Research on Parental Beliefs

Flavio Cunha
Department of Economics and Texas Policy Lab
Rice University

March 3, 2021

Hart and Risley, 1995

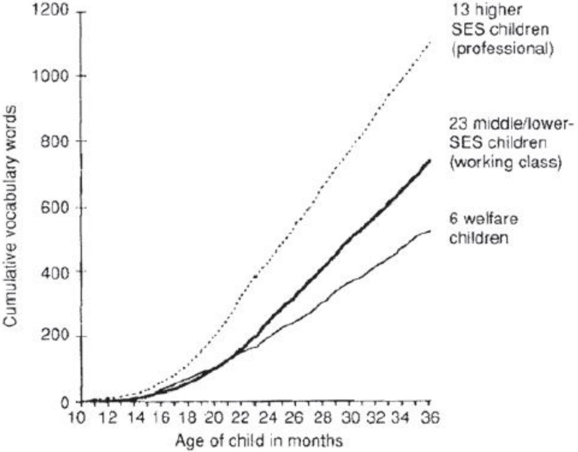
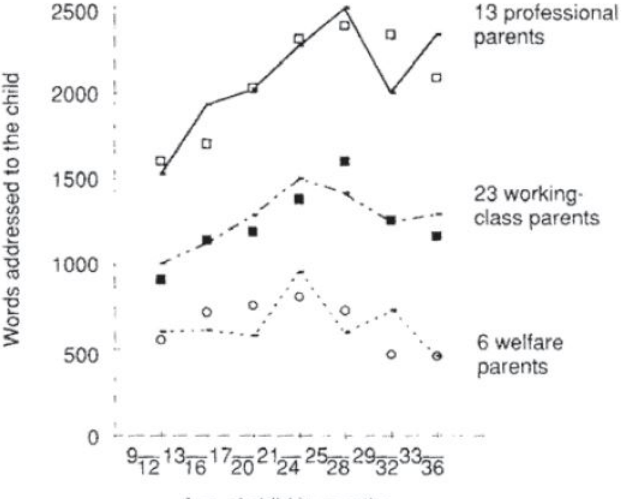
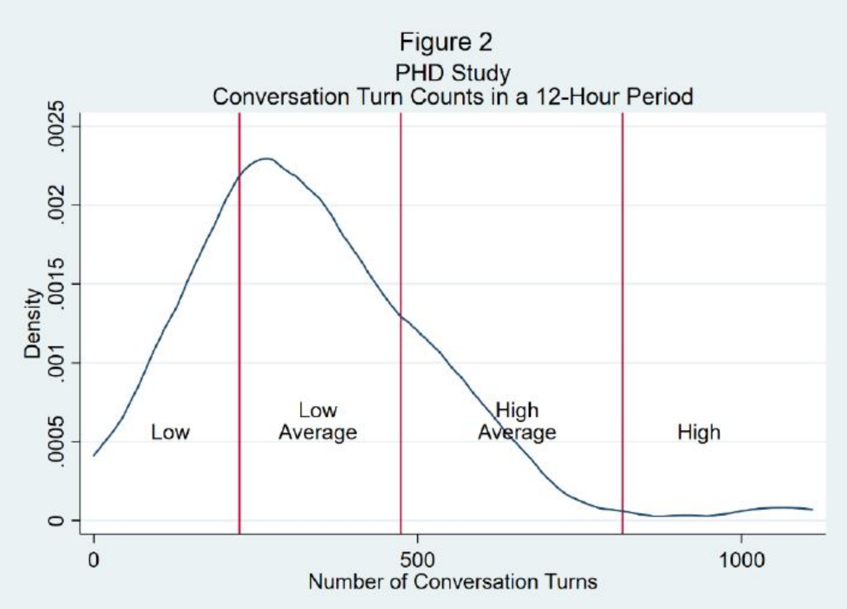


Figure 2. The widening gap we saw in the vocabulary growth of children from professional, working-class, and welfare families across their first 3

