will learn the MSD tasks used in the elicitation exercise. This feature of the data is true for both elicitation methods.

There are two important differences between the elicitation method that relies on mothers reporting probabilities and the one in which the mothers report age ranges. The first difference relates to how the elicited probabilities correlate with the difficulty of the MSD items. The probabilities derived from answers to the "how likely" questions are uncorrelated with the difficulty of the MSD item. In fact, the likelihood reported by mothers for any given MSD item is around 80% for Scenario 1, 60% for Scenario 2, 66% for Scenario 3, and 49% for Scenario 4.

Table 2
Basic Features of Raw Data

MSD Items ranked in ascending order of difficulty		NHANES	Probability Scenarios					Age ranges Scenarios				
Rank	Item Description		Obs.	1	2	3	4	Obs.	1	2	3	4
1	Child lets someone know that wearing wet pants bothers him/her.	0.99	90	0.78 (0.24)	0.55 (0.27)	0.70 (0.27)	0.51 (0.26)	303	0.64 (0.33)	0.55 (0.36)	0.50 (0.36)	0.43 (0.37)
2	Child speaks a partial sentence of 3 words or more.	0.72	544	0.81 (0.18)	0.63 (0.22)	0.61 (0.20)	0.45 (0.20)	757	0.60 (0.36)	0.44 (0.38)	0.42 (0.38)	0.31 (0.36)
3	Child counts 3 objects correctly.	0.39	544	0.84 (0.18)	0.67 (0.22)	0.62 (0.20)	0.47 (0.20)	757	0.41 (0.38)	0.32 (0.36)	0.26 (0.33)	0.19 (0.30)
4	Child knows own age and sex.	0.31	544	0.83 (0.19)	0.66 (0.23)	0.62 (0.21)	0.47 (0.21)	757	0.33 (0.36)	0.26 (0.33)	0.23 (0.31)	0.17 (0.29)
5	Child says first and last name together without someone's help.	0.26	544	0.80 (0.20)	0.64 (0.22)	0.60 (0.21)	0.46 (0.21)	757	0.31 (0.36)	0.24 (0.33)	0.22 (0.31)	0.17 (0.29)
6	Child says the names of at least 4 colors.	0.20	90	0.81 (0.23)	0.59 (0.28)	0.74 (0.22)	0.56 (0.27)	303	0.26 (0.31)	0.22 (0.29)	0.19 (0.28)	0.16 (0.26)
7	Child counts out loud up to 10.	0.07	90	0.80 (0.20)	0.58 (0.27)	0.75 (0.19)	0.53 (0.27)	303	0.24 (0.30)	0.20 (0.28)	0.19 (0.28)	0.16 (0.27)
8	Child draws a picture of a man/woman, 2 parts besides head.	0.02	90	0.71 (0.25)	0.51 (0.28)	0.67 (0.21)	0.48 (0.26)	303	0.15 (0.26)	0.15 (0.26)	0.14 (0.26)	0.13 (0.25)

Note: Standard errors in parentheses.

This issue can be illustrated by focusing on two MSD items. The first is "speak a partial sentence with three words or more" and the second is "say first and last name." According to the CNLSY/79 dataset, by age 24 months, 72% of the children will have already spoken a partial sentence with three words or more, but only 26% of them will have already said their first and last names. This difference indicates that "say first and last name" is more difficult for a two-year-old child than "speak partial sentence." When mothers are asked the likelihood that children will be able to do these tasks by age two years, their answers for a given scenario are about the same for both items. For example, the typical mother states that, for both items, the probability is around 80% and 45% in Scenarios 1 and 4, respectively. This suggests that mothers believe that these two items have about the same difficulty level, which contradicts the evidence from the CNLSY/79 dataset.

This issue exists, but is arguably far less serious, when mothers are asked to report age ranges. With respect to the two MSD items highlighted above, the transformation of age ranges into probabilities imply that in Scenario 1, 60% of children will know how to speak a partial sentence, but only 31% will say first and last name. Thus, the elicitation according to the age-range methodology suggests that "say first and last name" is more difficult than "speak partial sentence," which is consistent with the CNLSY/79 dataset.[8]

The second difference refers to the fact that both methods produce very different estimates of probabilities, and neither of them are consistent with the probabilities observed in the CNLSY/79 dataset. The issue is more serious for the easiest and hardest MSD items. For example, 99% of children will have learned the easiest MSD item ("let someone know that wearing wet pants/diapers bothers child"). No elicitation method comes close to this figure, even for the best scenarios of human capital at birth and investment. The same

