ELICITING MATERNAL EXPECTATIONS ABOUT THE TECHNOLOGY OF COGNITIVE SKILL FORMATION

Flávio Cunha
Irma Elo
Jennifer Culhane

Working Paper 19144
http://www.nber.org/papers/w19144

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2013

We thank Dalton Banks, Michelle Gifford, Debbie Jaffe, Snejana Nihtianova, Ben Sapp, 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 CESifo 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 INO1200013 from the Institute for New Economic Thinking. All remaining errors are ours. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation
Flávio Cunha, Irma Elo, and Jennifer Culhane
NBER Working Paper No. 19144
June 2013
JEL No. I10,I20,J24

ABSTRACT

In this paper, we formulate a model of early childhood development in which mothers have subjective expectations about the technology of skill formation. The model is useful for understanding how maternal knowledge about child development affects the maternal choices of investments in the human capital of children. Unfortunately, the model is not identified from data that are usually available to econometricians. To solve this problem, we conduct a study where mothers were interviewed to elicit maternal expectations about the technology of skill formation. We interviewed a sample of socioeconomically disadvantaged AfricanAmerican women. We find that the median subjective expectation about the elasticity of child development with respect to investments is between 4% and 19%. In comparison, when we estimate the technology of skill formation from the CNLSY/79 data, we find that the elasticity is between 18% and 26%. We use the model and our unique data to answer a simple but important question: What would happen to investments and child development if we implemented a policy that moved expectations from the median to the objective estimates that we obtain from the CNLSY/79 data? According to our estimates, maternal investments would go up by between 4% and 24% and the stocks of cognitive skills at age 24 months would subsequently increase between 1% and 5%. Needless to say, the impacts of such a policy would be even higher for mothers whose expectations were below the median.

Flávio Cunha
Department of Economics
University of Pennsylvania
160 McNeil Building
3718 Locust Walk
Philadelphia, PA 19104-6297
and NBER
fcunha@sas.upenn.edu

Irma Elo
University of Pennsylvania
Population Studies Center
Department of Sociology
3718 Locust Walk
Philadelphia, PA 19104
popelo@pop.upenn.edu

Jennifer Culhane
Children's Hospital of Philadelphia
34th Street and Civic Center Boulevard
Philadelphia, PA. 19104-4399
culhanej@email.chop.edu

1. Introduction

Research shows that the early environment that children face has long-term consequences for their human capital development (e.g., Karoly et al, 2005; Cunha and Heckman, 2007; Campbell et al, 2008; Hoddinott et al, 2008; Heckman et al, 2010; Almond and Currie, 2010). A central question of interest for public policy is whether early investments depend on the parental expectations about the importance of the environment for their children's development. If so, it is conceivable that policymakers can improve the outcomes of children through informational interventions that affect parental expectations.

The issue that low subjective expectations about returns may affect investments has been recognized in developmental psychology for over 50 years (Hunt, 1961; Vygostky, 1978). Our research is related to, but different from, that field's large body of literature that focuses on measuring maternal and paternal knowledge about child development. These studies show that the lower the parents' socio-economic status (SES), the lower their expectations about cognitive development (e.g., Epstein, 1979; Hess et al., 1980; Ninio, 1988; Mansbach and Greenbaum, 1999).

The gaps in beliefs about when children will master certain skills may arise for different reasons. First, the gaps may be a product of the differences in investments that arise even when parents have the same beliefs about the technology of skill formation. These differences in investments could be associated with differences in preferences, resources, or parental/offspring characteristics (e.g.,Behrman, Pollak, and Taubman, 1982; Becker and Tomes, 1986). Second, the gaps may be a product of the differences in beliefs holding investments fixed. Obviously, it is also possible that some of the differences in investments arise because of differences in beliefs as well as differences in preferences, resources, and parental/offspring characteristics. Unfortunately, the data collected so far in the different fields of the social sciences do not allow us to isolate the importance of these different factors.

Even though no studies quantify maternal beliefs about the technology of skill formation, a few studies do suggest that maternal beliefs play a significant role in the determination of investments. In a pioneering study, Hart and Risley (1995) documented the differences in how much parents of different SES talked to their babies. Children whose families were on welfare heard about 600 words per hour. In contrast, children from professional families heard over 2,000 words per hour. Hart and Risley (1995) also showed that the better the early language environment at home (as measured by the number of words or conversational turns) the better the language development of children, the higher their IQ, and the better they did in school.

A natural question that arises is: Why do high SES parents talk more to their children than low SES ones? The research by Rowe (2008) argues that the gaps in the early home language environment exist because poor, uneducated mothers do not know about the role it plays in determining the language and cognitive development of their children. Indeed, Suskind et al (2013) show that it is possible to improve the home language environment through an informational intervention with low SES mothers. The experimental intervention, known as the Thirty Million Word Project, had sizable impacts in the early home language environment as measured by conversational turns (about 30 per hour in the control group versus 45 in the treatment group). It also had a large impact on child language development as measured by the number of child vocalizations (about 120 per hour in the control group versus 180 per hour in the treatment group).

Attanasio and Kauffman (2009) show that the higher the subjective expectations of the returns to schooling, the more likely the decision to invest in education. In his study in the Dominican Republic, Jensen (2010) finds that students hold *subjective* expectations of returns to schooling that imply extreme underestimation of the *objective* returns to schooling. More important for the purposes of the current paper, Jensen shows that individuals react to new information: Students in randomly selected schools who were given information about the higher measured returns completed on average 0.20–0.35 more years of school over the next four years than those who were not.

Recent literature provides evidence that expectations about the technology of skill formation can be changed by public information campaigns, but it is important to keep in mind that the effectiveness of the campaigns may differ across SES groups. Aizer and Stroud (2010), for example, track 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.

Further evidence that public policy can alter expectations comes from developing countries. Field et al (2009) show that iodine deficiency is causally linked to deficits in cognitive development. In related research, Roy (2009) investigates how maternal investments are affected by maternal knowledge that iodized salt prevents early onset brain damage. Her research documents that mothers often elect not to use iodized salt, despite its being inexpensive, because they are simply unaware of its positive impact on child development. Roy also found that uptake of iodized salt increased after parents became aware of its benefit. As expected from Field et al's (2009) results, Roy also finds sizable improvements in children's cognitive development.

In this paper, we formulate a model of early childhood development in which mothers have subjective expectations about the technology of skill formation. The model is useful for understanding how maternal knowledge about child development affects maternal choices about investment in the human capital of children. Unfortunately, the model is not identified from data that are usually available to econometricians. 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 created a survey instrument to elicit maternal expectations about the technology of skill formation. In summary, in this survey we create scenarios of "high" and "low" levels of investments. For each investment scenario, we ask the respondents to provide the youngest and oldest age at which they believe a baby will learn how to do a set of tasks. The tasks are taken from the Motor-Social Development (MSD) Scale used in the Children of the National Longitudinal Survey of Youth/1979 (CNLSY/79) and the National Health and Nutrition Examination Survey (NHANES). By comparing the answers to the questions about the "high" investment scenarios with those of the "low" investment scenarios, we are able to estimate maternal subjective expectations about the technology of skill formation.

We interview a sample of socioeconomically disadvantaged, pregnant African-American women. We find that the median subjective expectation about the elasticity of child development with respect to investments is between 4% and 19% and our preferred set of estimates is clustered around the high values in this interval. In comparison, when we estimate the technology of skill formation from the CNLSY/79 data, also using the MSD scale, we find that the elasticity is between 21% and 36%.

We find that the subjective expectations are positively correlated with the child's health at birth, as proxied by variables such as birth weight, birth length, and gestational age. More important, we also document large heterogeneity in subjective expectations even after accounting for measurement error in a systematic fashion. In our preferred set of estimates, the 25th percentile is between 2% and 5% and the 75th percentile is between 24% and 40%. The distribution of beliefs is positively skewed, so even in this very disadvantaged sample, some mothers have very high expectations about the elasticity of child development with respect to investments.

We use the model and our unique data to answer the following question: Consider the median mother in our survey. What would happen to investments and child development if we implemented a policy that moved her subjective expectation to the objective estimates that we obtain from the CNLSY/79 data? According to our estimates, investments would go up by between 4% and 24% and the stocks of cognitive skills at age 24 months would increase between 1% and 5%. The impacts of such a policy would be even higher for mothers whose expectations were below the median.

The paper is organized as follows. In Section 2, we present a simple model of maternal subjective expectations about the technology of skill formation, and we illustrate the identification problems that arise from using this simple model. In Section 3, we introduce our methodology for estimating maternal subjective expectations about the technology of skill formation. We present the results from our survey in Section 4. In Section 5, we estimate the remaining components of the model presented in Section 2. This is done so that we can evaluate the economic importance of introducing a policy that moves maternal expectations.

2. Model

Consider a mother i who has just given birth to a child whose health is $q_{0,i}$. To make the setting as simple as possible, consider a static problem in which the mother's preferences depend on household consumption, c_i, and the child's cognitive development at the end of the period, $q_{1,i}$.[1] Suppose that maternal preferences are represented by the following Cobb-Douglas function:

$$u_i\left(c_i, q_{1,i}\right) = \ln c_i + \alpha_i \ln q_{1,i} \quad (1)$$

where α_i captures how much mother i values the child's cognitive development. The child's cognitive development at the end of the period is determined by the following technology of skill formation:

$$\ln q_{1,i} = \ln A + \rho \ln q_{0,i} + \gamma_i \ln x_i + \theta_i + \nu_i \quad (2)$$

[1] In the empirical application below, we measure $q_{1,i}$ by developmental tests around the time the child is 24 months old.

where A is an intercept, x_i is the maternal investment in the child's cognitive skills, θ_i is the maternal efficiency in producing skills, and ν_i are shocks. The parameter γ_i captures the elasticity of child development with respect to investments. Given preferences, the higher the value of γ_i, the higher the maternal investments in the cognitive skills of the child.[2]

Let y_i and π_i denote, respectively, the mother's income and the relative price of investment. In our simple set-up, the mother faces the following budget constraint:

$$c_i + \pi_i x_i = y_i \tag{3}$$

Finally, let $\mathcal{H}_i$ denote the mother's information set. The mother's problem is to choose consumption and investment that maximize:

$$E\left(u_i\left(c_i, q_{1,i}\right)\middle|\mathcal{H}_i\right) \tag{4}$$

subject to (1), (2), and (3). It is common to explicitly assume that the mother's information set contains the state variables that influence the choice of investment and consumption. For example, if we assume that at the time investment decisions are made the variable θ_i is known by the mother, but ν_i is not, then the state vector is $\left\{q_{0,i}, y_i, \theta_i, \pi_i\right\} \subset \mathcal{H}_i$. Typically, it is implicitly assumed that the parameter vector is $\{\alpha_i, A, \rho, \gamma_i\} \subset \mathcal{H}_i$. As a result, the expectation in (4) is with respect to the shock ν_i, which is not observed by the mother at the time of the investment choice. In particular, note that $E(\gamma_i|\mathcal{H}_i) = \gamma_i$. As a result, the optimal investment is given by:

$$x_i^* = \frac{\alpha_i\gamma_i}{1+\alpha_i\gamma_i}\frac{y_i}{\pi_i} \tag{5}$$

Given data in which the vector $D_i = \left(q_{0,i}, q_{1,i}, x_i, y_i, \pi_i\right)$ is observed, it is possible to identify the distribution of γ_i directly from the estimation of the technology of skill formation (2) (e.g., Heckman and Vytlacil, 1998). If we estimate (5), we clearly identify the share $s_i = \frac{\alpha_i\gamma_i}{1+\alpha_i\gamma_i}$. We can then recover the utility parameter α_i:

$$\alpha_i = \frac{s_i}{1-s_i\gamma_i} \tag{6}$$

The assumption that $\gamma_i \in \mathcal{H}_i$ guarantees that the parameters α_i and γ_i can be identified given data D_i. We can use the estimated parameters to answer questions such as what

[2] Cunha and Heckman (2007) work with a CES technology of skill formation. The simpler specification we adopt in this paper is motivated by the fact that Cunha, Heckman, and Schennach (2010) do not reject the Cobb-Douglas production function for cognitive skills at early stages of the lifecycle.

would happen to child development and investments if we had a program that increased maternal income or if we implemented a policy that subsidized the price of investments.

The model, however, cannot be used to understand informational issues about the technology of skill formation precisely because, by assumption, $\gamma_i \in \mathcal{H}_i$. In principle, one could assume that $\gamma_i \notin \mathcal{H}_i$ and $E(\gamma_i|\mathcal{H}_i) = \mu_{\gamma,i}$. However, one cannot separately identify the maternal expectation about the elasticity of cognitive development with respect to investment, $\mu_{\gamma,i}$, from maternal valuation of the child's cognitive development, α_i. The identification problem arises because investment choices depend not on γ_i, but on the maternal expectation of γ_i, which is $\mu_{\gamma,i}$:

$$x_i^* = \frac{\alpha_i \mu_{\gamma,i}}{1+\alpha_i \mu_{\gamma,i}} \frac{y_i}{\pi_i} \qquad (7)$$

Although we can still use the model to estimate the impact of a policy that increased income or subsidized investments – simply because we do not need to separately identify α_i from $\mu_{\gamma,i}$ to answer this question – we cannot test the assumption that $\gamma_i \in \mathcal{H}_i$. What we cannot do is to say by how much investments would change if we implemented a policy that moved $\mu_{\gamma,i}$ closer to γ_i.

In the context of this model, the main contribution of our research is to develop and implement a methodology to elicit $\mu_{\gamma,i}$. Clearly, to the extent that investments are partly determined beliefs $\mu_{\gamma,i}$, these variables are interesting by themselves. More important, if we add $\mu_{\gamma,i}$ to the data D_i, we would be able to separately identify heterogeneity in preferences from heterogeneity in beliefs.

3. Eliciting Expectations about the technology of skill formation

As we show in the model described in Section 2, a mother's decisions regarding how much to invest in her child depend partly on the subjective expectations that she holds about the technology of skill formation, in particular, her subjective expectations about the parameter γ. In order to quantify this relationship, it is obviously necessary to collect data on these expectations. Since at least the 1940s, economists and other social scientists have studied the usefulness of expectations (or "anticipation") data for understanding and forecasting firm investments, individual consumption of durable goods, individual choices about total fertility, and age of retirement (e.g., Hart, 1940; Ryder and Westoff, 1971; Rippe and Wilkinson, 1974; De Menil and Bhalla, 1975; Griliches, 1980; Jacobs and Jones, 1980; Koenig, Nerlov, and Oudiz, 1981; Wolpin and Gonul, 1985).

More recently, data on expectations have been collected to test or relax the assumption that individuals have rational expectations. In a series of papers, Dominitz and Manski (1996, 1997a,b) developed a useful framework for the measurement of subjective distributions and very successfully used it to measure subjective expectations about the returns to schooling, perception of job insecurity, and future income.[3] Kauffman and Pistaferri (2009) showed how data on subjective expectations on future earnings could be useful to separately identify the roles of information and insurance in models that try to quantify households' ability to smooth consumption across time and states of nature.

There is also a large literature on subjective expectations and their role in human capital investment decisions made by individuals. Stinebrickner and Stinebrickner (2007) collect longitudinal data on expectations and college performance and show that the decision to drop out of college is primarily driven by the student learning about his or her own academic ability. Zafar (2008) also constructs panel data on college students and establishes that the gap across genders in the choice of a college major is not explained by differences in expectations about academic ability but rather by differences in preferences.

Standard economic theory suggests that educational choices are partially determined by expectations about future earnings (e.g., Attanasio and Kauffman, 2009). As discussed above, Jensen (2010) shows that experimental information about returns to schooling affects schooling decisions made by individuals. Wiswall and Zafar (2012) build on this literature by showing evidence that subjective expectations react to new information. When combined, these findings suggest that human capital investment choices are partially determined by subjective beliefs about returns that can be affected by information that is made available to the decision maker. Interestingly, there is evidence that expectations and resources interact: Kauffman (2012) explores the heterogeneity in returns between individuals from poor and rich backgrounds in Mexico, and she finds that the former require higher expected returns than the latter in order to be induced to attend college. Interestingly, poor individuals with high expected returns are particularly responsive to changes in direct costs, which is consistent with their being credit constrained.

Finally, the literature in economics has used expectations data to understand decisions to participate in a crime, contraceptive choices, and retirement behavior. For example,

[3] Dominitz, Manski, and Heinz (2003) investigate expectations about future benefits from Social Security and Bruine de Bruin et al. (2011) build on the framework to measure consumer uncertainty about future inflation.

Lochner (2007) shows that individual heterogeneity in beliefs about the criminal justice system leads to differences in criminal participation. Delavande (2008) studies how beliefs about contraception methods help explain choices made by sexually active women. She finds that the contraceptive method chosen by a woman depends partially on her perception about that method's effectiveness. Van der Klaauw and Wolpin (2008) use the HRS data on survival expectations and retirement expectations to help estimate a dynamic stochastic model of retirement behavior under the assumption of rational expectations.

A remarkable feature of the data collected in these studies is the fact that there is a tight connection between information about expectations that is elicited and the economic model that is formulated to study the topic of interest. In this study, we would have liked to collect data on beliefs about γ directly, but this is impossible because mothers do not think about child development in the abstract framework presented in Section 2. As a proxy, we argue, however, that women do have implicit knowledge about the impacts of their actions, and the research design we describe below allows the analyst to translate this implicit knowledge into beliefs regarding the parameter γ.

3.1. Survey instrument

In order to elicit maternal subjective expectations of child development, we adapt the MSD instrument used in the CNLSY/79. 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 during the first four months of life (i.e., 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.

One major advantage of using the same items is 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 (2) that we employ in Section 5. As we now explain, although the questions are similar, they differ in two important details. In the MSD instrument, a mother provides yes/no answers to questions about child development. For example, one of the items in the MSD Scale for children who are 24 months old is: "Does your child speak a partial sentence of three words or more?" If the child has already spoken a partial sentence of three words or more, the mother chooses yes; otherwise, she chooses no. The key property of the instrument is that the tasks are described in language easily understood by the mothers and that the tasks are recognizable based on the daily interactions of mothers and their children.

The first difference is that in our instrument, which is designed to measure subjective expectations about the technology of skill formation, the mother is asked: "What do you think is the youngest age and the oldest age at which a child learns to speak a partial sentence of three words or more?" The respondent uses a sliding scale to indicate the age range in which she believes a child will develop these skills (Appendix Figure A1)[4].

There is another important difference. Because we are interested in measuring the expectation with respect to the parameter γ_i in the technology of skill formation, it is necessary for the respondents to provide answers to the above age-range question for different levels of investments. Thus, the second step is to create hypothetical scenarios of parental investments. We need at least two scenarios: one in which investment is "high" and another in which investment is "low." In fact, our survey instrument describes to the expectant mother four different scenarios of investments and the baby's health at birth. In the first scenario, the baby's health at birth is "good" $(\overline{q}_0)$ and the mother chooses a "high" level of investment $(\overline{x})$. In the second scenario, the mother also chooses a "high" level of investment $(\overline{x})$, but the baby's health at birth is "poor" $(\underline{q}_0)$. In the third scenario, the baby's health at birth is "good" $(\overline{q}_0)$, but the mother chooses a "low" level of investment $(\underline{x})$. Finally, in the fourth scenario, the baby's health at birth is "poor" $(\underline{q}_0)$ and the maternal choice of investment is low $(\underline{x})$.

It is important to emphasize that the levels of the two inputs in the technology of skill formation are invariant across groups of subjects in the survey. As we make clearer below, the variability in the beliefs about γ_i arises because of the heterogeneity in the age ranges provided by survey respondents.

Before answering the survey questions, the respondents watch a five-minute video that explains in detail the differences between the baby's "good" and "poor" health at birth. For example, in one version of the instrument, we designate "good" health as the one in which the baby is "normal" at birth, while the "poor" level of initial human capital

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

corresponds to a baby that is “small” at birth. As we explain to the mother, a “normal” baby is the one whose gestation lasts 9 months, weighs 8 pounds at birth, and is 20 inches long at birth. Conversely, the “small” baby is a baby that is born after 7 months of gestation, weighs only 5 pounds at birth, and is only 18 inches long at birth. The “normal” and “small” babies occupy extremely different positions in the distribution of health at birth: the “normal” baby is around the 60th percentile in the distribution, while the “small” baby is around the 1st percentile.