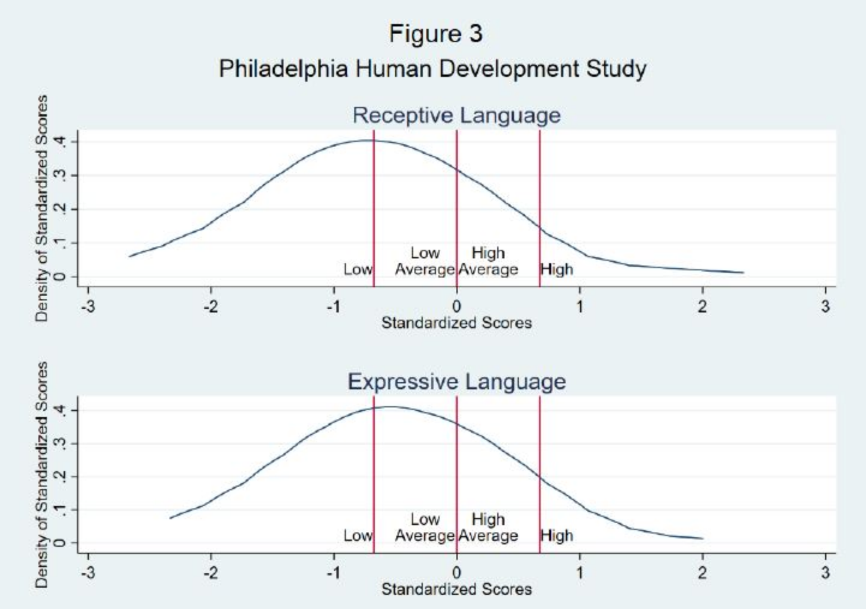
Hart and Risley, 1995



Cunha, Gerdes, and Nihtianova, 2020



Cunha, Gerdes, and Nihtianova, 2020



Cunha, Gerdes, and Nihtianova, 2020

Table 3
Correlation between HOME and Conversational Turns with Quartiles of Family Income
PHD Study

VARIABLES	(1) HOME Score	(2) Conversational Turns
Second quartile of family income	0.591***	0.093
	(0.170)	(0.164)
Third quartile of family income	0.933***	0.412**
	(0.164)	(0.191)
Fourth quartile of family income	1.068***	0.011
	(0.147)	(0.155)
Constant	-0.442***	-1.938***
	(0.140)	(0.478)
Observations	234	239
R-squared	0.196	0.094

Cunha, Gerdes, and Nihtianova, 2020

Table 2
Between and Within Sum of Squares as Fractions of Total Sum of Squares
Inner-City and Suburban Samples
PHD Study Data

Variable	Between	Within
Standardized HOME Score	11.2%	88.8%
Conversational Turn Counts (12 hours)	1.3%	98.7%
Standardized BSID Language Composite Score	14.3%	85.7%

Cunha, Gerdes, and Nihtianova, 2020

Table 4
Correlation between Language Development with HOME and LENA Measures
Dependent Variable: Standardized Bayley Scales of Infant Development Language Composite Score

Variables	Model 1	Model 2	Model 3	Model 4
	Panel A: Standardize HOME Score Only			
Standardized HOME Score	0.261***	0.099**	0.084**	0.070*
	(0.065)	(0.044)	(0.042)	(0.042)
	Panel B: Standardized Conversational Turns Only			
Standardized Conversational Turn Counts	0.193***	0.165***	0.153***	0.142***
	(0.066)	(0.058)	(0.057)	(0.055)
	Panel C: Standardized Adult Word Counts Only			
Standardized Adult Word Counts	0.100*	0.102*	0.101*	0.088
	(0.059)	(0.054)	(0.054)	(0.055)
	Panel D: Standardized TV Time			
Standardized TV Time	-0.170***	-0.013	-0.002	-0.004
	(0.059)	(0.057)	(0.057)	(0.057)
Demographic characteristics	N	Y	Y	Y
Dummy for Inner-City Sample	N	N	Y	Y
Dummies for Quartiles of Family Income	N	N	N	Y

Introduction

Two papers:

1. Cunha, Elo, and Culhane (CEC, Journal of Econometrics, 2020).
   - How do we estimate parental beliefs?
   - Do parental beliefs predict investments in early childhood?

2. Cunha, Gerdes, and Nihtianova (CGM, TPL Working Paper, 2020).
   - Are parental beliefs malleable?
   - Does an increase in parental beliefs cause an increase in parental investments?

Introduction: CEC

A static model of parental beliefs and early investments.

Survey instrument.

From parental reports to parental expectations about child development.

From parental expectations about child development to beliefs about the parameters of the technology of skill formation.

The Philadelphia Human Development Study.

Results.

A Static Model of Beliefs and Investments

Utility function:

$$U(c_i, h_{i,1}) = \ln c_i + \alpha_{i,1} \ln h_{i,1} + \alpha_{i,2} \ln x_i \quad (1)$$

c_i is the household consumption.