[8] It is possible to summarize the discussion above with simple OLS regressions in which the dependent variables are average probabilities specific for each elicitation method and scenario, reported in Table 2, and the independent variables are a constant and the MSD-item difficulty rank. It turns out that the coefficients on item difficulty are near zero and statistically insignificant for the probabilities generated by "how likely" questions. At the same time, the same coefficients are statistically significant for the age-range questions. We then do the same analysis using the average objective probabilities from the NHANES dataset. In this case, the coefficient of difficulty rank is also negative and statistically significant. This evidence suggests that the elicitation method that relies on mothers reporting age ranges produces a correlation pattern between item difficulty and probabilities that are closer to the ones observed in the NHANES dataset.

conclusion can be reached for the hardest MSD items ("count out loud up to 10" and "draw a picture of a man/woman"). Although very few children are able to do these tasks by age 24 months, none of the methods suggest probabilities that are near what is observed in the data, even when we only consider the worst scenarios of human capital at birth and investments. This conclusion partly explains why the PHD study discarded the easiest and hardest MSD items and only used those items for which the age-range elicitation method performs reasonably well in levels.[9]

Table 3 displays the summary statistics of the subjective expectations with respect to ψ. This table presents the results when we assume that $\Delta_0 = 10\%$ and $\Delta_1 = 90\%$.[10] The typical and median woman believes that the intercept term (the parameter $\mu_{\psi,0}$ in Equation [5]) is 0.115 and 0.108, respectively. There is enormous variability in the beliefs, as the variance is 0.035 (which implies a coefficient of variation of 1.62) and the interquartile range is 0.252, which is about twice the mean value.

Table 3
Maternal Beliefs about the Technology of Skill Formation

	25th percentile	Median	75th percentile	Mean	Variance
$\mu_{\psi,0}$	-0.015	0.101	0.236	0.115	0.035
	(0.009)	(0.008)	(0.009)	(0.007)	(0.002)
$\mu_{\psi,1}$	0.077	0.296	0.554	0.365	0.204
	(0.011)	(0.016)	(0.022)	(0.016)	(0.026)
$\mu_{\psi,2}$	0.065	0.166	0.285	0.192	0.046
	(0.006)	(0.007)	(0.010)	(0.008)	(0.005)
$\mu_{\psi,3}$	-0.008	0.094	0.335	0.190	0.320
	(0.007)	(0.010)	(0.024)	(0.020)	(0.051)

Note: Standard errors in parentheses.

[9] It is important to keep in mind that the elicitation method that uses "how likely" questions can perform better than the age-range questions when it comes to predicting investments. For this reason, the PhD study asked participants both types of questions.

[10] Table 3 shows point estimates and their standard errors, which are quite small. For this reason, we do not discuss our findings about standard errors.

The parameter $\mu_{\psi,1}$ represents maternal beliefs about the coefficient on human capital at birth. Above 90% of survey respondents believe that the higher the human capital at birth, the higher the human capital at age two years. However, there is considerable heterogeneity in beliefs even in this very homogenous sample. While the median mother believes that $\mu_{\psi,1}$ is about 30%, the mother at the 25th percentile believes this figure is about four times smaller. In contrast, the mother at the 75th percentile holds expectations about $\mu_{\psi,1}$ that are almost twice as large as the median mother's expectations.

According to the model presented in Section 3, an important parameter to drive maternal behavior with respect to investment is the parameter $\mu_{\psi,2}$. Approximately 10% of the sample believes that this parameter takes on negative values. In other words, they believe that investments hurt the chances that children will acquire new skills.[11] Again, heterogeneity in beliefs is important. The mean of maternal beliefs is about 19% and the variance is about 5%. These two figures suggest a coefficient of variation around 1.11.

The parameter $\mu_{\psi,3}$ is the term that captures deviations from the Cobb-Douglas function. About 70% of respondents believe that the technology of skill formation has more complementarity between human capital at birth and investments than the one implied by Cobb-Douglas parameterization. This finding has important consequences for how maternal investments respond to the child's human capital at birth. Again, the heterogeneity in beliefs is important for this parameter as well. While the mother at the 25th percentile believes that the production function is Cobb-Douglas, the median mother believes $\mu_{\psi,3}$ is around 10%, and the 75th-percentile mother's expectations are 33%.