The video also shows examples of activities that mothers do with the child. With the exception of breastfeeding, all of the activities are part of the Home Observation for the Measurement of Environment – Short Form (HOME-SF) instrument[5]: (a) soothing the baby when he/she is upset; (b) moving the baby’s arms and legs around playfully; (c) talking to the baby; (d) playing peek-a-boo with the baby; (e) singing songs with the baby; (f) telling stories to the baby; (g) reading books to the baby; and (h) taking the baby outside to play in the yard, park, or playground. 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, mothers spend more time doing these activities than in the “low” level. For example, in one version of the survey instrument, we say that 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 15th and 85th percentile of investments.

Our sample is divided into different groups of subjects in terms of the definition of what constitutes “high” and “low” investments or “good” and “poor” health at birth. As a result, we can and do investigate the sensitivity of answers with respect to variations in the definition of scenarios.

3.2. Estimating expectations

We now discuss how to transform the answer to the question asked in our instrument – “What do you think is the youngest age and the oldest age at which a child learns to do [an MSD task]?” – into a measurement of the subjective expectation of child development at age 24 months, a quantity we denote by $E\left(\ln q_{i,1} \middle| \theta_i, q_0, x\right)$. This expectation is conditional on the level of initial human capital and investment given to the respondent through the scenarios described above.

[5] Bradley and Caldwell (1980, 1984).

In order to go from the age range to $E\left(\ln q_{i,1} \mid \theta_i, q_0, x\right)$, we break the problem into three steps. In the first step, we transform the age range into the probability that a child will learn a given MSD task by age 24 months. In the second step, we transform this probability into an estimate of the child's skill.[6] This estimate contains information about $E\left(\ln q_{i,1} \mid \theta_i, q_0, x\right)$ but is potentially contaminated with measurement error. In the third step, we show how to address the measurement error in a flexible way.

3.2.1. Transforming age range into probability

Without loss of generality, consider the scenario in which both initial conditions and investment are "high." For this scenario, suppose that the survey respondent states that the youngest and oldest age at which a child will learn how to speak partial sentences of three words or more is $\underline{a}$ and $\overline{a}$ 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 $\underline{a}$ is a number Δ_0 (arbitrarily) close to zero and the probability after age $\overline{a}$ 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. Suppose, for example, that the relationship between probability and age is logistic. That is, let $p^S_{i,j,k}(a)$ denote the maternal subjective expectation that the child i will be able to do MSD item j (e.g., "speak a partial sentence of three words or more") under hypothetical scenario k by age a months. Under the logistic assumption, this probability is linked to the child's age according to the following parametric specification:

$$\ln \frac{p^S_{i,j,k}(a)}{1-p^S_{i,j,k}(a)} = r_{i,j,k,0} + r_{i,j,k,1} a \qquad (8)$$

Given Δ_0 and Δ_1, the parameters $r_{i,j,k,0}$ and $r_{i,j,k,1}$ are just identified from the data provided by the survey respondent. In fact, it is possible to show that:

$$\hat{r}_{i,j,k,0} = \frac{\overline{a}\ln\left(\frac{\Delta_0}{1-\Delta_0}\right) - \underline{a}\ln\left(\frac{\Delta_1}{1-\Delta_1}\right)}{(\overline{a}-\underline{a})}, \quad \hat{r}_{i,j,k,1} = \frac{\ln\left(\frac{\Delta_1}{1-\Delta_1}\right) - \ln\left(\frac{\Delta_0}{1-\Delta_0}\right)}{(\overline{a}-\underline{a})}.$$

Given the knowledge of the parameters $\hat{r}_{i,j,k,0}$ and $\hat{r}_{i,j,k,1}$, we can invert the logistic function (8) to predict the probability at age 24 months:

[6] This second step is not strictly necessary. The reason we do it is that when we objectively estimate the technology of skill formation, we do so by transforming the MSD scores (which is a sum of fifteen yes/no answers, yes coded as one and no coded as zero) into a metric of "mental" development that is measured in months.

$$p_{i,j,k}^{s}(24) = \frac{e^{\hat{r}_{i,j,k,0}+\hat{r}_{i,j,k,1}\times 24}}{1+e^{\hat{r}_{i,j,k,0}+\hat{r}_{i,j,k,1}\times 24}}.$$

For concreteness, Figure 1 illustrates this algorithm for two different scenarios of investments. In both scenarios, the baby's health at birth is "good." When investment is "high," suppose that a respondent states that the lowest and highest ages are 18 and 28 months, respectively. If we choose $\Delta_0 = 0.005$ and $\Delta_1 = 0.995$, then the interpolation under the logistic assumption implies that the probability at age 24 months is around 0.75 (Figure 1, solid curve). For comparison, when investment is "low," suppose that the same respondent reports that the lowest and highest ages are 20 and 30 months, respectively. Using the same values for Δ_0 and Δ_1, the higher age range implies a lower probability of learning how to "speak a partial sentence of three words or more" at age 24 months, of around 0.25 (Figure 1, dashed curve).

3.2.2. Transforming probability into a measure of expected development

To see how we can derive an error-ridden measure of maternal expectation of development at age 24 months, $q_{i,j,k}^{s}$, from the probability obtained in the previous step, $p_{i,j,k}^{s}(24)$, we explore the information from the National Health and Nutrition Examination Survey (NHANES) data set.[7] An important feature of the MSD instrument is that it asks an item about children who are at very different ages. 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. This large variation in age allows us to estimate the fraction of children who can perform the same task at each age a, a quantity that we denote by $\pi_{j,a}$. We can then estimate how this probability evolves with age by adopting the following "flexible" logistic specification:

$$\ln \frac{\pi_{j,a}}{1-\pi_{j,a}} = g_j(a) + \psi_{j,a}$$

where $g_j(a)$ is monotonically increasing in a and $\psi_{j,a}$ is an error term that is orthogonal to age a. For illustration purposes, Figure 2 (right panel) shows the data and the resulting logistic prediction using the logistic specification for the MSD item "speak a partial sentence of three words or more." Clearly, the function $g_j(a)$ provides a very good fit of the data.

[7] In principle, we could implement the procedure we describe in the following paragraphs in the CNLSY/79 data set. However, as we discuss in Appendix B, the sample for which the MSD score is observed in the CNLSY/79 data is not representative of the children born to the NLSY/79 respondents, while the NHANES data set is.

The interpretation of the function $g_j(a)$ is straightforward: If we had 100 children who are a months old, we would expect $100\frac{e^{g_j(a)}}{1+e^{g_j(a)}}$ of them to be able to "speak a partial sentence of three words or more." Conversely, consider a group of 100 children, all of whom have the same unknown age. Suppose that a fraction p of children in this group could "speak a partial sentence of three words or more." Would it be possible to estimate the age of this group of children from the information above? The answer is yes! Given the monotonicity of the function $g_j(a)$, we can invert it to obtain an estimate of the age of the children in the group. The estimator would be

$$\hat{a} = g_j^{-1}\left[\ln\left(\frac{p}{1-p}\right)\right]. \quad (9)$$

It turns out that when we use the probability $p_{i,j,k}^{s}(24)$ derived in subsection 3.2.1 in the right-hand side of (9) above, we obtain in the left-hand side of (9) the error-ridden measure of maternal expectations of child development at age 24 months, $q_{i,j,k}^{s}$. That is, $q_{i,j,k}^{s} = g_j^{-1}\left[\ln\left(\frac{p_{i,j,k}^{s}(24)}{1-p_{i,j,k}^{s}(24)}\right)\right]$. Conveniently, $q_{i,j,k}^{s}$ is measured in age in months, which is the same metric we will use in Section 5 to measure skills when objectively estimating the parameters of the technology of skill formation (2).

Importantly, the higher the subjective probability that the mother reports for a given item j and scenario k, the higher the corresponding quality $q_{i,j,k}^{s}$. Figure 2 illustrates the mechanics of the argument. Again, consider the hypothetical survey respondent in subsection 3.2.1. As discussed above, her answers imply probabilities around 0.75 and 0.25 for the "high" and "low" investment scenarios, respectively. As shown in Figure 2, 25% of the children who are about 16 months old and 75% of the children who are about 22 months old have already learned "how to speak a partial sentence of three words or more." Thus, when investment is "high," the mother expects the 24-month-old child to have the skills of the typical 22-month-old child; when investment is "low," she expects the 24-month-old child to attain the development level of a typical 16-month-old child.

3.2.3. Accounting for measurement error

In order to elicit maternal expectations with respect to γ_i, it is sufficient to ask the mother about the age ranges for just one MSD item (say, "partial sentence of three words or more"). In what follows, let $k = 1$ denote the situation in which baby's health at birth is "good" and investment is "high" and let $k = 3$ denote the scenario in which baby's health at birth is also "good," but investment is "low." In this case, we would estimate the maternal subjective expectation of γ_i, $\mu_{\gamma,i}(\overline{q}_0)$, using the following relationship $\mu_{\gamma,i}(\overline{q}_0) = \frac{\ln q^S_{i,j,1} - \ln q^S_{i,j,3}}{\ln \overline{x} - \ln \underline{x}}$. For example, if we use the age ranges provided by our hypothetical survey respondent in the equation above, we would conclude that $\mu_{\gamma,i}(\overline{q}_0) = \frac{\ln 22 - \ln 16}{\ln 3 - \ln 1} \approx 29\%$.

There are at least three reasons to include more than one MSD item in the instrument for elicitation of expectations. First, we can investigate how the respondents' answers vary across MSD items for a fixed scenario. For example, the top right panel in Figure 3 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 3 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 former is a more difficult item than the latter.

If the respondents understand the survey instrument, we would expect them to assign a lower probability to items that are more difficult. This is the case depicted in the top left panel of Figure 3: Fixing the scenario in which the baby’s health at birth is “good” and investments is “high,” this hypothetical respondent reports age ranges 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). As we document below, this is qualitatively what happens with the participants in our survey.

It is also possible that respondents report similar age ranges for the same scenario across different items. Such a possibility is depicted at the bottom half of Figure 3. 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. Fortunately, this is not what happens in our data.

As anticipated from the discussion above, if the instrument contains multiple MSD items, we can also investigate the importance of measurement error in subjects’ responses. For example, suppose that for each respondent i and scenario k we have J different measurements of implicit subjective expectations of child development. If we knew that all of these J measurements are equally informative, we could just average over the J

answers for the same scenario k to minimize the role of measurement error in the analysis. In this case, the maternal expectations about child development for scenario k would be defined by:

$$E\left(\ln q_{i,1} \,\middle|\theta_i, q_0, x\right) = \frac{1}{J}\sum_{j=1}^{J} \ln q_{i,j,k}^{s}. \qquad (10)$$

However, it is reasonable to expect that measurement error is better captured by a richer specification. For example, the respondents may have a harder time answering some items and an easier time answering others. This suggests that not all items are equally informative. For this reason, we also investigate the following specification for the measurement-error model:

$$\ln q_{i,j,k}^{s} = \chi_{0,j} + \chi_{1,j} E\left(\ln q_{i,1} \,\middle|\theta_i, q_0, x\right) + \xi_{i,j} + \epsilon_{i,j,k} \qquad (11)$$

where $\xi_{i,j}$ is the measurement error associated with item j and $\epsilon_{i,j,k}$ is the measurement error associated with item j and scenario k. The measurement system (14) can be estimated via maximum likelihood and the implicit subjective expectations about child development $E\left(\ln q_{i,1} \,\middle|\theta_i, q_0, x\right)$ can be predicted via the Bartlett method. This approach gives higher weights to items that have higher loading $\chi_{1,j}$ or lower variance of measurement error $\epsilon_{i,j,k}$.

Finally, knowledge of $E\left(\ln q_{i,1} \,\middle|\theta_i, q_0, x\right)$ allows us to estimate the expected beliefs about the technology of skill formation. Assume that the information provided by the four hypothetical scenarios we constructed does not affect the respondents' expectations about future shocks that will be realized after the child is born, so that $E(\nu_i|\theta_i, q_0, x) = 0$. Under this assumption, note that the following relationship holds for the first and third scenarios:

$$E\left(\ln q_{i,1} \,\middle|\theta_i, \overline{q}_0, \overline{x}\right) = \ln A + \rho \ln \overline{q}_0 + \mu_{\gamma,i} \ln \overline{x} + \theta_i$$

$$E\left(\ln q_{i,1} \,\middle|\theta_i, \overline{q}_0, \underline{x}\right) = \ln A + \rho \ln \overline{q}_0 + \mu_{\gamma,i} \ln \underline{x} + \theta_i$$

Clearly, we can estimate $\mu_{\gamma,i}\left(\overline{q}_0\right)$ for every mother i from:

$$\mu_{\gamma,i}\left(\overline{q}_0\right) = \frac{E\left(\ln q_{i,1}|\theta_i, \overline{q}_0, \overline{x}\right) - E\left(\ln q_{i,1}|\theta_i, \overline{q}_0, \underline{x}\right)}{\ln \overline{x} - \ln \underline{x}} \qquad (12)$$

Expression (12) states that $\mu_{\gamma,i}\left(\overline{q}_0\right)$ is the subjective expectation about the elasticity of child development with respect to parental investment. Note that, so far, we have used only two of the four scenarios that we have. It is also possible to construct a second estimate of beliefs from the second and fourth scenarios:

$$\mu_{\gamma,i}\left(\underline{q}_0\right) = \frac{E\left(\ln q_{i,1}\middle|\theta_i,\underline{q}_0,\overline{x}\right)-E\left(\ln q_{i,1}\middle|\theta_i,\underline{q}_0,\underline{x}\right)}{\ln\overline{x}-\ln\underline{x}} \tag{13}$$

If parents believe that the technology of skill formation is Cobb-Douglas, then we should not reject the null hypothesis that $\mu_{\gamma,i}(\overline{q}_0) = \mu_{\gamma,i}\left(\underline{q}_0\right)$.

4. Results

In this section, we describe the empirical results from our analysis of the Maternal Knowledge of Infant Development Survey (MKIDS). To focus on the important results, we have placed a detailed explanation about the study procedures as well as important features of the data in Appendix A.

The analysis in this paper focuses on the 335 black participants (Appendix Table A1). The study participants are young: Around 80% of the sample is at most 25 years old. With respect to education, around 20% of the respondents are high-school dropouts or have received a GED. The fraction of participants who have a high-school diploma is 38%, which is just slightly higher than the fraction with some college experience: 31%. Few participants have a two-year or four-year college degree.

The sample is economically disadvantaged. The median income is $1500/month, which puts the sample under the 20th percentile in the US distribution of household income.[8] [9] Another indication of the prevalence of poverty in the group is that close to 90% of the respondents are on Medicaid. This contrasts with a figure of around 16% for the overall US population (US Census Bureau, 2011). Finally, the vast majority of the respondents are single.

4.1 Subjective expectations about the technology of skill formation

Before we report our findings about maternal expectations, we briefly describe basic features of the data. Typically, the youngest and oldest ages provided by respondents vary in predictable ways (Appendix Table A2). Holding constant the health at birth, the youngest and oldest age ranges take on higher values when investment is "low".

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

[9] Unfortunately, we did not collect information on the number of adult earners in the household.

Conversely, holding constant investment, age ranges are higher when health at birth is "poor". More important, age ranges are also higher for more MSD items that are more difficult (as explained in Section 3.2.3). As shown in Appendix Table A3, the probabilities at age 24 months, $p_{i,j,k}^{s}(24)$, also exhibit the same qualitative features.

Given probabilities $p_{i,j,k}^{s}(24)$ we can estimate our error-ridden measures of the natural log child development at age 24 months for a scenario of investment and health at birth, $\ln q_{i,j,k}^{s}$. The estimation of the statistical model (11) suggests that measurement error is, indeed, important in our context. As shown in Appendix Table A5, measurement error accounts for 40% to 70% of the total variance in responses. In any case, estimation of the measurement error model allows us to predict the subjective expected natural log of human capital at age 24 months, $E\left(\ln q_{i,1} \left|\theta_i, q_0, x\right.\right)$ which we plot on Appendix Figure A2. Again, *ceteris paribus*, higher x or higher q_0 imply higher expectations about child development at age 24 months.

Table 1 displays the summary statistics of the subjective expectations with respect to γ. The top part of Table 1 presents the results when $E\left(\ln q_{i,1} \left|\theta_i, q_0, x\right.\right)$ is determined by the simple average as defined in equation (10). The bottom part of the panel reports our findings when $E\left(\ln q_{i,1} \left|\theta_i, q_0, x\right.\right)$ is estimated from the measurement-error model (11). The typical and median woman believes that by increasing time spent with children by 100%, increases the stock of skills at age 24 months by approximately 8.8% and 4.5%, respectively. If we focus on the results from the measurement-error model (11), the figures are very similar. The mean and median expectations are 7.4% and 3.9%, respectively.

We find evidence of large heterogeneity in expectations. For example, the 75th percentile expectations are between 21% (in the factor model) and reach 23% (in the simple average model). In contrast, the 25th percentile has negative expectations: -4.5% in the simple average model and -3.4% in the factor model. As we will see in Section 4.3, the negative expectations are sensitive to the logistic approximation.

An advantage of defining scenarios for health at birth is that we can investigate whether respondents believe that the technology of skill formation follows a Cobb-Douglas specification. Interestingly, we find that the expectations about returns are even lower for the scenarios in which q_0 is "poor". In the simple average model, when the baby's health at birth is "good," the mean and median respondent expectation is around 13% and 7%, respectively. In contrast, when the baby's health at birth is "poor", median and mean expectations are 5% and 3%, respectively. The findings are basically the same when we focus on the factor model in the bottom panel of Table 1. Interestingly, the respondents on the 25th percentile do not have negative expectations for the "good" health scenarios, but they do so for the "poor" health scenarios.

The mean and median expectations are lower if we focus only on the MSD items that are primarily capturing cognitive (or language) skills. In this case, median expectation is between 0% and 2%, while the mean is 5% and 8%. In contrast, when we use only the motor items, median expectations are slightly higher and lie somewhere between 3% and 5% and mean expectations are located between 7% and 10%.[10]

4.2 Method of interview and definition of scenarios

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, Tylleskar, 2005). In our context, one could be worried about respondents who hold very low expectations would report higher beliefs because they understand that this is a more socially desirable answer.[11] As reported in Appendix Table A6, we find that the women below the median in the ACASI sample tend to report even lower expectations than the same group in the CAPI sample. This is consistent with respondents giving socially desirable answers, but the differences are not statistically significant.

Next, we investigate how expectations vary with respect to different definitions of what constitutes "low" versus "high" levels of investments. For this part of the analysis, it is important to clarify that the data was collected via ACASI. As explained above, the main group of respondents was shown a video in which a mother spent either six or two hours per day interacting with the baby. In contrast, a second group of respondents was told that a mother spent either four or three hours per day interacting with the baby. Table 2 shows the results. If we focus on the median expectation, we conclude that the qualitative and quantitative findings are robust to different definitions of what

[10] It is also possible to consider combinations of items within scenarios. The results here depend on whether we account for measurement error. When we do, we find that the difference between cognitive and motor items is almost exclusively driven by the scenarios in which the baby's health at birth is "good."

[11] There is also evidence that ACASI may generate lower quality data because the respondent may have difficulty understanding by himself or herself the instructions on the survey instrument (Estes et al., 2010, Elst et al, 2009).

constitutes “high” and “low” investments. For example, median expectations overall items and scenarios is around 7% for the respondents that were told six versus two hours of interaction per day and 7.5% for the respondents who were informed four versus three hours of investments per day. A similar conclusion arises when we focus on the expectations when the baby’s health is “good” (roughly 14% for the “six versus two” and 12% for the “four versus three”). The main difference in the median expectations relates to the scenario when health is “poor”: It is negative for the “four versus three” group, but positive for the “six versus two” group. Again, the difference in median expectations between “good” and “poor” health at birth indicate that respondents do not believe that the technology is Cobb-Douglas. Their beliefs are that the production function exhibits more complementarity between q_0 and x than the one implied by the Cobb-Douglas specification.

Although the results relating to mean expectations are qualitatively robust, there are important quantitative discrepancies. A possible explanation for the discrepancy in the means in the way scenarios are presented may be due to a combination of small sample sizes and the large effect of outliers in the means (but not the medians).[12] Another explanation is that some respondents believe that the returns decay fast as the number of hours increase. These explanations are not mutually exclusive and they operate by increasing the expectations in the right tail of the distribution.

We next investigate what happens with differences in expectations when “poor” health at birth is defined closer to the midpoint of the distribution of observed health at birth. A third set of respondents were told that by “good” health we meant a baby whose gestation lasted 9 months, weighed 8 pounds at birth and was 20 inches long at birth. On the other hand, by “poor” health we meant a baby whose gestation lasted only 8.5 months, weighed 7 pounds at birth, and was 19 inches long at birth. The respondents were also told that the mother could spend four (“high” investment) or three (“low” investment) hours per day interacting with the baby.

