Subjective Beliefs, Uncertainty, and Costs of Investment in Early Childhood

[JOB MARKET PAPER]

[CLICK HERE FOR LATEST VERSION]

Marcos Lee

Rice University

November 11, 2023

I study how parents perceive the uncertainty of the returns to investment of the child development process and how this uncertainty impacts their actual investment in children. To do so, I develop an elicitation procedure of parental subjective belief distributions about the technology of skill production and their subjective investment costs that is guided by a model of parental investment. I collect data using this procedure and extend existing measurement error correction methods that are necessary in the belief estimation. I show that parents hold downward-biased mean beliefs about returns to investment. Moreover, parents who have higher mean beliefs also hold lower levels of uncertainty about their beliefs. Both mean beliefs and uncertainty correlate with actual investment measures. Finally, I estimate a model of parental investment with reference-dependent preferences and show that even though parents hold low mean beliefs, they have a strong incentive to invest if their child is at risk of having a developmental delay.

Email: marcos.lee@rice.edu. I am very grateful to Flavio Cunha for his support at every stage of this project, and to Rossella Calvi, Isabelle Perrigne, Xun Tang, and participants at the Rice and Bocconi University Brown Bags for the many helpful comments and discussions. This research has been funded by the Rice Social Sciences Dissertation Grant, and has obtained ethical approval from the Rice IRB, protocol IRB-FY2022-269.

1. Introduction

Research by developmental psychologists, sociologists and economists has shown that the development of cognitive and non-cognitive skills throughout early childhood is one of the most important processes in our lives. It impacts not only education attainment and socio-emotional behavior throughout adolescence, but also has long-term consequences for lifetime earnings. One of the key drivers of this process is parental investments. Therefore, understanding why some parents invest in their children more than others is a fundamental question to investigate policies targeting lifetime inequality. Investment gaps that happen early in the childhood can lead to skill gaps in the future that are difficult to remedy (Cunha, Heckman, and Schennach 2010; Heckman et al. 2010).

To help understand how these investment gaps arise, it is crucial to uncover the determinants of parental investment. Past research has highlighted the role of family resources (Dahl and Lochner 2012; Caucutt et al. 2020) and parental characteristics such as maternal education and cognitive skills (Currie and Moretti 2003; Aizer and Stroud 2010; Arendt, Christensen, and Hjorth-Trolle 2021). Recent papers in economics highlight the role of parental information about the process of child development (Cunha, Elo, and Culhane 2022; Boneva and Rauh 2018; Attanasio, Cunha, and Jervis 2019; Dizon-Ross 2019). While this body of work show that low-income parents often underestimate the returns on investment, and that this is correlated with lower investment, it does not consider the uncertainty parents have about their beliefs.

This paper aims to address this gap by investigating how belief uncertainty about the technology of skill formation affect the parental investment decisions in their children. I develop a new methodology that can recover parental belief distributions about the returns to investment in children, including their uncertainty. The methodology and the identification strategy is carefully guided by a theoretical model of parental investment. I find that a one standard deviation increase in mean beliefs predicts a 20% standard deviation increase in daily investment hours, while the same increase in uncertainty predicts a –13% standard deviation decrease in investment. Finally, I estimate a model of parental investment with reference dependent preferences and subjective beliefs, and show that even though parents hold low mean beliefs, they have a strong incentive to invest if their child is at risk of being at a developmental delay.

Investing in your child is an uncertain process that requires consistent resource commitment for several years with uncertain outcomes. Understanding how belief uncertainty about the returns to investment impacts investment in early childhood is important for many reasons. First, it remains unclear in which direction belief uncertainty affects investment. On one hand, the literature on schooling decisions for middle and high school students show that an increased perceived earnings risk or unemployment risk cause parents and children to invest less (Attanasio and Kaufmann 2014; Wiswall and Zafar 2015). On the other hand, in a model where agents are uncertain about the returns to investment and there is learning, prior beliefs with large uncertainty (ie., high variance) means that there is still a lot to learn about the true nature of returns to investment (Mira 2007; Badev and Cunha 2012). This paper introduces this discussion into the early childhood environment, but I do not impose any structure on whether the risk aversion or the learning aspect dominates. Instead, I examine the first order correlations between mean and uncertainty of beliefs to tease out this relationship.

Second, if uncertainty is an important factor in determining parental investment, it is essential to consider this factor when designing policies that target parental beliefs. There have been many successful parenting interventions which increase investment that focuses in educating parents about the significance of spending quality time with their children, whether through activities such as reading, playing, or engaging in conversations (see, the Jamaica Home Visiting Program and the Nurse-Family Partnership Program). Nevertheless, recent interventions capable of measuring parental beliefs have had mixed results. Some demonstrate that while parents' subjective mean returns to investment increase, their actual investments do not change, while others see an increase in investment but no change in mean beliefs (Attanasio, Cunha, and Jervis 2019; List, Pernaudet, and Suskind 2021).[1] These mixed results indicate that there may be other factors at play. One plausible explanation is that these interventions may be changing parental uncertainty in non-trivial ways. Only by developing the methodology to measure parental uncertainty can we understand how uncertainty is affected by these interventions.

To study parental beliefs, I cannot rely on existing data since it is not possible to estimate subjective preferences or beliefs from observed data alone (Manski 2004). Therefore, I design a survey to elicit the subjective belief distribution and investment decisions from parents. The survey is guided by a theoretical model of parental invest-

[1] In particular, List, Pernaudet, and Suskind (2021) designed two interventions, one involving parents watching a brief video on the importance of investments, and the other featuring an intensive home visitation instruction and feedback program. Their findings revealed that while parental beliefs changed in both interventions, only the latter led to increased investment. On the other hand, Attanasio, Cunha, and Jervis (2019) designs an information and education intervention, but while investments changed after the program, their measured mean beliefs did not change.

ment under subjective beliefs of the technology of skill production. For a sample of parents, I present two sets of experiments where parents are presented with specific hypothetical scenarios about the environment faced by a hypothetical child. I then estimate beliefs, costs, and preferences while correcting for measurement error.

In the first experiment, parents are presented with a hypothetical scenario of either a baby with normal or poor health and high or low investment. They are then asked to choose what are the youngest, oldest, and most likely ages (in months) that a child, under that specific scenario, will be able to perform some specific activities. Given the exogenous variation of variables in the scenarios, this age-range question allows me to obtain a measure of parental belief about the skill a child will have in the future in terms of age in months. This is a key aspect, since skill is a variable without a scale. By phrasing the questions in terms of ages, it is possible to measure a child's skill as the developmental delay (or advance) the parent believes a child would develop given the health and investment scenario. With this information and under some statistical assumptions about the beliefs held by parents, I estimate the mean belief and uncertainty about the return to investment. This methodology also deals with the problem of measurement error which is common in this type of elicitation surveys.

The second experiment's objective is to elicit the subjective cost of investment from parents. I do not model parental investment as a monetary expenditure, but define investment as daily hours in active interaction between parent and child. Research has shown that during the early childhood stage of children's lives monetary investments are not as important as active engagement and interaction between parents and children (Carneiro and Ginja 2016; Fiorini and Keane 2014; Bono et al. 2016; Carneiro and Rodrigues 2009; Kalil, Ryan, and Corey 2012).[2] Given this definition, one must be careful in defining the price of one unit of investment. I define the price of investment as the amount of leisure time the parent is willing to give up to spend one hour in active interaction with their child. Parents are presented with a hypothetical scenario of initial child's skill, income level, and working hours of a hypothetical family.[3] They are asked to answer a series of willingness-to-pay questions about trading one hour with their child and one hour of leisure away from their child.

I collect data through Qualtrics from a sample of women between the ages of 18 and

[2]It is important to note that parents spend the majority of their time with children in non-active childcare, where the child is not the main focus of the activity (Kalil et al. 2023). In my elicitation design, I stress to respondents this difference and ask them to only consider their time in "active interaction".

[3]These variables are chosen due to the specific class of parental investment models that I use to motivate the problem.

40, where the women have at least one child and the oldest child could not be older than 5 years old. This particular is chosen because one would like for the respondents to have some degree of knowledge about child development. Individuals that are too far removed from thinking about the child development process may be more likely to give random answers since they have never thought too much about this topic. I first document several regularities and patterns predicted by the model. Respondents report that they believe children complete harder activities at higher ages. Moreover, respondents report higher ages under scenarios where the child have poor health and low investment, demonstrating that they understand the tradeoffs in child development. These patterns are consistent with the model and provide evidence that respondents are answering the questions in a meaningful way.

I find that mothers in this sample have low mean beliefs, which is consistent with past research (e.g., Cunha, Elo, and Culhane (2013); Boneva and Rauh (2018)). Mothers believe that a 10% increase in investment leads to a 1.01% increase in their child's human capital, which is about half the returns when estimating using objective data (see, Attanasio, Cunha, and Jervis 2019). I find that uncertainty is low across the sample, with a coefficient of variation close equal to 0.7. Moreover, parents with higher mean beliefs hold lower uncertainty. One standard deviation increase in mean beliefs predict a –28% decrease in uncertainty. This suggests that individuals who are more pessimistic about the returns to investment are also more uncertain about their beliefs. I then regress measures of actual time investments from mothers on their estimated beliefs. Similar to other work, I find that higher mean beliefs predict higher actual investment. However, I uncover that higher uncertainty is correlated with lower actual investment. A one standard deviation increase in mean beliefs predicts a 23% standard deviation increase in daily investment hours, while the same increase in uncertainty predicts a –14% standard deviation decrease in investment.

The combination of data from the belief elicitation and stated choice experiments allow me to estimate a model of parental investment in children that takes into account subjective beliefs and costs. I incorporate reference-dependent preferences, in which parental investment depends on how they view their child development relative to specific developmental milestones. Crucially, parents use their own subjective beliefs to compare the perceived development. Under this model, I can study how parents value avoiding developmental delays. I find that parents strongly value their child skill relative to leisure and household consumption. Moreover, when estimating the preference for large developmental delays, I find that parents have a strong incentive to invest if their child is at risk of being at a developmental delay. Overall, even though parents have low beliefs about the returns to investment, they still value their child skill and have a strong incentive to invest if their child is at risk of being at a developmental delay.

Counterfactual analysis shows that both increasing mean beliefs and uncertainty about the returns to investment also increases investment. This seemingly contradicts the reduced form evidence, which shows that higher uncertainty is correlated with lower investment. However, it is important to consider that the reduced form evidence is not a causal statement. Indeed, given that individuals with higher mean beliefs are also ones with lower uncertainty, it could be that the correlational regressions are simply reproducing the effect of mean beliefs.

When breaking down the effect of increasing uncertainty on investment, I show that there is substantial heterogeneity in investment changes. The individuals that increase their investment due to the increase in uncertainty are the ones that hold very low mean beliefs, have higher opportunity costs of investment, and therefore do not invest much in their child. On the other hand, those with high mean beliefs, low costs, and high baseline investment reduce their investment due to the increase in uncertainty.

This paper contributes to different strands of literature. First, it is part of a large literature that highlights the importance of eliciting subjective expectations in economics and their uses in estimating economic models (Manski 1993, 2004; Delavande 2008; Jensen 2010; Zafar 2013; Almås, Attanasio, and Jervis 2023).

Second, it is closely related to papers that highlight the role of parental information about the process of child development in early childhood (Cunha, Elo, and Culhane 2013, 2022; Boneva and Rauh 2018; Attanasio, Cunha, and Jervis 2019; Dizon-Ross 2019; List, Pernaudet, and Suskind 2021).[4] This paper contributes to this literature by eliciting not only parental expectations but also their subjective uncertainty about the technology of skill formation. To the best of my knowledge, I am the first to explicitly design an elicitation procedure of subjective uncertainty with respect to the returns to investment in children.

Third, this paper also contributes to the literature on uncertainty and risk aversion in educational investments (Giustinelli 2016; Attanasio and Kaufmann 2014; Carneiro and Ginja 2016; Tabetando 2019; Sovero 2018; Tanaka and Yamano 2015; Basu and Dimova 2022). These papers adopt different approaches and find different conclusions. For example, Tabetando (2019) finds that for a sample of parents in Uganda, poorer

[4] There is an extensive literature on examining how beliefs impact educational decisions in adolescence in the context of high school, university, and major choice (Delavande and Zafar 2018; Patnaik et al. 2022; Wiswall and Zafar 2015; Stinebrickner and Stinebrickner 2014).

households are more risk averse and invest less in schooling compared to others. The authors suggests this is due to lack of access to credit. Sovero (2018) uses Mexican survey data and find that more risk averse mothers not only spend more on their son's schooling but also that their sons have higher weight, an indication of higher parental investment. In a closely related study, Conti, Giannola, and Toppeta (2022) examine how parents choose to allocate their time across different types of investment when there are varying health risks associated with each option. I focus on the uncertainty parents have about their own beliefs and how it impacts their investment.

Fourth, I contribute to the literature that incorporates reference-dependent preferences in the context of parental investment in children (Wang et al. 2022; Kinsler and Pavan 2021). I differ from the previous literature in that I allow parents to have subjective beliefs about the returns to investment. On the other hand, I do not consider the role of peers as a reference point, but instead use developmental milestones as a reference point. I illustrate how this model can be used to study how parents value avoiding developmental delays.

Finally, I also contribute to the literature that studies the importance of time investment in children and their determinants, especially in early childhood (Kalil et al. 2023; Folbre et al. 2005; Kalil, Ryan, and Corey 2012; Price 2008; Bono et al. 2016; Guryan, Hurst, and Kearney 2008; Schoonbroodt 2018; Conti, Giannola, and Toppeta 2022; Del Boca, Flinn, and Wiswall 2014). I model investment in early childhood as consisting of hours of active interaction spent between parents and child, and estimate the implicit cost of time for parents. Previous research either estimates this opportunity cost using observational data, which implicitly assumes a uniform cost to all parents, or they force a specific cost structure to parents[5]. My findings show that working and educated mothers have larger implicit costs of time, and that mothers with higher implicit costs of time also invest less in their children. While I do not establish a causal relation between these findings, they contradict previous findings that labor supply and education are positively correlated with investment in children. Therefore, it is important to consider this channel more carefully.

The remainder of this paper is structured as follows. Section 2 describes the economic model and specifies the technology of skill formation, as well as the concepts of subjective expectation and uncertainty. Section 3 describes the survey instrument, how the collected data identify the model primitives and the estimation method. Section 4

[5]For example, Cunha, Elo, and Culhane (2013) asks parents to think that 1 hour of daily investment costs $15 dollars.

describes the data. Section 5 discusses the results, while Section 6 concludes.

2. Model

Consider a mother who must decide how much to invest in her child within a one-period model.[6] Let y_i denote her income, x_i denote her time investment in her child, and c_i denote her consumption. She faces a budget constraint given by

$$c_i + p_i x_i = y_i, \tag{1}$$

where p_i is the relative price of parental investment. Moreover, she faces a time constraint where time investment x_i is bounded by the total time available:

$$x_i + h_{l_i} + h_{w_i} = 16. \tag{2}$$

Thus, the mother has 16 hours per day, assuming she sleeps 8 hours, to allocate between her child (x_i), leisure (h_{l_i}), and work (h_{w_i}). I assume that work is not a choice variable but given exogenously.

Let $\theta_{i,0}$ and $\theta_{i,1}$ denote the stock of human capital of child i at birth and at 24 months. Let x_i denote the investment in human capital made by the mother between birth and $t = 1$. Finally, let ξ_i denote a shock to the development process unknown to the parents. I assume that the technology of skill formation is given by a Cobb-Douglas formulation:

$$\ln \theta_{i,1} = \delta_0 + \delta_1 \ln \theta_{i,0} + \delta_2 \ln x_i + \xi_i. \tag{3}$$

The above equation describes the *objective* process of skill formation.