$h_{i,1}$ is the child's human capital at the end of the period.

x_i is the early investment.

A Static Model of Beliefs and Investments

The budget constraint

$$c_i + p_i x_i = y_i \tag{2}$$

p_i is the relative price of the investment good.

y_i is the household income.

A Static Model of Beliefs and Investments

The technology of skill formation:

$$\ln h_{i,1} = \psi_0 + \psi_1 \ln h_{i,0} + \psi_2 \ln x_i + \psi_3 \ln h_{i,0} \ln x_i + \nu_i \tag{3}$$

$h_{i,0}$ is the child's human capital at the end of the period.

A Static Model of Beliefs and Investments

Let Ψ_i denote the parent's information set.

Define $\mathbf{E}\left(\psi_j \mid h_{i,0}, x_i, \Psi_i\right) = \mu_{i,j}$.

Assume, although not necessary, that $\mathbf{E}\left(\nu_i \mid h_{i,0}, x_i, \Psi_i\right) = 0$.

Then, the perceived technology of skill formation is:

$$\mathbf{E}\left(\ln h_{i,1} \mid h_{i,0}, x_i, \Psi_i\right) = \mu_{i,0} + \mu_{i,1} \ln h_{i,0} + \mu_{i,2} \ln x_i + \mu_{i,3} \ln h_{i,0} \ln x_i \tag{4}$$

The Parent's Problem

The parent's problem is to maximize the expected utility subject to:

- The budget constraint.
- The perceived technology of skill formation.

The solution is:

$$x_i = \left[\frac{\alpha_{i,1} \left(\mu_{i,2} + \mu_{i,3} \ln h_{i,0} \right) + \alpha_{i,2}}{1 + \alpha_{i,1} \left(\mu_{i,2} + \mu_{i,3} \ln h_{i,0} \right) + \alpha_{i,2}} \right] \frac{y_i}{p_i} \tag{5}$$

We cannot separately identify $\alpha_{i,1}$, $\alpha_{i,2}$, $\mu_{i,1}$, $\mu_{i,2}$ if we only observe "typical" variables in datasets: $h_{i,0}$,$h_{i,1}$,x_i, p_i, and y_i.

Do we need to separately identify all these terms?

It depends on the question. Define the "parameter" κ_i:

$$\kappa_i = \left[\frac{\alpha_{i,1} \left(\mu_{i,2} + \mu_{i,3} \ln h_{i,0} \right) + \alpha_{i,2}}{1 + \alpha_{i,1} \left(\mu_{i,2} + \mu_{i,3} \ln h_{i,0} \right) + \alpha_{i,2}} \right] \tag{6}$$

Consider two questions:

Question 1: *What will happen with investments if we have an exogenous increase in household income that does not impact preferences, information set, or $h_{i,0}$?*

Question 2: *What will happen with investments if the parents participate in a parenting program that raises $\mu_{i,2}$ by 10% but does not impact preferences, household income, or $h_{i,0}$?*

Steps in the Construction of the Survey Instrument

IRT analysis of the Motor-Social Development (MSD) Scale.

Create hypothetical scenarios of h_0 and x.

Adapt MSD items for the elicitation of expected development given scenarios for h_0 and x.

MSD Instrument

SECTION 3: MOTOR AND SOCIAL DEVELOPMENT

PART B: (22 MONTHS - 3 YEARS, 11 MONTHS)

MOTHER/GUARDIAN:

If ______________________________ is **at least 22 months** old, but **not yet 4 years** old, please answer these 15 questions.
Child's Name

1. Has your child ever let someone know, without crying, that wearing wet (soiled) pants or diapers bothered him/her? YES.... 1 NO..... 0 72/

2. Has your child ever spoken a partial sentence of 3 words or more? YES.... 1 NO..... 0 73/

3. Has your child ever walked upstairs by himself/herself without holding on to a rail? YES.... 1 NO..... 0 74/

4. Has your child ever washed and dried his/her hands

Item Response Theory

Let a_i denote the child's age at the time of the MSD assessment.