An important result is that the median respondent believes that the technology is locally Cobb-Douglas. That is, the median expectation is not different across scenarios of “good” and “poor” health at birth. However, the mean expectations are distinct in the same way as before: mean expectations are higher for “good” health. The differences are in the

[12] Note that when “high” and “low” investments are four and three hours per day, the denominator in equations (12) and (13) is around 0.288. In contrast, when “high” and “low” investments are six and two hours per day, the same denominator is around 1.10, a number almost four times larger.

right tail of the distribution of expectations. However, we do not know if the discrepancies in the means would disappear if we had a larger sample size.

4.3. Robustness check: Logistic assumption and target age

In this subsection, we examine the robustness of our findings with respect to the interpolation that uses the logistic distribution. Consider the scenario in which q_0 is "good", and x is "high" and suppose that the survey respondent states that the youngest and oldest age at which a child will learn how to speak partial sentences of three words or more is $\underline{a}$ and $\overline{a}$ months, respectively. Then, we envelop the logistic distribution between $\underline{a}$ and $\overline{a}$ by defining the following two triangular distributions. The "upper" triangular distribution is the one in which the mode is set arbitrarily close to $\underline{a}$, while the "lower" triangular distribution is the one in which the mode is set arbitrarily close to $\overline{a}$. Figure 4 plots the logistic as well as the upper and lower triangular distributions for a hypothetical respondent who provides $\underline{a} = 18$ and $\overline{a} = 28$ months. Let $F_{UT}(a)$, $F_{LT}(a)$, and $F_L(a)$ denote, respectively, the upper triangular, lower triangular, and logistic distribution. Note that for any $a \in [\underline{a}, \overline{a}]$ we have $F_{LT}(a) \le F_L(a) \le F_{UT}(a)$.

Table 3 displays the results from our robustness analysis.[13] The triangular distributions tend to produce larger mean and median estimates. For example, the median expectation is 4.5% for the logistic and around 7% for the triangular distributions. More important, there is an important difference in the left tail distribution of beliefs: The 25th percentile is negative for the logistic (-3.4%), but positive for the lower and upper triangular distributions (3.3% and 2.1%, respectively). In contrast, note that there are negligible differences for the 75th percentile.[14]

Another assumption we have maintained so far relates to the target age, which in our exercise so far has been set at 24 months. A large literature in developmental psychology documents that parents tend to overestimate the age at which children develop skills (e.g., Epstein, 1979; Hess et al., 1980; Ninio, 1988; Mansbach and Greenbaum, 1999). In fact, our findings reported in Appendix Table A3 are consistent with the evidence from

[13] We also consider a uniform interpolation.

[14] The sensitivity of the results is distinct for the scenarios of health at birth. While they increase mean expectations by 50% when $q_0 = \overline{q}_0$, they produce much higher means when $q_0 = \underline{q}_0$ (an increase of over 100%). Although the median also increases more for $q_0 = \underline{q}_0$, the increase is not enough to equalize median expectations across scenarios of health at birth.

developmental psychology. If many women believe that the age ranges are above 24 months across all scenarios, then our methodology will generate mean beliefs that are zero. Clearly, this will produce downwardly biased expectations.

In order to investigate the importance of the target age, we re-analyze our data taking ages 28, 32, and 36 months as target ages. We also consider four different interpolation methods we discussed above: the logistic, the uniform, the upper, and lower triangular distributions. The results are also displayed in Table 3. Focusing on the logistic distribution, we find evidence that age 24 months provides expectations that are too low: if we increase the target age from 24 to 28 months, then median expectation almost doubles from 4.5% to 8.7%. If we set the target age at 32 months, the median expectation increases to 10.7%. However, there is evidence that further increases in the target age tend to reduce returns. The reason is similar to the problem at age 24 months: Many respondents think that children will develop skills before age 36 months across all scenarios.

It turns out that the right tail of the distribution is sensitive to the changes to the assumptions about the target age. The 75th percentile increases from 22% to around 33% in the lower triangular distribution and from 24% to about 40% in the upper triangular distribution. Interestingly, although the left tail is sensitive to the interpolation method, it is not so with respect to the target age.

Instead of asking respondents to provide age ranges, it is possible to elicit probabilities directly from respondents. For example, one could ask "How likely is it that a baby will be able to speak a partial sentence of 3 words or more by age 2 years?". In this case, the respondent provides a number from 0 to 100 for each scenario of investment and the baby's health at birth. Interestingly, the answers from the direct elicitation of probabilities generate returns that are more similar to the ones produced by the upper and lower triangular distributions at ages 28 and 32 months than the logistic distribution at any age. However, the method of direct elicitation of probabilities is not without problems. As shown in Appendix D, the probabilities across MSD items for a given scenario are roughly independent from item difficulty. As discussed in Section 3.2.3, this answer pattern tends to produce substantial measurement error and may be an indication that respondents are not consistent in their evaluations across items.

Considering all target ages and interpolation methods, the median expectation across all scenarios and items is between 4.5% (logistic distribution at 24 months) and 19.1% (upper triangular distribution at 32 months). Given the wide variability in the median expectation, and the fact that the results vary by interpolation method and target age, it is worthwhile to explore whether there are other differences that are produced by these different approaches.

It is possible that one of the above methods produces expectations that are more closely correlated with actual investments. Unfortunately, the MKIDS study did not collect any investment data.[15] Instead, we return to the measurement-error model (11) and evaluate the performance of the different methods in terms of the total variance that can be attributed to $E(\ln q_{i,1} | \theta_i, q_0, x)$.

Table 4 displays these results. We consider two different ways of looking at the data. First, we simply look at the average share of the total variance that is attributed to the variance of $E(\ln q_{i,1} | \theta_i, q_0, x)$. For example, the "upper triangular distribution at 24 months" produces an estimate of $E(\ln q_{i,1} | \theta_i, q_0, x)$ that explains, on average, 50% of the total variance in the women's answers. In comparison, an estimate of $E(\ln q_{i,1} | \theta_i, q_0, x)$ generated by the logistic distribution at age 24 months – the method we have employed throughout the paper – explains approximately 39% of the total variance. This result implies that the responses, when analyzed under the logistic distribution at 24 months, are much less "reliable" than when we use the upper triangular distribution at 24 months. In fact, the models that use the logistic assumption tend to perform the worst when we consider this metric.

Another way to consider the model's performance is to examine the model's fit based on the log-likelihood. Although the models are not nested, they do estimate the same number of parameters. These comparisons are also presented in Table 4. Remarkably, the models that use the logistic interpolation also tend to have the worst performance in terms of model fitting. On the other hand, the upper triangular and uniform distributions around age 28 and 32 months seem to perform the best.

If we take the above analysis at its face value, we can draw two conclusions from the exercise. First, the logistic distribution produces estimates of expected developmental outcomes that explain the smallest share of variance. This is true regardless of the target age. Second, the uniform, the upper, and the lower triangular distributions at age 28 and 32 months are among the best performing methods. This result is important because

[15] It is common to measure investment by visiting the mother and child at home and conducting an assessment of the home environment. Our respondents were recruited and interviewed in the clinics where they received their prenatal care. This fact, combined with the fact that most of our sample consists of primiparous women, makes it impossible to measure investments in the human capital of children.

these alternatives tend to produce higher median expectations. For this reason, we conclude that the women's median expectations range from approximately 4.5% to 19.1% and that the preferred estimates are closer towards the end of this interval. Moreover, median expectations tend to be higher for the scenarios in which $q_0 = \overline{q}_0$, which indicates that the median respondent does not believe that the technology of skill formation is Cobb-Douglas. In what follows, we now turn to the estimation of the other components of the model presented in Section 2.

5. Estimation of the Technology of Skill Formation and Preferences

5.1 Preferences

Let $x(\mu_{\gamma,i})$ denote the parental investments when expectations are $\mu_{\gamma,i}$. An important question we would like to answer is: What is the elasticity of investments with respect to expectations $\mu_{\gamma,i}$? Given the model described in Section 2, it is easy to show that the elasticity is determined by the following equation:

$$elasticity = \frac{\frac{\partial x(\mu_{\gamma,i})}{\partial \mu_{\gamma,i}}}{\frac{x(\mu_{\gamma,i})}{\mu_{\gamma,i}}} = \frac{1}{1+\alpha_i \mu_{\gamma,i}} \quad (15)$$

Note that in order to calculate this elasticity, we only need to know the parameters α_i and $\mu_{\gamma,i}$. Clearly, if we observed x_i, $\mu_{\gamma,i}$, y_i, and π_i – which we do not – we could estimate α_i from equation (7).

In order to estimate α_i we need to follow a different route. Our approach is to elicit the preference parameter by stated-choice data. In our survey, we first told the respondent to assume that the baby's health at birth is "good." 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).

In order to link investment to time (i.e., the age of the child), we prepared a three-minute video in which we explain 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[16]. We illustrate 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."

For each combination of prices and income, we ask the respondents the following question (Appendix Figure A3):

> "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 $\$\pi$ 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 $h_{i,m,n}$ hours per day when the price is $\$\pi_m$ and income is $\$y_n$ then her monthly expenditure is $\$\pi_m h_{i,m,n}$ and the share of income allocated to investment is $s_{i,m,n} = \frac{\$\pi_m h_{i,m,n}}{\$y_n}$. Note that variability in the share $s_{i,m,n}$ across respondents i arises strictly because of variability in choices $h_{i,m,n}$ (all respondents face the same set of prices and incomes). 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%. The range of share values is between 4.9% and 11.75%. In comparison, Lino (2012) reports shares of investment around 7% for low-income parents. Given the estimated shares, we manipulate equation (7) to estimate α_i. Note that our estimate of parental valuation takes into account the heterogeneity in expectations about the elasticity parameter γ. Appendix Table A6 displays summary statistics of $\hat{\alpha}_i$ for the best interpolation and target

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

age combinations (see Table 4 and Section 4.7).[17] When we account for heterogeneity in beliefs, we find that the typical woman has α_i between 0.94 and 1.32, while the median woman's α_i is between 0.39 and 0.46. Clearly, the distribution is positively skewed and with a large standard deviation, indicating substantial heterogeneity in the women's valuation.

5.2. Objective estimation of the technology of skill formation

In this section, we rely on the CNLSY/79 data. Appendix B provides a description of the data set and summary statistics for the variables and the sample used in these analyses.

In order to objectively estimate the technology of skill formation, we assume that the dependent variable in (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, it is necessary to transform the raw score produced from the simple summation of maternal answers into a scale measured in time (i.e., age in months). Although most scales provide the equivalency table between scores and age in months, to the best of our knowledge this is not the case for the MSD scale. In order to derive this scale, we conduct an Item-Response Theory (IRT) analysis of the MSD data. The estimation of the IRT model allows us to classify the MSD items according to their level of difficulty (which is denoted by an item-specific intercept) as well as their informational content (which is captured by an item-specific factor loading). With this information, we can construct a scale of cognitive skills that is measured in age equivalent scores, which we refer to as "mental age" of development (see Appendix C).[18]

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 Schennach (2010), we factor analyze the items of the HOME-SF scale. In their analysis, the scale of the factor is set by the number of children's books in the household. Although this is a valid metric, this is 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

[17] To produce the estimates of $\hat{\alpha}_i$ in Table 14, we set to missing the 14 individuals who have non-positive values of $\mu_{\gamma,i}$. Clearly, the situation in which $\mu_{\gamma,i} = 0$ is problematic for the elasticity equation (15). Because the shares $\hat{s}_i$ are positive, negative values of $\mu_{\gamma,i}$ would imply negative values of $\hat{\alpha}_i$.

[18] IRT is also helpful with measurement error. Note that IRT is the equivalent for factor analysis when the measurements take on discrete values. Another advantage is that the IRT analysis allows us to estimate the cognitive scores of children whose mothers do not provide answers to all of the 15 MSD items.

instrument in a metric of time, i.e., the child's age. Details of the procedure are also described in Appendix C.

Finally, $q_{0,i}$ is measured by the child's health at 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.[19]

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:

$$\ln q_{1,i,l} = \ln A + \rho \ln q_{0,i,l} + \gamma_i \ln x_{i,l} + R_{i,l}\beta + \theta_i + \nu_{i,l} \tag{14}$$

where the index l denotes the birth order of the child and $R_{i,l}$ are 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). Note that in formulation (14), both γ_i and θ_i are constant across children l.

Table 5 shows the estimated γ_i of the technology (14).[20] In all of the regressions we show in Table 5, 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

---

[19] 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. There is reason to suspect that the differences in the quantity of investments do not completely capture differences in quality of investments. Kalil, Ryan, and Corey (2012) show that educated mothers not only spend more time with their children, but are also more likely to dedicate time to activities that best suit their children's developmental needs. It is possible that the differences are causal: Currie and Moretti (2002) and Carneiro, Meghir, and Parey (2013) explore exogenous variations in college-attendance costs to show that maternal schooling raises investments in children.

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

child's birth, a dummy variable for the child's gender, and dummy variables for the child's birth order.

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.[21] For this sample, the elasticity of skills with respect to investment (i.e., the parameter γ) 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 find γ 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.[22] The higher values of γ 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 closer around 24 months are the same families that have high values of γ.

Next, we investigate whether γ varies across mothers. We proceed by dividing the sample by maternal observable characteristics: maternal race/ethnicity, education, and skills. Interestingly, we find little evidence for difference in estimates of γ_i by maternal race or maternal education. We find some evidence that mothers with higher stocks of cognitive or socio-emotional skills tend to have higher values for γ, but this difference is not statistically significant. For the vast majority of the specifications we find that the estimates of γ_i are larger as we restrict the data to age ranges that are close to 24 months.

In the context of this paper, it is especially important to consider variability in γ because maternal expectations may be more closely associated with the mother's own γ_i than

---

[21] 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.

[22] 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.

with moments of the distribution of γ_i across the population, such as $E[\gamma_i]$. To tackle this question, we implement the procedure described by Arellano and Bonhomme (2012). [23] These authors derive estimators for moments of correlated random coefficients in panel data under the assumption of strict exogeneity of investments, an assumption we have maintained in our analysis.[24] A drawback of the approach is that it is necessary to observe at least 3 children for each mother. The sample, in this case, is rather small: there are only 303 mothers who satisfy this requirement. For this sample, $E[\gamma_i] \approx 19\%$.

The analysis reveals important heterogeneity in γ_i. For example, a mother who is half a standard deviation above the mean has $\gamma_i \approx 0.36$. In comparison, if we look at the overall sample for children who are between 19 and 29 months, the point estimate of γ is 26%. This heterogeneity in γ_i indicates that some mothers are very efficient in translating investments into cognitive skills of children.

5.3 Quantifying the impact of moving median expectations

Finally, we use all of this information to answer the following question: Take a respondent whose expectation is exactly at the median and consider the point estimates of the technology of skill formation that are located above the median expectation. What would happen to investments and child development if we were able to move her median expectations to the (higher) objective estimate we obtained from the CNLSY/79? Table 6 shows the estimated elasticity for the interpolation methods and target age ranges that are among the most reliable (and which also produce the highest estimates for median expectations). Clearly, the answer depends on the difference between $\mu_{\gamma,i}$ and γ. For the low value of $\gamma = 0.199$, the policy that equates $\mu_{\gamma,i}$ to γ produces little changes in investments (between 3.6% and 24.3%, depending on the interpolation and target age) and child development (between 0.7% and 5%). Obviously, for higher values of γ, policies that increase maternal beliefs have higher impacts on investment and child development. For example, if $\gamma = 0.257$, then the policy increases investments by at least 31.4% and child development by at least 7.4%. Importantly, these changes are estimated for the median expectation. The model implies even larger effects for the women whose expectations are below the median.

[23] See Appendix E for details on our implementation of the Arellano-Bonhomme (2012) estimator.
[24] As shown by Arellano and Honore (2001), in general the respondent-specific parameters are not point identified when strict exogeneity does not hold.

Conclusion

In this paper, we presented a simple model in which mothers have subjective expectations about the technology of skill formation. We show that the model can be used to evaluate the impact of policies that affect maternal knowledge about the importance of investments for developing the human capital of children. In order to be empirically useful, it is necessary to separately identify heterogeneity in expectations from heterogeneity in beliefs.

We propose to solve this problem by collecting data on subjective expectations about the technology of skill formation. We survey a sample of socio-economically disadvantaged, pregnant African-American women. By comparing the subjective expectations with the objective estimates of the technology of skill formation, we find evidence that our respondents may underestimate the elasticity of child development with respect to investments.

We also elicit data that allows us to estimate the parameters that describe parental preferences. We do so to evaluate the impact of a policy that would move expectations from the median value in our sample to the objective estimate based on the CNLSY/79. We find that investments would increase by at least 6.9% and that the children's stocks of cognitive skills at age 24 months would increase by at least 1.4%. The values are higher for mothers whose beliefs are below the median.

In future work, we will follow the respondents longitudinally and see if measures of expectations are correlated with parental investments once we account for other state variables that may be correlated with beliefs and investments, such as maternal skills, family income, and others. This will be an important step to validate the measures of beliefs and preferences we propose in this paper.

References

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

Almond, Douglas and Janet Currie (2010): "Human Capital Development Before Age Five", "Handbook of Labor Economics", Volume 4: Orley Ashenfelter and David Card, editors, Ch. 15: 1315-1486, Amsterdam: North Holland.

Arellano, Manuel and Stéphane Bonhomme (2012): "Identifying Distributional Characteristics in Random Coefficients Panel Data Models," *Review of Economic Studies*, Vol. 79(3): 987-1020.

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

Becker, Gary S. and Nigel Tomes (1986): "Child endowments and the quantity and quality of children," Journal of Political Economy, Vol. 84(4), pp. S143–S162.

Behrman, Jere R., Robert A. Pollak, Robert and Paul Taubman (1982): "Parental Preferences and Provision for Progeny, " Journal of Political Economy, University of Chicago Press, vol. 90(1): 52-73.

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.

Bruine de Bruin, Wandi, Charles F. Manski, Giorgio Topa, and Wilbert van der Klaauw (2011): "Measuring Consumer Uncertainty about Future Inflation," Journal of Applied Econometrics, Vol. 26(3): 454-478.

Campbell, Frances A., Barbara H. Wasik, Elizabeth Pungello, Margaret Burchinal, Oscar Barbarin, Kirsten Kainz, Joseph J. Sparling, Craig T. Ramey (2008): "Young adult outcomes of the Abecedarian and CARE early childhood educational interventions," *Early Childhood Research Quarterly*, Vol. 23(4): 452-466.

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 and James J. Heckman (2007): "The Technology of Skill Formation," *American Economic Review*, Vol. 97(2): 31-47.

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 (2008): "Pill, Patch or Shot? Subjective Expectations and Birth Control Choice," *International Economic Review*, Vol. 49(3): 999-1042.

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.

de Menil, George, and Surjit S. Bhalla (1975): "Direct Measurement of Popular Price Expectations," American Economic Review, Vol.65(1): 169-180.

Dominitz, Jeff and Charles F. Manski (1996): "Eliciting Student Expectations of the Returns to Schooling," *Journal of Human Resources*, Vol. 31(1), pp. 1-26.

Dominitz, Jeff and Charles F. Manski (1997a): "Using Expectations Data to Study Subjective Income Expectations," *Journal of the American Statistical Association*, Vol. 92(439): 855-867.

Dominitz, Jeff and Charles F. Manski (1997b): "Perceptions of Economic Insecurity: Evidence from the Survey of Economic Expectations," *Public Opinion Quarterly*, Vol. 61(2): 261-287.

Dominitz, Jeff, Charles F. Manski, and Jordan Heinz (2003): "Will Social Security Be There for You? How Americans Perceive Their Benefits," NBER Working Paper 9798.

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.

Estes, Larissa J., Linda E. Lloyd, Michelle Teti, Sheela Raja, Lisa Bowleg, Kristi L. Allgood, and Nancy Glick (2010): "Perceptions of Audio Computer-Assisted Self-Interviewing (ACASI) among Women in an HIV-Positive Prevention Program," PLoS ONE Vol. 5(2): e9149. doi:10.1371/journal.pone.0009149.

Field, Erica, Omar Robles, and Maximo Torero (2009): "Iodine Deficiency and Schooling Attainment in Tanzania", *American Economic Journal: Applied Economics*, Vol. 1(4): 140-169.

Griliches, Zvi (1980): "Expectations, Realizations, and the Aging of Young Men, " Research in Labor Economics, JAI Press, 1980.

Hart, Albert G. (1940): "Anticipations, Uncertainty, and Dynamic Planning," University of Chicago Press, Chicago, IL.