Parental preferences depend on household consumption, leisure, child development at the end of the period, $u_i(c_i, h_{l_i}, \theta_{i,1}; \boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ denotes the vector of parental preferences. The parents maximize their expected utility conditional on the agent's information set Ω_i at the time of the decision. Parents know their own preferences $\boldsymbol{\alpha}$, their income y_i, the price of investment goods p_i, their work hours h_{w_i}, and their child initial stock of human capital, $\theta_{i,0}$. However, they do not know the parameters of the objective technology function, $\delta_0, \delta_1, \delta_2$. Parents have beliefs about these parameters, and I denote this belief distribution by $G_i(\cdot)$. Therefore, the information set of the parents is the set $\Omega_i = \{\boldsymbol{\alpha}, p_i, y_i, h_{w_i}, \theta_{i,0}, G_i(\cdot)\}$.

[6] I will use "parent" and "mother" interchangeably, as this is a single-agent model.

The distribution $G_i(\cdot)$ has a mean equal to $E[\ln \theta_{i,1}|\Omega_i]$ and a variance equal to $Var(\ln \theta_{i,1}|\Omega_i)$. We can write the *subjective* maternal expectation and uncertainty about the skill of technology formation as

$$\begin{aligned} (4) \qquad & \mu_{i,\theta_1} \equiv E[\ln \theta_{i,1}|\Omega_i] = \mu_{i,\delta_0} + \mu_{i,\delta_1} \ln \theta_{i,0} + \mu_{i,\delta_2} \ln x_i, \\ & \sigma^2_{i,\theta_1} \equiv Var(\ln \theta_{i,1}|\Omega_i) = \sigma^2_{i,0} + \sigma^2_{i,\delta_1} \ln \theta^2_{i,0} + \sigma^2_{i,\delta_2} \ln x^2_{i,0} + \sigma_{i,\delta_1,\delta_2} \ln \theta_{i,0} \ln x_i, \end{aligned}$$

where $\mu_{i,\delta_k} = E[\delta_k|\Omega_i]$, $\sigma^2_{i,\delta_k} = Var(\delta_{i,k}|\Omega_i)$, and $\sigma^2_{i,0} = Var(\delta_{i,0} + \varepsilon_i|\Omega_i)$. These definitions derive directly form using the expectation and variance operators on (3). However, the subjective uncertainty assumes that: (i) the production shock ξ_i is uncorrelated with all other variables; and (ii) the covariances between δ_1, δ_0 and δ_2, δ_0 are equal to zero. I define μ_{i,δ_2} as the parental subjective expectation of returns to investment. Note that $\sigma^2_{i,0}$ contains the variation due to the natural heterogeneity of the child development process. Some children may develop faster or slower than others independently of the inputs, and parents have a belief related to this. I focus on σ^2_{i,δ_2}, which I define as the uncertainty parents have about the returns to investment.

Parents maximize their expected utility function conditional on their information set:

$$(5) \qquad \max_{x_i} \{E[u_i(c_i, h_{l_i}, \theta_{i,1}; \boldsymbol{\alpha})|\Omega_i]\}$$

subject to budget constraint (1), the time constraint (2), and technology of skill formation (3). The optimal investment function depends on parental preferences, their income, the price of investment, their working hours, and on their subjective expectation and uncertainty of the returns to investment, μ_{i,δ_2} and σ^2_{i,δ_2}:

$$x^*_i = f(\boldsymbol{\alpha}, y_i, p_i, h_{w_i}, \mu_{i,\delta_2}, \sigma^2_{i,\delta_2}, \sigma_{i,\delta_1,\delta_2}).$$

Previous papers consider models that only depended on $\boldsymbol{\alpha}$, y_i, p_i and δ_2, which imply that parents have complete knowledge about the technology of skill formation (e.g., Del Boca, Flinn, and Wiswall (2014)). Without additional information about subjective beliefs, it is not possible to estimate a model that does not assume complete knowledge. Cunha, Elo, and Culhane (2013) and Attanasio, Cunha, and Jervis (2019) assume a Cobb-Douglas utility function on consumption and child's skill which by construction removes the uncertainty of beliefs from the optimal investment function. This formulation of a parental investment model which incorporates subjective beliefs about the technology of skill production generalizes models in past research.

3. Experimental Design

In this section, I present the survey instruments, how they are used to identify model parameters, and the estimation procedures. I first describe the process of elicitation of the subjective belief distribution of parents. Then, i Finally, I describe the elicitation of the subjective expectation and uncertainty of the technology of skill formation from parents.

3.1. Elicitation of Subjective Beliefs

My objective is to elicit from parents their subjective expectation and uncertainty about the technology of skill formation function. The model described in Section 2 includes latent variables such as the child's skill. It is necessary to develop a mapping that can translate a latent variable with no meaningful metric, i.e., the child's skill θ, to an observable and easily interpretable cardinal metric that is meaningful with respect to child development. Similarly, we need an observable metric that can map the latent skill to the maternal belief of child development. Therefore, we can communicate with parents using these observable variables and in turn use them to estimate our model.

Before going into detail on how I elicit the subjective belief distribution of parents, it is useful to explain how to measure the *objective* development of a child in a cardinal metric. The most common measure of early childhood development is through developmental milestone assessments by caregivers or clinics. Examples include the *Bayley Scales of Infant and Toddler Development*, the *Motor and Social Development Scale*, and the *Ages and Stages Questionnaires*. They all involve evaluating whether a child at a specific age can perform activities or have achieved milestones that are appropriate for their age. They provide standardized information on children's development across multiple domains, including motor, language, and cognitive development.

I follow Cunha, Elo, and Culhane (2022) to develop the belief elicitation module. I give a brief overview of their methodology. They use the Motor and Social Development (MSD) scale from the National Health and Nutrition Examination Study 1988 (NHANES). The MSD asks mothers to answer 15 out of 48 questions regarding the motor, language, and numeracy development of their child, conditional on their age. For example, mothers with children ages 0 to 3 months answer different questions than mothers with children ages 22 to 47 months. Each question asks whether their child can perform a specific activity appropriate for their age. Moreover, the number of items a child is able to perform increases with their age, which is helpful in anchoring the latent child development.

They estimate an item response theory (IRT) model using the MSD instrument to fix the cardinal metric of development into developmental age, or "age-equivalent score". The IRT model has many uses. First, it reduces the dimensionality of the MSD module by estimating which items are most salient with respect to child development. Second, it creates a mapping between an observable variable, the MSD items, to the latent variable θ in a cardinal metric. Third, as shown later, it can be used to translate parent's answers to the belief instrument into the latent variable θ.

Let a_i denote the child i's age when their MSD answers are measured. Let κ_i denote child i's development relative to other children at the same age. Then, $\kappa = 0$ means that the child's development is typical for their age, $\kappa_i > 0$ means that they are advanced for their age, and $\kappa < 0$ means that they are delayed for their age. Following Cunha, Elo, and Culhane (2022), I estimate the following IRT model:

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

where $d^*_{i,j}$ is the unobserved latent variable related to individual i and MSD item j. We observe $d_{i,j}$, which is a binary variable equal to 1 if $d^*_{i,j} \geq 0$ and 0 otherwise. The parameter $b_{j,0}$ decreases with the difficulty of item j, while parameter $b_{j,1}$ increases as item j's difficulty decreases with age. Finally, parameter $b_{j,2}$ increases the more information item j contains about θ_i.

Assume that $\eta_{i,j}$ is i.i.d. normally distributed $N(0,1)$, and that θ follows a mixture of two Normal distributions. Then, normalize $b_{2,j} = 1$ for one of the MSD items and the mean of θ_i to be zero. The estimation is done via maximum likelihood.

From the estimated parameters of Equation 6, MSD items that are salient and relevant to the child development process are selected. Additionally, these items are used to develop the belief instrument survey. Finally, given the answers from respondents, Equation 6 is used to translate parental beliefs into the estimated cardinal metric obtained from the IRT.

3.1.1. Subjective Belief Instrument

I now describe the process to elicit the subjective expectation and uncertainty of the technology of skill formation from parents. The main idea is to ask parents their beliefs about the typical ages a child is able to perform certain activities under a specific level of initial human capital and parental investment. These activities are milestones that are used in the MSD-NHANES to assess developmental progress in children.

The general idea of the instrument is to tell the respondent to imagine a scenario of initial human capital of the baby θ_0 and a level of investment x.[7] Then, I ask the respondent to answer what they think is the minimum, most likely, and maximum age a baby under this situation would learn how to perform a specific activity from the MSD. The respondent will see different combinations of initial human capital and investment and for each one of the chosen MSD activities.

Ideally I would ask parents about all milestones for 24 month-old children, but to avoid respondent fatigue I use the 4 most salient ones as estimated from the IRT model in 6, namely[8]

1. Speak a partial sentence of 3 words;
2. Count 3 objects correctly;
3. Say first and last name together;
4. Know age and sex.

I first describe to the respondent what does it mean for a baby to have normal or poor health in the context of this survey. A normal health baby is one whose gestation lasted 9 months and that weighed 8 pounds and measured 20 inches in length. A poor health baby is one whose gestation lasted 7 months, weighed 5 pounds and measured 18 inches.[9]

[7]The full instrument can be seen in Appendix A.

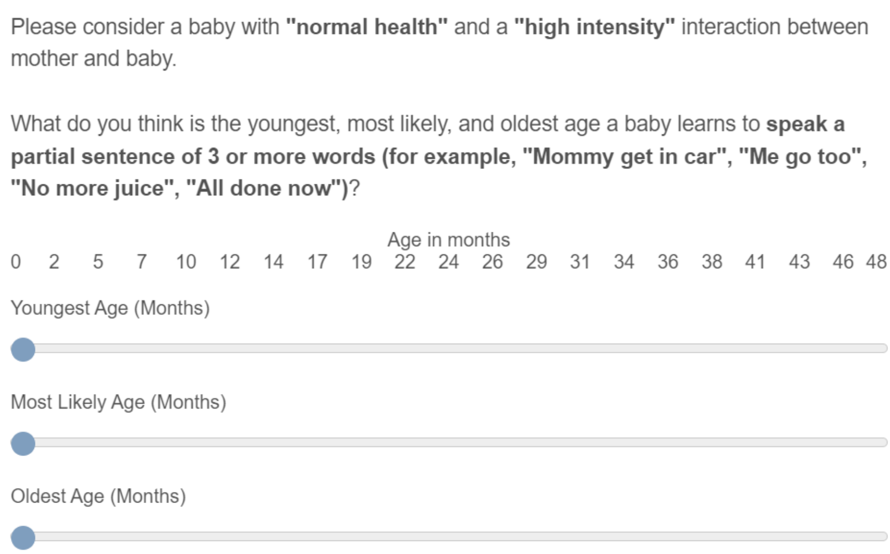
[8]The parameters of the IRT model give estimates of the importance and relevance of each item in explaining the variability of the latent variable. By examining which items have the larger factor loadings estimates, the researcher can choose items with more explaining power. See Cunha, Elo, and Culhane (2022) for a detailed description of how these items were chosen.

[9]These numbers are the same used in Cunha, Elo, and Culhane (2022), and were obtained from the Children of the National Longitudinal Survey of the Youth/1979 (CNLSY). They estimate a factor model using gestation, weight at birth, and height at birth as measures. The low scenario describes a premature birth: gestation lasts seven months (percentile 1 in the CNLSY/79), the birth weight is five pounds (percentile 4), and the length at birth is 18 inches (percentile 11). The normal scenario describes children born in a normal term. The gestation lasts nine months (percentile 85), weighs eight pounds (percentile 69), and the length at birth is 20 inches (percentile 85). Given factor score estimates and predicted scores, they compute the implied value of the latent variable θ_0 under a cardinal scale.

Then, I describe two types of interactions that a parent has with their baby. The first one is called active interaction, where the main and only focus of the parent is in their baby. Examples of activities during this time are: (a) soothing the baby, (b) playing or singing songs to the baby, (c) feeding, nursing, bathing, attending to health needs, among others. I contrast this with passive interaction or leisure, where the parent can still be with the baby but the baby is not the main focus. For example, going grocery shopping with the baby, browsing social media, household chores, etc. I then define a high intensity interaction is one in which the parent spends 6 hours in active time, while a low intensity is one where the parent spends 2 hours.[10]

A scenario is a pair of either normal $\bar{\theta}_0$ or poor health $\underline{\theta}_0$, and high $\bar{x}$ or low $\underline{x}$ intensity interaction. Given a scenario (θ_0, x), the respondent is asked what they think is minimum, most likely, and maximum age in months that a baby will learn how to do activity j. They are presented with 3 sliders, one for each answer, that ranges from 0 to 48 months. Figure 1 below depicts the instrument as shown to the respondents.

FIGURE 1. Subjective Belief Instrument



3.1.2. Age range to belief distribution

From the subjective belief instrument, I obtain in months the minimum age $\underline{a}_{i,j,k}$, most likely age $\dot{a}_{i,j,k}$, and oldest age $\bar{a}_{i,j,k}$ a respondent i believes a child under scenario k

[10] These numbers are the 25th and 90th percentile of the distribution of time spent with their child in the Panel Study of Income Dynamics time diaries. Differently from the initial health data, hours of investment is already a cardinal measure, so there is no need in transforming them.

will learn how to do MSD item j. I assume that these replies are a random variable that follows a distribution $H_{a_{i,j,k}}(\cdot)$ with mean $\mu_{a_{i,j,k}}$ and variance $\sigma^2_{a_{i,j,k}}$. The mean $\mu_{a_{i,j,k}}$ represents the average age individual i believes a child will learn MSD item j under scenario k, while the variance $\sigma^2_{a_{i,j,k}}$ represents the degree of uncertainty from the individual about j under k. From here on, I omit the subscripts i, j, k for conciseness, and only include when necessary to avoid ambiguity.

I need to transform the age range responses distribution, $H_{a_i}(\cdot)$ to the human capital subjective belief distribution, $G_{i,\ln\theta_1}(\cdot)$, which has mean $E[\ln\theta_{i,1}|\Omega_i]$ and variance $Var(\ln\theta_{i,1}|\Omega_i)$. Intuitively, for the mean I use the average age a child learns MSD item j in the population and compare it to the mean age individual i believes, implied by H_{a_i}, and rescale it to the human capital metric. For the variance, I do not use any population information since someone's belief is not related to the population heterogeneity in completing MSD items. In turn, I preserve the spread and shape implied by H_{a_i} to construct $G_{i,\ln\theta_1}(\cdot)$. I now describe the process in detail.

The mean μ_a does not yet represent the belief of child development $E[\ln\theta_1|\Omega_i]$. Each MSD item have different difficulties and in the population most children learn to count 3 objects at a different age than most children learn to say their first and last name together. Therefore, we must take into account that parents giving the same age ranges about these items imply different estimates of $E[\ln\theta_1|\Omega_i]$.

Let $\bar{\theta}_j$ denote the average age a child learns MSD item j. I would like to transform μ_a into a measure of child development in the same metric as the IRT equation, that is, the developmental delay (or lead) in months relative to a target age: $\ln a + \delta_j$. Conceptually, this means comparing μ_a to the average age children learn MSD item j in the population and compute $\delta_j = \mu_a - \bar{\theta}_j$.