For completeness, Panels B and C in Appendix Table A3 report how these beliefs vary as we change the values for the parameters Δ_0 and Δ_1. The distribution of beliefs changes very little as we change the values of these parameters. Thus, we forego discussion about this issue. For the remainder of this paper, we focus our analysis on the case in which $\Delta_0 =$

[11] Interestingly, the HOME-SF contains one item that reads as follows: "Some parents spend time teaching their children new skills while other parents believe children learn best on their own. Which most closely describes your attitude?" Approximately 7% of the black mothers report that the statement that "children learn best on their own" best describes their beliefs.

10% and $\Delta_1 = 90\%$ because this generates the best fit of the data, as indicated by the values of the log likelihood reported in that table.

It is possible that beliefs about the technology of skill formation vary with respect to different definitions of what constitutes "low" versus "high" levels of investments or human capital at birth. Next, we analyze the results from the experimentation about framing scenarios of investments and human capital at birth.

Before we proceed, it is important to remark that all of the data in the baseline scenario were collected via the computer-assisted personal interviewing (CAPI) technique in which an interviewer asked the study participants the questions and entered the answers in the electronic survey instrument. In contrast, all of the data in the alternative scenarios was collected via audio computer-assisted self-interview (ACASI). In this technique, the questionnaire is self-administered, and the computer displays each question and its answer alternatives while presenting a prerecorded interviewer's voice that reads these to the respondent, who listens privately through headphones. Researchers interested in eliciting information about sensitive information (e.g., sexual behavior) worry about face-to-face interviewing methods because it may induce study participants to report what is socially desirable (Waruru, Nduati, and Tylleskar, 2005). In our context, one could be worried that respondents who hold very low expectations would report higher beliefs because they understand that this is a more socially desirable answer.

~~As shown in Appendix Table A2, the main difference between the baseline scenario and the alternative scenario #1 is the technique of the interview. In particular, note that while there are very small differences in the definition of the scenarios for human capital at birth, there are no differences in the definition of the scenarios for investments. As a result, we argue that differences in beliefs between baseline and alternative scenario #1 – if they exist – are probably due to differences in interviewing techniques.~~

Table 4 presents the analysis of the sensitivity of beliefs with respect to the definitions of scenarios. To be sure, we restrict the analysis to the MKIDS sample, as this form of experimentation was restricted to this particular study. Moreover, we show the results

generated under the assumption that $\Delta_0 = 10\%$ and $\Delta_1 = 90\%$, but the qualitative conclusions are invariant to the values taken by that these two parameters.

To construct Table 4, we estimated a seemingly unrelated regression model (SURE, see Zellner, 1962) in which the dependent variables were the beliefs about the parameters $\mu_{\psi,l}$ for $l = 0,1,2,3$, and the regressors were an intercept and a dummy variable for each alternative description of the scenarios for investments. Therefore, the coefficients on these dummy variables capture differences in beliefs relative to baseline.

Table 4
Alternative Definition of Scenarios and Maternal Beliefs

Regressors	Dependent variables				F-Test
	$\mu_{\psi,0}$	$\mu_{\psi,1}$	$\mu_{\psi,2}$	$\mu_{\psi,3}$	(p-value)
Intercept (baseline)	0.018	0.147	0.112	0.070	-
	(0.017)	(0.043)	(0.022)	(0.062)	-
Dummy for alternative scenario #1	0.067	-0.027	-0.032	-0.081	1.080
	(0.037)	(0.094)	(0.048)	(0.136)	(0.364)
Dummy for alternative scenario #2	0.280	0.469	0.175	0.424	33.910
	(0.028)	(0.071)	(0.037)	(0.103)	(0.000)
Dummy for alternative scenario #3	0.206	0.027	0.051	0.091	6.750
	(0.041)	(0.104)	(0.054)	(0.152)	(0.000)

Note: Standard errors in parentheses, except in the F-test column where we report p-values.

First, note we cannot reject the null, and we conclude that CAPI and ACASI methods generate similar estimates of beliefs. We arrive at this conclusion by testing that the coefficient on "dummy for alternative scenario #1" is equal to zero. This is the case if we test the coefficients on each equation separately or if we perform a joint F-test in which the null hypothesis is that the four coefficients on the "dummy for alternative scenario #1" are all equal to zero. The F statistic is 1.08 with a corresponding p-value equal to 0.364.