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., Seong H. Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Q. Yavitz (2010): "Analyzing Social Experiments as Implemented: A Reexamination of the Evidence From the High Scope Perry Preschool Program," *Quantitative Economics*, Vol. 1(1): 1-46.

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.

Hess, Robert D., Keiko Kashiwagi, Hiroshi Azuma, Gary Price, and W. P. Dickson (1980): "Maternal Expectations for Mastery of Developmental Tasks in Japan and the United States," *International Journal of Psychology*, Vol. 15(2), pp. 259-271.

Hoddinott, John, John A. Maluccio, Jere R Behrman, Rafael Flores and Reynaldo Martorell (2008): "The Impact of Nutrition During Early Childhood on Income, Hours Worked, and Wages of Guatemalan Adults" *Lancet* Vol. 371 (February): 411-416.

Hunt, Joseph M. (1961): "Intelligence and experience," New York, Ronald Press.

Jacobs, R. L., and R. A. Jones (1980): "Price Expectations in the United States: 1947-1975," American Economic Review, Vol.70(1): 269-276.

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

Kauffman, Katja M. (2012): "Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns," Bocconi University Working Paper.

Kaufmann, Katja and Luigi Pistaferri (2009):"Disentangling Insurance and Information in Intertemporal Consumption Choices," American Economic Review P&P, Vol. 99(2): 387-92.

Kalil, Ariel, Rebecca Ryan, and Michael Corey (2012): "Diverging Destinies: Maternal Education and the Developmental Gradient in Time with Children," *Demography*, Vol. 49(4): 1361-1383.

Karoly Lynn A., M. Rebecca Kilburn, and Jill S. Cannon (2005): "Early Childhood Interventions: Proven Results, Future Promise," RAND Corporation, Santa Monica, CA.

Koenig, H., M. Nerlov, and G. Oudiz (1981): "On the Formation of Price Expectations: An Analysis of Business Test Data by Log-linear Probability Models," European Economic Review, Vol. 16(1):103-138.

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.

Lochner, Lance (2007): "Individual Perception of the Criminal Justice System," *American Economic Review*, Vol. 97(1), pp. 444-460.

Mansbach, Ivonne K. and Charles W. Greenbaum (1999): "Developmental Maturity Expectations of Israeli Fathers and Mothers: Effects of Education, Ethnic Origin, and Religiosity," *International Journal of Behavioral Development*, Vol. 23(3): 771-797.

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.

Rippe, Richard D. and Maurice Wilkinson (1974): "Forecasting Accuracy of the McGraw-Hill Anticipatory Data", Journal of the American Statistical Association, Vol. 69(348):849-858.

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.

Roy, Shalini (2009): "Can Awareness Campaigns Promote Early Childhood Investment? Evidence from the Philippines," University of Pennsylvania Ph.D. Dissertation.

Ryder, Norman. B and Charles F. Westoff (1971): "Reproduction in the United States, 1965", Princeton University Press, Princeton, NJ.

Stinebrickner Todd R. and Stinebrickner Ralph (2007): "Learning About Academic Ability and the College Drop-Out Decision", University of Western Ontario Working Paper.

Van der Elst, Elisabeth M., Haile Selassie Okuku, Phellister Nakamya, Allan Muhaari, Alun Davies, R. Scott McClelland, Matthew A. Price, Adrian D. Smith, Susan M. Graham, Eduard J. Sanders (2009): "Is Audio Computer-Assisted Self-Interview (ACASI) Useful in Risk Behaviour Assessment of Female and Male Sex Workers, Mombasa, Kenya?", PLoS ONE Vol. 4(5): e5340. doi:10.1371/journal.pone.0005340.

Van der Klaauw, Wilbert and Kenneth I. Wolpin (2008): "Social Security and the Savings and Retirement Behavior of Low Income Households," *Journal of Econometrics*, Vol. 145(1-2): 21-42.

Vygotsky, Lev S. (1978): "Mind in Society," MIT Press, Cambride, MA.

Wallsten, Thomas, S., David V. Budescu, Amnon Rapoport, Rami Zwick, and Barbara Forsyth (1986): "Measuring the Vague Meanings of Probability Terms," *Journal of Experimental Psychology: General*, Vol. 115(4): 348-365.

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.

Wiswall, Matt and Basit Zafar (2012): "Determinants of College Major Choices: Identification from an Information Experiment", Arizona State University Working Paper.

Wolpin, Ken and Gonul, Fusun (1985): "On the Use of Expectations Data in Micro Surveys: The Case of Retirement, " Center for Human Resource Research at Ohio State University working paper.

Zafar, Basit (2008): "College Major Choice and the Gender Gap," Staff Report no. 364, Federal Reserve Bank of New York.

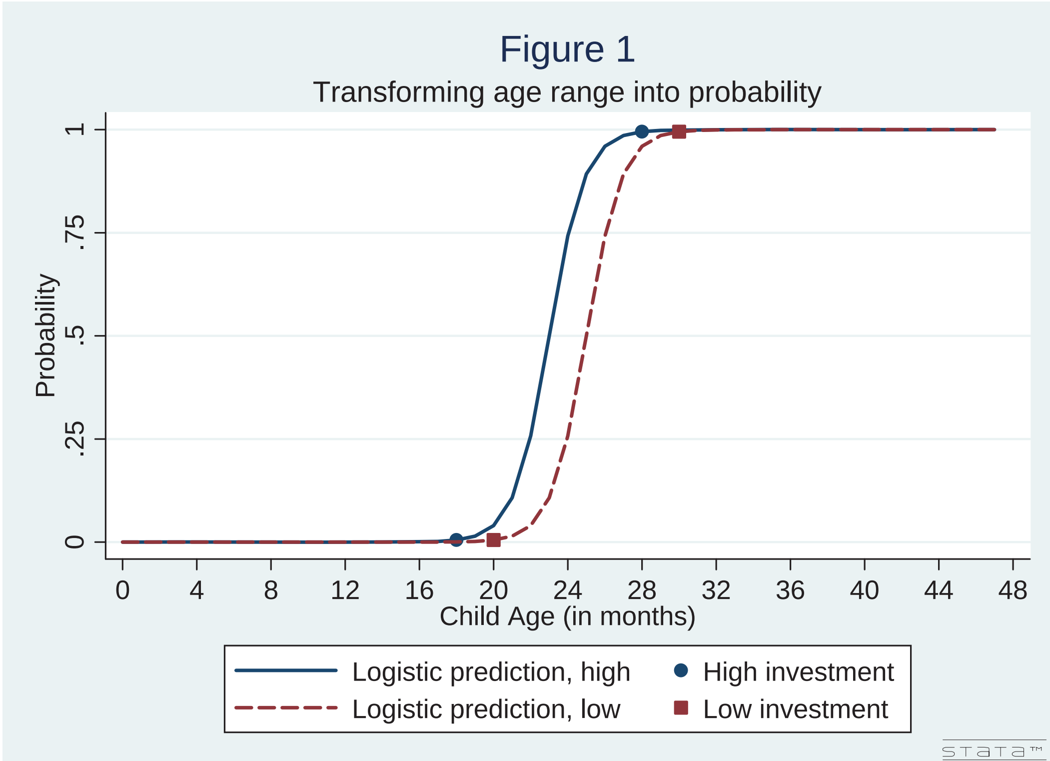


Figure 1
Transforming age range into probability
Probability
0
.25
.5
.75
1
0 4 8 12 16 20 24 28 32 36 40 44 48
Child Age (in months)
Logistic prediction, high
Logistic prediction, low
High investment
Low investment

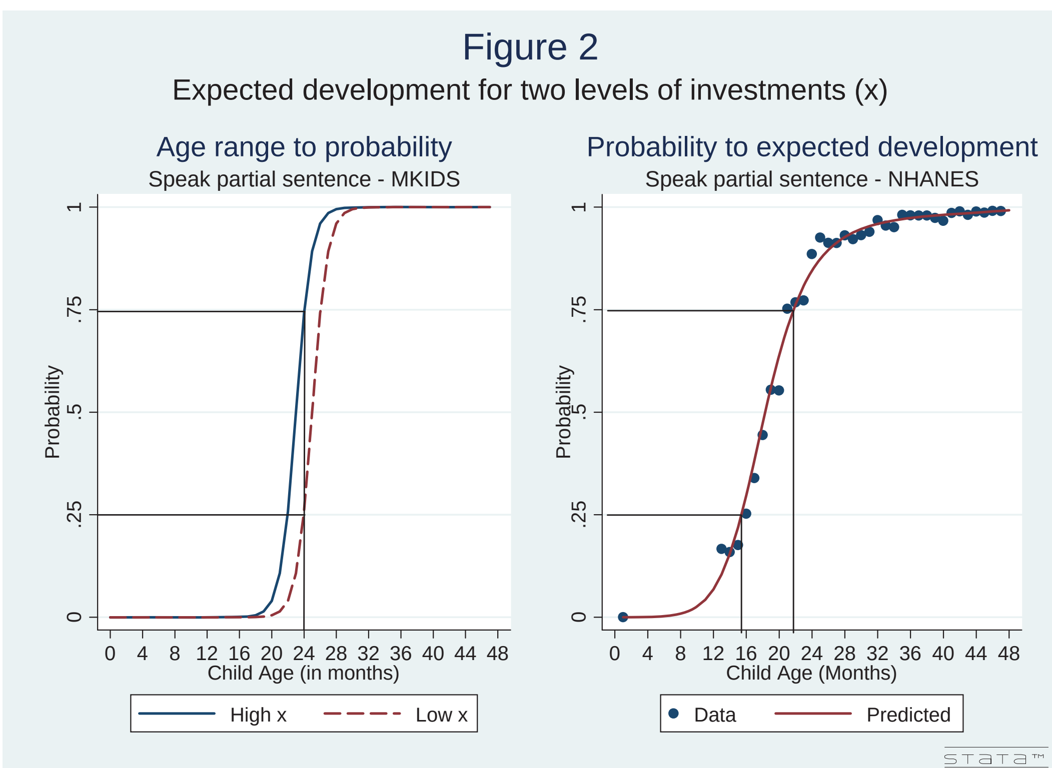


Figure 2
Expected development for two levels of investments (x)
Age range to probability
Speak partial sentence - MKIDS
Probability
0
.25
.5
.75
1
0 4 8 12 16 20 24 28 32 36 40 44 48
Child Age (in months)
High x
Low x
Probability to expected development
Speak partial sentence - NHANES
Probability
0 4 8 12 16 20 24 28 32 36 40 44 48
Child Age (Months)
Data
Predicted

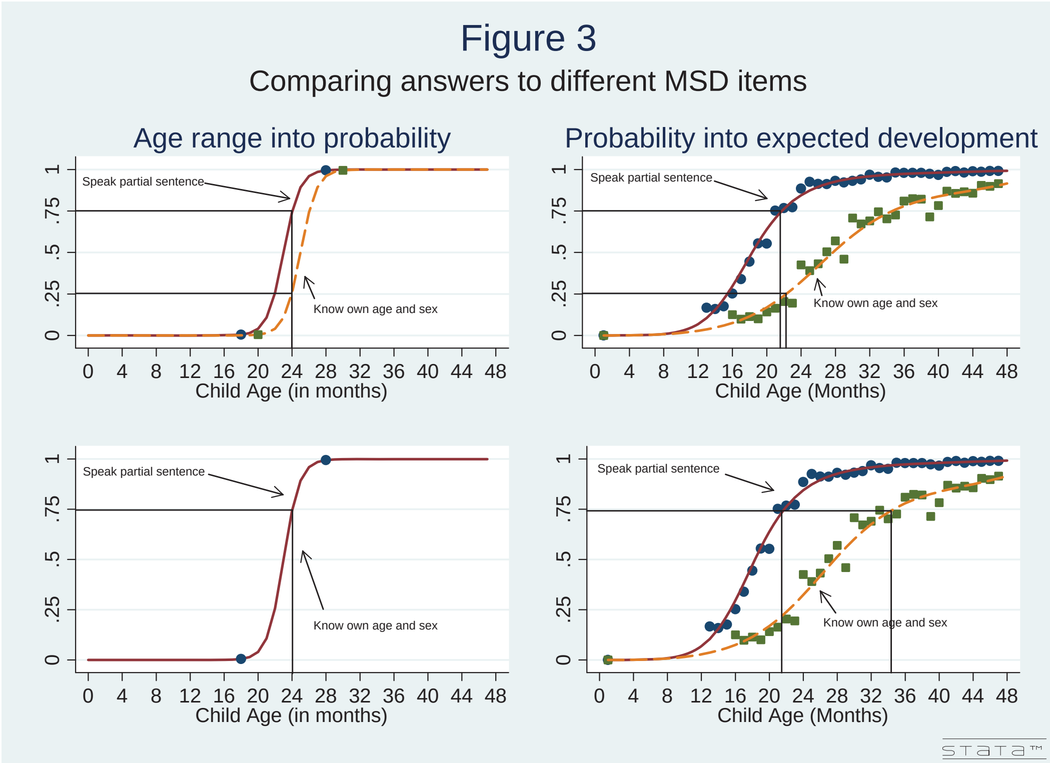


Figure 3
Comparing answers to different MSD items
Age range into probability
Probability into expected development
Speak partial sentence
Know own age and sex
Child Age (in months)
Child Age (Months)
0
.25
.5
.75
1
0 4 8 12 16 20 24 28 32 36 40 44 48
stata™

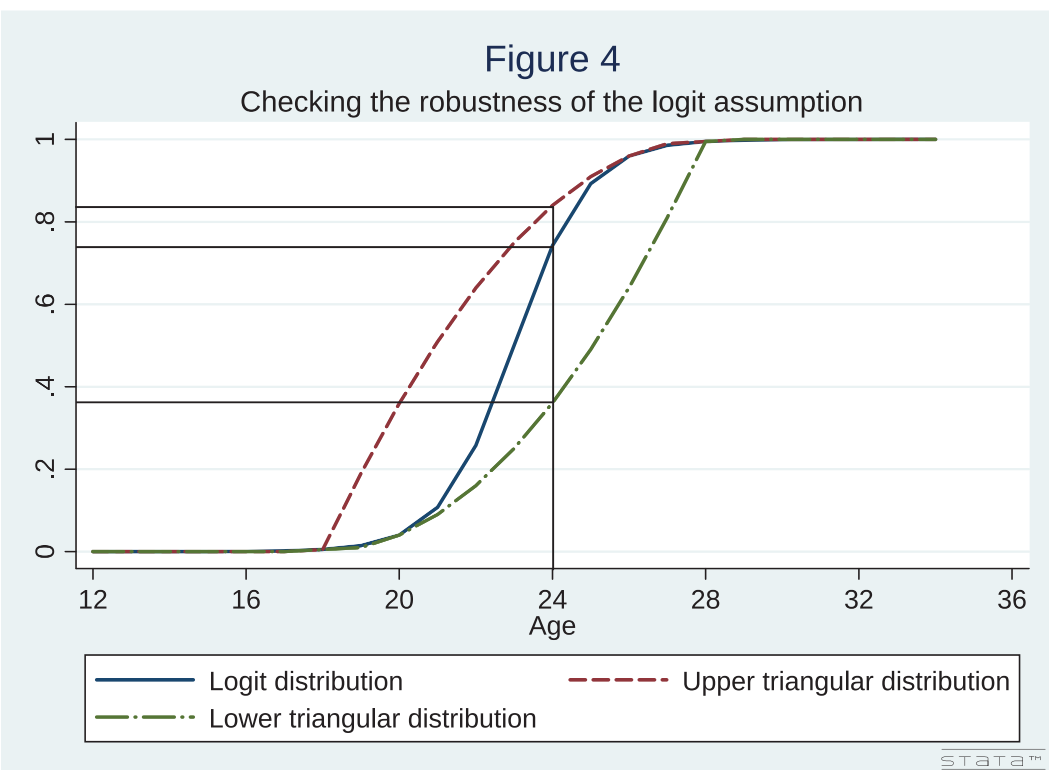


Figure 4
Checking the robustness of the logit assumption
0
.2
.4
.6
.8
1
12 16 20 24 28 32 36
Age
Logit distribution
Upper triangular distribution
Lower triangular distribution
stata™

Table 1

Subjective expectations about the technology of skill formation

Accounting for measurement error by averaging across items

	Mean	Median	25th Percentile	75% Percentile	Std Deviation
Overall items and Scenarios[1]	8.8%	4.5%	-4.5%	23.0%	32.9%
"Good" versus "poor" heath at birth					
Overall items, only scenarios with "good" health at birth	12.7%	7.2%	0.0%	29.1%	36.5%
Overall items, only scenarios with "poor" health at birth	4.9%	2.9%	-7.2%	21.1%	33.6%
Primarily motor vs. primarily cognitive items					
Cognitive items, all scenarios	7.6%	0.0%	-0.9%	20.8%	33.4%
Motor items, all scenarios	9.9%	4.6%	-8.0%	23.7%	38.7%
Primarily motor vs. primarily cognitive items by health at birth					
Cognitive items, "good" health at birth	10.3%	0.0%	0.0%	26.1%	35.6%
Cognitive items, "poor" health at birth	4.9%	0.0%	-1.8%	17.0%	34.5%
Motor items, "good" health at birth	14.8%	9.0%	0.0%	35.7%	43.8%
Motor items, "poor" health at birth	4.9%	0.0%	-6.1%	22.3%	41.2%

Accounting for measurement error by estimating factor model

	Mean	Median	25th Percentile	75% Percentile	Std Deviation
Overall items and Scenarios[1]	7.4%	3.9%	-3.4%	21.0%	32.9%
"Good" versus "poor" heath at birth					
Overall items, only scenarios with "good" health at birth	11.8%	6.7%	1.3%	29.4%	36.1%
Overall items, only scenarios with "poor" health at birth	3.0%	0.2%	-8.2%	15.8%	32.8%
Primarily motor vs. primarily cognitive items					
Cognitive items, all scenarios	5.1%	2.1%	-4.0%	16.3%	36.4%
Motor items, all scenarios	6.9%	3.1%	-5.0%	19.0%	35.4%
Primarily motor vs. primarily cognitive items by health at birth					
Cognitive items, "good" health at birth	10.2%	4.2%	-0.4%	24.7%	39.2%
Cognitive items, "poor" health at birth	0.0%	-4.4%	-7.4%	13.3%	36.9%
Motor items, "good" health at birth	16.6%	10.4%	3.4%	36.1%	39.9%
Motor items, "poor" health at birth	-2.8%	-4.1%	-13.9%	7.3%	36.6%

[1]Health at birth is "good" if (1) the baby's weight at birth is 8 pounds, the baby's length at birth is 20 inches, and the gestational age is 9 months. Health at birth is "poor" if the baby's weight at birth is 5 pounds, the baby's length at birth is 18 inches, and the gestational age is 7 months. When investment is "high" the mother spends 6 hours/day interacting with the baby. In contrast, when investment is "low" the mother spends only 2 hours/day interacting with the baby.

Table 2

Subjective beliefs about the technology of skill formation

How do maternal mean beliefs change with definitions of health conditions at birth and investments?