I construct $\bar{\theta}_j$ from the IRT model in equation (6). I do so because it corrects for the associated measurement error that is present in the MSD instrument used in NHANES. From (6), I know that:

$$Prob(d_{i,j} = 1|\ln\theta_{i,1}) = \Phi(b_{j,0} + b_{j,1}\Big(\ln a_i + \frac{b_{j,2}}{b_{j,1}}\xi_i\Big)) = \Phi(b_{j,0} + b_{j,1}\ln\theta_{i,1}),$$

where $\Phi(\cdot)$ denotes the cdf of $N(0,1)$. I compute $\bar{\theta}_j$ by backing out $\ln\bar{\theta}_j$ implied by when $Prob(d_{i,j} = 1) = 0.5$, or the mean age that children learn j[11]. Then, I compute the parental belief of developmental delay $\delta_{i,j,k} = \mu_{a_{i,j,k}} - \theta_j$. Finally, I choose the target age to be 24 months. Therefore, the measure of subjective expectation for parent i, MSD item j,

[11]This probability gives the mean because of the normality assumption

under scenario k is given by

$$\ln \theta_{i,j,k,1} = \ln(24 - \delta_{i,j,k}). \tag{7}$$

In the case of the mean belief, I use the population distribution (through the IRT model) to change the location of the age completion distribution $H_a(\cdot)$ to reflect how different MSD items have different mean completion ages in the population. However, in the case of the variance of the belief distribution, I do not need to use the population distribution since the variance of the population distribution reflects the population heterogeneity in completing specific activities, which has no relation to the uncertainty from parents.

The variance $Var(\ln \theta_1|\Omega_i)$ is composed of two terms: (i) the subjective uncertainty about the nature of the returns to investment and to initial skill; (ii) and the belief about heterogeneity of the child development process $Var(\varepsilon_i|\Omega_i)$. Similarly, the age range distribution G_{a_i} contains information about both components. Consider two parents $i = \{A, B\}$ that give the same answer for $\dot{a}_{i,j,k}$, the most likely age for MSD item j, but parent A believes that the minimum and maximum ages are 20 and 30, while B believes it is 18 and 32. The implied variance $\sigma^2_{a_{i,j,k}}$ for parent A will be smaller than for parent B, but it does not necessarily mean parent B is more uncertain than parent A. Parent B may believe that MSD item j is harder, and therefore there is more heterogeneity in the population about which age children are able to accomplish it. This heterogeneity is captured by the constant in equation (4), which includes $Var(\varepsilon_i|\Omega_i)$.

I use the Interquartile Range of the $H_a(\cdot)$ distribution to maintain the shape and dispersion from $H_a(\cdot)$ to $G_\theta(\cdot)$.[12] The Interquartile Range (IQR) is the difference between the 75th and 25th percentiles of the distribution. The IQR of a distribution H is defined as:

$$IQR_H = H^{-1}(0.75) - H^{-1}(0.25).$$

Then, to obtain the estimate $\sigma^2_{\theta_{i,1}}$ of $Var(\ln \theta_1|H_i)$, I compute the value of IQR for the distribution H_{a_i} and impose the IQR of distribution $G_{\theta_{i,1}}(\cdot)$ to be the same. Then, I can compute the variance of $G_{\ln \theta_{i,1}}(\cdot)$ implied by the estimate $\mu_{\theta_{i,1}}$ and $IQR_{G_{\theta_1}} \equiv IQR_{H_{a_i}}$.

[12]The use of the Interquartile Range as a measure of dispersion of beliefs is common in the literature of household surveys. See Bruine de Bruin et al. (2023).

3.2. Identification and estimation of the subjective expectation and uncertainty from belief distribution

While the discussion in the previous sections referred to general distributions of regarding human capital and subjective beliefs, in practice I need to parametrize all distributions to achieve identification. Crucially, I need to define a specific distribution for: (i) $H_{a_{i,j,k}}(\cdot)$, the distribution of beliefs about the age a child will learn MSD item j under scenario k, and (ii) $G_i(\cdot)$, the distribution of beliefs about the future human capital $\ln\theta_1$.

I assume that $H_{a_{i,j,k}}(\cdot)$ is a triangular distribution with mean $\mu_{a_{i,j,k}}$ and variance $\sigma^2_{a_{i,j,k}}$. This assumption allows the mode of the distribution to range from anywhere within the support as opposed to the Normal distribution. A triangular distribution is defined by three points: $\underline{a}_{i,j,k}$, $\dot{a}_{i,j,k}$, and $\bar{a}_{i,j,k}$, which are the minimum, most likely, and maximum age a child will learn MSD item j under scenario k. The mean and variance of a triangular distribution are given by:

$$\mu = \frac{\underline{a} + \dot{a} + \bar{a}}{3} \quad \text{and} \quad \sigma^2 = \frac{\underline{a}^2 + \dot{a}^2 + \bar{a}^2 - \underline{a}\dot{a} - \underline{a}\bar{a} - \dot{a}\bar{a}}{18}.$$

Meanwhile, I assume that $G_i(\cdot)$ is a normal distribution with mean $E[\ln\theta_1|\Omega_i]$ and variance $Var[\ln\theta_1|\Omega_i]$. I perform numerous tests using different combinations of distributions for $H_{a_{i,j,k}}(\cdot)$ and $G_i(\cdot)$, and the results are robust to these choices. These results can be found in the appendix.

The identification of the parameters from Equation 4 follow from Cunha, Elo, and Culhane (2022). To illustrate, note that we can write the parameter μ_{i,δ_2} as:

$$\mu_{i,\delta_2}(\bar{\theta}_0) = \frac{E[\ln\theta_{i,1}|\bar{\theta}_0, \bar{x}] - E[\ln\theta_{i,1}|\bar{\theta}_0, \underline{x}]}{\ln\bar{x} - \ln\underline{x}}, \tag{8}$$

where $\bar{\theta}_0$, $\bar{x}$ correspond to the scenario of high investment and high skill, and so on. The moment from Equation 8 can be directly computed from the elicitation instrument. The parameter is overidentified since there is an additional moment $\mu_{i,\delta_2}(\underline{\theta}_0)$ which conditions on the low skill scenario $\underline{\theta}_0$.

A similar argument can be made to identify σ^2_{i,δ_2}. One can show that:

$$\sigma^2_{i,\delta_2}(\bar{\theta}_0) = \frac{Var(\ln\theta_{i,1}|\bar{\theta}_0, \bar{x}) - Var(\ln\theta_{i,1}|\bar{\theta}_0, \underline{x})}{(\ln\bar{x})^2 - (\ln\underline{x})^2}.$$

However, these moments assume that the collected data has no measurement error. There is extensive evidence that measurement error is a problem in surveys. I now show how to estimate μ_{i,δ_2} and σ^2_{i,δ_2} while addressing measurement error. I assume that the measurement error of $\ln \theta_{i,j,k,1}$ and $\sigma^2_{i,j,k}$ is given by

$$\text{(9)} \qquad \mu_{\theta_{i,j,k,1}} = \overbrace{\mu_{i,\delta_0} + \mu_{i,\delta_1}(\ln \theta_0)_{j,k} + \mu_{i,\delta_2}(\ln x)_{j,k}}^{E[\ln \theta_{i,1}|\Omega_i]} \quad + \eta_{i,j,k,1},$$

$$\text{(10)} \qquad \sigma^2_{\theta_{i,j,k,1}} = \underbrace{\sigma^2_{i,0} + \sigma^2_{i,\delta_1}(\ln \theta_0)^2_{j,k} + \sigma^2_{i,\delta_2}(\ln x)^2_{j,k} + \sigma_{i,\delta_1,\delta_2} \ln x_{j,k} \ln \theta_{0,j,k}}_{Var(\ln \theta_{i,1}|\Omega_i)} \quad + \eta_{i,j,k,2}.$$

Given that both $\mu_{\theta_{i,j,k,1}}$, and $\sigma^2_{\theta_{i,j,k,1}}$ are measures constructed from the same set of responses, it is very likely that the measurement errors $\eta_{i,j,k,1}$ and $\eta_{i,j,k,2}$ are correlated to each other. Therefore, previous approaches to estimating these models, such as factors models or single-equation random coefficient models as in Swamy (1970) are not appropriate. Instead, I extend the Swamy (1970) estimator to accommodate a system of equations with correlated measurement errors within individuals. I refer to this estimator as the System Random Coefficients (SRC) estimator. An added problem is the fact that the parameters of equation (10) are variances, and therefore need to be strictly positive. I modify the SRC estimator to impose a non-negativity constraint on parameters.[13]

3.3. Stated choice instrument

I now describe the second instrument, intended to elicit the individual's opportunity cost of investment. The instrument consists of a series of stated choice experiments. I create a series of hypothetical scenarios of monthly household income, initial baby's health, and how many hours the individual spends at work. Then, I ask the respondent to answer a question: how much they would be willing to pay to spend one hour of leisure instead of taking care of the baby.

I first describe in detail the difference between *active and passive interaction* between parent and baby. I follow an extensive literature that distinguishes the quality of inter-

[13]An alternative would be to allow a more flexible measurement error assumption in Equation 10. However, this would require specific parametric assumptions about the distribution of measurement errors. The SRC estimator only requires to assume a specific covariance structure of the random coefficients and additive separability, while being silent on the actual distribution of the random coefficients. On the other hand, factor model approaches would require imposing parametric distributional assumptions on the random coefficients.

action and define *active interaction* as one in which the parent's main focus is in the child.[14] I list several examples of active interaction to the parent, such as (i) soothing the baby when he/she is upset; (ii) playing peek-a-boo with the baby; (iii) singing songs with the baby; (iv) feeding, nursing, bathing, attending to health needs; among others. In contrast, I give examples of passive interaction with the baby, such as (i) grocery shopping with baby; (ii) browsing social media apps on smartphone with baby at your side; (iii) nap time for baby; (iv) household chores (cleaning, cooking, etc) while baby is at your side; among others.

I then describe a situation in which the parent would like to spend one hour away from the baby every weekday of the month. I highlight that this one hour would be for personal leisure. A friend offers to take care of the baby during this 1 hour, and while they will not be engaged in active interaction, the baby will be safe. I ask the individual to choose the highest hourly rate they would be willing to pay to their friend for the whole month. They choose out of a slider that ranges from $0 to $30. As they move the slider, they can see how much the hourly rate means in monthly expenses, assuming 20 weekdays in a month. If the person would rather not spend one hour away, I ask them to choose $0. The intention is to elicit their willingness-to-pay for one hour of leisure and not the price they would pay for childcare. Therefore, I emphasize that this is a friend and not a professional caretaker such as a nanny.

The exact wording of the questions are as follows:

> Think about the time you have available outside of work during weekdays. Imagine that for 1 month during weekdays (20 days), you want to spend 1 hour of leisure time away from your baby. A friend offers to take care of your baby during this 1 hour for the month, in exchange for a payment. Your friend will take good care of your baby, but they will not be engaged in active interaction with your baby.

[14] See Folbre et al. (2005); Bono et al. (2016); Guryan, Hurst, and Kearney (2008); Schoonbroodt (2018).

FIGURE 2. Stated Choice Instrument - Willingness to Pay

For each question, I establish the following scenarios. The monthly household income can be Y = {\$2000, \$4000, \$6000}, the initial baby's health can be either $H = \{normal, poor\}$, and the working hours of the individual can be $HR = \{0, 4, 8\}$. These variables are chosen according to the model in Section 2. The household income is chosen to reflect different percentiles of income in the population. A \$2,000 household income puts the household around the poverty line as established by the United States Federal Government, while a household income of \$6,000 puts the household around the median income.

3.4. Model Identification and Estimation

I describe the model identification and estimation strategy. I assume that the respondent's utility is linearly separable in the endogenous variables $c_i, h_{l_i}, \theta_{i,1}$. By following the steps described in the previous section, I can estimate for each individual the parameters of their subjective skill production function. I use them as inputs in the estimation procedure. The remaining parameters left are the preference parameters $\boldsymbol{\alpha}$, and the subjective price of one hour of investment, p_i.

Let k denote an element on the set of possible scenarios $K = Y \times H \times HR$. Denote $X = \{0, 1, 2, 3, 4, 5, 6, 7\}$. Additionally, denote by $wp_{i,k}$ the maximum willingness-to-pay for one hour of leisure for individual i under scenario k. I asssume that $wp_{i,j}$ is a noisy measure of the true subjective price of one hour of investment, p_i. I estimate p_i under a flexible factor model:

$$wp_{i,k} = p_i + \gamma_i x_k + \varepsilon_{i,k}, \tag{11}$$

where x_k denote the scenario variables, γ_i is the vector of coefficients associated with the scenarios, and $\varepsilon_{i,k}$ is a measurement error. I estimate the parameters of the model using the Swamy (1970) estimator.

The optimal investment choice implied by the model is $x(\alpha)$, while the econometrician observes x^*. Due to measurement error, these values are different from each other. Let η_x denote the associated measurement error. Assume that:

$$x^* = x(\alpha)e^{\eta_x}, \tag{12}$$

where η_x is normally distributed with mean $\mu_{\eta_x} = -\frac{\sigma^2_{\eta_x}}{2}$ and variance $\sigma^2_{\eta_x}$.

Given these assumptions, the likelihood function is given by, where ϕ is the pdf of the standard normal distribution:

$$l(\alpha) = \sum_i^N \Big(\ln \phi(\ln x_i^* - \ln x_i(\alpha)) \Big). \tag{13}$$

To construct the model implied optimal investment choice $x(\alpha)$, one must compute the indirect utility function over a fine grid of possible values of x. I define the space of possible values of x as the interval $X = (0.0, 20.0]$[15] For a given value of $\alpha = \bar{\alpha}$, I construct a grid over this interval composed of 100 equidistant points, and compute the indirect utility function for each value. Then, I find the value $x^1(\bar{\alpha})$ which gives the maximum utility. I then construct a smaller interval around $x^1(\bar{\alpha})$ with 100 equidistant points, and proceed with the same algorithm and obtain $x^2(\bar{\alpha})$. I repeat this process until the difference between any two steps is smaller than 10^{-10}. This process is repeated for each new guess of α.

4. Data

I collect the data using Qualtrics, a company that holds online panels of individuals who are willing to take surveys for a small fee. My sample consists of 711 women between the ages of 18 and 40 who had at least one child, but no child older than 5 years old. In this section, I present the answers for all survey segments and descriptive statistics. I also present evidence that each survey segment has consistent participant answers and is not completely random.

[15]Optimal investment is measured in terms of daily hours. The lower bound of the interval is set to 10^{-6}

4.1. Sample Characteristics and Actual Investment

The survey participants answer the questions from the subjective belief instrument and the stated choice instrument. Additionally, participants answer several demographic questions and report actual investment in their oldest child. Table 1 describes the sample characteristics. The average age of participants is 27.8, with about 1.4 children on average, while the average age of children is 2.14. The sample is also relatively well educated, with 44.4% of the sample having at least a 2 year college degree. Personal and household income are distributed evenly across the categories. The sample is also relatively diverse, with 21.5% of individuals being Hispanic, 27.3% being Non-Hispanic Black, and 38% being Non-Hispanic White. Around 74% of women are working on average 6.8 hours a day, while 31.2% are in school.

In Table 2, I present statistics of actual investment. The participants are asked how many hours they spend with their child on a typical weekday and weekend day on reading, talking, playing inside, and playing outside. Table 2 shows that the average participant spends 4.7 hours on a typical weekday and 5.4 hours on a typical weekend day with their child. The average participant spends 0.5 hours reading, 1.8 hours talking, 1.5 hours playing inside, and 0.9 hours playing outside on a typical weekday. On a typical weekend day, the average participant spends 0.7 hours reading, 2.1 hours talking, 1.7 hours playing inside, and 1 hour playing outside. The table also shows that the standard deviation of the hours of investment is relatively large, suggesting that there is significant heterogeneity in the sample. Finally, the maximum number of hours spent on any given activity imply participants spending more than physically possible in care, suggesting that these respondents did not understand or answer the question truthfully. Therefore, I drop all values above the 95th percentile.