We find mixed results about whether the definition of scenarios matter for beliefs or not. On one hand, the coefficients on the dummies for alternative scenario #2 imply uniformly

higher beliefs for all parameters ψ. The results for "dummy for alternative scenario #2" are particularly striking because both the separate and the joint tests reject the null hypothesis of no difference.

On the other hand, the results based on the analysis of the coefficients on the "dummy for alternative scenario #3" are not as conclusive. True, the F statistic is large enough to reject the null hypothesis that the coefficients are all equal to zero, but it is possible that this result is likely driven by the large differences regarding $\mu_{\psi,0}$. The regressions involving the beliefs about $\mu_{\psi,1}$, $\mu_{\psi,2}$, and $\mu_{\psi,3}$ do not show differences between the baseline and alternative scenario #3.

The main lesson from this analysis is that it is necessary to more deeply investigate how the framing of scenarios affects the elicitation of beliefs.

4.3 Preferences

Appendix Figure A4 plots the demand function of investment for each level of income (left panel) and the Engel curve for each level of price (right panel). Clearly, the demand for investments is a decreasing function of prices, and as income rises, so does the amount of investments chosen by the respondents.

We can estimate shares for each respondent i from: $\hat{s}_i = \frac{1}{9}\sum_{m=1}^{3}\sum_{n=1}^{3} s_{i,m,n}$. In our sample, the mean and median shares of expenditure on investments are around 8%. In comparison, Lino (2012) reported shares of investment at around 7% for low-income parents. Given the estimated shares, we estimate Equation (13) to recover factor scores $\alpha_{i,1}$ and $\alpha_{i,2}$ for each mother i. Table 5 displays summary statistics for the parameters that describe preferences. When we account for heterogeneity in beliefs, we find that the typical woman has $\alpha_{i,1}$ equal to 3% and $\alpha_{i,2}$ close to 8%. There is heterogeneity in preferences. The coefficient of variation for $\alpha_{i,1}$ is about 50%, and the one for $\alpha_{i,2}$ is a

little over 20%. These figures are much lower than the coefficients of variation found for beliefs about the technology of skill formation.[12]

4.4 Objective Estimation of the Technology of Skill Formation

We rely on the CNLSY/79 data to objectively estimate the technology of skill formation - Equation (1). Appendix B provides a description of the data set and summary statistics for the variables and the sample used in these analyses.

Table 5
Maternal Beliefs about the Technology of Skill Formation

	25th percentile	Median	75th percentile	Mean	Variance
$\alpha_{i,1}$	0.026	0.031	0.040	0.031	0.00024
	(0.000)	(0.000)	(0.001)	(0.000)	(0.00001)
$\alpha_{i,2}$	0.067	0.078	0.094	0.079	0.00031
	(0.000)	(0.001)	(0.001)	(0.000)	(0.00001)

Note: Standard errors in parentheses.

In order to objectively estimate the technology of skill formation, we assume that the dependent variable in Equation (2), $q_{1,i}$, is the child's cognitive development around age 24 months, which in the CNLSY/79 is measured by the MSD scale. In order to maintain comparability with the analysis in Section 4.2, we transform maternal answers into a scale measured in time (i.e., age in months) using the IRT model estimated with the NHANES dataset.

Correspondingly, x_i is investment during the first 24 months of the child's life. In the CNLSY/79, investment is measured by the HOME-SF. As in Cunha, Heckman, and

[12] The preference parameters are correlated. An increase in $\alpha_{i,1}$ by one standard deviation is associated with an increase over half a standard deviation in $\alpha_{i,2}$. The preference parameters $\alpha_{i,1}$ and $\alpha_{i,2}$ are negatively and weakly correlated with $\mu_{i,\psi,0}$. One standard deviation increase in $\mu_{i,\psi,0}$ is associated with a 10% standard deviation reduction in $\alpha_{i,1}$ or $\alpha_{i,2}$. For $j = 2,3$, one standard deviation increase in $\mu_{i,\psi,j}$ is associated with approximately a 15% and 25% increase in $\alpha_{i,1}$, respectively. For $\alpha_{i,2}$, these figures are 8% and 15%, respectively. A small difference is that $\alpha_{i,1}$ is weakly correlated with $\mu_{i,\psi,1}$ (one standard deviation increase in $\mu_{i,\psi,j}$ is associated with an increase in $\alpha_{i,1}$ that is 10% of its standard deviation), while $\alpha_{i,2}$ is not.