Accounting for measurement error by estimating factor model

Health conditions at birth	Healthy	Gestation: 9 months Birth weight: 7 pounds Birth length: 20 inches		Healthy	Gestation: 9 months Birth weight: 7 pounds Hospital time: 3 days		Normal	Gestation lasts 9 months Birth weight: 8 pounds Birth length: 20 inches	
	Not Healthy	Gestation: 7 months Birth weight: 5 pounds Birth length: 19 inches		Not Healthy	Gestation: 7 months Birth weight: 5 pounds Hospital time: 7 days		Small	Gestation: 8.5 months Birth weight: 6 pounds Birth length: 19 inches	
Investments	High	6 hours/day		High	4 hours/day		High	4 hours/day	
	Low	2 hours/day		Low	3 hours/day		Low	3 hours/day	
	Number of observations = 42			Number of observations = 71			Number of observations = 32		
	Mean	Median	Std Dev	Mean	Median	Std Dev	Mean	Median	Std Dev
Overall items and Scenarios	7.2%	6.9%	28.1%	18.7%	7.6%	48.4%	22.6%	8.1%	41.1%
Overall items, only scenarios with high health conditions at birth	8.3%	13.8%	37.9%	29.8%	11.9%	59.0%	28.9%	8.4%	49.5%
Overall items, only scenarios with low health conditions at birth	6.1%	1.1%	29.2%	7.7%	-3.2%	49.2%	16.3%	8.2%	42.4%

Table 3

Subjective beliefs about the technology of skill formation

Checking sensitivity of the logit assumption

Not accounting for measurement error

Overall items and Scenarios

	Mean	Median	25th Percentile	75th Percentile	Std Dev
			Logistic		
Target age is 24 months	8.8%	4.5%	-3.4%	21.0%	32.9%
Target age is 28 months	12.6%	8.7%	-3.0%	28.9%	32.2%
Target age is 32 months	14.6%	10.7%	-3.0%	30.1%	29.0%
Target age is 36 months	12.8%	7.6%	-2.1%	27.8%	24.8%
			Uniform		
	Mean	Median	25th Percentile	75th Percentile	Std Dev
Target age is 24 months	7.4%	3.9%	-3.4%	21.0%	32.9%
Target age is 28 months	22.8%	17.2%	4.9%	36.7%	23.0%
Target age is 32 months	21.7%	19.2%	4.2%	33.0%	21.6%
Target age is 36 months	21.0%	16.7%	5.7%	32.8%	19.3%
			Lower Triangular		
	Mean	Median	25th Percentile	75th Percentile	Std Dev
Target age is 24 months	15.8%	7.2%	3.3%	22.3%	22.9%
Target age is 28 months	20.8%	15.7%	4.3%	32.1%	22.7%
Target age is 32 months	21.7%	17.8%	4.2%	33.7%	22.1%
Target age is 36 months	21.7%	17.6%	3.9%	32.8%	22.3%
			Upper Triangular		
	Mean	Median	25th Percentile	75th Percentile	Std Dev
Target age is 24 months	15.2%	7.1%	2.1%	24.3%	23.5%
Target age is 28 months	24.1%	18.3%	4.8%	39.9%	25.3%
Target age is 32 months	22.9%	19.1%	4.1%	40.0%	23.7%
Target age is 36 months	21.1%	15.2%	4.4%	35.7%	21.2%
			Direct Elicitation of Probability		
	Mean	Median	25th Percentile	75th Percentile	Std Dev
Target age is 24 months	26.1%	17.5%	4.8%	40.9%	35.3%

Table 4

Comparison of different interpolation method with respect to measurement error

Interpolation Method	Mean Share of Signal to Total Variance Ratio	Median Expectation	Interpolation Method	Log likelihood	Median Expectation
Upper triangular at 24 months	50.0%	7.1%	Lower triangular at 24 months	-8362.76	7.2%
Lower triangular at 32 months	48.9%	17.8%	Lower triangular at 28 months	-8526.64	15.7%
Lower triangular at 28 months	48.7%	15.7%	Lower triangular at 32 months	-9303.08	17.8%
Upper triangular at 28 months	48.4%	18.3%	Uniform at 32 months	-9953.78	19.2%
Upper triangular at 32 months	48.2%	19.1%	Uniform at 28 months	-10480.21	17.2%
Uniform at 32 months	47.8%	19.2%	Uniform at 36 months	-10480.21	16.7%
Uniform at 28 months	47.6%	17.2%	Upper triangular at 32 months	-11734.71	19.1%
Lower triangular at 36 months	47.4%	17.6%	Upper triangular at 24 months	-11939.96	7.1%
Lower triangular at 24 months	46.3%	7.2%	Upper triangular at 28 months	-12742.74	18.3%
Uniform at 36 months	45.3%	16.7%	Lower triangular at 36 months	-13262.94	17.6%
Upper triangular at 36 months	45.3%	15.2%	Upper triangular at 36 months	-13307.42	15.2%
Logistic at 28 months	41.4%	8.7%	Logistic at 36 months	-14877.54	7.6%
Logistic at 32 months	41.4%	10.7%	Uniform at 24 months	-14897.94	3.9%
Uniform at 24 months	38.6%	3.9%	Logistic at 24 months	-14897.94	4.5%
Logistic at 24 months	38.6%	4.5%	Logistic at 32 months	-14935.06	10.7%
Logistic at 36 months	37.5%	7.6%	Logistic at 28 months	-15041.75	8.7%

Table 5

Objective estimation of the technology of skill formation

Estimates of γ for full sample and selected subsamples

Dependent variable: Natural log of skills around age 24 months[1]

	13 to 35 Months	16 to 32 Months	19 to 29 Months
Full Sample			
Full Sample	18.0%*** (1.99%)	19.9%*** (2.70%)	25.7%*** (3.84%)
Analysis by race/ethnicity	13 to 35 Months	16 to 32 Months	19 to 29 Months
Hispanic subsample only	18.3%*** (4.21%)	18.9%*** (5.45%)	34.4%*** (7.64%)
Black subsample only	18.7%*** (3.17%)	21.7%*** (4.42%)	20.1%*** (6.33%)
Non-Hispanic, non-black subsample only	16.9%*** (3.80%)	19.7%*** (5.45%)	31.1%*** (8.82%)
Analysis by maternal education at first birth			
Mother is high-school dropout at birth of the first child	18.1%*** (3.16%)	20.2%*** (4.64%)	20.4%*** (5.80%)
Mother is at least high-school graduate at birth of the first child	17.8%*** (2.71%)	18.6%*** (3.69%)	29.3%*** (5.30%)
Analysis by maternal cognitive skills			
Maternal AFQT is in bottom quartile	16.3%*** (3.19%)	18.9%*** (4.23%)	16.5%*** (4.99%)
Maternal AFQT is in 2nd quartile or higher	20.0%*** (2.84%)	20.9%*** (3.99%)	29.4%*** (5.79%)
Analysis by maternal score on Rotter's locus of control scale			
Maternal score on Rotter's locus of control is in top quartile[2]	19.6%*** (2.44%)	20.0%*** (3.28%)	23.5%*** (4.71%)
Maternal score on Rotter's locus of control is in 3rd quartile or lower[2]	13.9%*** (4.14%)	20.4%*** (5.67%)	17.6%** (6.94%)
Analysis by maternal score on Rosenberg's self esteem scale			
Maternal score on Rosenberg's self esteem scale is in bottom quartile[3]	13.9%*** (4.16%)	21.4%*** (6.29%)	18.2%* (9.94%)
Maternal score on Rosenberg's self esteem scale is in 2nd quartile or higher[3]	18.6%*** (2.37%)	19.1%*** (3.02%)	27.2%*** (4.33%)

Robust standard errors in parentheses. All regressions have dummy variables for: (i) the child's gender, (ii) birth order, (iii) age at the time of measurement of the dependent variable, (iv) year of birth 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]The Rotter locus of control scale measures the extent to which individuals believe that they can control events that affect them. In the NLSY/79, it takes on values between 4 and 16. Low values indicate that individuals tend to believe that they can control the events, while high values suggest that individuals believe that events are beyond their control.

[3]The Rosenberg self esteem scale measures an individuals self esteem. In the NLSY/79, it takes on values between 9 and 30. Low values indicate lack of self esteem.

Table 6
Moving median expectations close to objective estimates

Interpolation method	Median expectation (μ_γ)	Target elasticity (γ)	Change in investments	Change in child development at age 24 months
Lower triangular at 32 months	17.8%	19.9%	10.6%	2.2%
Upper triangular at 32 months	19.1%	19.9%	3.6%	0.7%
Lower triangular at 28 months	15.7%	19.9%	24.3%	5.0%
Upper triangular at 28 months	18.3%	19.9%	7.9%	1.6%
Lower triangular at 32 months	17.8%	25.7%	40.4%	9.5%
Upper triangular at 32 months	19.1%	25.7%	31.4%	7.4%
Lower triangular at 28 months	15.7%	25.7%	58.1%	13.7%
Upper triangular at 28 months	18.3%	25.7%	36.9%	8.7%
Lower triangular at 32 months	17.8%	28.3%	53.8%	19.5%
Upper triangular at 32 months	19.1%	28.3%	43.8%	15.9%
Lower triangular at 28 months	15.7%	28.3%	73.3%	26.6%
Upper triangular at 28 months	18.3%	28.3%	50.0%	18.1%

A The MKIDS Data

In this Appendix, we describe the MKIDS data. To collect the data we describe next, we clarify that we first obtained IRB approval from the University of Pennsylvania, Drexel College of Medicine, and the Children's Hospital of Philadelphia.

A.1 Procedures and sample used in this paper

The sample is recruited from four prenatal clinics affiliated with a university hospital in Philadelphia, PA. Eligibility criteria include women who were currently pregnant, at least 18 years of age, English-speaking, and had at most only one previous live birth. Appendix Table A1 shows other demographic characteristics of our sample.

The recruitment procedures consisted of the following: every week clinic staff released to the study coordinator a list of the date, time, and location of prenatal appointments of potentially eligible study participants. Once a potential participant registered at the clinic, the interviewer approached her to explain the study and screen for eligibility. If eligible, the participant was asked to provide written informed consent. Over1300 subjects were approached of which 539 were deemed eligible. Of these women, 535 agreed to participate. Subjects who completed the entire survey received $25 for their participation. The interview was conducted in a private office at the prenatal clinic while the respondents waited for their prenatal care visit.

A.2 Features of the data

Appendix Table A2 reports the average youngest and oldest age reported by the respondents for each scenario of investment and the baby's health at birth. In order to facilitate the presentation of our results, we sort the MSD items in increasing order of difficulty. Note that maternal answers are consistent in that they report higher ages for more difficult items.

Appendix Table A3 shows the summary statistics for the probabilities at age 24 months for each MSD item and scenario. As we explained in subsection 3.2.1 the probabilities $p^{s}_{i,j,k}(24)$ are derived from the age range information provided by the respondent under the assumption of logistic distribution. Column (1) displays the NHANES objective probability that 24-month-old children have already learned a certain MSD task. For example, 88.5% of 24-month-old children can "speak a partial sentence of 3 words or more." However, only 13.5% of 24-month-old children "can count out loud up to 10."

Column (2) shows the corresponding average probabilities for each MSD item when the initial stock of human capital and parental investment are set according to the values in Scenario 1. We make two observations: First, these probabilities tend to decrease with the difficulty of the task. This qualitative feature is consistent with the ordering in the NHANES. However, it is also clear that there are substantial discrepancies in a percentage of children who are expected to be able to complete the task by age 24 months: women's subjective probabilities in MKIDS under Scenario 1, which is the one in which children have "good" health at birth and high levels of investments, tend to be lower than the estimated probabilities in NHANES. Nevertheless, we find reassuring that the ordering provided by the MKIDS participants is consistent with the ordering we estimate from the NHANES dataset.

It is also interesting to compare column (2) with the corresponding columns for Scenarios 2, 3, and 4. As we explained above, the difference between Scenarios 1 and 3 and Scenarios 2 and 4 is in the values taken by x. In contrast, Scenarios 1 and 4 differ in terms of both q_0 and x. Interestingly, parents understand that the technology of skill formation is increasing in both inputs. For example, according to parents, the average probability that the child will be able to "speak a partial sentence of 3 words" by age 24 months is 31% for Scenario 1 (column 2), 22% for Scenario 2 (column 4), about 19% for Scenario 3 (column 3), and 16% for Scenario 4 (column 5).

A crucial step in the analysis is the one in which the probabilities $p^s_{i,j,k}(24)$ are transformed into measurements of expected human capital at age 24 months $q^s_{i,j,k}$. This is done by the inverse of the function $g_j(a)$ in equation (12). We rely on the NHANES data set to do so. Using the notation from Subsection 3.2.2, $\pi_{j,a}$ is the fraction of children who can perform MSD task j at age a. We specify the function $g_j(a)$ as a polynomial of degree five at age, a. That is, we estimate the following equation:

$$\ln \frac{\pi_{j,a}}{1-\pi_{j,a}} = \sum_{n=0}^{5} \kappa_n \, a^n + \psi_{j,a}$$

Appendix Table A4 shows the estimated coefficients κ_n, $n = 1, \dots, 5$. The high R^2 denotes a very good model fit, which is important because the estimated model is used to transform $p^s_{i,j,k}(24)$ into $q^s_{i,j,k}$. This is done by numerically inverting the function $g_j(a)$ (see equation 9).

We report our findings from the measurement-error model specified in equation (11). Perhaps not surprisingly, we find substantial evidence of measurement error. Appendix

Table A5 displays the decomposition of the variance of $\ln q_{i,j,k}^{s}$ into information about the latent variable $E\left(\ln q_{i,1} \left|\theta_i, q_0, x\right.\right)$, item-specific measurement error, $\xi_{i,j}$, and measurement uniqueness, $\epsilon_{i,j,k}$. The measurement signal – the part of the variance on $\ln q_{i,j,k}^{s}$ that is due to $E\left(\ln q_{i,1} \left|\theta_i, q_0, x\right.\right)$ – varies between just below 30% to close to 60%. Clearly, not all items are equally informative. There are eight MSD items that contain a signal between 40% and 60% of total variance; six items' signal captures between 30% and 40% of total variance; and the signal of one item represents less than 30% of total variance of $\ln q_{i,j,k}^{s}$.

We use the measurements $\ln q_{i,j,k}^{s}$ and the measurement-error model (11) to estimate the subjective expected natural log of human capital at age 24 months, $E\left(\ln q_{i,1} \left|\theta_i, q_0, x\right.\right)$. This is done by estimating the Bartlett scores, which are unbiased estimates of the factors. Appendix Figure A2 plots the marginal densities of the four $E\left(\ln q_{i,1} \left|\theta_i, q_0, x\right.\right)$ (one for each scenario). First, the marginal densities are positively skewed. The global modal is situated around the value $\ln 12 \approx 2.5$, which implies that a large group of women expect human capital at age 24 months to be equivalent to human capital of the typical child who is 12 months old, even in the best scenario. There is also evidence of a local mode, situated around $\ln 24 \approx 3.5$, which implies the existence of a large number of women with expectations at age 24 months around the human capital of the typical 24-month-old child.

Second, we focus on the scenarios in which the baby's health at birth is "good" (the densities depicted in solid lines in Appendix Figure A2). Clearly, the density in which the investment is high is located to the right of the density in which the investment is low. This property derives directly from the maternal responses regarding age ranges: mothers believe that children will learn tasks at earlier ages in the high-investment scenario. The difference between high and low investments is much smaller in the scenarios in which the baby's health at birth is "poor".

A.3 Method of Interview

We examine whether the way the survey is implemented affects the results. Table A6 compares subjective expectations for a sample using computer-assisted personal interviews (CAPI) with those of a sample using audio version of the computer-assisted personal interviews (ACASI).[25] The results are qualitative consistent across survey methods, but the respondents on the left tail of the median tend to report lower values under ACASI than under CAPI.

B The CNLSY/79 Data

B.1 Description of the data and sample used in this paper

The CNLSY/79 consists of all children born to NLSY/79 female respondents.[26] These children have been independently followed and interviewed every two years since 1986. To this day, there are roughly 11,500 children enrolled in the study, but not all children are measured in all the interviews. As shown in Appendix Table B1, there are over 2,000 Hispanic, over 3,000 black, and over 6,000 non-black and non-Hispanic (henceforth, white) children in the CNLSY/79. More important for this paper, many children were born many years before 1983. As a result, we do not observe early measures of investments or skills for around 30% and 40% of the white and black sample, respectively. In other words, the most that one can hope for is to obtain the full data for 60% of the black children and 70% of the white children. Unfortunately, even these rates are not possible because of other missing data issues.

There is also substantial missing information on proxy measures of health at birth – birth weight, birth length, gestational age, and number of days in the hospital after birth. They are important not only for the objective estimation of the technology of skill formation but also for the creation of scenarios for the elicitation of subjective beliefs. The measures of health at birth are reported by the mother and the information is available for 1,366 Hispanic, 1,760 black, and 3,961 white children, respectively.

In order to estimate the technology of skill formation, we also need data on skills and investments around age 24 months. If we were to use only the children whose skills are measured between 22 and 26 months, the sample would be too small (184 Hispanic, 242 black, and 588 white children). In order to have a larger sample size, we consider intervals centered at age 24 months. The cost of this choice is that we add individuals whose skills are measured with different MSD items. The items at younger ages focus on dimensions of skills that may be differentially affected by parental investments. If we include the individuals whose skills and investments are measured between (and including) 13 and 35

---

[25] Note that the definitions of the baby's health at birth are slightly different. The "poor" health at birth in the CAPI survey was 18 inches long at birth versus 19 inches in the ACASI interview; the "good" health at birth in the CAPI survey was 8 pounds at birth versus 7 pounds in the ACASI survey.

[26] The NLSY/79 respondents are a representative sample of U.S. women born between 1957 and 1964.

months, we have 913 Hispanic, 1,182 black and 2,626 white children. If we make the width a little narrower and include only the children whose skills and investments were measured between (and including) 16 and 32 months, then our sample is smaller: 685 Hispanic, 882 blacks, and 1,948 whites. Finally, if we make the band even narrower and consider only the individuals with data collected between (and including) 19 and 29 months, then our sample size is substantially smaller and contains only 435 Hispanic, 548 black children, and 1,260 white children.

B.2. Summary statistics

Appendix Table B2 contains descriptive statistics used in the estimation of the technology of skill formation. In our sample, the typical child's skill level is close to 25.5 months. Given that the typical child is interviewed at age 27.8 months, this implies a mental development delay of around 2.3 months.

Appendix Table B2 also shows differences between the skills of Hispanic, black, and white children. According to the MSD instrument, black children tend to have more skills at age 24 months than the observationally equivalent Hispanic and white children. However, the black-white difference is neither large nor statistically significant after we control for the child's gender, birth order, and age at date of the measurement, as well as the age of the mother at the birth of the child (Appendix Table B3). The same conclusion holds even after we have added dummies for maternal education, a measure of maternal skills, and a variable that contains information on permanent family income.

Interestingly, Appendix Table B3 shows large differences across gender and birth order. The typical female has a developmental advantage of around two and half months. The second-born child is 1.5 months less developed than his older sibling, and this difference is larger for children with higher birth order. Columns 2 and 3 in Appendix Table B3 display the heterogeneity of skills around age 24 months across and within families. Interestingly, even within families there is a large discrepancy between boys and girls and among siblings of different birth order.

Health at birth is scaled and located by gestational age. As expected, the health at birth of the typical child in our sample analysis is around 9 months. There is a small difference in this dimension across races and black children tend to be worse off than Hispanic or white children in this dimension. If we control for the child's gender, birth order, and the mother's age at the time of the birth of the child, we conclude that the difference between white and black children is roughly about four days as measured on the scale of health at birth (Appendix Table B3). About a quarter of this difference can be accounted for by the difference in maternal education, maternal skills, and permanent family income across the groups.

The gender gap in health at birth is the opposite to those found in skills at age 24 months: the girls' stocks of q_0 at birth are lower than boys'**.** This is true whether we look at variation across or within families. The birth-order gaps are mostly small and not statistically significant. (Appendix Table B3, columns 4-6).

Appendix Table B2 also shows that the typical child receives a little under two months of investment per year. However, there is large heterogeneity: the child in the 90th percentile receives almost twice the amount of investments as the child in the 10th percentile. The black-white gap in investments is around a half month per year (in favor of white children) once we account for the child's gender, birth order, and age at interview as well as maternal age at the birth of the child (Appendix Table B3). Furthermore, only 30% of this difference is explained by maternal education, maternal skills, and permanent family income.

As shown in Appendix Table B3, girls tend to receive slightly more investments. The birth-order gap is substantially larger: the first-born child receives between 4.7 and 6.5 more days of investments per year than the second-born. The third-born is allocated at least one fewer week of investments per year than the first-born child. The variation within families confirms differences between genders or among siblings of different birth order.

We finish this subsection by comparing the variability of residuals of key variables (skills around 24 months, health at birth, and investments) across and within families. This is important because we rely on variation within families to estimate the technology of skill formation. Appendix Figure B1 plots the histograms of the cognitive skill residuals around age 24 months, health at birth, and investments. In order to produce these residuals, we regress the key variable of interest against dummies for the child's gender, birth order, and age at the time of the interview, as well as maternal age at the time of the birth of the child. The top row shows the histogram of the residuals across all families. The bottom row shows the histogram of deviation of the residuals with respect to the family average residual. Interestingly, our estimates indicate that 20% to 30% of the variability of residualized key variables is attributed to within-family variability, while the bulk of the variability (70% to 80%) is due to unmeasured factors that vary across families.

C Measurement of Child Development and Investments in the CNLSY/79

C.1 Construction of the measure of cognitive skills at age 24 months

Several problems arise in the use of the MSD scale as a measure of early human capital. First, Cunha, Heckman, and Schennach (2010) find evidence that the MSD is contaminated with measurement error. One reason why the MSD has measurement error may be that it has relatively few items. For example, the Bayley Scale of Infant Development has a total of 64 items – 32 items measure mental development and the other 32 items measure motor development. In comparison, the MSD has only 15 items. Another reason that may generate measurement error is the fact that the MSD is based on maternal reports. In contrast, the Bayley Scale of Infant Development involves direct observation of the child by a trained expert in child development.