I estimate a linear regression of hours of actual investment on the individual's socio-economic variables. All continuous variables are standardized. The results are presented in Table 3. The table shows who those that currently work or go to school spend less time interacting with their child than those who do not. Additionally, the number of hours spent on a typical weekend is positively correlated with the number of hours worked during the weekday. This is consistent with the idea that parents face a time constraint on investment, but compensate on weekends by investing more.

Table 3 also shows that those with a high school degree or some college spend more time interacting with their child than those with less education, while Non-Hispanic Black and Hispanics spend less time interacting with their child than Non-Hispanic Whites. Finally, the table shows that those with a higher household income spend more

TABLE 1. Sample Composition

Statistic	N	Mean	St. Dev.
Age of Respondent	711	27.799	5.741
Number of Children	710	1.380	0.616
Age of Children	711	2.146	1.320
Working	711	0.737	0.441
Daily Work Hours	524	6.876	2.584
In School	711	0.312	0.464
Ethnicity			
Hispanic	711	0.215	0.411
Non-Hispanic White	711	0.381	0.486
Non-Hispanic Black	711	0.273	0.446
Other	711	0.131	0.337
Marital Status			
Single	707	0.306	0.461
Married or Cohabitating	707	0.644	0.479
Divorced, Separated, Widowed	707	0.051	0.220
Education			
Dropout or GED	711	0.093	0.290
High School or Some College	711	0.463	0.499
2 or 4 Year College	711	0.444	0.497
Personal Income			
Less than $25,000	690	0.283	0.451
$25,000 to $49,999	690	0.223	0.417
$50,000 to $99,999	690	0.354	0.478
More than $100,000	690	0.141	0.348
Household Income			
Less than $25,000	699	0.195	0.396
$25,000 to $49,999	699	0.209	0.407
$50,000 to $99,999	699	0.319	0.466
More than $100,000	699	0.278	0.448

Note: This table presents descriptive statistics for the sample of respondents. There are 711 individuals, but some questions were not answered by all respondents. For Daily Work Hours, only those who respond that they work provide answers.

Table 2. Actual Investment Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
On A **Typical Weekday**, How Many Hours Do You Spend with your Child on							
Reading	707	0.493	1.215	0.000	0.033	0.500	25.000
Talking	706	1.789	3.656	0.000	0.083	2.000	50.000
Playing Inside	706	1.512	3.737	0.000	0.083	1.667	70.000
Playing Outside	699	0.902	4.502	0.000	0.033	1.000	105.000
Total	711	4.655	9.914	0.000	0.283	5.292	178.667
On A **Typical Weekend**, How Many Hours Do You Spend with your Child on							
Reading	704	0.656	1.468	0.000	0.033	0.750	25.000
Talking	703	2.141	3.795	0.000	0.092	3.000	40.000
Playing Inside	703	1.672	2.605	0.000	0.083	2.000	26.667
Playing Outside	696	1.012	1.488	0.000	0.050	1.188	15.000
Total	711	5.410	7.481	0.000	0.283	7.583	60.000

Note: This table presents descriptive statistics for the actual investment measures. There were a total of 711 individuals, but some questions were not answered by all respondents.

time interacting with their child than those with a lower household income. Overall, the collected investment measures show patterns that are consistent with the literature.

4.2. Subjective Belief Instrument Responses

I present survey participants‘ responses to the age range questions. Table 4 shows the mean and standard deviation of the youngest, most likely, and oldest ages for each activity asked for each scenario of investment and health. The activities are ordered by difficulty according to the IRT model. I highlight two data features. First, we see that the mean age responses tends to increase with the difficulty of the activity for all types of ages. This is consistent with participants paying attention to each activity and understanding that they have different difficulty levels for a child.

Second, we see that as the scenario moves from high investment and normal health to low investment and poor health, the mean age responses tend to increase. This is also seen in Figure 3, where I show histograms for the activity of speaking a partial sentence.[16] The histograms show that the distribution of age responses is shifted to the right as the scenario moves from high investment and normal health to low investment and poor health. This is consistent with the idea that participants are responding to the scenario and not just randomly answering the questions.

4.3. Willingness to Pay

Table 5 displays descriptive statistics for the stated choice instrument. I present the mean and standard deviation of the maximum willingness-to-pay for the each scenario of health, household income, and daily working hours.

The table shows evidence of the consistency of the participants answers. The standard deviation in the sample across all scenarios remains constant, indicating that the variability in answers is similar regardless of the scenarios. There is no discernible difference across health scenarios. The means of each combination of work hour and income are similar whether the health of the baby is good or poor. This can be confirmed in Table 6, where the willingness-to-pay is regressed on the scenario variables. The coefficients for baby health are all not statistically significant than zero.

On both Tables, there is a strong gradient in work hours and household income. As they work more and earn more, their willingness-to-pay also increases. For example, conditional on good health, in the scenario of 0 working hours and $2,000 income, the

[16] The histograms for the other activities are similar and are presented in Appendix B.

TABLE 3. Correlations of Investment with Demographics

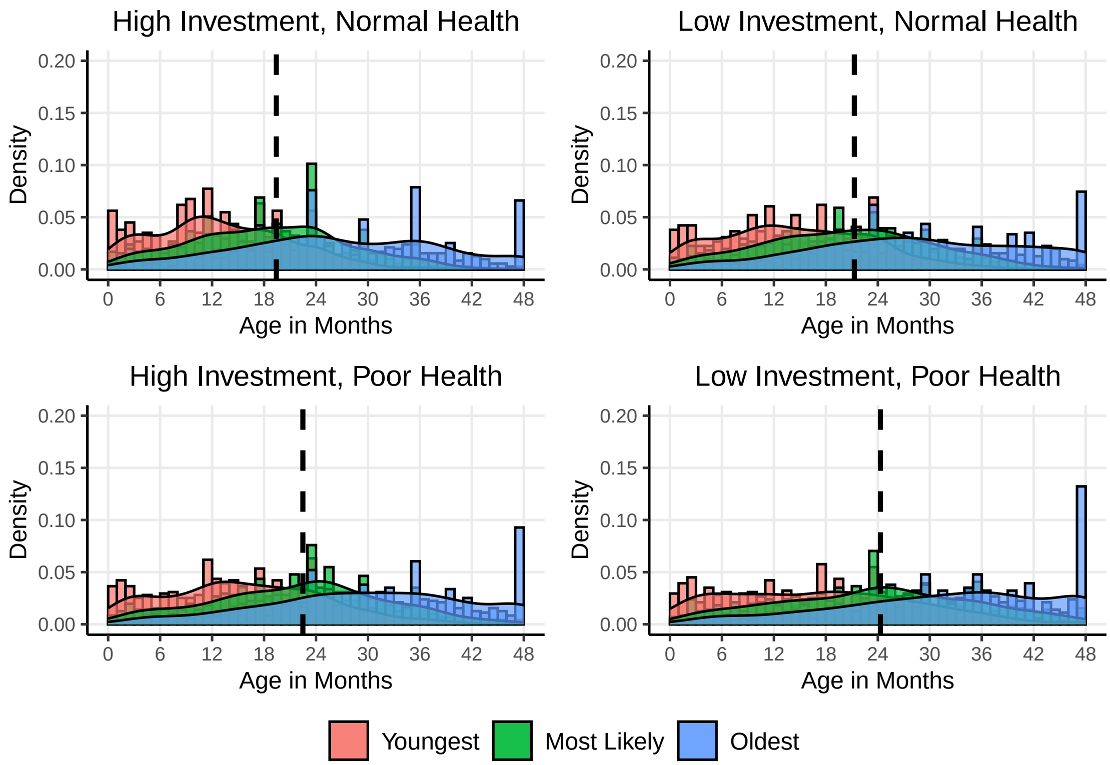
	Weekday Hours	Weekend Hours
Age $\leq$ 30	0.077	−0.040
	(0.087)	(0.091)
# Children	−0.051	−0.076
	(0.058)	(0.061)
Oldest Child $\leq$ 3	0.147*	0.052
	(0.084)	(0.093)
Works	−0.650***	−0.666***
	(0.150)	(0.129)
Working Hours	0.016	0.062***
	(0.016)	(0.015)
In School	−0.265***	−0.313***
	(0.080)	(0.084)
Ethnicity (Omitted: Non-Hisp White)		
Non-Hisp Black	−0.325***	−0.287***
	(0.102)	(0.106)
Hispanic	−0.127	−0.176*
	(0.108)	(0.103)
Other	−0.127	−0.162
	(0.117)	(0.118)
Marital Status (Omitted: Single)		
Married	0.124	0.214**
	(0.099)	(0.098)
Separated	−0.053	−0.077
	(0.142)	(0.141)
Education (Omitted: Dropout)		
High School	0.299**	0.184
	(0.119)	(0.137)
College Degree	0.343***	0.232
	(0.130)	(0.148)
Household Income (Omitted: \$0-\$25,000)		
\$25-\$50,000	0.287**	0.305***
	(0.128)	(0.117)
\$50-\$100,000	0.263**	0.294***
	(0.120)	(0.112)
\$100,000+	0.014	0.074
	(0.129)	(0.124)
Constant	0.007	0.007
	(0.215)	(0.210)
Observations	659	661
Note:	*p<0.1; **p<0.05; ***p<0.01	

TABLE 4. Mean and Standard Deviation of Responses to the Belief Instrument

	High Investment and Normal Health					
	Youngest Age		Most Likely Age		Oldest Age	
	Mean	SD	Mean	SD	Mean	SD
Partial Sentence	13.16	7.99	18.68	9.25	26.45	11.99
Age and Sex	15.47	8.75	21.84	9.60	29.39	11.18
First and Last Name	18.33	10.63	24.97	10.79	32.10	11.83
Counts 3 Objects	16.76	10.08	23.13	10.56	30.84	12.12
	Low Investment and Normal Health					
	Youngest Age		Most Likely Age		Oldest Age	
	Mean	SD	Mean	SD	Mean	SD
Partial Sentence	14.69	8.53	20.67	9.86	28.54	12.09
Age and Sex	17.34	9.56	23.87	10.14	30.90	11.47
First and Last Name	20.09	11.43	26.66	11.30	33.40	11.87
Counts 3 Objects	18.19	10.68	24.87	10.89	32.27	12.09
	High Investment and Poor Health					
	Youngest Age		Most Likely Age		Oldest Age	
	Mean	SD	Mean	SD	Mean	SD
Partial Sentence	15.78	9.21	22.00	10.12	29.68	11.82
Age and Sex	17.96	10.11	24.44	10.34	31.71	11.53
First and Last Name	20.83	11.56	27.40	11.44	34.19	11.85
Counts 3 Objects	18.76	10.99	25.40	11.32	32.81	12.16
	Low Investment and Poor Health					
	Youngest Age		Most Likely Age		Oldest Age	
	Mean	SD	Mean	SD	Mean	SD
Partial Sentence	17.55	10.94	23.86	11.32	31.52	12.32
Age and Sex	20.38	11.81	26.96	11.68	34.13	12.22
First and Last Name	22.35	13.08	28.72	12.38	35.40	12.19
Counts 3 Objects	21.03	12.43	27.69	12.12	34.55	12.49

Note: This table presents the mean and standard deviation of the youngest, most likely, and oldest ages for each activity asked for each scenario of investment and health. The activities are ordered by difficulty according to the IRT model, with the easiest activity being “Partial Sentence” and the hardest being “Counts 3 Objects”.

FIGURE 3. Distribution of Ages by Scenario for "Speak a Partial Sentence of 3 Words"



Note: This figure displays the histogram and density of the youngest, most likely, and oldest ages for the activity of speaking a partial sentence of 3 words. The histograms are colored by the type of age. The dashed line represents the mean of the most likely age.

TABLE 5. Descriptive Statistics - Willingness-to-Pay

	Conditional on Good Health					
	Household Income					
	$2,000		$4,000		$8,000	
	Mean	SD	Mean	SD	Mean	SD
Work 0 Hours	8.32	7.58	9.50	7.91	10.86	8.26
Work 4 Hours	11.29	6.90	12.41	6.51	13.47	6.90
Work 8 Hours	12.51	6.73	14.27	6.51	15.36	7.25
	Conditional on Poor Health					
	Household Income					
	$2,000		$4,000		$8,000	
	Mean	SD	Mean	SD	Mean	SD
Work 0 Hours	9.12	8.44	9.74	8.30	11.22	8.95
Work 4 Hours	11.42	7.74	12.34	7.62	13.55	8.05
Work 8 Hours	12.58	8.13	13.56	7.95	15.05	8.69

This table presents the mean and standard deviation of the maximum willingness-to-pay for the each scenario of health, household income, and daily working hours.

mean response is of \$8.32. Then, at 8 working hours, the mean response increases ot \$12.51. Conversely, at \$6,000 income, the mean increases to \$10.86. This is again confirmed in Table 6. Increasing the working hours by 1 hour increases the willingness-to-pay by \$0.51, while increasing the income by \$1,000 increases the willingness-to-pay by \$0.59.

These patterns are consistent with a model where parents value leisure. As their available hours of leisure decrease, they value it more. However, income is a significant factor, since as it increases, their budget constraint expands and they are able to allocate more resources to "buy" leisure. These correlations give support to the model presented in Section 2.

TABLE 6. Correlations of Investment and Willingness to Pay and Scenarios

	(1)	(2)	(3)	(4)
Hours of Work	0.51***			0.51***
	(0.02)			(0.02)
Household Income (in \$ thousands)		0.59***		0.59***
		(0.05)		(0.05)
Baby Health			-0.06	-0.06
			(0.15)	(0.15)
Constant	9.99***	9.66***	12.07***	7.64***
	(0.12)	(0.20)	(0.11)	(0.23)

Note: *p<0.1; **p<0.05; ***p<0.01. Robust standard error in parenthesis. The four columns show the correlation between willingness to pay and scenario variables. Hours of Work can be 0, 4, or 8. Household Income can be \$2,000, \$4,000, or \$6,000. Baby Health can be good or poor.