Let $d_{i,j}^*$ denote the following latent variable:

$$d_{i,j}^* = b_{0,j} + b_{1,j}\left(\ln a_i + \frac{b_{2,j}}{b_{1,j}}\theta_i\right) + \eta_{i,j} \tag{7}$$

θ_i is deviation from typical development for age. We observe:

$$d_{i,j} = \begin{cases} 1, & \text{if } d_{i,j}^* \geq 0 \\ 0, & \text{if } d_{i,j}^* < 0 \end{cases} \tag{8}$$

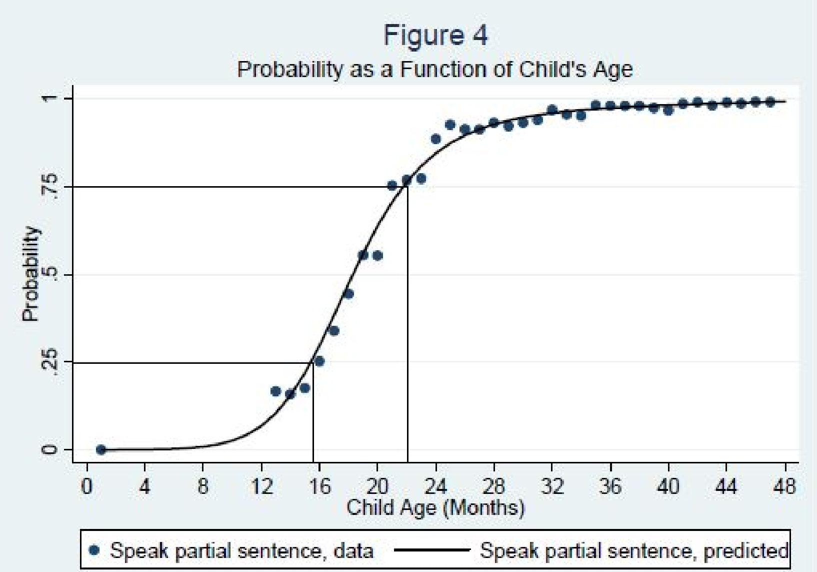


Figure 4
Probability as a Function of Child's Age
Probability
0
.25
.5
.75
1
0
4
8
12
16
20
24
28
32
36
40
44
48
Child Age (Months)
Speak partial sentence, data
Speak partial sentence, predicted

Scenarios (see CEC, 2013)

Panel C: Definition of the Scenarios				
	Human capital at birth		Investments	
Scenario number	Short description	Full description	Short description	Full description
1	"High"	Pregnancy lasted 9 months, birth weight is 8 pounds, birth length is 20 inches.	"High"	6 hours per day doing HOME Scale activities.
2	"Low"	Pregnancy lasted 7 months, birth weight is 5 pounds, birth length is 18 inches.	"High"	6 hours per day doing HOME Scale activities.
3	"High"	Pregnancy lasted 9 months, birth weight is 8 pounds, birth length is 20 inches.	"Low"	2 hours per day doing HOME Scale activities.
4	"Low"	Pregnancy lasted 7 months, birth weight is 5 pounds, birth length is 18 inches.	"Low"	2 hours per day doing HOME Scale activities.

Fig. 1. This figure provides detailed information about both forms of the elicitation instrument. Panel A reproduces the elicitation items in the subjective probability form. Panel B displays the elicitation items in the age range form. Panel C describes the scenarios of human capital at birth and investments. The study participants watched a short video describing these scenarios.

MSD Adaptation

Two types of elicitation questions:

How likely is it that a baby will [MSD Item] by age 24 months if scenario is (h_0, x)?

What are the youngest and oldest ages a baby will [MSD Item] if scenario is (h_0, x)?

The instrument uses four MSD items (see CEC, 2013):

1. Speaks partial sentence with three words or more.
2. Learns to say his/her first and last names together without anyone's help.
3. Counts three objects correctly.
4. Learns his or her own age and sex.

From answers to expectation about the natural log of human capital

In ”How Likely” questions, parents supply a probability between 0% and 100%.

In ”Age Range” questions, parents supply numbers between 0 and 48 months.

We transform the answers into error-ridden measures of $\mathbf{E}\left(\ln h_{i,1} \mid h_{i,0}, x_i, \Psi_i\right)$.

We account for measurement error.

We investigate the construct’s predictive validity.

Do parental beliefs at pregnancy predict investments one year later?