To mitigate the problem of measurement error, we conduct item response theory (IRT) analysis and treat the mother's responses as repeated binary indicators of her child's ability. To understand how IRT can help mitigate measurement error, note that the child's raw score in the MSD is simply the number of "yes" answers that the mother reports for the child. If all items were of identical difficulty, then this simple average is the best that one could achieve to estimate a child's ability. To the extent that the items differ in their level of difficulty, a weighted average can provide a more precise estimate of the child's ability by assigning a higher weight to more difficult items. IRT analysis provides a way to obtain these weights and makes it possible to estimate the child's ability as precisely as possible.

Let $d_{i,j}^{*}$ denote the latent variable that is determined according to:

$$d_{i,j}^{*} = b_{0,j} + b_{1,j} q_{i,1} + \varepsilon_{i,j}.$$

The variable $q_{i,1}$ is the child's (latent) cognitive ability around age 24 months and is independent of the error term $\varepsilon_{i,j}$, which is i.i.d. across children i and items j. The variable $d_{i,j}^{*}$ is not observed. Instead, we observe $d_{i,j} = 1$ if, and only if, $d_{i,j}^{*} \geq 0$ and $d_{i,j} = 0$, otherwise. Thus:

$$Pr\left(d_{i,j} = 1 \,\middle|\, q_{i,1}\right) = 1 - Pr\left(\varepsilon_{i,j} \leq -b_{0,j} - b_{1,j} q_{i,1} \,\middle|\, q_{i,1}\right) = 1 - F_{\varepsilon}\left(-b_{0,j} - b_{1,j} q_{i,1}\right)$$

where F_ε is the distribution of $\varepsilon_{i,j}$. The parameter $b_{0,j}$ represents the location of item j. In the case of cognitive testing, the higher the value of $b_{0,j}$, the higher the difficulty of the item. The parameter $b_{1,j}$ represents the discrimination of the item: when $b_{1,j}$ is high, children with low ability have a much smaller chance of correctly performing a given task j than children with high ability.

Estimation of the IRT model is done via maximum likelihood. Assume that $q_{i,1}$ is i.i.d. across children and let f_{q_1} denote its density function. The problem is to choose parameter vectors $b_0 = (b_{0,1}, \dots, b_{0,J})$ and $b_1 = (b_{11}, \dots, b_{1,j})$ as well as a density function f_{q_1} that solves:

$$\max_{b_0,b_1,f_{q_1}} \prod_{i=1}^{I} \prod_{j=1}^{J} \int [F_\varepsilon(-b_{0,j} - b_{1,j}q_{i,1})]^{1-d_{i,j}} [1 - F_\varepsilon(-b_{0,j} - b_{1,j}q_{i,1})]^{d_{i,j}} f_{q_1}(q_1) dq_1 \quad \text{(B1)}$$

In principle, it would be possible to use the CNLSY/79 data to estimate the parameters b_0 and b_1 and the density function f_{q_1} in (B1). Although the CNLSY/79 sample is representative of the children of the women born between 1957 and 1964, it is not a representative sample of U.S. children in general. For this reason, we use the data from the National Health and Nutrition Examination Survey 1988-1994 (NHANES). Interestingly, the MSD scale is used in both studies.

In our empirical implementation, we assume that $\varepsilon_{i,j} \sim N(0,1)$ and $f_{q_1}(q_1) = \sum_{k=1}^{K} \omega_k \phi(q_1, \mu_{q_1,k}, \sigma^2_{q_1,k})$ where $\phi(q_1, \mu_{q_1,k}, \sigma^2_{q_1,k})$ is the density of a normal random variable with mean $\mu_{q_1,k}$ and variance $\sigma^2_{q_1,k}$. The term ω_k is the weight of the element k and satisfies $\sum_{k=1}^{k} \omega_k = 1$. To fix the location of q_1, we impose the restriction that $\sum_{k=1}^{k} \omega_k \mu_{q_1,k} = 0$. To fix the scale of q_1, we set $b_{1,36} = 1$. Appendix Table C1 presents the estimated parameters $\hat{b}_0$ and $\hat{b}_1$ for the 48 items of the MSD scale. Note that there is substantial heterogeneity across items in terms of difficulty and discrimination power.

Next, we estimate $q_{i,1}$ for each child in the NHANES data set. This is done by solving for each child i the following problem:

$$\hat{q}_{i,1} = \operatorname{argmax} \prod_{j=1}^{J} [F_\varepsilon(-b_{0,j} - b_{1,j}q_{i,1})]^{1-d_{i,j}} [1 - F_\varepsilon(-b_{0,j} - b_{1,j}q_{i,1})]^{d_{i,j}} \quad \text{(B2)}$$

If $d_{i,j}$ were continuous measures and $f_{q_1}(q_1)$ was the normal density, then the solution of (B2) would be given by the Bartlett scores of the factor $q_{i,1}$. The advantage of the Bartlett scores is that they produce unbiased estimates of the factors, which is one of the reasons we choose this approach.

Importantly, $\hat{q}_{i,1}$ addresses measurement error. Items that are more difficult or have more discriminating power obtain a higher weight. However, the IRT analysis per se does not produce estimates of cognitive skills that have a natural metric. To do so, we explore the fact that children as young as two months old and as old as 47 months old were assessed by NANHES. As one would expect, there is a monotonic relationship between the MSD score and the child's age at the time of the test (Appendix Figure C1). It is this monotonic relationship that we explore to transform the test score in an estimated "age of development."

For each age (measured in months), let $q_a^{median} = median(q_{i,1} | age = a)$. That is, q_a^{median} is the median MSD score among children who were exactly a months old at the time of the MSD test. Appendix Figure C2 plots the relationship between age a (in the horizontal axis) and median score q_a^{median}. Interestingly, as shown in Appendix Figure C2, the empirical relationship is very well approximated by the following function:

$$q_a^{median} = \zeta_0 + \zeta_1 \ln a$$

where ζ_0 and ζ_1 are estimated by OLS. Note that the quantity $\frac{\hat{q}_{i,1} - \zeta_0}{\zeta_1}$ is the child's MSD score written in terms of developmental age. In other words, a child whose score is $\hat{q}_{i,1}$ has the skills of children who are exactly $e^{\frac{\hat{q}_{i,1} - \zeta_0}{\zeta_1}}$ months old. In the estimation of the technology of skill formation (2), we use $\ln q_{i,1} = \frac{\hat{q}_{i,1} - \zeta_0}{\zeta_1}$ as the dependent variable in our objective estimation of the technology of skill formation in Section 5.2.

C.2 Construction of investment data

To obtain a measure of investment in the metric of time, we proceed in two steps. First, we estimate the distribution of time spent investing in children from the Child Development Supplement of the Panel Study of Income Dynamics - Child Development Supplement (PSID-CDS). The PSID-CDS asks parents to report their children time diaries for two days of the week (one weekday and one weekend day, both picked randomly). It is possible to use this information to construct a measure of parental investments in hours of interaction with their child per day. However, for consistency with the measures of cognitive skills, we measure investments in number of months per year.[27] [28] We approximate the density of time investment time from the PSID-CDS data with a mixture of normal densities. Let $F_X(x)$ denote the distribution of time investment estimated from the PSID-CDS sample.

We then impose the same distribution on the factor that is extracted from the HOME-SF items from the CNLSY/79. Specifically, let $M_{i,j}$ denote an item of the HOME-SF scale. We assume that the relationship between observed $M_{i,j}$ and latent investment x_i is given by the following equation:

$$M_{i,j} = b_{0,j}^M + b_{1,j}^M x_i + \varepsilon_{i,j}^M$$

where $\varepsilon_{i,j}^M \sim N\left(0, \sigma_{M,j}^2\right)$ is measurement error, $x_i \sim F_X$ is independent from $\varepsilon_{i,j}^M$ and F_X is the distribution of investment estimated from the PSID-CDS data on time investments.

D Direct elicitation of probabilities

We now describe an alternative way to elicit expectations about the parameter γ_i. The advantage of the approach described here is that it does not require interpolation assumptions to predict the probability at a given target age. As originally conceived, the MSD scale are maternal self-reports indicating, with yes/no answers, whether a child has reached a particular developmental milestone by a given age. We have changed the instrument so that, instead of yes/no answers, the mother reports the subjective probability that a developmental milestone will be reached at a given age (e.g., 24 months) depending on a level of investments and the infant's characteristics at birth. For example, we show the mother what a high level of investment means that the mother spends 4 hours a day interacting with the child. We tell the mother that the child is born normal weight at birth, normal length at birth, and normal gestation length. We then ask the respondent to provide the subjective probability using a sliding scale on the computer screen that ranges from 0 to 100. This process is repeated for three other combinations of investments and child health at birth (See Figure D1).

---

[27] The *stocks* of cognitive skills and health at birth are measured in months. The *flow* of investment is measured in months per year.

[28] The conversion is simple: h hours per day with the child is equivalent to $365h$ hours per year. Note that there are $365/12$ days in a month. As a result, there are $24 \times \frac{365}{12} = 730$ hours in a month. Thus, $365h$ hours per year are equivalent to $\frac{365h/year}{730/month} = \frac{h}{2}$ months per year. For example, it is equivalent to say that a mother spends 2 hours per day or 1 month per year with the child.

We chose to use sliding scales for two reasons. First, they allow us to combine verbal and numerical representations of probabilistic statements. Evidence in cognitive psychology shows that subjects best communicate their beliefs when they are given access to verbal expressions of probabilistic statements (Wallsten et al. 1986). The labels representing probabilities in Figure D1 (e.g., "Toss-up") were chosen according to the findings described in Hamm (1991). Second, Delavande, Giné, and McKenzie (2011) show that individuals report probabilities more accurately when their beliefs are represented with visual instruments.

In our implementation of this approach, our video explains to the mother that the "high" investment is four hours and the "low" investment is three hours. In terms of the baby's health at birth, we defined the "healthy" baby is the one whose gestation lasts nine months, weighs seven pounds at birth, and stays at most three days at the hospital. We define the "not healthy" baby the one whose gestation was short (only seven months), the weight at birth was only five pounds, and that the baby had to stay in the hospital for seven days after birth.

As shown in Section 4.3, the direct elicitation of probabilities produces expectations about γ_i that are close to the ones generated by the upper and lower triangular distributions at target ages 28 and 32 months. Our analyses of measurement error and model fit indicate that these methods are among the top performing ones when we use the elicitation of expectations through the age-range approach. Although the direct elicitation of probabilities does not require any assumption about the interpolation method, it does have other problems. As discussed in Section 3.2.3, if respondents understand the instrument well, we would expect them to report lower probabilities for items that are more difficult once we hold constant investment and the baby's health at birth. Unfortunately, as shown in Table D1, this is not what happens with the direct elicitation of probabilities. Remarkably, the probabilities across items fluctuate around 75% for Scenario "1", 56% for Scenario "2", 70% for Scenario "3" and 50% for Scenario "4". Interestingly, the findings confirm our conclusion that respondents understand that the production function is monotonic, but they also have different returns for

E Implementation of the Arellano-Bonhomme Procedure

In what follows, suppose that each mother has L children. Let $\Gamma_i = (\theta_i \quad \gamma_i)'$ and $\mathrm{B} = (\rho \quad \beta)'$. Define the matrices X_i and Z_i as:

$$X_i = \begin{bmatrix} 1 & x_{i,1} \\ \vdots & \vdots \\ 1 & x_{i,L} \end{bmatrix} \quad Z_i = \begin{bmatrix} \ln q_{0,i,1} & R_{i,1} \\ \vdots & \vdots \\ \ln q_{0,i,L} & R_{i,L} \end{bmatrix}$$

We rewrite (8) as:

$$\ln q_{i,1} = Z_i \mathrm{B} + X_i \Gamma_i + \nu_{i,l}.$$

Note that $\ln q_{0,i,j}$ is in the matrix Z_i, which has a common parameter vector B. Although it would be interesting to allow the parameter ρ to vary across i, this is costly: As we increase the number of parameters that vary across individuals, we also need to increase the number of children observed for each mother. So, when ρ, γ, and θ vary across mothers i, it is necessary to observe at least four children for each mother. Unfortunately, there are very few respondents who satisfy this condition in the CNLSY/79 data set.

Let $Q_i = I - X_i(X_i'X_i)^{-1}X_i'$. The estimator for B is:

$$\widehat{\mathrm{B}} = [\sum_{i=1}^{I} Z_i' Q_i Z_i]^{-1} \left[\sum_{i=1}^{I} Z_i' Q_i \ln q_{i,1}\right].$$

We use $\widehat{\mathrm{B}}$ to obtain an estimator for Γ_i:

$$\hat{\Gamma}_i = (X_i'X_i)^{-1} X_i' \left(\ln q_{i,1} - Z_i \widehat{\mathrm{B}}\right).$$

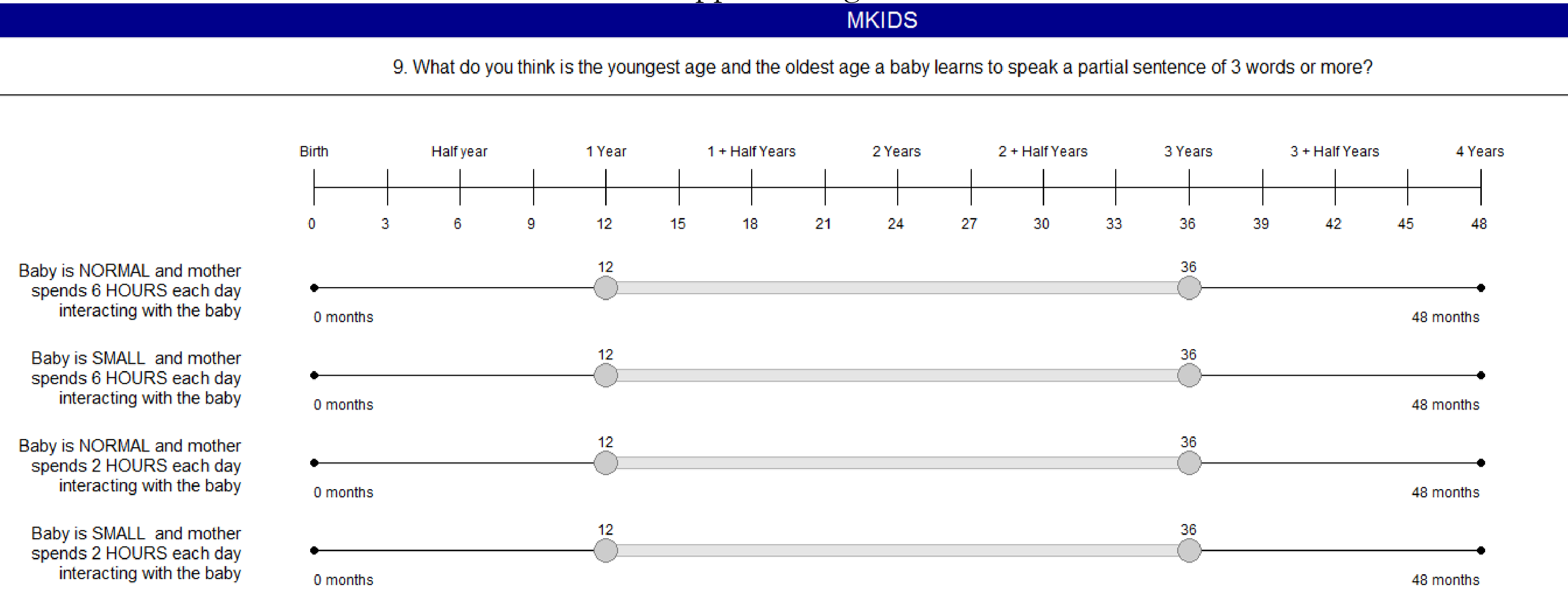
Clearly, we can use $\hat{\Gamma}_i$ to estimate $\Gamma = E[\Gamma_i]$:

$$\hat{\Gamma} = \frac{1}{I} \sum_{i=1}^{I} \hat{\Gamma}_i.$$

Next, define the variance of orthogonal (to X_i) residuals as $\hat{\sigma}^2 = \frac{1}{I(T-2)} \sum_{i=1}^{I} \left(\ln q_{i,1} - Z_i \widehat{\mathrm{B}}\right)' Q_i \left(\ln q_{i,1} - Z_i \widehat{\mathrm{B}}\right)$. Then, an estimator of $\Sigma_\Gamma = Var(\Gamma_i)$ is given by:

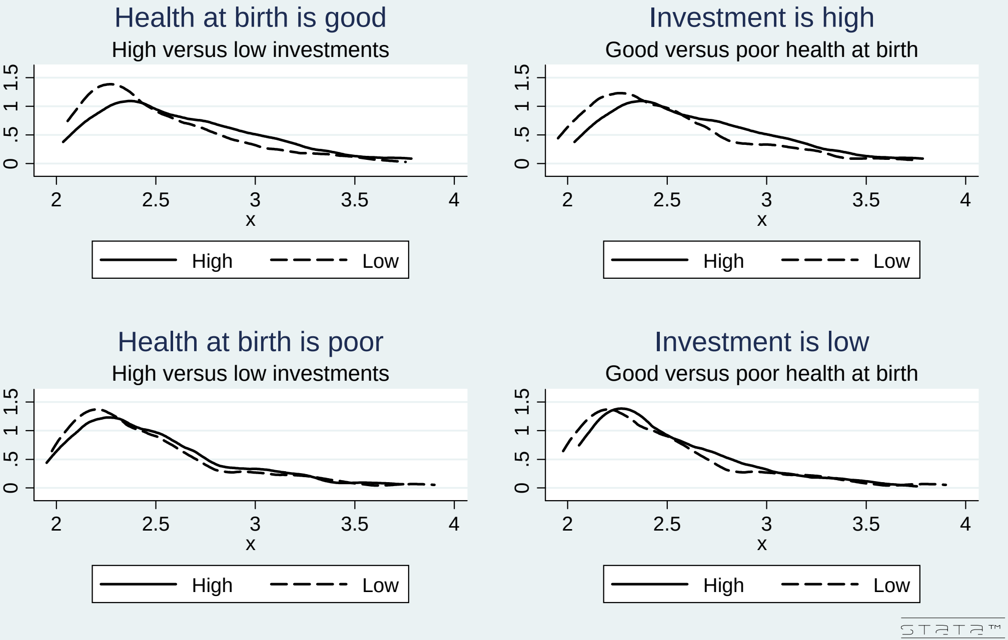
$$\Sigma_\Gamma = \frac{1}{I} \sum_{i=1}^{I} \left(\hat{\Gamma}_i - \hat{\Gamma}\right) \left(\hat{\Gamma}_i - \hat{\Gamma}\right)' - \hat{\sigma}^2 \frac{1}{I} \sum_{i=1}^{I} (X_i'X_i)^{-1}.$$

Appendix Figure A1



Appendix Figure A2

Kernel density of expected human capital at age 24 months



Appendix Figure A3

MKIDS

5. Suppose that your household income is $2,000 per month and that for each hour per day that the mother spends interacting with the child she has to spend $45 per month on educational goods. Consider the following four options:

OPTION A: Mom spends 2 hours per day interacting with the child, $90 per month on educational goods, and $1910 on household goods.

OPTION B: Mom spends 3 hours per day interacting with the child, $135 per month on educational goods, and $1865 on household goods.

OPTION C: Mom spends 4 hours per day interacting with the child, $180 per month on educational goods, and $1820 on household goods.

OPTION D: Mom spends 5 hours per day interacting with the child, $225 per month on educational goods, and $1775 on household goods.

What option do you prefer, A, B, C, or D?