5. Results

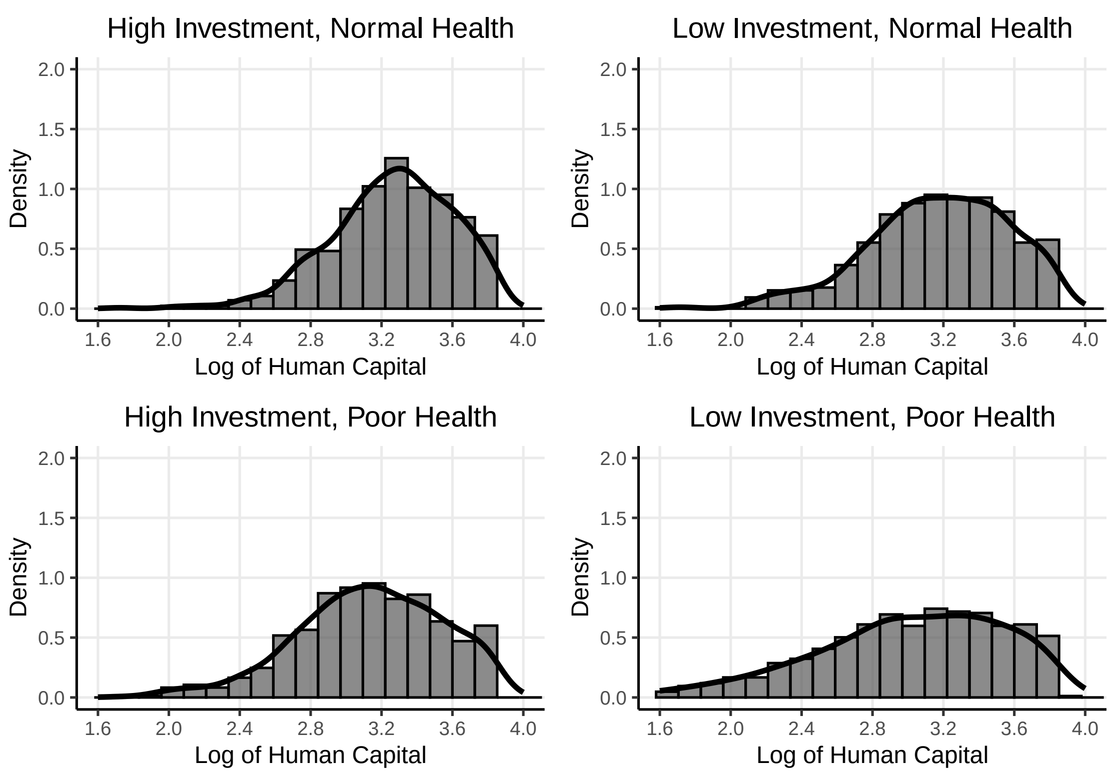
In this section, I first present the estimation results of the subjective production function parameters and the subjective cost of care. Then, I discuss how they are related to each other, to demographic variables, and to actual investment measures. I then illustrate the use of parental subjective expectation by estimating an investment model with reference-dependent preferences and discuss its interpretation.

5.1. Parental Beliefs

I now present the estimates of the subjective expectation and uncertainty about the future human capital of the child. These estimates assume that the age range distribution $H(\cdot)$ follows a triangular distribution, while the distribution of the log of human capital $G(\cdot)$ follows a normal distribution.

In Figure 4, I plot the histogram and distribution of $E[\ln \theta_1|\Omega_i]$ for each of the four possible hypothetical scenarios. For each scenario, there are four estimates of $E[\ln \theta_1|\Omega_i]$, one for each activity j. Therefore, I average $E[\ln \theta_1|\Omega_i]$ across all four activities. I find that the distribution of $E[\ln \theta_1|\Omega_i]$ is more dispersed and has a lower mean for the low investment and poor health scenario, while it is more concentrated around a higher mean for the high investment and normal health scenario. This indicates that parental beliefs are positively correlated with investment and health, as predicted by the model.

FIGURE 4. Distribution of Error-ridden Expectation of Log of Human Capital by Scenario



In Figure 5, I plot the distribution of the error-ridden variance of the log of human capital. I find that the distribution is more concentrated for the low investment and poor health scenario, while it is more dispersed around a lower mean for the high investment and normal health scenario. Therefore, parents are more uncertain about the future human capital of their child under a scenario of higher returns.

FIGURE 5. Distribution of Error-ridden Variance of Log of Human Capital by Scenario

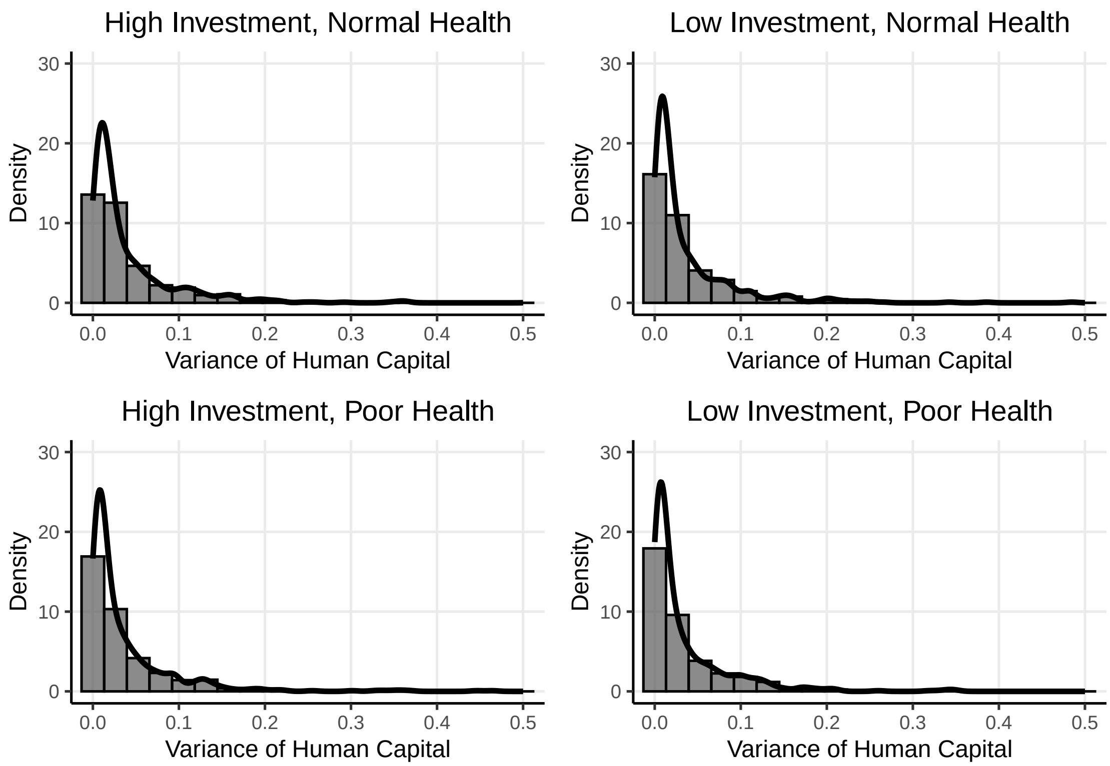


Table 7 shows the mean and standard deviation of $E[\ln\theta_1|\Omega_i]$ and $Var(\ln\theta_1|\Omega_i)$ and summarizes the above discussion. The mean values of $E[\ln\theta_1|\Omega_i]$ decrease as the scenarios move from the best to the worst, while its standard deviation increases. For the "best" scenario, where there is high investment and normal health, parents believe that the child's skill will be 3.260, which translates to a developmental age of 26 months. Note that the skill is anchored at the age of 24 months, so parents are slightly optimistic under this scenario. For the "worst" scenario of low investment and poor health, parents believe the child's skill will be 2.990, or 19.8 months, a developmental delay of about 4 months.

In contrast, the mean and standard deviation of $Var(\ln\theta_1|\Omega_i)$ decrease as the scenarios move from the best to the worst, while the standard deviation increases. This is consistent with parents having more pessimistic and less uncertain beliefs about the future human capital of their child when they invest less and when their child is in poor health.

TABLE 7. Summary Statistics for $E[\ln \theta_1|H_i]$ and $Var(\ln \theta_1|H_i)$ by Scenario

	$E[\ln \theta_1\|H_i]$		$Var(\ln \theta_1\|H_i)$	
Scenario	Mean	St. Dev.	Mean	St. Dev.
High Investment, Normal Health	3.260	0.341	0.042	0.058
Low Investment, Normal Health	3.175	0.392	0.037	0.052
High Investment, Poor Health	3.131	0.414	0.038	0.055
Low Investment, Poor Health	2.990	0.562	0.035	0.049

5.2. Estimates of the Subjective Production Function and Subjective Cost

Table 8 presents the estimates of the subjective production function parameters, μ_{δ_k} and $\sigma^2_{\delta_k}$, from (9) and (10) using the SRC estimator. Additionally, I estimate individual-level coefficients, μ_{i,δ_k} and σ^2_{i,δ_k}, and their standard errors.

Panel A of Table 8 displays the aggregate estimates. All parameters are statistically significant at the 1% level. I focus the discussion on the estimates of the subjective returns to investment parameters, δ_2, as they are the most relevant for the analysis. I find that the subjective mean of the returns to investment parameter, i.e. μ_{δ_2}, is 0.101. On average mothers believe that a 10% increase in investments would lead to a 1.01% increase in the child human capital by 24 months. As a comparison, Cunha, Elo, and Culhane (2022) report a subjective elasticity of investment of around 0.17, while Cunha, Elo, and Culhane (2013) report an objective elasticity of investment of around 0.26. Therefore, mothers in this sample are more pessimistic about the returns to investment than the previous literature. The estimated subjective variance, $\sigma^2_{\delta_2}$, is 0.005. Together with the mean estimate, these parameters give a coefficient of variation ($CV = \frac{\sigma}{\mu}$) of 0.70 thereby suggesting a low degree of uncertainty in beliefs of the average parent in the sample. At the individual level, about 39% of the sample have a coefficient of variation higher than 1.0.

In Panel B, I conduct a significance test for each individual-level parameter. I report the percentage of the estimates whose p-values are lower than 10% confidence region. Focusing on the subjective returns to investment parameters, I find that 43.62% of the estimates of μ_{i,δ_2} are statistically significant at the 10% level, while 28.49% of the estimates of σ^2_{i,δ_2} are statistically significant at the 10% level. The percentage value for the mean estimates are in line with Cunha, Elo, and Culhane (2022).

However, the percentage of significant coefficients for the variance estimates is low. This is likely due to the fact that the variance estimates are more affected by measurement errors that are not fully corrected by the methodology I propose. On the other hand, it could also be due to model misspecification. For example, the full variance specification, assuming uncorrelation of production shocks to all parameters, would be:

$$Var(\ln \theta_1 | H_i) = (\sigma^2_{i,\delta_0} + \sigma^2_{i,\varepsilon}) + \sigma^2_{i,\delta_1} \ln \theta^2_{i,0} + \sigma^2_{i,\delta_2} \ln x^2_i + \sigma_{\delta_0,\delta_1} \ln \theta_{i,0} + \sigma_{\delta_0,\delta_2} \ln x_i + \sigma_{\delta_1,\delta_2} \ln \theta_{i,0} \ln x_i.$$

The estimation of this model is not feasible due to the nature of the estimation strategy.

While theoretically feasible, the scenarios ($Z = \{\theta_0, x\}$) are simplified to be variables with only two distinct values and thus the within-individual variation of right-hand side variables is very small. Therefore, there would be severe multicollinearity in estimating the full model. Nevertheless, as I show in the next section, these estimates contain information that is relevant to demographic variables and real investment.

TABLE 8. Estimates of Mean Subjective Production Function Parameters

Panel A: Aggregate Estimates			
	$E[\ln\theta_1 \mid H_i]$		$Var(\ln\theta_1 \mid H_i)$
μ_{δ_0}	2.893***	$\sigma^2_{\delta_0}$	0.024***
	(0.0279)		(0.0016)
μ_{δ_1}	0.063***	$\sigma^2_{\delta_1}$	0.002***
	(0.0044)		(0.0002)
μ_{δ_2}	0.101***	$\sigma^2_{\delta_2}$	0.005***
	(0.0080)		(0.0007)
		$\sigma_{1,2}$	−0.003***
			(0.0006)
Panel B: Individual Estimates			
Parameter	% Significant	Parameter	% Significant
μ_{i,δ_0}	100.00%	σ^2_{i,δ_0}	64.09%
μ_{i,δ_1}	48.07%	σ^2_{i,δ_1}	27.60%
μ_{i,δ_2}	43.62%	σ^2_{i,δ_2}	28.49%
		$\sigma_{i,1,2}$	19.44%

Note: Robust standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. This table shows aggregate estimates of equations (9) and (10) and their individual-level predicted estimates. The percentage of significant estimates is calculated using a 10% significance level, and the null hypothesis is that the coefficient is equal to zero.

Figures 6 and Table 9 display the distribution and summary statistics of the estimated individual level parameters of the subjective beliefs of the production function and the subjective costs.[17] The mean values of the production function parameters are by construction equal to the estimates in Table 8. The standard deviation and selected percentiles of the distribution of the individual-level parameters are also reported and give an idea of the heterogeneity in beliefs. Focusing on the parameters related to the returns to investment, i.e. δ_2, I find that in general the mean beliefs exhibit variation

[17] I present the estimates of the factor model that produces the price coefficients in Appendix B.

across individuals, with the standard deviation of μ_{i,δ_2} being 0.099 and a coefficient of variation equal to 1.0. The standard deviation of the variance estimates, σ^2_{i,δ_2}, is 0.008, which produces a coefficient of variation of 1.6, indicating a larger heterogeneity.

The mean price estimate shows that on average mothers price one hour of investment compared to leisure at around \$12.17, with the overall distribution of prices not being very dispersed, as the 25th and 75 percentiles are close to the mean. The price estimate is elicited in the context of tradeoff of leisure and care on a typical weekday. It implies a monthly cost of 1 hour care of children everyday on weekdays at around \$270.

TABLE 9. Individual Level Coefficient Distributions

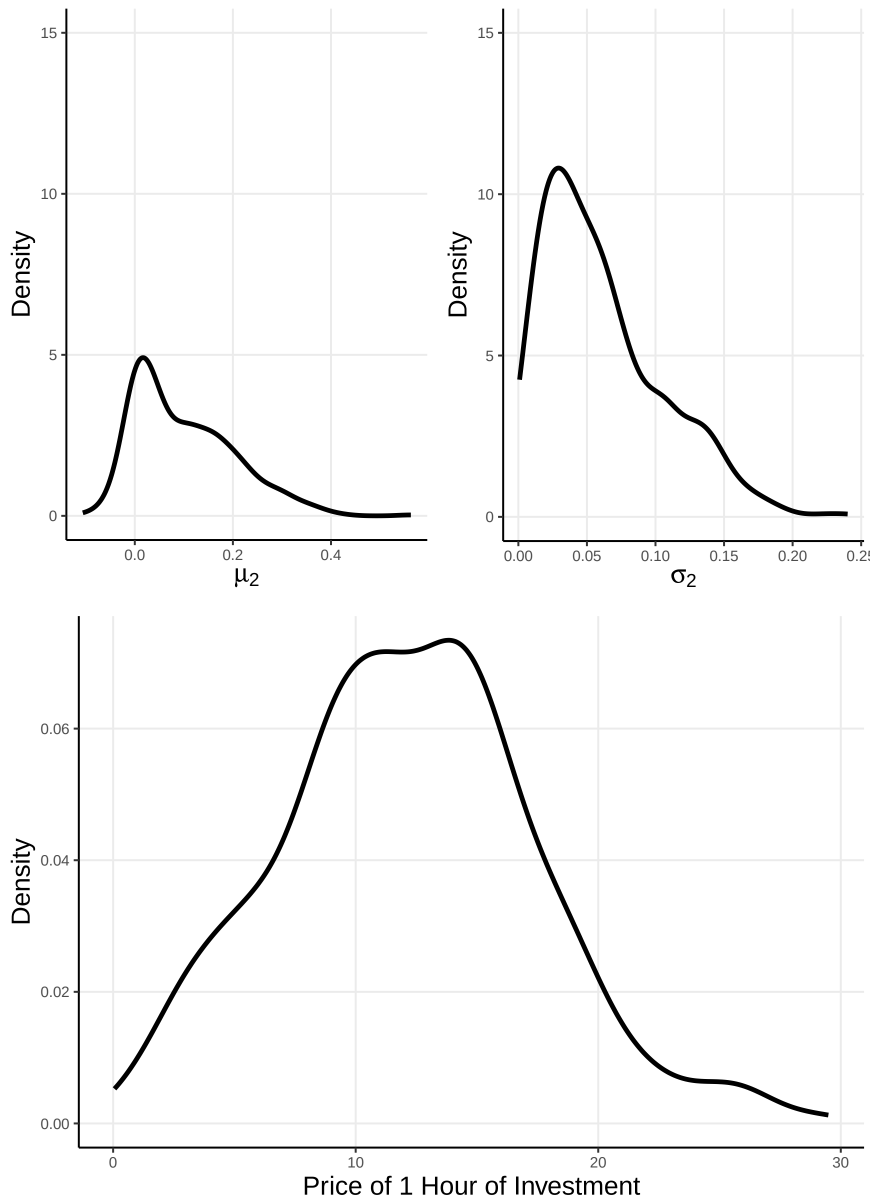
Variable	Mean	St. Dev.	25th Percentile	Median	75th Percentile
Price	12.164	5.291	8.600	12.019	15.416
μ_{i,δ_0}	2.904	0.600	2.480	2.935	3.398
μ_{i,δ_1}	0.063	0.061	0.014	0.050	0.099
μ_{i,δ_2}	0.101	0.102	0.017	0.079	0.166
$\sigma^2_{i,0}$	0.024	0.029	0.002	0.012	0.038
σ^2_{i,δ_1}	0.002	0.002	0.0002	0.001	0.002
σ^2_{i,δ_2}	0.005	0.008	0.001	0.002	0.007
$\sigma_{i,\delta_1,\delta_2}$	−0.003	0.006	−0.005	−0.001	0.0002

5.3. Relationship between Subjective Variables, Demographics, and Actual Investment

I now investigate how observable characteristics of mothers relate to their subjective beliefs. I focus on the subjective returns to investment parameters, μ_{i,δ_2} and σ^2_{i,δ_2}, and the price of care, p_i. I standardize these variables to have a mean equal to 0 and a standard deviation equal to 1 to keep the relationships comparable. Then I run a linear regression on the demographic variables described in the Data section. I report the results in Table 10.