Schennach (2010), we factor analyze the items of the HOME-SF scale. In their analysis, the scale of the factor was set by the number of children's books in the household. Although this is a valid metric, this was not convenient for the current study. To maintain consistency with the analysis in Section 4 above, it is necessary to set the location and scale of the instrument in a metric of time (months per year). Details of the procedure are also described in Appendix B.

Finally, $q_{0,i}$ is measured by the child's health at the time of birth. Among other information, the CNLSY/79 data set asks parents to report the child's weight and length at birth, the length of the gestation, and the number of days that the child spent in the hospital after birth. In order to produce a scalar variable, we factor analyze the four measures above and extract one factor. The location and scale of the factor are set by the gestation length. This is convenient because gestation length is measured in number of months, which is the same unit used for cognitive skills around 24 months.[13]

[13] Tables B1-B3 in Appendix B describe in detail summary statistics for the CNLSY/79 variables that we use for the estimation of the technology of skill formation (2). For example, the stocks of skills for the typical Hispanic, black, and white children around 24 months are, respectively, 24, 26.4 and 25.6 months. The black-white difference is not statistically significant. The advantage of black children in the MSD scale arises partly due to the fact that they exhibit superior performance in motor items. In terms of investments, the typical white child tends to receive around 2.2 months of investments per year, while the median black child receives only 1.5 months per year. This difference is statistically significant even after we account for the differences in family backgrounds of children.

Table 6
Objective Estimation of the Technology of Skill Formation
Dependent variable: Natural log of skills around age 24 months[1]

	Maternal Fixed Effect Procedure			
	13 to 35 Months	16 to 32 Months	19 to 29 Months	22 to 26 Months
	(1)	(2)	(3)	(4)
Natural logarithm of health at birth[2]	0.563***	0.509***	0.600**	0.19
	(0.138)	(0.165)	(0.304)	(0.488)
Natural logarithm of investments[3]	0.180***	0.199***	0.257***	0.283***
	(0.020)	(0.027)	(0.038)	(0.081)
Constant	1.822***	1.999***	1.702**	0.92
	(0.308)	(0.368)	(0.676)	(1.703)
Observations	4,721	3,515	2,243	1,014
R-squared	0.734	0.624	0.516	0.61
Number of Mothers	3,042	2,542	1,814	915

Robust standard errors in parentheses. All regressions have dummy variables for: (i) the child's gender, (ii) the child's birth order, (iii) child's year of birth, (iv) the child's age at the time of the measurement of the MSD score, and (v) maternal age at the time of the child's birth. *** p<0.01, ** p<0.05, * p<0.1

[1]Skills are measured by the Motor-Social Development Scale and are scaled in "mental" age of development.

[2]Health at birth is captured by factor analyzing weight at birth, length at birth, gestational age, and number of days in hospital. The scale and location of the factor are determined by gestational age.

[3]Investments are measured by the components of the HOME-SF instrument. We factor analyze the items and set the location and the scale of the factor in months/year of direct engagement between mother and child.

We use within-family variation to estimate the parameters of the technology of skill formation. Thus, in the empirical application that follows, we consider the following parameterization of the technology of skill formation:

$$\ln q_{1,i,l} = \psi_{0,i} + \psi_1 \ln q_{i,0} + \psi_2 \ln x_i + \psi_3 \ln q_{i,0} \times \ln x_i + R_{i,l}\beta + \nu_{i,l} \qquad (14)$$

where the index l denotes the birth order of the child and $R_{i,l}$ the observed characteristics of child l (e.g., the child's gender, birth order, year of birth, and the age at the time of the MSD test).

Table 6 shows the estimated parameters of the technology of Equation (14). In all of the regressions we show in Table 6, we control for the child's age at the time of the interview, the child's year of birth (to account for cohort effects), dummy variables for maternal age at the time of the child's birth, a dummy variable for the child's gender, and dummy

variables for the child's birth order. [14] We start by showing the results when we use the least restricted sample: we include all children whose age at the time of the MSD measurement is between 13 and 35 months.[15] For this sample, the elasticity of skills with respect to investment (i.e., the parameter ψ_2) is 18%. This means that a 10% increase in investments translates into a 1.8% increase in skills at age 24 months. Column (2) restricts the age range of children at the time of the interview to 16 and 32 months. Interestingly, we find that the elasticity parameter is about 10% higher (around 20%). Column (3) displays the results when we work with an even more restricted sample: we only include the children who are between 19 and 29 months old. We findψ_2 to be significantly higher in this sample: the elasticity in the overall sample is 26%, which is about 43% higher than when we work with the least restricted sample.[16] The higher values of ψ_2 may be due to the fact that the components of the MSD instrument applied to older children focus on developmental dimensions that are more affected by parental investments. Another possibility is that the families for which we observe child development more closely at around 24 months are the same families that have high values of ψ_2.