OPTION A OPTION B OPTION C OPTION D

Appendix Table A1
Demographic Characteristics
MKIDS African-American Participants
Number of observations = 335 of which 201 (60%) are primiparous

	Age Group				
	18-20	21-25	26-30	31-35	36+
Overall (Fraction)	35.52	46.57	12.24	3.87	1.80
Primiparous	47.76	38.31	10.95	1.50	1.49
Non-primiparous	17.16	58.96	14.19	7.47	2.24

	Education Level				
	Dropout or GED	High School Graduate	Some College	2-Year College Graduate	4-Year College Graduate
Overall (Fraction)	19.40	38.51	30.75	4.48	6.87
Primiparous	18.91	39.30	30.35	2.99	8.46
Non-primiparous	20.15	37.31	31.34	6.72	4.48

	Household income ($/month)				
	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
Overall ($/Month)	319.00	800.00	1500.00	2300.00	3200.00
Overall - Location in US Distribution[1]	2	6	19	28	40
Primiparous	209.00	800.00	1500.00	2500.00	3500.00
Non-primiparous	360.00	705.00	1500.00	2028.00	2900.00

	Type of Medical Insurance		
	Private Insurance	Medicaid	Other
Overall (Fraction)	7.76	87.46	4.78
Primiparous	9.95	86.57	3.48
Non-primiparous	4.48	88.81	6.72

	Marital Status		
	Single	Cohabiting or married	Separated or divorced
Overall (Fraction)	80.54	18.26	1.20
Primiparous	83.08	15.92	1.00
Non-primiparous	76.69	21.80	1.50

[1]US Census Bureau

Appendix Table A2

Average lowest and highest age by MSD item and scenario[1]

	Health at birth is "good"[2]				Health at birth is "poor"[2]			
	Investment is "high"[3] Scenario 1		Investment in "low"[3] Scenario 3		Investment is "high"[3] Scenario 2		Investment in "low"[3] Scenario 4	
	Lowest age	Highest age	Lowest age	Highest age	Lowest age	Highest age	Lowest age	Highest age
MSD 30: Let someone know, without crying, that wearing wet or soiled pants or diapers bothers him or her?	14.58	28.39	18.52	32.74	16.81	30.32	21.11	34.16
MSD 35: Speak a partial sentence of 3 words or more?	22.26	35.69	26.63	39.02	24.66	37.88	29.09	40.53
MSD 37: Walk upstairs by himself/herself without holding on to a rail?	22.47	36.15	25.42	38.16	25.39	38.09	28.29	39.86
MSD 44: Wash and dry his/her hands without any help except for turning the water on and off?	24.03	37.34	27.37	39.53	26.08	38.54	29.53	41.22
MSD 39: Walk upstairs by himself/herself with no help, stepping on each step with only one foot?	24.23	37.44	27.22	39.99	26.69	39.15	29.68	41.27
MSD 38: Count 3 objects correctly?	25.08	37.72	28.70	40.84	26.54	38.63	30.50	42.26
MSD45: Dress himself/herself without any help except for tying shoes?	29.96	42.27	32.87	43.77	31.77	42.76	35.00	44.37
MSD 46: Go to the toilet alone?	25.86	38.20	29.04	40.53	28.05	39.84	31.53	42.23
MSD 40: Know his/her age and sex?	25.35	38.70	28.97	40.91	27.16	39.99	31.10	42.35
MSD 43: Do a somersault without help from anybody?	28.78	40.49	31.66	42.45	31.04	41.84	34.09	43.72
MSD 42: Pedal a tricycle at least 10 feet?	30.16	41.44	32.92	43.68	32.04	42.49	34.98	45.01
MSD 41: Say the names of at least 4 colors?	26.06	38.85	30.23	41.84	28.24	39.86	32.08	43.30
MSD 36: Say his/her first and last name together without someone's help?	27.30	39.37	30.97	41.81	28.64	40.42	32.88	43.00
MSD 47: Count out loud up to 10?	27.96	40.58	31.42	42.96	29.11	41.28	33.03	43.78
MSD 48: Draw a picture of a man or woman with at least 2 parts of the body besides a head?	33.42	43.35	35.85	44.59	34.16	43.78	37.13	45.34

[1]MSD items are listed in descending order with respect to the probability that children age 24 months are able to execute the task.

[2]Heath at birth is "good" if the baby weighs 8 pounds at birth, is 20 inches long at birth, and the gestational age is 9 months. It is "poor" if the baby weighs 5 pounds at birth, is 18 inches long at birth, and the gestational age is 7 months.

[3]Investment is "high" when mother spends 6 hours/day interacting with the child and is "low" when the mother spends 2 hours/day.

Appendix Table A3

Average probability by age 24 months by MSD item and scenario[1]

		Health at birth is "good"[2]		Health condition at birth is "poor"[2]	
		Scenario 1	Scenario 3	Scenario 2	Scenario 4
	NHANES	Investment is "high"[3]	Investment in "low"[3]	Investment is "high"[3]	Investment in "low"[3]
MSD 30: Let someone know, without crying, that wearing wet or soiled pants or diapers bothers him or her?	92.52%	66.32%	41.86%	55.26%	38.48%
MSD 35: Speak a partial sentence of 3 words or more?	88.54%	31.29%	19.51%	22.44%	16.06%
MSD 39: Walk upstairs by himself/herself with no help, stepping on each step with only one foot?	62.11%	24.57%	15.55%	21.16%	16.19%
MSD45: Dress himself/herself without any help except for tying shoes?	47.92%	21.92%	14.98%	19.44%	16.34%
MSD 40: Know his/her age and sex?	42.55%	18.37%	17.20%	19.26%	20.50%
MSD 41: Say the names of at least 4 colors?	32.98%	13.71%	13.09%	18.21%	20.39%

[1]MSD items are listed in descending order with respect to the probability that children age 24 months are able to execute the task.

[2]Heath at birth is "good" if the baby weighs 8 pounds at birth, is 20 inches long at birth, and the gestational age is 9 months. It is "poor" if the baby weighs 5 pounds at birth, is 18 inches long at birth, and the gestational age is 7 months.

[3]Investment is "high" when mother spends 6 hours/day interacting with the child and is "low" when the mother spends 2 hours/day.

Appendix Table A4

Logistic Approximation of NHANES MSD Items by Age

Panel A

VARIABLES	Let someone know, without crying, that wearing wet or soiled pants or diapers bothers him or her?	Speak a partial sentence of 3 words or more?	Say his/her first and last name together without someone's help?	Walk upstairs by himself/herself without holding on to a rail?	Count 3 objects correctly?
Intercept	-11.05***	-9.921***	-10.10***	-10.18***	-9.860***
	(0.29)	(0.39)	(0.46)	(0.40)	(0.36)
Age	2.010***	0.742***	0.995***	1.018***	0.667***
	(0.10)	(0.16)	(0.19)	(0.22)	(0.20)
Age squared	-0.148***	-1.150E-02	-0.0679***	-0.0437*	-1.430E-02
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Age cubed	0.00565***	4.870E-05	0.00281***	1.330E-03	2.470E-04
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age to the fourth	-9.53e-05***	-1.430E-06	-5.43e-05***	-2.400E-05	-4.920E-06
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age to the fifth	5.78e-07***	3.100E-08	3.87e-07***	1.800E-07	5.300E-08
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	17	37	37	34	34
R-squared	0.999	0.994	0.986	0.992	0.993

Panel B

	Walk upstairs by himself/herself with no help, stepping on each step with only one foot?	Know his/her age and sex?	Say the names of at least 4 colors?	Pedal a tricycle at least 10 feet?	Do a somersault without help from anybody?
Intercept	-10.07***	-10.15***	-9.819***	-10.08***	-10.46***
	(0.49)	(0.39)	(0.34)	(0.41)	(0.41)
Age	0.884***	1.012***	0.627**	0.911***	1.319***
	(0.27)	(0.22)	(0.26)	(0.30)	(0.31)
Age squared	-3.500E-02	-0.0609**	-1.890E-02	-4.720E-02	-0.0756**
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Age cubed	1.090E-03	0.00235**	6.040E-04	1.700E-03	0.00249*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age to the fourth	-2.160E-05	-4.49e-05**	-1.380E-05	-3.210E-05	-4.33e-05*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age to the fifth	1.740E-07	3.23e-07***	1.250E-07	2.370E-07	3.03e-07**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	34	34	31	31	31
R-squared	0.985	0.99	0.994	0.992	0.991

Panel C

	Wash and dry his/her hands without any help except for turning the water on and off?	Dress himself/herself without any help except for tying shoes?	Go to the toilet alone?	Count out loud up to 10?	Draw a picture of a man or woman with at least 2 parts of the body besides a head?
Intercept	-11.18***	-10.51***	-11.00***	-8.731***	-9.714***
	(0.99)	(0.89)	(0.93)	(0.39)	(0.77)
Age	2.114**	1.386E+00	1.937**	-5.660E-01	5.340E-01
	(0.93)	(0.83)	(0.87)	(0.36)	(0.78)
Age squared	-1.540E-01	-9.360E-02	-0.157*	0.0887**	-3.350E-02
	(0.09)	(0.08)	(0.09)	(0.04)	(0.08)
Age cubed	5.490E-03	3.440E-03	0.00614*	-0.00315**	1.600E-03
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age to the fourth	-9.150E-05	-6.040E-05	-0.000107*	4.33e-05*	-3.510E-05
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age to the fifth	5.740E-07	4.030E-07	6.85e-07*	-1.900E-07	2.790E-07
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	27	28	28	28	26
R-squared	0.988	0.983	0.988	0.996	0.994

Standard errors in parentheses

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

Appendix Table A5
Decomposition of measurement error

MSD Item	Scenario	Total Variance	Share of variance due to signal	Share of variance due to item measurement error	Share of variance due to item and scenario measurement error
MSD 30: Let someone know, without crying, that wearing wet or soiled pants or diapers bothers him or her?	1	0.6885	12.22%	52.23%	35.55%
	2	0.6171	12.07%	58.28%	29.65%
	3	0.6469	11.32%	55.59%	33.09%
	4	0.7197	11.23%	49.97%	38.80%
MSD 35: Speak a partial sentence of 3 words or more?	1	0.6156	27.17%	27.90%	39.26%
	2	0.5561	26.64%	30.88%	36.93%
	3	0.4675	31.15%	36.74%	25.62%
	4	0.4866	33.03%	35.30%	24.79%
MSD 36: Say his/her first and last name together without someone's help?	1	0.5457	50.05%	27.83%	24.84%
	2	0.4572	52.91%	33.22%	16.75%
	3	0.4614	51.54%	32.91%	18.35%
	4	0.5324	49.28%	28.52%	24.88%
MSD 37: Walk upstairs by himself/herself without holding on to a rail?	1	0.7349	33.23%	41.14%	25.62%
	2	0.6726	32.16%	44.95%	22.89%
	3	0.6769	31.41%	44.67%	23.91%
	4	0.7517	31.21%	40.23%	28.56%
MSD 38: Count 3 objects correctly?	1	0.5915	45.90%	22.09%	34.38%
	2	0.4753	50.60%	27.50%	24.52%
	3	0.4240	55.75%	30.82%	16.31%
	4	0.4459	58.50%	29.31%	15.21%
MSD 39: Walk upstairs by himself/herself with no help, stepping on each step with only one foot?	1	0.7827	33.59%	38.31%	29.32%
	2	0.6864	33.93%	43.68%	23.62%
	3	0.6770	33.82%	44.29%	23.12%
	4	0.6754	37.41%	44.40%	19.55%
MSD 40: Know his/her age and sex?	1	0.6609	44.91%	30.26%	29.00%
	2	0.6001	43.80%	28.33%	31.94%
	3	0.5543	46.62%	23.13%	34.58%
	4	0.6095	46.79%	26.11%	31.45%
MSD 41: Say the names of at least 4 colors?	1	0.4614	53.52%	20.71%	26.06%
	2	0.3982	54.93%	24.01%	21.37%
	3	0.3798	56.61%	25.17%	18.54%
	4	0.4006	59.23%	23.86%	17.24%
MSD 42: Pedal a tricycle at least 10 feet?	1	0.5850	37.53%	36.44%	23.98%
	2	0.5556	35.00%	38.37%	24.71%
	3	0.5084	37.60%	41.93%	18.42%
	4	0.5732	36.80%	37.19%	24.00%
MSD 43: Do a somersault without help from anybody?	1	1.0449	23.34%	44.37%	32.28%
	2	0.9270	23.30%	50.01%	26.68%
	3	0.8430	25.19%	54.99%	19.80%
	4	1.0300	22.75%	45.01%	32.22%
MSD 44: Wash and dry his/her hands without any help except for turning the water on and off?	1	1.0351	48.68%	36.23%	29.88%
	2	0.9623	46.38%	38.97%	28.74%
	3	0.8998	48.76%	41.68%	24.37%
	4	0.9265	52.25%	40.48%	23.14%
MSD 45: Dress himself/herself without any help except for tying shoes?	1	0.7304	56.77%	29.11%	27.32%
	2	0.6617	55.51%	32.14%	25.26%
	3	0.7090	50.93%	29.99%	30.92%
	4	0.8167	48.78%	26.04%	36.53%
MSD 46: Go to the toilet alone?	1	0.9599	52.08%	24.46%	39.15%
	2	0.8429	52.53%	27.85%	35.43%
	3	0.7411	58.73%	31.68%	27.27%
	4	0.7908	60.73%	29.68%	27.87%
MSD 47: Count out loud up to 10?	1	0.2819	34.77%	20.49%	24.63%
	2	0.2519	34.47%	22.94%	22.66%
	3	0.2320	36.79%	24.90%	17.03%
	4	0.2620	35.94%	22.05%	21.22%
MSD 48: Draw a picture of a man or woman with at least 2 parts of the body besides a head?	1	0.3671	21.61%	21.21%	40.88%
	2	0.3179	22.10%	14.02%	47.20%
	3	0.3315	20.84%	18.16%	45.28%
	4	0.3677	20.73%	22.81%	40.82%

Appendix Table A6

Using expectations data

Moments of α from stated choice data

Parental valuation of child development	Mean	Median	Std Error
Lower triangular at 32 months	1.06	0.43	1.41
Upper triangular at 32 months	1.32	0.46	2.17
Lower triangular at 28 months	1.11	0.46	1.71
Upper triangular at 28 months	0.94	0.39	1.30

Not using expectations data

Parental valuation of child development	Mean	Median	Std Error
Assuming $\gamma = \mu_\gamma$ = 0.199	0.45	0.43	0.11
Assuming $\gamma = \mu_\gamma$ = 0.257	0.35	0.33	0.09
Assuming $\gamma = \mu_\gamma$ = 0.283	0.32	0.30	0.08

Appendix Table A7
Objective Estimation of the Technology of Skill Formation
Random Coefficient Model for γ and θ
Dependent variable: Natural log of skills around age 24 months[1]

	Overall Sample 13 to 35 Months	
Natural logarithm of health at birth[2]	0.535	
	(0.410)	
	Mean	Variance
Maternal specific natural logarithm of investments (γ)[3]	0.190**	0.110
	(0.094)	(0.57)
Maternal specific intercept (θ)	1.949**	0.071
	(0.913)	(0.20)
Observations	303	
Number of Mothers	985	

Standard errors are estimated by the bootstrap method. The regression has dummy variables for the child's age at the time of the measurement of the dependent variable and the child's birth order.

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

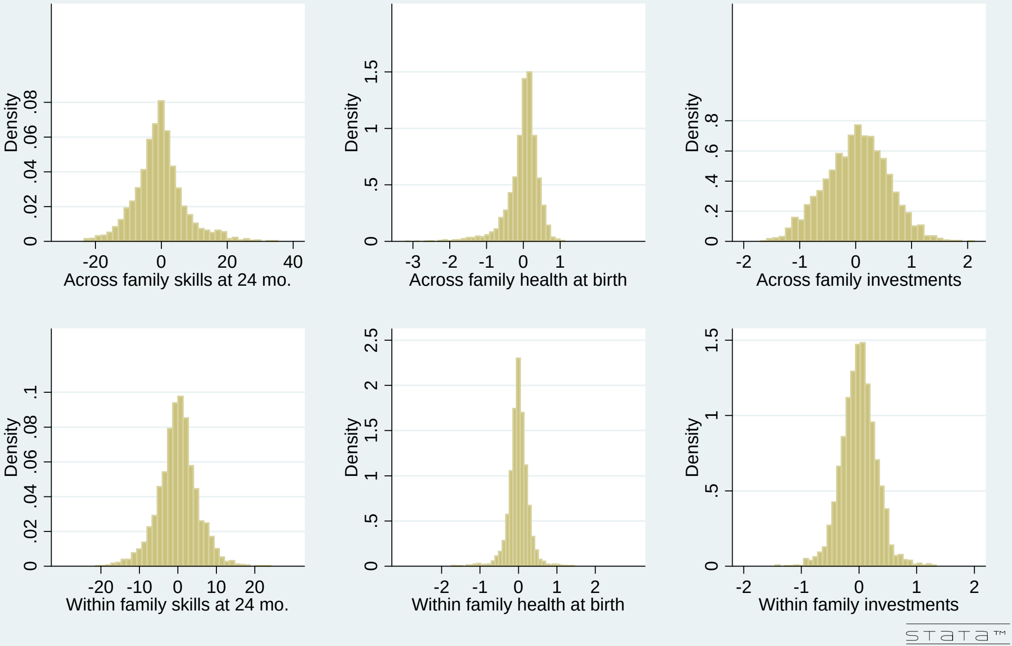
[1]Skills are measured by the Motor-Social Development Scale and are scaled in "mental" age of development.

[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.

Appendix Figure B1

Histograms of skills, health at birth, and investments



Appendix Table B1

CNLSY/79 Black and White Sample

	Hispanic		Black		White[1]		Total	
	Number	%	Number	%	Number	%	Number	%
Total Sample	2,209	100.00%	3,187	100.00%	6,095	100.00%	11,491	100.00%
Children born in 1983 or later.	1,520	68.81%	1,942	60.94%	4,259	69.88%	7,721	67.19%
Children born in 1983 or later and measures of health at birth are reported.	1,366	61.84%	1,760	55.22%	3,961	64.99%	7,087	61.67%
Children born in 1983 or later, measures of health at birth are reported, and skills and investments between ages 13 and 35 months are observed.	913	41.33%	1,182	37.09%	2,626	43.08%	4,721	41.08%
Children born in 1983 or later, measures of health at birth are reported, and skills and investments between ages 16 and 32 months are observed.	685	31.01%	882	27.67%	1,948	31.96%	3,515	30.59%
Children born in 1983 or later, measures of health at birth are reported, and skills and investments between ages 19 and 29 months are observed.	435	19.69%	548	17.19%	1,260	20.67%	2,243	19.52%
Children born in 1983 or later, measures of health at birth are reported, and skills and investments between ages 22 and 26 months are observed.	184	8.33%	242	7.59%	588	9.65%	1,014	8.82%

[1]Non-Hispanic and non-black children.

Appendix Table B2

Descriptive Statistics

	Overall Sample N = 4,721		Hispanic Sample N = 913		Black Sample N = 1,182		White Sample N = 2,626	
	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
Skills around age 24 months[1]	25.50	10.96	24.07	10.31	26.42	10.85	25.59	11.18
Health conditions at birth[2]	9.04	0.44	9.07	0.44	8.95	0.48	9.08	0.41
Investments[3]	1.87	0.71	1.59	0.74	1.46	0.66	2.15	0.58
Age skill is measured	27.80	4.75	27.89	4.91	27.12	4.82	28.08	4.64
Maternal age at birth	0.49	0.50	0.49	0.50	0.50	0.50	0.49	0.50
Child is female	0.35	0.48	0.29	0.45	0.28	0.45	0.41	0.49
Child is first born	0.35	0.48	0.32	0.47	0.34	0.47	0.37	0.48
Child is second born	0.18	0.39	0.21	0.41	0.21	0.41	0.16	0.37
Child is third born	0.07	0.26	0.10	0.29	0.10	0.30	0.05	0.22
Child is fourth born	0.04	0.20	0.08	0.27	0.06	0.24	0.02	0.13
Child is fifth born (or above)	24.59	6.62	24.81	6.63	25.11	6.57	24.28	6.62

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

[2]Health conditions at birth are captured by factor analyzing weight at birth, length at birth, gestational age, and number of days at 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.