First, I find that younger mothers tend to have lower mean beliefs and to be more uncertain about the return to investment in children. Second, Non-hispanic Black and Hispanic mothers are more pessimistic and more uncertain about the return to investment than White mothers. Third, I find that married mothers are less uncertain

FIGURE 6. Distribution of Individual Level Coefficients



about the return to investment, and price investments much lower than non-married ones. Fourth, I find a strong education gradient in mean beliefs. Finally, while income is not significantly correlated with beliefs, there is a strong income gradient in the cost of care.

While no causal relations can be extracted from these results, they suggest certain patterns that are consistent with the literature. For example, the racial difference in mean beliefs is consistent with the findings of Cunha (2014). The positive education gradient in beliefs is conceptually similar to the use of parental education as a proxy for parental knowledge of the benefits of early investment. Lower and more uncertain beliefs for younger mothers suggests that mothers learn throughout the process of raising children, although an extension incorporating learning in the production function is beyond the scope of this paper. Similarly, it is not surprising that the price of care is more costly for higher income mothers, as they are more likely to be employed and have higher opportunity costs of time. Finally, the negative correlation between marriage and uncertainty and price shows that family structure can be an important determinant in belief formation. This link has been noted in the literature on children's outcomes.

Next, I explore how the estimated beliefs relate to each other. I regress the subjective mean on subjective uncertainty and cost of care, and control for the observable characteristics. I also estimate the same model but using the subjective uncertainty as the dependent variable to obtain the correlation between cost of care as well. Table 11 reports the results. I find that mothers that have lower beliefs about the returns to investment tend to be more uncertain as well. This result holds as we add demographic controls and the price of care. Further, mothers that have higher costs of care tend to have lower mean beliefs and be more uncertain.

Finally, I explore how the estimated beliefs relate to actual investment. Two different measures of time investment are collected in the data, the number of hours spent with the child on a typical weekday and weekend. Therefore, I estimate two regressions using these two measures as the dependent variables, and controlling for observable variables. Table 12 reports the correlation between the subjective returns to investment parameters and the actual investment measures. The first column uses only the mean belief as the independent variable, as in the previous literature, and I progressively add the subjective variance, the price of care, and demographic controls.

I find that the subjective mean of the returns to investment parameter, i.e. μ_{δ_2}, is positively correlated with the actual investment measures. This finding is consistent

TABLE 10. Correlation of Subjective Variables and Socio-Economic Variables

	Dependent variable:		
	μ_{i,δ_2}	σ^2_{i,δ_2}	Price
	(1)	(2)	(3)
Age $\leq$ 30	-0.184^{*}	0.277^{***}	0.109
	(0.107)	(0.085)	(0.097)
# Children	−0.080	0.185^{**}	0.028
	(0.067)	(0.087)	(0.074)
Oldest Child $\leq$ 3	−0.007	0.072	−0.169
	(0.111)	(0.103)	(0.117)
Ethnicity (Omitted Group: Non-Hisp White)			
Non-Hisp Black	-0.240^{**}	0.328^{**}	0.100
	(0.119)	(0.134)	(0.119)
Hispanic	-0.257^{**}	0.149	0.113
	(0.118)	(0.106)	(0.122)
Other	−0.122	0.051	0.012
	(0.139)	(0.107)	(0.130)
Marital State (Omitted Group: Single)			
Married	0.163	-0.232^{**}	-0.417^{***}
	(0.105)	(0.114)	(0.112)
Separated	−0.064	−0.167	−0.072
	(0.207)	(0.217)	(0.211)
Education Level (Omitted Group: Dropout/GED)			
High School	0.292^{**}	−0.019	0.111
	(0.135)	(0.173)	(0.171)
College Degree	0.396^{***}	−0.021	0.106
	(0.149)	(0.183)	(0.181)
Household Income (Omitted Group: $0-$25,000)			
$25-$50,000	0.148	−0.126	0.313^{**}
	(0.139)	(0.162)	(0.149)
$50-$100,000	0.094	−0.141	0.427^{***}
	(0.137)	(0.154)	(0.149)
$100,000+	−0.172	−0.151	0.485^{***}
	(0.159)	(0.170)	(0.164)
Constant	−0.069	−0.313	−0.196
	(0.233)	(0.266)	(0.248)
Observations	539	539	539

Note: Robust standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All continuous variables are standardized.

TABLE 11. Correlation between beliefs and cost of care

	μ_{δ_2}			
	(1)	(2)	(3)	(4)
$\sigma^2_{\delta_2}$	-0.335^{***}		-0.313^{***}	-0.286^{***}
	(0.033)		(0.031)	(0.033)
Price		-0.190^{***}	-0.141^{***}	-0.142^{***}
		(0.041)	(0.039)	(0.042)
	$\sigma^2_{\delta_2}$			
	(1)	(2)	(3)	(4)
μ_{δ_2}	-0.335^{***}		-0.317^{***}	-0.276^{***}
	(0.036)		(0.037)	(0.037)
Price		0.155^{***}	0.095^{***}	0.091^{**}
		(0.037)	(0.036)	(0.038)
Observations	542	542	542	539
Demographics	No	No	No	Yes

Note: Robust standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All continuous variables are standardized.

with the previous literature (see, Cunha, Elo, and Culhane (2013, 2022); Boneva and Rauh (2018)). A one standard deviation increase in mean beliefs leads to a 23% standard deviation increase in daily hours of investment, whether on weekdays or weekends. However, higher uncertainty predicts lower investments, with a one standard deviation increase associated with a 14% standard deviation decrease in investments. We see a similar pattern in the price of care, but the reduction of investment is larger for weekends than for weekdays.

Overall, these patterns provide reassurance that the elicitation methods provided meaningful data since results for mean beliefs follow the previous literature. Furthermore, the results for the subjective uncertainty suggests that increased uncertainty may play an important role in influencing investment, and future interventions that target belief improvements may find useful to also measure uncertainty.

TABLE 12. Correlation of Real Investment with Beliefs

	(1)	(2)	(3)	(4)
		Weekday Hours		
μ_{δ_2}	0.354***	0.302***	0.282***	0.229***
	(0.041)	(0.044)	(0.044)	(0.047)
$\sigma^2_{\delta_2}$		−0.152***	−0.138***	−0.138***
		(0.033)	(0.032)	(0.035)
Price			−0.136***	−0.127***
			(0.040)	(0.040)
		Weekend Hours		
μ_{δ_2}	0.374***	0.311***	0.298***	0.245***
	(0.041)	(0.044)	(0.044)	(0.047)
$\sigma^2_{\delta_2}$		−0.185***	−0.176***	−0.133***
		(0.034)	(0.034)	(0.035)
Price			−0.090**	−0.079*
			(0.044)	(0.043)
Observations	542	542	542	539
Demographics	No	No	No	Yes

Note: Robust standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All continuous variables are standardized.

5.4. An Application to a Model With Reference Dependent Preferences

Previous research either ignore subjective beliefs or assume specific functional forms for the utility function that rules out higher order beliefs. The stated choice data together with the subjective belief estimates for individuals allows me to estimate a flexible model of parental investment in children. In this section, I estimate a model of parental investment in children with reference dependent preferences that incorporates subjective uncertainty on the decision making of individuals. Parental preferences depend on household consumption, leisure, child development at the end of the period, and the relative development of the child compared to a reference level of development, θ_{ref}. I assume that the mother's preferences are given by:

$$
\begin{aligned}
u_i(c_i, h_{l_i}, \theta_{i,1}) = & \alpha_1 \ln c_i + \alpha_2 \ln h_{l_i} + \alpha_3 \ln \theta_{i,1} + \\
& \alpha_4 (\ln \theta_{ref} - \ln \theta_{i,1}) \mathbb{1}\{(\ln \theta_{i,1} \leq \ln \theta_{ref})\}
\end{aligned}
$$

The parameter α_1 captures the preference for consumption and α_2 captures the preference for leisure. The coefficient α_3 capture how parents value the end-of-period human capital of their child, while α_4 denotes how parents value their child's skill being below the reference point. The sign of α_4 determines the shape of the indifference curves. If $\alpha_4 > 0$, the indifference curves are convex.

The reference point θ_{ref} is the level of development that the parents consider to be "desirable" or "satisfactory". In a similar context of parental child investment, Wang et al. (2022) and Kinsler and Pavan (2021) show that parents use their local peers as a reference point[18] In my context, this definition of a "local" peer is not feasible. Since child development is anchored in developmental age, a natural reference point is a threshold for developmental delay.

The expected utility function which parents maximize is:

$$
\begin{aligned}
E[u_i(c_i, h_{l_i}, \theta_{i,1}) | \Omega_i] = & \alpha_1 \ln c_i + \alpha_2 \ln h_{l_i} + \alpha_3 E[\ln \theta_{i,1} | \Omega_i] + \\
& \alpha_4 \int_{\theta_{i,1}} (\ln \theta_{ref} - \ln \theta_{i,1}) \mathbb{1}\{(\ln \theta_{i,1} \leq \ln \theta_{ref})\} dG_i,
\end{aligned}
$$

where $G_i(\cdot)$ is the distribution of subjective beliefs. Given the assumption that $G_i(\cdot)$ is a

[18] In Wang et al. (2022), the reference is the average development of children in the village, while in Kinsler and Pavan (2021) it is the peers in the same school and classroom.

Normal distribution, it follows that the integral can be simplified to:[19]

$$\int_{\theta_{i,1}} (\ln\theta_{ref} - \ln\theta_{i,1})\mathbb{1}\{(\ln\theta_{i,1} \leq \ln\theta_{ref})\}dG_i =$$

$$(\ln\theta_{ref} - \mu_{i,\theta_1})\Phi\left(\frac{\ln\theta_{ref} - \mu_{i,\theta_1}}{\sigma_{i,\theta_1}}\right) + \sigma_{i,\theta_1}\phi\left(\frac{\ln\theta_{ref} - \mu_{i,\theta_1}}{\sigma_{i,\theta_1}}\right),$$

where Φ and ϕ denote the cdf and pdf of a standard Normal distribution, respectively, and μ_{i,θ_1} and σ_{i,θ_1} are the mean and standard deviation of the subjective beliefs of the production function. Then, the estimation of the model parameters $\boldsymbol{\alpha}$ follows the procedure described in section 3.4.

Table 13 displays the estimation of the preference parameters of the investment model. Note that I use the subjective production function parameters and the subjective cost of care as inputs in the estimation procedure. Therefore, the estimates of the preference parameters take into account that different individuals face different costs of care and have different beliefs about the returns to investment.

TABLE 13. Model Estimates

Parameter	Estimates		
	$\theta_{ref} = 18$	$\theta_{ref} = 21$	$\theta_{ref} = 24$
α_2	1.768***	1.917***	4.991***
	(0.001)	(0.001)	(0.001)
α_3	-0.787***	-0.773***	-0.812***
	(0.001)	(0.002)	(0.001)
α_4	18.023***	16.929***	15.951***
	(0.003)	(0.001)	(0.002)

Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. This table shows the estimates of the preference parameters of the investment model for different reference points.

I estimate the model for different reference points θ_{ref}. The elicitation of the subjective beliefs targets the developmental age of 24 months. In other words, the activities that are used to elicit beliefs from parents were typical for children around 2 years old, and the anchoring of the beliefs to a cardinal metric used the 24 months milestone.

[19] Note that the integral involves the expected value of a Truncated Normal Distribution. Specifically,

$$\int_{\theta_{i,1}} (\ln\theta_{ref} - \ln\theta_{i,1})\mathbb{1}\{(\ln\theta_{i,1} \leq \ln\theta_{ref})\}dG_i = E[\ln\theta_{i,1}|\Omega_i, \ln\theta_{i,1} \leq \ln\theta_{ref}]Pr(\ln\theta_{i,1} \leq \ln\theta_{ref})$$

.

Then, the reference point θ_{ref} should be a developmental age around 24 months. I estimate the model using $\theta_{ref} = \{18, 21, 24\}$ months. A $\theta_{ref} = 18$ indicates that the reference point is a developmental age of 18 months, which is a large delay in development. The estimate of α_4 will determine the preference of the mother for investing in their child so that they reach the reference point or go beyond it.

The preference for consumption, α_1 is fixed to 1 for identification. All parameters are statistically significant. The preference for leisure, α_2 is positive for all reference points. The preference for the child's future skill, α_3 is negative. This may seem counterintuitive, as it means that parents dislike their child's human capital. However, note that α_4 is a large and positive number. Therefore, the negative sign of α_3 implies that parents have a strong preference for investing in their child when they are below the reference point, but that incentive disappears as they cross it. This is consistent with the literature on reference dependent preferences, where individuals dislike being below a reference point, but are indifferent to being above it. Additionally, note that the magnitude of α_4 decreases as the reference value increases. This highlights that parents' preferences for being above the reference point are stronger when the reference point is lower, that is, they are more concerned about their child being in a situation of severe developmental delay.

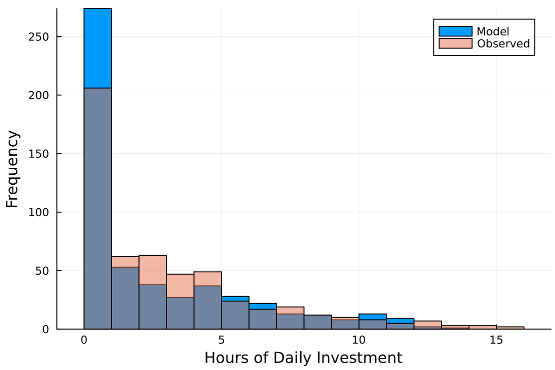
Figure 7 shows the distribution of daily investment hours that is implied by the model under the estimated parameters and compares it to the observed investment. In general, the estimated parameters using references $\theta_{ref} = 18$ and $\theta_{ref} = 21$ have a good fit, but the one using $\theta_{ref} = 24$ cannot replicate larger investments.