"How Likely"

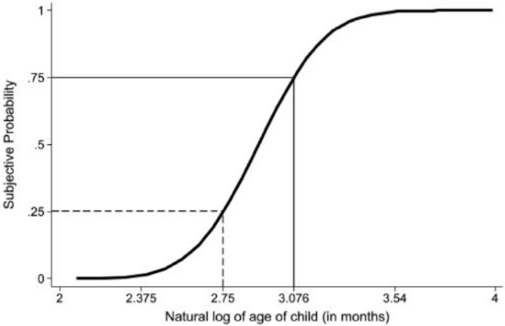


Fig. 2. This figure shows how we use the IRT model to relocate and rescale maternal subjective probability reports (shown in the vertical axis) to error-ridden measures of the expectation of the natural log of human capital at age two years (shown in the horizontal axis) for two scenarios of investments ("high" vs. "low") when human capital at birth is "high". When the investment is "low", the mother reports that there is a 25% chance that the child will learn how to speak a partial sentence with three words or more by age 24 months. When the investment is "high", the mother reports that the probability is 75% by age 24 months. These probabilities correspond to $2.75 = ln\ 16$ and $3.076 = ln\ 22$, respectively. According to the IRT model, 25% of the 16-month-old children and 75% of the 22-month-old children speak a partial sentence with three words or more.

"Age Ranges"

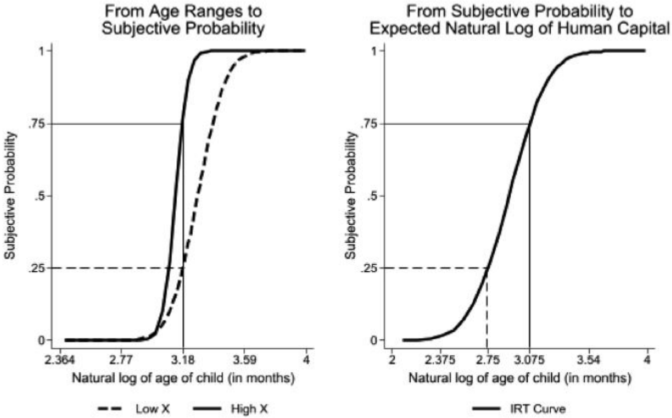


Fig. 3. This figure shows the two steps involved in transforming age ranges to error-ridden measures of expectation of the natural log of human capital at age 24 months. In the first step, which we show in the left panel, we use maternal reports of the age ranges and assumptions about the interpolating function and the parameters Δ_0 and Δ_1. For this figure, we assume that the interpolating function is the normal CDF, $\Delta_0 = 10\%$ and $\Delta_1 = 90\%$. We show the transformation from age ranges to subjective probability for two scenarios of investments ("high" vs. "low") when human capital at birth is "high". When the investment is "low", the mother reports age ranges equal to 22 and 32 months. As a result, the probability that the child will learn how to "speak a partial sentence with three words or more" by age 24 months is 25%. When the investment is "high", the age ranges are 21 and 25 months, which corresponds to a probability of 75%. The right panel shows the transformation from subjective probability to

”Age Ranges”

We construct ”bounds” (upper and lower triangular, see CEC, 2013, 2020).

We consider different points of ”arbitrarily” close to zero (see Delavande, 2015).

Identification if no Measurement Error: GMM

Let $\ln h^L_{i,j,k}$ and $\ln h^A_{i,j,k}$ denote subjective beliefs for parent i, MSD item j, and scenario k as elicited by the ”how likely” and ”age ranges” instruments, respectively.

If there is no measurement error, then:

$$\ln h^L_{i,j,k} = \mathbf{E} \left(\ln h_{i,1} \middle| h_{0,k}, x_k \Psi_i \right)$$

$$\ln h^A_{i,j,k} = \mathbf{E} \left(\ln h_{i,1} \middle| h_{0,k}, x_k \Psi_i \right)$$

Identification if no Measurement Error: GMM

$$M_{i,1} = \underbrace{\frac{\Delta \mathbf{E}\left(\ln h_{i,1} | \overline{h}_0, \Psi_i\right)}{\Delta \ln x \times \Delta \ln h_0}}_{\text{difference between scenarios 1 and 3}} - \underbrace{\frac{\Delta \mathbf{E}\left(\ln h_{i,1} | \underline{h}_0, \Psi_i\right)}{\Delta \ln x \times \Delta \ln h_0}}_{\text{difference between scenarios 2 and 4}} = \mu_{i,3}$$