4.5 Quantifying the Importance of MSE

In this Section, we simulate the impact on maternal investments of a policy that changed MSE to the estimated values of the parameter vector ψ estimated in the CNLSY/79. We begin by comparing the importance of preference parameters and MSE in explaining heterogeneity in investments. Table 7 is generated by simulating maternal investments for the situation in which preference parameters, MSE, income, and prices were set at their median levels. In the first row, we investigate how investments change if we set the

[14] To focus on the parameter of interest, Appendix Table B4 reports our estimates for the other parameters in (18) for the full sample regression.

[15] We choose ages 13, 16, 19, and 22 as the cutoff ages owing to the structure of the MSD instrument. As explained in Section 3.1, Part E of the MSD instrument is given to children who are at least 13 and at most 15 months old. The parents of children who are at least 16 and at most 18 months old respond to Part F. Part G is assigned to the parents of children who are between 19 and 21 months old. Finally, Part H is answered by parents whose children are at least 22 and at most 47 months. The end date is determined so that age 24 months is the center of the interval.

[16] If we only include the respondents whose development is measured between 22 and 26 months, our estimate for γ is 28%. However, the sample size becomes too small to be decomposed in the smaller subsamples presented in Table 5.

parameter α_1 at the 75th percentile (which corresponds to a 28% change in α_1), while maintaining everything else at the 50th percentile. As shown in the first row, investments would increase by 1.6%, which implies an elasticity of 5.8%.

If we move the parameter α_2 from the median to the 75th percentile, the move corresponds to a change of 21.4%. Investments, in this case, increase by 18.3%. Thus, the elasticity is high, at 85.2%, indicating of the variability in investments is as yet unexplained.

Table 7

Comparative Statics of Investments

	Median	75th percentile	% Change in investments	% Change in parameter	Elasticity
α_1	1.70	1.73	0.02	28.0%	5.8%
α_2	1.70	2.01	0.18	21.4%	85.2%
$\mu_{\psi,2}$	1.70	1.77	0.04	72.0%	5.8%
$\mu_{\psi,3}$	1.70	1.70	0.00	257.1%	0.1%
$\mu_{\psi,3}$	1.70	1.86	0.09	257.1%	3.6%

This table shows the comparative statics of optimal investments in relation to preference and belief parameters. Each row shows what happens to investments as we move one parameter and fix the other parameters at the median value. In the last row, we replace the human capital at birth from the mean value to the value at the first percentile.

Next, we investigate how MSE affect investments. If we increase $\mu_{\psi,2}$ by 72% (from median to 75th percentile), investments increase by 4.1%, which implies an elasticity of 5.8%. If we change $\mu_{\psi,3}$ from median to the third quartile—which is equivalent to an increase of 257.1%—investments change by only 0.2%, which indicates negligible elasticities. However, this is driven by the point at which the log of natural log of child development is evaluated (which is at the median value). If we evaluate the elasticity at the first percentile of q_0, the elasticity is 3.6%.

Finally, we use our data to answer the following question. Suppose we were to carry out an intervention that set MSE exactly equal to the parameters of the technology of skill

formation as estimated from the CNLSY/79. What would be the impact of such intervention on investment?

As shown in Table 8, our estimates suggest that such policy would increase investments by at least 4% and at most 12%. In other words, although many mothers underestimate the importance that investments play in the development of their children's skills, the model suggests that the channel through which MSE affects investments—namely, maternal preferences for child development—is not strong enough to be a major determinant of investment.

Table 8

Maternal Beliefs and Technology

Cases	Factual investment	Counterfactual investment	% Change	Effect size
$\mu_{\psi,2} = 0.267$ $\mu_{\psi,3} = 0.000$	1.84	1.92	4.4%	10.3%
$\mu_{\psi,2} = 0.454$ $\mu_{\psi,3} = 0.000$	1.84	2.05	11.7%	26.9%

The model can be used to generate estimates of the effect sizes of policies that change maternal beliefs. The effect sizes are between 10% and 27%. If the effect size is 10% and we assume that the intervention treats parents separately and that there is no attrition, it would necessitate a sample size of 3,140 parents, half in the treatment group and the other half in the control group. If the effect size is 27%, then it is necessary to have 532 study participants, half in each group.