I now simulate a policy that increases the mean belief of the returns to investment, μ_{δ_2}. I simulate the policy for the different references points, and report the results in Table **??**. I omit results for $\theta_{ref} = 24$ given the poor fit. This kind of simulation replicates the types of interventions that target parental beliefs. However, these interventions do not measure the uncertainty of the parents. Therefore, I will also simulate the policy coupled with changes in uncertainty.

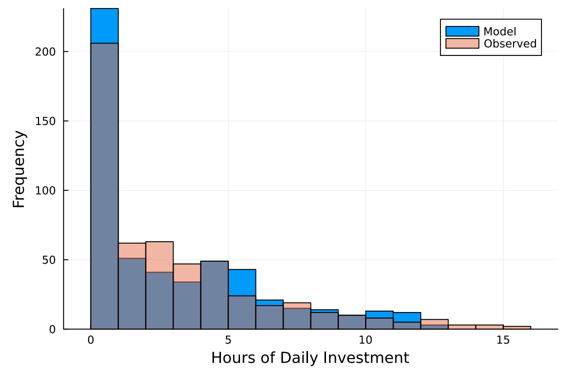
I first increase all individual beliefs about the returns to investment, μ_{i,δ_2}, by 10% for all individuals. I find that the increase in the mean belief leads to between 2% and 3% increase in the investment in children. The same is done for the uncertainty about returns to investment, σ^2_{i,δ_2}, and it also leads to a similar increase. A joint increase of μ_{i,δ_2} and decrease of σ^2_{i,δ_2} offsets and produces a negligible change in investment.

This results contradicts the reduced form evidence presented in the previous section. In Table 2, I find that higher uncertainty is associated with lower investment. However,

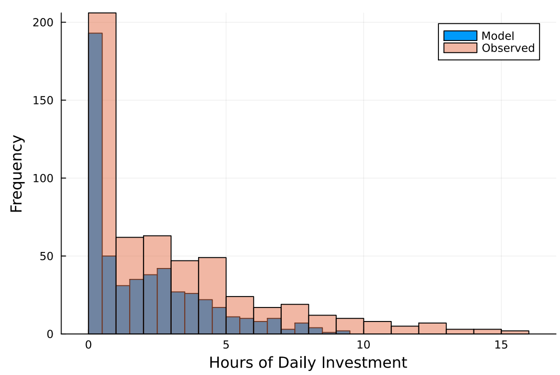
FIGURE 7. Model Predicted and Observed Hours of Investment



A. $\theta_{ref} = 18$



B. $\theta_{ref} = 21$



C. $\theta_{ref} = 24$

the model predicts that increasing uncertainty leads to higher investment. However, it is important to consider that the reduced form evidence is not a causal statement. Indeed, given that individuals with higher mean beliefs are also ones with lower uncertainty, it could be that the correlational regressions are simply reproducing the effect of mean beliefs.

TABLE 14. Percent Change in Daily Investment Hours Due to Change in Beliefs

	$\theta_{ref} = 18$	$\theta_{ref} = 21$
10% μ_{i,δ_2}	1.80	2.70
10% σ_{i,δ_2}	2.58	1.79
10% μ_{i,δ_2} and -10% σ_{i,δ_2}	-1.06	0.21

To further understand the effect of uncertainty on investment, I break down the increase in σ^2_{i,δ_2} in Table 15. I simulate the policy for individuals with different mean beliefs, different costs of care, and different baseline investment. I report the results in Table 15. The first column shows the percentage of individuals for which the increase in uncertainty leads to an increase or decrease in investment. About 39% observe an increase in investment, while 21% reduce their investments, with the remaining 40% not changing. The second column shows the average change in investment in hours. The third column shows the average mean belief of the individuals that increase their investment, while the fourth column shows the average price of care. Finally, the last column shows the average baseline investment of the individuals that increase their investment.

TABLE 15. Breakdown of Change in Investment Due to a 10% Increase in σ^2_{i,δ_2}

	% by Sign	Nominal Change (Hours)	μ_2	Price	Baseline $\hat{x}$
Increase in Investment	38.92	0.02	0.03	13.15	0.34
Decrease in Investment	21.60	-0.03	0.16	11.55	5.09
No Change in Investment	39.48	0.00	0.13	11.34	3.24

I find that the individuals that increase their investment due to the increase in uncertainty are the ones that hold very low mean beliefs, have higher opportunity costs of investment, and therefore do not invest much in their child. On the other hand, those with high mean beliefs, low costs, and high baseline investment reduce their investment due to the increase in uncertainty.

6. Conclusion

This paper develops a methodology that elicits both mean beliefs and belief uncertainty about the parameters of the skill production function in early childhood. The methodology and data collection procedure is motivated by a model of parental investment in children in which parents have do not have full knowledge of the skill production function. I elicit the belief distribution and allow for a flexible distribution of beliefs.

Additionally, I explore a new measure of parental subjective costs of investment in children. Recognizing that time investments represent opportunity costs for parents, I ask parents to report the amount of money they would be willing to pay to trade off investment and leisure. Together with the subjective belief distribution, I use this measure to estimate a model of parental investment in children which takes into account subjective beliefs.

I show that the collected data exhibit patterns that are consistent with the proposed theoretical model. Respondents report lower ages for the youngest, most likely, and oldest age of learning under the scenario of high investment and normal health, consistent with a model where children reach developmental milestones faster under more investment. They also report higher opportunity costs of investment under scenarios where they work more hours a day and have higher household income.

I estimate parental beliefs and find that parents in this sample have low mean beliefs about the returns to investment, and uncertainty is relatively large. Moreover, individuals that have higher mean beliefs also tend to have lower subjective uncertainty. I also find that both mean beliefs and uncertainty correlate with actual time investment measures, but while higher mean beliefs predict higher investment, higher uncertainty predicts lower investment.

To illustrate the use of this data, I combine the literature on subjective beliefs and reference points and estimate a model of parental investment that incorporates these two aspects. I find that parents strongly value their child skill even when holding low mean beliefs. Moreover, they have a strong incentive to invest if their child is at risk of being at a developmental delay.

In general, my findings indicate that belief heterogeneity in returns to investment is important in predicting actual time investment in children, and uncertainty about beliefs can be a relevant target for policy interventions. Moreover, the methodology developed in this paper is flexible enough to be transported to other contexts of child investments. I also explore how the opportunity costs of time investment can be elicited from parents to empirically price the subjective cost of care of mothers. Since I find that more pessimist parents tend to be more uncertain about their beliefs, it suggests that information interventions may have larger gains for more uncertain parents.

Several extensions and follow ups are possible. First, an overlooked aspect is the extent that risk aversion and intertemporal preferences impacts child investment. There are several method of elicitation of these features, but they are usually done in an abstract setting of money lotteries. If one wishes to study these aspects, it is necessary to develop context-specific elicitation procedures.

Second, eliciting the subjective cost of care presents several extension opportunities. The subjective cost of child care is poorly understood. The cost that women face in child care is multidimensional. Accurately estimating it is key to understanding why policies that subsidize child care, or ones that provide information interventions, do not translate into increases in investment. As Guryan, Hurst, and Kearney (2008) and Schoonbroodt (2018) point out, it is necessary to separate the costs of child care in at least two dimensions: (i) as a tradeoff against foregone earnings, and therefore, labor supply, and (ii) as a tradeoff against leisure and housework, usually "outside" working hours. In my elicitation method, I hold labor supply fixed and frame cost as a trade-off against leisure. A more flexible approach would be to elicit an additional cost framed as foregone earnings.

Finally, future research can conduct randomized controlled trials and use the methodology developed in this paper to explore the effects of information interventions on parental uncertainty, their subjective costs of cares, and if they can mediate changes in actual investment in children.

References

Aizer, Anna, and Laura Stroud. 2010. "Education, Knowledge and the Evolution of Disparities in Health." Working Paper 15840, National Bureau of Economic Research.

Almås, Ingvild, Orazio Attanasio, and Pamela Jervis. 2023. "Economics and Measurement: New measures to model decision making." Working Paper 30839, National Bureau of Economic Research.

Arendt, Jacob Nielsen, Mads Lybech Christensen, and Anders Hjorth-Trolle. 2021. "Maternal education and child health: Causal evidence from Denmark." *Journal of Health Economics* 80: 102552.

Attanasio, Orazio, Flávio Cunha, and Pamela Jervis. 2019. "Subjective Parental Beliefs. Their Measurement and Role." Working Paper 26516, National Bureau of Economic Research.

Attanasio, Orazio, and Katja Kaufmann. 2014. "Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender." *Journal of Development Economics* 109: 203–216.

Badev, Anton, and Flávio Cunha. 2012. "Learning about the Technology of Skill Formation." Working Paper, University of Pennsylvania.

Basu, Arnab, and Ralitza Dimova. 2022. "Household Preferences and Child Labor in Rural Ethiopia." *Journal of African Economies*: forthcoming.

Boneva, Teodora, and Christopher Rauh. 2018. "Parental Beliefs about Returns to Educational Investments—The Later the Better?" *Journal of the European Economic Association* 16 (6): 1669–1711.

Bono, Emilia Del, Marco Francesconi, Yvonne Kelly, and Amanda Sacker. 2016. "Early Maternal Time Investment and Early Child Outcomes." *The Economic Journal* 126 (596): F96–F135.

Bruine de Bruin, Wändi, Alycia Chin, Jeff Dominitz, and Wilbert van der Klaauw. 2023. "Chapter 1 - Household surveys and probabilistic questions." In *Handbook of Economic Expectations*, edited by Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw, 3–31: Academic Press.

Carneiro, Pedro, and Rita Ginja. 2016. "Partial Insurance and Investments in Children." *The Economic Journal* 126 (596): F66–F95.

Carneiro, Pedro, and Margarida Rodrigues. 2009. "Evaluating the Effect of Maternal Time on Child Development Using the Generalized Propensity Score." Working Paper, IZA.

Caucutt, Elizabeth, Lance Lochner, Joseph Mullins, and Youngmin Park. 2020. "Child skill production: Accounting for parental and market-based time and goods investments." Working Paper, NBER.

Conti, Gabriella, Michele Giannola, and Alessandro Toppeta. 2022. "Parental Beliefs, Perceived Health Risks, and Time Investment in Children: Evidence from COVID-19." SSRN Scholarly Paper 4291702, IFS.

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

Cunha, Flávio. 2014. "Gaps in Early Investments in Children." Working Paper, Rice University.

Cunha, Flávio, Irma Elo, and Jennifer Culhane. 2013. "Eliciting Maternal Subjective Expectations about the Technology of Cognitive Skill Formation." WorkingPaper 19144, National Bureau of Economic Research.

Cunha, Flávio, Irma Elo, and Jennifer Culhane. 2022. "Maternal Subjective Expectations About the Technology of Skill Formation Predict Investments in Children One Year Later." *Journal of Econometrics* 231 (1): 3–32.

Currie, Janet, and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *The Quarterly Journal of Economics* 118 (4): 1495–1532.

Dahl, Gordon B., and Lance Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." *American Economic Review* 102 (5): 1927–1956.

Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall. 2014. "Household Choices and Child Development." *The Review of Economic Studies* 81 (1): 137–185.

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

Delavande, Adeline, and Basit Zafar. 2018. "University Choice: The Role of Expected Earnings, Non-pecuniary Outcomes and Financial Constraints." *Journal of Political Economy* 127 (5): 2343–2393.

Dizon-Ross, Rebecca. 2019. "Parents' Beliefs about Their Children's Academic Ability: Implications for Educational Investments." *American Economic Review* 109 (8): 2728–2765.

Fiorini, Mario, and Michael Keane. 2014. "How the Allocation of Children's Time Affects Cognitive and Noncognitive Development." *Journal of Labor Economics* 32 (4): 787 – 836.

Folbre, Nancy, Jayoung Yoon, Kade Finnoff, and Allison Sidle Fuligni. 2005. "By What Measure? Family Time Devoted to Children in the United States." *Demography* 42 (2): 373–390.

Giustinelli, Pamela. 2016. "Group Decision Making with Uncertain Outcomes: Unpacking Child–Parent Choice of the High School Track." *International Economic Review* 57 (2): 573–602.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney. 2008. "Parental Education and Parental Time with Children." *Journal of Economic Perspectives* 22 (3): 23–46.

Heckman, James, Seong Hyeok Moon, Rodrigo Pinto, Peter Savelyev, and Adam Yavitz. 2010. "Analyzing Social Experiments as Implemented: A Teexamination of the Evidence From the HighScope Perry Preschool Program." *Quantitative Economics* 1 (1): 1–46.

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

Kalil, Ariel, Susan Mayer, William Delgado, and Lisa A. Gennetian. 2023. "Education Gradients in Parental Time Investment and Subjective Well-being." Working Paper.

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

Kinsler, Josh, and Ronni Pavan. 2021. "Local Distortions in Parental Beliefs over Child Skill." *Journal of Political Economy* 129 (1): 81–100. Publisher: The University of Chicago Press.

List, John A., Julie Pernaudet, and Dana L. Suskind. 2021. "Shifting Parental Beliefs About Child Development to Foster Parental Investments and Improve School Readiness Outcomes." *Nature Communications* 12 (1): 5765.

Manski, Charles F. 1993. "Adolescent Econometricians: How Do Youth Infer the Returns to

Schooling?" In *Studies of Supply and Demand in Higher Education*, 43–60: University of Chicago Press.
Manski, Charles F. 2004. "Measuring Expectations." *Econometrica* 72 (5): 1329–1376.
Mira, Pedro. 2007. "Uncertain Infant Mortality, Learning, and Life-cycle Fertility." *International Economic Review* 48 (3): 809–846.
Patnaik, Arpita, Joanna Venator, Matthew Wiswall, and Basit Zafar. 2022. "The Role of Heterogeneous Risk Preferences, Discount Rates, and Earnings Expectations in College Major Choice." *Journal of Econometrics* 231 (1): 98–122.
Price, Joseph. 2008. "Parent-Child Quality Time Does Birth Order Matter?" *Journal of Human Resources* 43 (1): 240–265.
Schoonbroodt, Alice. 2018. "Parental child care during and outside of typical work hours." *Review of Economics of the Household* 16 (2): 453–476.
Sovero, Veronica. 2018. "Risk Preferences and Child Investments: Evidence From Mexico." *Review of Economics of the Household* 16 (4): 1027–1047.
Stinebrickner, Ralph, and Todd Stinebrickner. 2014. "Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model." *Journal of Labor Economics* 32 (3): 601–644.
Swamy, P. A. V. B. 1970. "Efficient Inference in a Random Coefficient Regression Model." *Econometrica* 38 (2): 311–323.
Tabetando, Rayner. 2019. "Parental risk aversion and educational investment: panel evidence from rural Uganda." *Review of Economics of the Household* 17 (2): 647–670.
Tanaka, Yuki, and Takashi Yamano. 2015. "Risk and Time Preference on Schooling: Experimental Evidence from a Low-Income Country." Working Paper 14-24, National Graduate Institute for Policy Studies.
Wang, Fan, Esteban Puentes, Jere R. Behrman, and Flávio Cunha. 2022. "You are what your parents expect: Height and local reference points." *Journal of Econometrics*: 105269.
Wiswall, Matthew, and Basit Zafar. 2015. "Determinants of College Major Choice: Identification Using an Information Experiment." *The Review of Economic Studies* 82 (2): 791–824.
Zafar, Basit. 2013. "College Major Choice and the Gender Gap." *Journal of Human Resources* 48 (3): 545–595.