Appendix Table B3

Descriptive Regressions of Key Variables

	Skills around age 24 months[1]			Health at birth[2]			Investments[3]		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	OLS	FE	OLS	OLS	FE	OLS	OLS	FE
Child is black	2.017***	2.108***		-0.121***	-0.0975***		-0.0975***	-0.124***	
	(0.368)	(0.454)		(0.023)	(0.027)		(0.035)	(0.038)	
Child is white	1.722***	1.741***		0.012	-0.015		0.473***	0.337***	
	(0.308)	(0.385)		(0.019)	(0.023)		(0.030)	(0.034)	
Child is female	2.464***	2.490***	1.936***	-0.0513***	-0.0568***	-0.0361**	0.0707***	0.0877***	0.0733***
	(0.209)	(0.247)	(0.29)	(0.013)	(0.015)	(0.015)	(0.016)	(0.018)	(0.017)
Child is second born	-1.523***	-1.478***	-2.001***	0.019	0.0299*	-0.018	-0.200***	-0.142***	-0.0846***
	(0.241)	(0.289)	(0.40)	(0.014)	(0.016)	(0.020)	(0.017)	(0.020)	(0.024)
Child is third born	-2.378***	-2.262***	-3.567***	0.0345*	0.0440*	-0.040	-0.380***	-0.270***	-0.159***
	(0.291)	(0.372)	(0.64)	(0.018)	(0.023)	(0.035)	(0.024)	(0.030)	(0.041)
Child is fourth born	-2.843***	-2.635***	-4.150***	0.0896***	0.0979***	-0.047	-0.493***	-0.339***	-0.141**
	(0.456)	(0.543)	(0.95)	(0.025)	(0.032)	(0.053)	(0.037)	(0.045)	(0.060)
Child is fifth born (or above)	-4.120***	-3.601***	-5.396***	-0.033	-0.015	-0.181**	-0.713***	-0.512***	-0.137
	(0.625)	(0.764)	(1.21)	(0.043)	(0.048)	(0.075)	(0.054)	(0.059)	(0.089)
Constant	20.20***	12.90***	25.66***	8.994***	8.737***	9.010***	2.214***	0.417*	2.256***
	(1.038)	(2.988)	(1.67)	(0.076)	(0.178)	(0.079)	(0.094)	(0.247)	(0.139)
Dummies for age at interview	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for maternal age at birth	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for maternal education[4]	No	Yes	No (FE)	No	Yes	No (FE)	No	Yes	No (FE)
Maternal skill[5]	No	Yes	No (FE)	No	Yes	No (FE)	No	Yes	No (FE)
Permanent family income[6]	No	Yes	No (FE)	No	Yes	No (FE)	No	Yes	No (FE)
Observations	4,721	3,421	4,721	4,721	3,421	4,721	4,721	3,421	4,721
R-squared	0.574	0.585	0.65	0.036	0.044	0.03	0.391	0.444	0.171

Robust standard errors in parentheses

*** 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.

[4]We divide women into two levels of education measured at the time of the mother's first birth: high-school dropout and GED (omitted), high-school graduate or above.

[5]AFQT (percentiles) and the scores on Rotter Locus of Control and Rosenberg Self-Esteem Scale.

[6]Permanent family income is the average family income between 1986 and 2010.

Appendix Table B4
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)	(5)	(8)
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

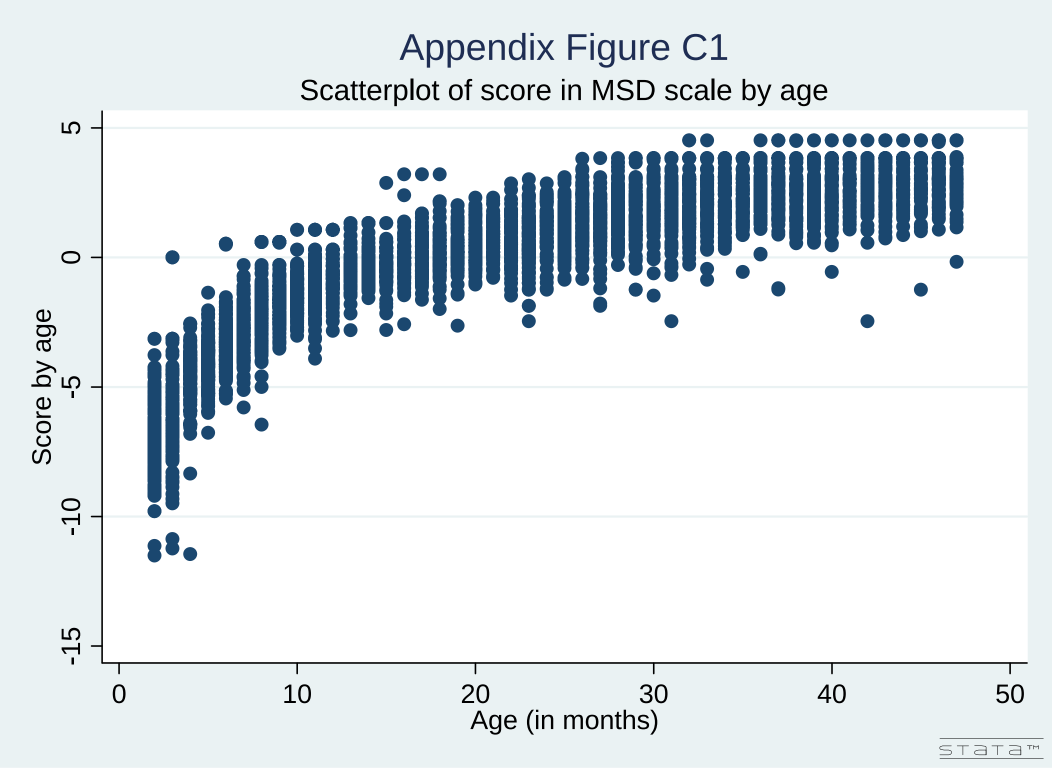
Robust standard errors in parentheses. All regressions have dummy variables for: (i) the child's gender, (ii) the child's birth order, (iii) child's year of birth, (iv) the child's age at the time of the measurement of the MSD score, and (v) maternal age at the time of the child's birth.

*** 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.

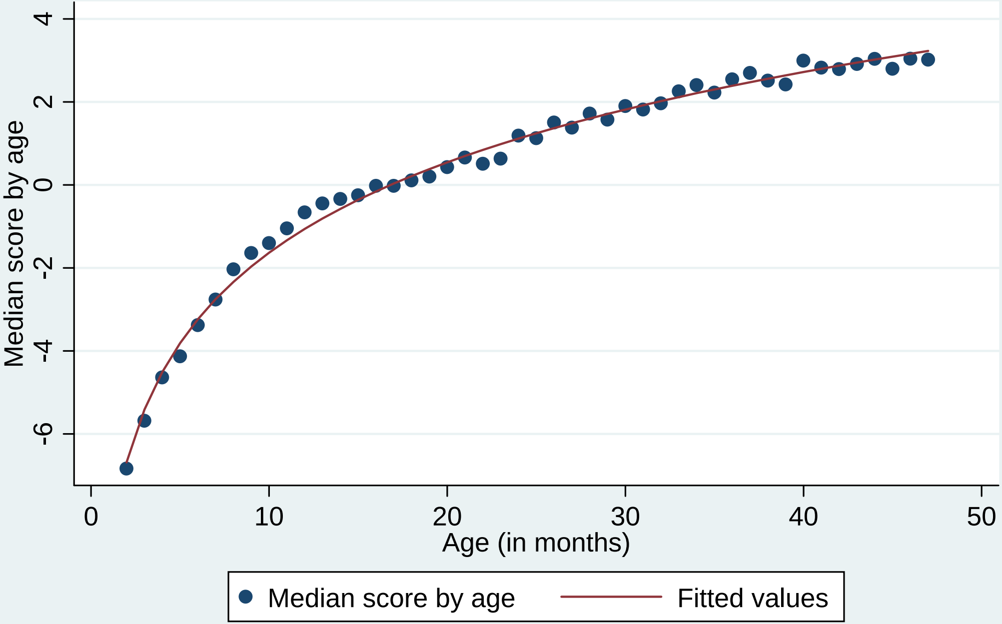
[2]Health at birth is captured by factor analyzing weight at birth, length at birth, gestational age, and number of days in hospital. The scale and location of the factor are determined by gestational age.

[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.


Appendix Figure C1
Scatterplot of score in MSD scale by age
Score by age
5
0
-5
-10
-15
0
10
20
30
40
50
Age (in months)
stata™

Appendix Figure C2

Median score in MSD Scale by Age

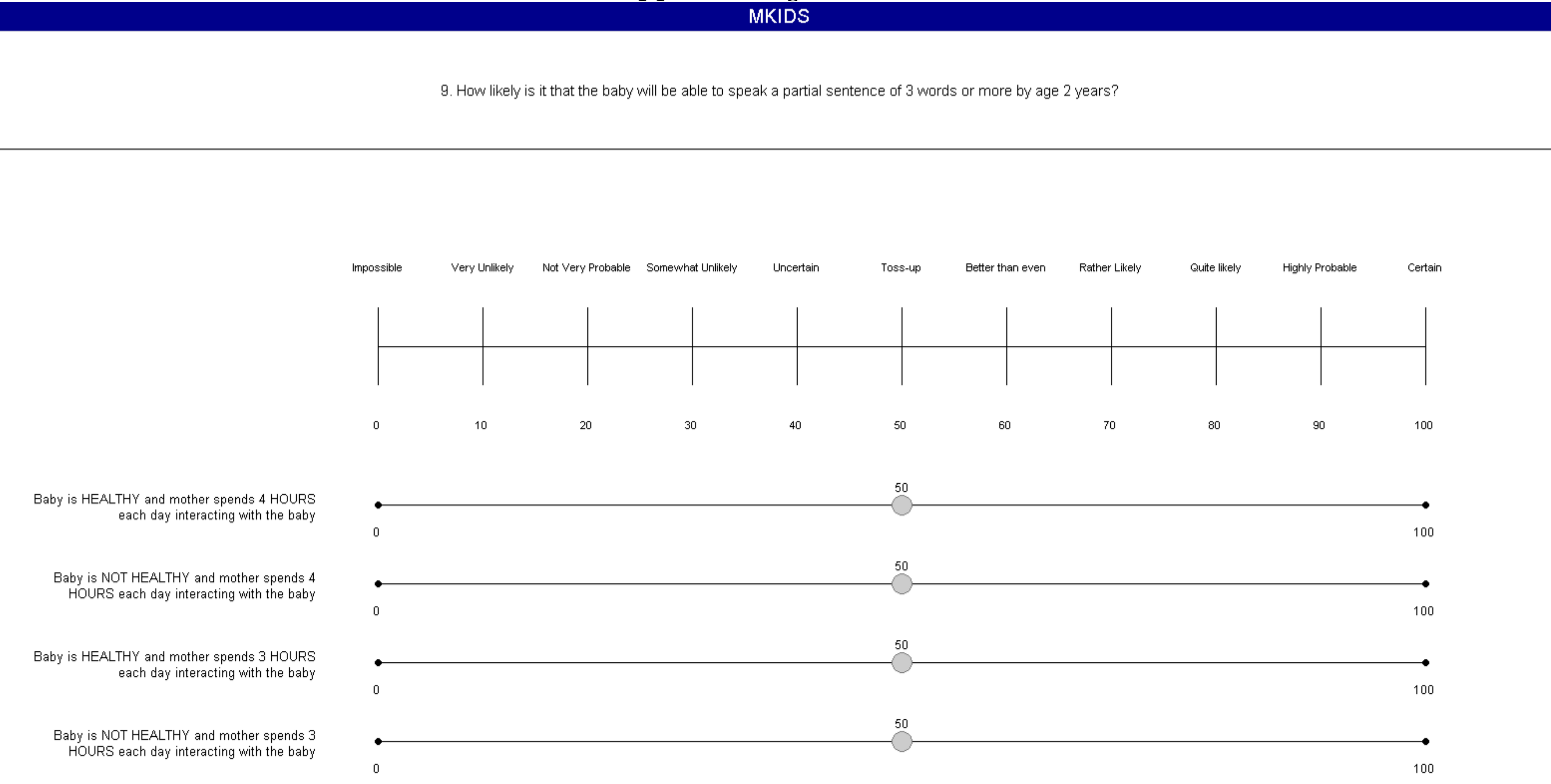


Appendix Table C1
Item Response Theory Analysis - Motor-Social Development Scale
National Health And Nutrition Examination Survey - 1988-1994 (NHANES)

Motor-Social Development Scale Item Description		Item Location Parameter	Item Discrimination Parameter
When lying on stomach on a flat surface, has child lifted head off the surface for a moment?	Mean	4.1648	0.2860
	Std Error	0.5797	0.0950
When lying on stomach, has child ever turned head from side to side?	Mean	4.1266	0.3000
	Std Error	1.1818	0.1724
Have child's eyes ever followed a moving object?	Mean	5.0000	0.3500
	Std Error	0.0000	0.0000
Have child's eyes ever followed a moving object all the one from one side to the other?	Mean	4.4841	0.4326
	Std Error	1.0742	0.1515
When lying on stomach, has child ever raised head and chest from surface while resting weight on lower arms or hands?	Mean	2.7998	0.3517
	Std Error	0.4542	0.0828
Has child ever turned head around to look at something?	Mean	4.5187	0.4515
	Std Error	1.0898	0.1544
While lying on back and being pulled up to a sitting position, has child ever held head stiffly so that it did not hang back as he/she was pulled up?	Mean	2.2546	0.2669
	Std Error	0.3699	0.0661
Has your child ever laughed out loud without being tickled or touched?	Mean	3.2142	0.3757
	Std Error	0.2666	0.0634
Has your child ever held in one hand a moderate size object such as a block or a rattle?	Mean	4.3798	0.5883
	Std Error	0.3061	0.0725
Has your child ever rolled over on his/her own on purpose?	Mean	4.1177	0.6528
	Std Error	0.2883	0.0888
Has your child ever looked around with eyes for a toy which was lost or not nearby?	Mean	3.4538	0.6233
	Std Error	0.2394	0.0832
Has child ever smiled at someone when that person talked to or smiled at (but did not touch) him or her?	Mean	3.0725	0.1825
	Std Error	0.1828	0.0421
Has your child ever seemed to enjoy looking in the mirror at himself/herself?	Mean	3.2592	0.4272
	Std Error	0.1596	0.0557
Has your child ever picked up small objects such as raisins or cookie crumbs using only thumb and first finger?	Mean	1.9412	0.5713
	Std Error	0.1162	0.0694
Has your child ever sat alone with no help except for leaning forward on his/her hands or with just a little help from someone else?	Mean	3.3750	0.7811
	Std Error	0.2010	0.0984
Has your child ever said recognizable words such as mama or dada?	Mean	1.9950	0.6228
	Std Error	0.1362	0.0792
Has your child ever shown that he/she knows the names of common objects when somebody names them out loud?	Mean	1.0182	0.3798
	Std Error	0.1081	0.0534
Has your child ever walked at least 2 steps with one hand held or holding on to something?	Mean	2.4166	1.0086
	Std Error	0.1635	0.1253
Has your child ever sat for 10 minutes without any support at all?	Mean	3.2201	0.9782
	Std Error	0.1635	0.1184
Has your child ever crawled when left lying on stomach?	Mean	3.1257	0.9306
	Std Error	0.1579	0.1123
Has your child ever been pulled from a sitting to standing position and supported own weight with legs stretched out?	Mean	2.7923	0.7174
	Std Error	0.1373	0.0877
Has your child ever waved good-bye without help from another person?	Mean	1.6965	0.7900
	Std Error	0.1014	0.0936
Has your child ever pulled himself/herself to a standing position without help from another person?	Mean	3.2154	1.1867
	Std Error	0.2024	0.1571
Has your child ever stood alone on his/her feet for 10 seconds or more without holding on to anything or another person?	Mean	2.6768	1.5675
	Std Error	0.2246	0.2017

Motor-Social Development Scale Item Description		Item Location Parameter	Item Discrimination Parameter
Has your child ever walked at least 2 steps without holding on to anything or another person?	Mean	2.2288	1.7460
	Std Error	0.2287	0.2364
Has your child ever shown that he/she wanted something by pointing, pulling, or making pleasant sounds rather than crying or whining?	Mean	1.6588	0.4191
	Std Error	0.0841	0.0599
Has your child ever said the name of a familiar object, such as a ball?	Mean	0.8479	0.7794
	Std Error	0.0910	0.1095
Has your child ever crawled up at least 2 stairs or steps?	Mean	1.7643	0.8350
	Std Error	0.1117	0.1159
Has your child ever said 2 recognizable words besides mama or dada?	Mean	1.1173	0.7985
	Std Error	0.0940	0.1106
Has your child ever let someone know, without crying, that wearing wet (soiled) pants or diapers bothers him or her?	Mean	0.3771	0.6227
	Std Error	0.0677	0.0853
Has your child ever walked up at least 2 stairs with one hand held or holding the railing?	Mean	1.0166	1.0534
	Std Error	0.1061	0.1355
Has your child ever run?	Mean	1.2824	1.6217
	Std Error	0.1687	0.2402
Has your child ever made a line with a crayon or pencil?	Mean	1.0777	0.6667
	Std Error	0.0701	0.0862
Has your child ever fed himself/herself with a spoon or fork without spilling much?	Mean	0.8683	0.7454
	Std Error	0.0742	0.0942
Has your child ever spoken a partial sentence of 3 words or more?	Mean	-0.1658	1.2167
	Std Error	0.1116	0.1471
Has your child ever said first and last name together without someone's help?	Mean	-2.1577	1.1441
	Std Error	0.1263	0.1373
Has your child ever walked upstairs by himself/herself without holding on to a rail?	Mean	-0.1634	0.6580
	Std Error	0.0725	0.0806
Has your child ever counted 3 objects correctly?	Mean	-1.2248	1.2132
	Std Error	0.1277	0.1491
Has your child ever walked up stairs by himself/herself with no help, stepping on each step with only one foot?	Mean	-0.4666	0.6267
	Std Error	0.0723	0.0768
Does your child know own age and sex?	Mean	-1.5018	0.9739
	Std Error	0.1060	0.1172
Has your child ever said the names of at least 4 colors?	Mean	-1.5706	0.9296
	Std Error	0.1160	0.1150
Has your child ever pedaled a tricycle at least 10 feet?	Mean	-0.9240	0.6395
	Std Error	0.0822	0.0796
Has your child ever done a somersault without help from anybody?	Mean	-0.2674	0.3339
	Std Error	0.0548	0.0438
Has your child ever washed and dried hands without any help except for turning the water on and off?	Mean	0.3236	0.5845
	Std Error	0.0826	0.0780
Has your child ever dressed himself or herself without any help except for tying shoes (and buttoning the backs of dresses)?	Mean	-0.6462	0.5930
	Std Error	0.0830	0.0751
Has your child ever gone to the toilet alone?	Mean	-0.5887	0.7900
	Std Error	0.1010	0.1005
Has your child ever counted out loud up to 10?	Mean	-2.0115	0.8207
	Std Error	0.1246	0.1032
Has your child ever drawn a picture of a man or woman with at least 2 parts of the body besides a head?	Mean	-3.0515	0.8036
	Std Error	0.1525	0.1031

Appendix Figure D1



Appendix Table D1

Average probability by age 24 months by MSD item and scenario[1]

	NHANES	Health condition at birth is "high"[2] Scenario 1 Investment is "high"[3]	Health condition at birth is "high"[2] Scenario 3 Investment in "low"[3]	Health condition at birth is "low"[2] Scenario 2 Investment is "high"[3]	Health condition at birth is "low"[2] Scenario 4 Investment in "low"[3]
MSD 30: Let someone know, without crying, that wearing wet or soiled pants or diapers bothers him or her?	92.52%	78.56%	70.11%	54.56%	51.40%
MSD 35: Speak a partial sentence of 3 words or more?	88.54%	81.66%	73.71%	58.79%	53.58%
MSD 37: Walk upstairs by himself/herself without holding on to a rail?	80.00%	75.90%	69.20%	51.39%	51.52%
MSD 44: Wash and dry his/her hands without any help except for turning the water on and off?	75.79%	77.37%	70.26%	58.71%	54.81%
MSD 39: Walk upstairs by himself/herself with no help, stepping on each step with only one foot?	62.11%	75.51%	71.49%	55.01%	51.87%
MSD 38: Count 3 objects correctly?	50.00%	80.47%	75.04%	60.22%	54.93%
MSD45: Dress himself/herself without any help except for tying shoes?	47.92%	73.52%	68.58%	54.15%	50.65%
MSD 46: Go to the toilet alone?	43.75%	78.69%	69.03%	56.79%	51.56%
MSD 40: Know his/her age and sex?	42.55%	82.03%	75.96%	58.83%	56.41%
MSD 43: Do a somersault without help from anybody?	42.11%	67.43%	61.93%	47.17%	44.07%
MSD 42: Pedal a tricycle at least 10 feet?	33.33%	72.34%	66.43%	53.65%	48.88%
MSD 41: Say the names of at least 4 colors?	32.98%	80.73%	74.08%	58.61%	55.64%
MSD 36: Say his/her first and last name together without someone's help?	26.04%	76.92%	69.62%	59.30%	54.48%
MSD 47: Count out loud up to 10?	13.54%	80.28%	75.02%	58.38%	53.49%
MSD 48: Draw a picture of a man or woman with at least 2 parts of the body besides a head?	2.08%	70.96%	67.52%	51.39%	48.27%

[1]MSD items are listed in descending order with respect to the probability that children age 24 months are able to execute the task.

[2]Heath condition at birth is "high" if the baby weighs 8 pounds at birth, is 20 inches long at birth, and the gestational age is 9 months. It is "low" if the baby weighs 5 pounds at birth, is 18 inches long at birth, and the gestational age is 7 months.

[3]Investment is "high" when mother spends 6 hours/day interacting with the child and is "low" when the mother spends 2 hours/day.