Conclusion

In this paper we investigate the important of MSE on the degree to which a mother invests in the development of her child. To accomplish our goal we collected primary data on 777 poor, urban, African-American pregnant women to elicit their MSE as well as preference

parameters. Using the CNLSY/79 objective estimates of the technology of skill formation were obtained through a fixed-effect and instrumental-variable technique. We formulate a simple model in which mothers have MSE about the technology of skill formation. We show that the model can be used to evaluate the impact of policies that affect MSE to develop the human capital of children. We use the primary data to separately identify heterogeneity in preferences from heterogeneity in MSE. Our results indicate that most women underestimate the parameter values of the technology of skill formation. Although we see this persistent underestimation, this does not appear to explain a large proportion of the variability in investments. We also see that investments would increase by about 12% and that the children's stocks of cognitive skills at age 2 would increase by about 5.4% if we set MSE equal to the estimated parameters of the technology of skill formation. The impact on investment and child development is higher for mothers whose MSE is below the median value and for mothers who have high preference towards child developmental outcomes which is captured by the parameter α_1 in our model. Although most mothers do not believe that the technology of skill formation follows a Cobb-Douglas specification, our results indicate that this departure has small effects on investment choices. In particular, investment choices do not vary greatly by human capital at birth, as would be predicted by the Cobb-Douglas functional form. However, we see that investments have high elasticity with respect to the preference parameter α_2. Unfortunately, our model does not shed any light on the nature of what this parameter is measuring.

A large body of work often referred to as Home Visitation Programs have as objectives increasing knowledge and changing attitudes and beliefs. Overall, these programs have been largely unsuccessful in changing behavior. However, Suskind et al (2013) showed that it is possible to improve children's language environment through a randomized home-based intervention that provides quantitative linguistic feedback to the mother and that this feedback influences the mother's linguistic behavior. In this work, the authors used the Language Environment Analysis (LENA) technology which provides counts of adult words (AWC) as well as conversation turns (CTC). Mothers in the treatment group showed a significant and prolonged increase from mean baseline to mean post-intervention

for AWC (effect size is 36%) and CTC (effect size is 25%). These figures are consistent with the predictions of our model. Additionally, pre and post evaluation showed that mothers in the treatment group had a significant increase in knowledge about the importance of the home language environment on the child's development. This work shows that interventions that affect MSE can translate into behavior change. Potentially, this program was successful because of its specificity, quantifiable targets, and low participant burden. However, these attributes are not common in home visitation programs, the vast majority of which show very little efficacy (add Gomby et al, 1999; Heckman, 2015).

Future work should focus on specifying the elements that are captured by the parameter α_2. Our model suggests that a policy that moves a mother from XXX to XXXX percentile in the distribution of this parameter would lead to an 85% increase in maternal investments. This is a very large change in investments. Because we do not understand the elements that are captured by this parameter we do not know how to design the next generation of interventions, whether such interventions would change maternal investments, and if they would be cost effective.

References

Aizer, Anna and Laura Stroud (2010): "Education, Medical Knowledge and the Evolution of Disparities in Health," NBER Working Paper 15840.

Attanasio, Orazio P. and Katja M. Kaufmann (2009): "Educational Choices, Subjective Expectations, and Credit Constraints," NBER Working Paper 15087.

Bradley, Robert H., and Bettye M. Caldwell (1980): "The Relation of Home Environment, Cognitive Competence, and IQ Among Males and Females," *Child Development*, Vol. 51(4): 1140–1148.

Bradley, Robert H., and Bettye M. Caldwell (1984): "The Relation of Infants' Home Environments to Achievement Test Performance in First Grade: A Follow-Up Study," *Child Development*, Vol. 55(3): 803–809.

Carneiro, Pedro, Costas Meghir, and Matthias Parey (2013): "Maternal Education, Home Environments and Child Development," *Journal of European Economic Association*, Vol. 11(IS1):123-160.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach (2010): "Estimating the Technology of Cognitive and Non-Cognitive Skill Formation," *Econometrica*, Vol. 78(3): 883-931

Currie, Janet and Enrico Moretti (2002): "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings and Longitudinal Data," NBER Working Paper 9360.

Delavande, Adeline, Xavier Giné, and David McKenzie (2011): "Eliciting Probabilistic Expectations with Visual Aids in Developing Countries: How Sensitive Are Answers to Variations in Elicitation Design?," *Journal of Applied Econometrics*, Vol. 26(3): 479-497.