Appendix A. Survey Instrument

A.1. Belief Elicitation - Instructions

Throughout this section, we will refer to different scenarios of health of the baby and hours of interaction the mother spends with the baby.

1) A "normal health" baby is one whose gestation lasted 9 months, weighed 8 pounds, and measured 20 inches at birth. A "poor health" baby is one whose gestation lasted 7 months, weighed 5 pounds, and measured 18 inches at birth.

2) A "high intensity" interaction is one in which the mothers spends 6 hours a day with the baby in active interaction, while a "low intensity" one the mother spends 2 hours a day with the baby in active interaction. These interactions includes activities such as:

- soothing the baby when he/she is upset;
- moving the baby's arms and legs around playfully;
- playing peek-a-boo with the baby;
- singing songs with the baby;
- speaking to the baby;
- feeding, nursing, bathing, attending to health needs;

We would like you to consider a hypothetical scenario involving a mother and her baby. In this scenario, the baby's health can either be good or poor, and the mother's interaction with the baby can be either high intensity or low intensity. After considering these factors, we would like you to determine the youngest, most likely, and oldest age (in months) at which the baby in this specific situation will learn to perform a certain activity.

To illustrate, let us consider the example of a baby with "normal health" and a "low intensity" interaction between mother and baby. In this case, we would like you to provide your personal belief on the youngest, most likely, and oldest ages at which this baby will learn to walk at least 5 steps by itself. To help you understand the youngest, most likely, and oldest age, we suggest imagining 10 identical babies, with some learning to perform the activity at different ages. In this way, the youngest age would be the earliest at which any of the babies learn the activity, the most likely age would be the age at which most of the babies learn the activity, and the oldest age would be the latest at which any of the babies learn the activity.

In total, there will be 16 questions of this nature, each pertaining to a different scenario and activity. While there are no right or wrong answers, we ask that you carefully consider each situation and activity before giving us your honest personal belief.

A.2. Stated Choice - Instructions

In this section, we will refer to the active interaction time a mother spends with her baby as hours of mother-child interaction. Here are some examples of activities a mom does during active time interactions:

- soothing the baby when he/she is upset;
- moving the baby's arms and legs around playfully;
- playing peek-a-boo with the baby;
- singing songs with the baby;
- speaking to the baby;
- feeding, nursing, bathing, attending to health needs;

What is important to highlight is that active interaction time is one where the main and sole focus of the mother is in the baby.

When the mother is at home but not in active interaction time, we call this leisure or passive interaction time. The mother can still be together with the baby, but the baby is not the main focus of the activity. Here are some examples of activities a mom does during passive time interactions:

- Grocery shopping with baby;
- Browsing social media apps on smartphone with baby at your side;
- Nap time for baby;
- Household chores (cleaning, cooking, etc) while baby is at your side;
- Exercising at home or at the gym;

- Watching TV;

In all these activities, although the mother may check on the baby every 5-10 minutes, she is not exclusively focused on the baby during these activities.

We will also refer to the health of the baby.

A “normal health” baby is one whose gestation lasted 9 months, the baby weighed 8 pounds and measured 20 inches at birth. A “poor health” baby is one whose gestation lasted 7 months, the baby weighed 5 pounds and measured 18 inches at birth.

In this section of the survey, we will ask you to imagine yourself in a new household, composed of you, a partner, and a hypothetical baby (that is, not one of your current children). We will present different situations of household income and health of the baby.

Then, we will ask you to imagine a situation where you want to spend 1 hour away from your baby after work every day for one month (not on weekends). For example, you may want to set a time for your personal rest, or you want to practice a hobby. You ask a friend to look after your baby for that hour. Your friend is a careful person who will ensure that your baby is well taken care of, but they will not engage in active interaction with your baby.

We will ask you to choose the highest hourly rate you would be willing to pay your friend during weekdays for one month under different working hour situations. We assume that there are 20 weekdays in the month. If you would not be willing to stay 1 hour away from the baby, please select $0 dollars.

For example, suppose that when you work 8 hours a day every day, the highest rate you would pay your friend is $10. This is equal to 1 hour times 20 weekdays in a month times $10/hour you are willing to pay, which is equal to $200 per month. The answer would look like this:

After you answer this question, we will ask that you choose how many hours you would like to spend with your child outside your work hours in active interaction.

We will ask you these questions while varying the health of the baby, the household income, and how many hours you work every day.

We know these questions are not easy to answer. Note that there is no right or wrong, good or bad, answer, we are just interested in what you personally think. Please try to consider each scenario carefully and tell us what you personally believe is the best option. We ask that you make an effort to thoughfully answer all questions.

Appendix B. Tables and Figures

FIGURE A1. Distribution of Ages by Scenario for "Count 3 Objects Correctly"

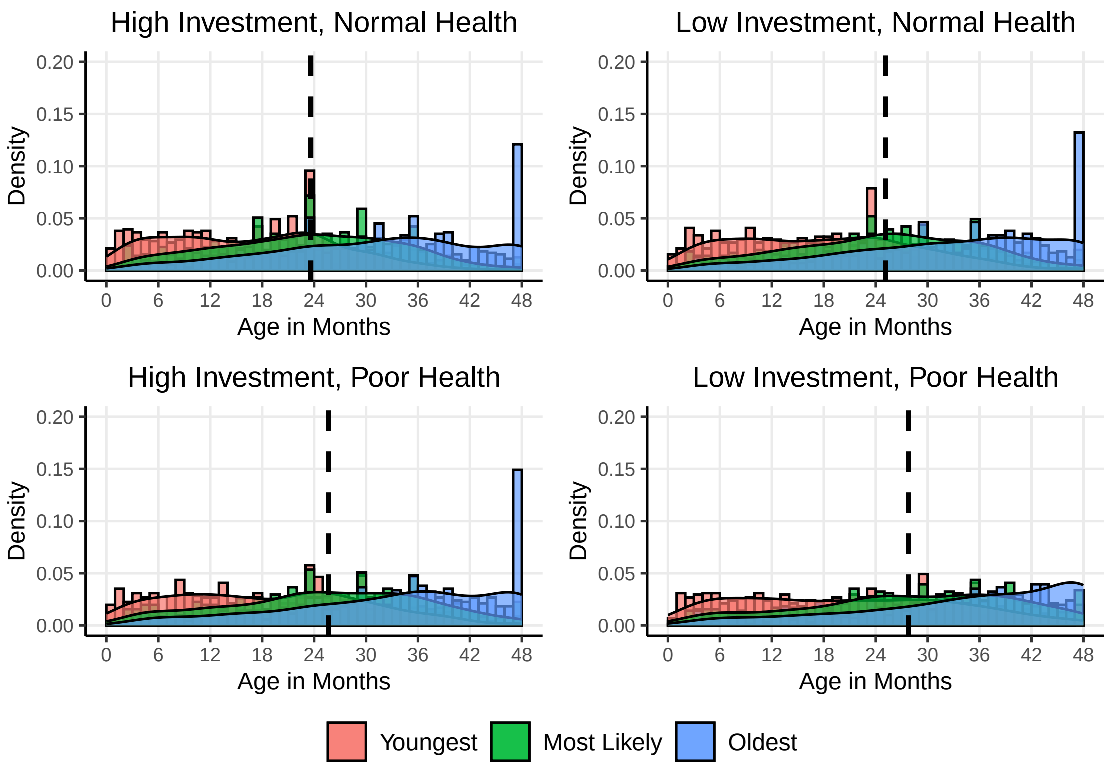


FIGURE A2. Distribution of Ages by Scenario for "Say First and Last Name Together"

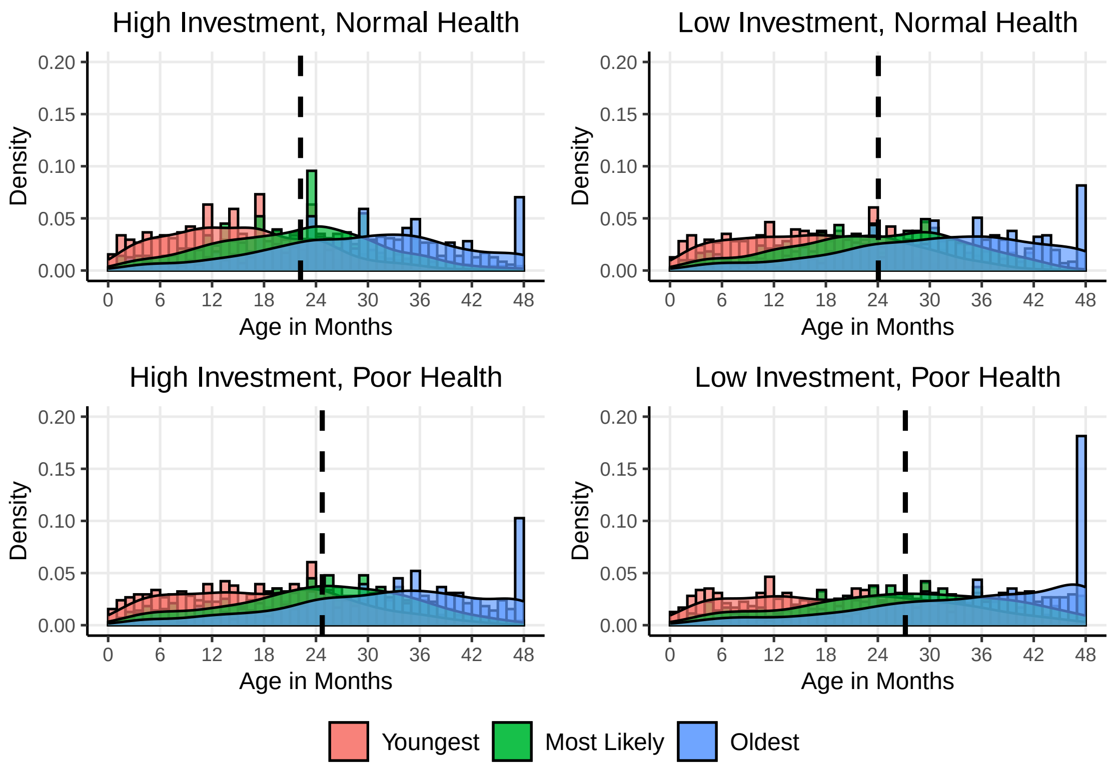


FIGURE A3. Distribution of Ages by Scenario for "Know Own Age and Sex"

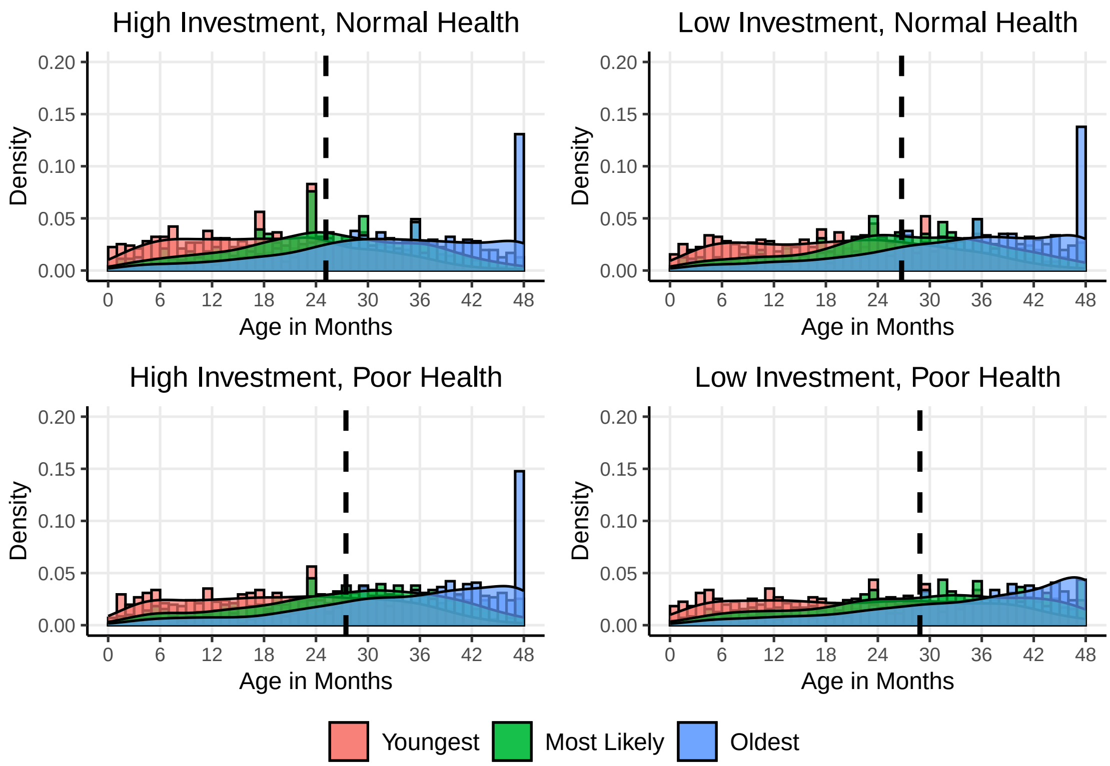


Table A1. Descriptive Statistics Investment

	Good Health, Household Income of $2,000					
	Work 8 Hours		Work 4 Hours		Work 0 Hours	
	Mean	SD	Mean	SD	Mean	SD
Willingness to Pay ($)	12.40	6.72	11.10	6.89	8.14	7.51
Hours of Investment	3.24	1.96	3.84	1.86	4.57	2.40
	Good Health, Household Income of $4,000					
	Work 8 Hours		Work 4 Hours		Work 0 Hours	
	Mean	SD	Mean	SD	Mean	SD
Willingness to Pay ($)	14.15	6.51	12.29	6.56	9.32	7.91
Hours of Investment	3.27	2.02	3.93	1.78	4.61	2.36
	Good Health, Household Income of $6,000					
	Work 8 Hours		Work 4 Hours		Work 0 Hours	
	Mean	SD	Mean	SD	Mean	SD
Willingness to Pay ($)	15.28	7.24	13.30	6.90	10.55	8.16
Hours of Investment	3.38	1.96	3.94	1.79	4.61	2.34
	Poor Health, Household Income of $2,000					
	Work 8 Hours		Work 4 Hours		Work 0 Hours	
	Mean	SD	Mean	SD	Mean	SD
Willingness to Pay ($)	12.42	8.14	11.24	7.73	8.87	8.39
Hours of Investment	3.78	2.10	4.31	1.97	4.76	2.39
	Poor Health, Household Income of $4,000					
	Work 8 Hours		Work 4 Hours		Work 0 Hours	
	Mean	SD	Mean	SD	Mean	SD
Willingness to Pay ($)	13.40	8.00	12.18	7.63	9.51	8.31
Hours of Investment	3.75	2.06	4.36	1.91	4.80	2.37
	Poor Health, Household Income of $6,000					
	Work 8 Hours		Work 4 Hours		Work 0 Hours	
	Mean	SD	Mean	SD	Mean	SD
Willingness to Pay ($)	14.92	8.74	13.47	8.12	11.03	8.98
Hours of Investment	3.78	2.15	4.28	2.01	4.82	2.34

This table shows

Table A2. Estimates of Willingness to Pay System

Panel B: Aggregate Estimates	
p_i	12.106***
	(0.242)
Work Hours	1.702***
	(0.118)
Income	1.053***
	(0.083)
θ_0	0.010
	(0.109)
Panel B: Individual Estimates	
Parameter	% Significant
p_i	99.427%

Note: Robust standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. This table shows aggregate estimates of the willingness to pay equations 11 and the individual-level predicted estimates of p_i. The percentage of significant estimates is calculated using a 10% significance level, and the null hypothesis is that the coefficient is equal to zero.