Epstein, Ann S. (1979): "Pregnant Teenagers' Knowledge of Infant Development," High/Scope Educational Research Foundation Working Paper, available at http://www.eric.ed.gov/PDFS/ED176875.pdf.

Hart, Betty and Todd R. Risley (1995): "Meaningful Differences in the Everyday Experience of Young American Children," Paul H. Brookes Publishing Co.

Heckman, James J. and Edward Vytlacil (1998): "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return is Correlated with Schooling," *Journal of Human Resources*, Vol. 33(4): 974-987.

Jensen, Robert (2010): "The (Perceived) Returns to Education and the Demand for Schooling," Quarterly Journal of Economics, Vol. 125(2): 515-548.

Lino, Mark (2012): "Expenditure on Children by Families, 2011," U.S. Department of Agriculture, Center for Nutrition Policy and Promotion. Miscellaneous Publication No. 1528-2011.

Manski, Charles F. (2004): "Measuring Expectations," *Econometrica*, Vol 72(5): 1329-1376.

Ninio, Anat (1988): "The Effects of Cultural Background, Sex, and Parenthood on Beliefs about the Timetable of Cognitive Development in Infancy," *Merrill-Palmer Quarterly*, Vol. 34(4), pp. 369-388.

Rowe, Meredith L. (2008): "Child-directed speech: Relation to socioeconomic status, knowledge of child development, and child vocabulary skill," Journal of Child Language, Vol. 35(1): 185-205.

Waruru, Anthony K., Ruth Nduati and Thorkild Tylleskar (2005): "Audio computer-assisted self-interviewing (ACASI) may avert socially desirable responses about infant feeding in the context of HIV," *BMC Medical Informatics and Decision Making*, Vol. 5 (24) doi:10.1186/1472-6947-5-24.

We interviewed a sample of socioeconomically disadvantaged, pregnant African-American women. We found that the subjective expectation about the elasticity of child development with respect to investments depends on the child's human capital at birth. If the child's human capital at birth is at the 25th percentile, the median parent believes the elasticity is 28.5%. For a child at the 30th percentile in the distribution of human capital, the median parent believes the elasticity is 30%. In comparison, when we estimated the technology of skill formation from the Children of the National Longitudinal Survey of Youth (CNLSY)/79 data, also using the Motor Social development (MSD) scale, we found that the elasticity can be as high as 45%.

In this section, we start by presenting the MSD scale, which played a major role in the development of the elicitation questionnaire. Then, we present the survey questionnaire items that were used to elicit expectations. We show how to transform maternal answers to these items into maternal expectations about child development (the left-hand side variable in Equation [2]). The next step shows how to use maternal expectations data to recover

expectations about the parameters of the technology of skill formation. Finally, we show our procedure to identify the parameters of the utility function.

3.1 The Motor Social Development Scale

The MSD scale played an important role in our analysis, so we briefly explain it in this section. This scale was used in the CNSLY/1979 and in the National Health and Nutrition Examination Study 1988 (NHANES). In the MSD instrument, mothers answer 15 out of 48 items regarding motor, language, and numeracy development. These items are divided into eight components (parts A through H) that a mother completes contingent on the child's age. Part A is appropriate for infants aged zero through three months, and the most advanced section, Part H, is addressed to children between the ages of 22 and 47 months. All items are dichotomous (scored *"no" is equal to zero* and *"yes" is equal to one*) and the total raw score for children of a particular age is obtained by a simple summation (with a range 0 to 15) of the affirmative responses in the age-appropriate section. Because the age at which children learn how to do given tasks varies considerably across children, one MSD item may be present in many of the parts of the instrument. For example, the MSD item "speak a partial sentence of three words or more" is asked about children who are between 13 and 47 months. Thus, this particular item is a member of parts E, F, G, and H of the MSD instrument.

Two key properties of the MSD instrument that make it appealing to our goal is that the tasks are described in language easily understood by the mothers, and the tasks are recognizable based on the daily interactions of mothers and their children. In fact, this is one important reason why we took the MSD instrument as a starting point in the development of the questionnaire to elicit maternal expectations.

Another important reason to start from the MSD instrument is that we ensure that comparability is maintained. The set of items used to elicit maternal subjective expectations about child development is the same one used to measure actual child development in the objective estimation of the technology of skill formation, Equation (5), that we employ in Section 4.