$$M_{i,2} = \frac{\ln \overline{h}_0 \Delta \mathbf{E}\left(\ln h_{i,1} | \underline{h}_0, \Psi_i\right)}{\Delta \ln x \times \Delta \ln h_0} - \frac{\ln \underline{h}_0 \Delta \mathbf{E}\left(\ln h_{i,1} | \overline{h}_0, \Psi_i\right)}{\Delta \ln x \times \Delta \ln h_0} = \mu_{i,2}$$

$$M_{i,3} = \frac{\ln \bar{x}_0 \Delta \mathbf{E}\left(\ln h_{i,1} | \underline{x}, \Psi_i\right)}{\Delta \ln x \times \Delta \ln h_0} - \frac{\ln \underline{x} \Delta \mathbf{E}\left(\ln h_{i,1} | \bar{x}, \Psi_i\right)}{\Delta \ln x \times \Delta \ln h_0} = \mu_{i,1}$$

Identification with Measurement Error

With measurement error, then:

$$\ln h^{L}_{i,j,k} = \mathbf{E}\left(\ln h_{i,1} \middle| h_{0,k}, x_k \Psi_i\right) + \epsilon^{L}_{i,j,k}$$

$$\ln h^{A}_{i,j,k} = \mathbf{E}\left(\ln h_{i,1} \middle| h_{0,k}, x_k \Psi_i\right) + \epsilon^{A}_{i,j,k}$$

Remember that:

$$\mathbf{E}\left(\ln h_{i,1} \middle| h_{0,k}, x_k \Psi_i\right) = \mu_{i,0} + \mu_{i,1} \ln h_{0,k} + \mu_{i,2} \ln x_k + \mu_{i,3} \ln h_{0,k} \ln x_k$$

Identification with Measurement Error

So:

$$\ln h^L_{i,j,k} = \mu_{i,0} + \mu_{i,1} \ln h_{0,k} + \mu_{i,2} \ln x_k + \mu_{i,3} \ln h_{0,k} \ln x_k + \epsilon^L_{i,j,k}$$

$$\ln h^A_{i,j,k} = \mu_{i,0} + \mu_{i,1} \ln h_{0,k} + \mu_{i,2} \ln x_k + \mu_{i,3} \ln h_{0,k} \ln x_k + \epsilon^A_{i,j,k}$$

Two ways we can approach this problem.

1. Random coefficient regression model.
2. Factor model.

Random coefficient regression model

Define $Z_k = (1, \ln h_{0,k}, \ln x_k, \ln h_{0,k} \ln x_k)$ and $\mu_i = \mu + u_i$. Assume:

$$\mathbf{E}\left(u_i | Z_k \right) = 0$$

$$\mathbf{E}\left(u_i u_i' | Z_k \right) = \Gamma$$

The first assumption states that deviations from mean belief are orthogonal to scenarios, which is reasonable as scenarios are set in exogenous fashion.

The second assumption allows for non-diagnonal variance-covariance matrix of deviations from mean beliefs.

We never assume that $\mu = \psi$.

Random coefficient regression model

Define $\omega^f_{i,j,k} = Z_k u_i + \epsilon^f_{i,j,k}$ for $f \in \{A, L\}$. Assume:

$$\mathbf{E}\left(\omega^f_{i,j,k}\middle| Z_k\right) = 0$$

$$\mathbf{E}\left(\omega^f_{i,j,k}\omega^L_{i,j,k}{}'\middle| Z_k\right) = Z_k \Gamma Z_k' + \sigma^f_i I$$

The first assumption states that measurement error is orthogonal to scenarios.

The second assumption allows the variance of measurement error to vary across individuals, but does not allow for correlation across measurement errors. We will relax it in the factor approach.

Factor Approach

Consider the model:

$$\ln h^f_{i,j,k} = \mu_{i,0} + \mu_{i,1} \ln h_{0,k} + \mu_{i,2} \ln x_k + \mu_{i,3} \ln h_{0,k} \ln x_k + \epsilon^f_{i,j,k}$$

Note that:

- μ_i are the latent factors.
- $Z_k = (1, \ln h_{0,k}, \ln x_k, \ln h_{0,k} \ln x_k)$ are the factor loadings.
- $\epsilon_{fi,j,k}$ are the uniquenesses.

Factor Approach

Assumption 2: $\epsilon^f_{i,j,k} = \alpha_k \lambda^f_{i,j}$

Assumption 3: $\lambda^f_{i,j}$ is uncorrelated with $\lambda^{f'}_{i,j'}$ for $f \neq f'$ or $j \neq j'$.

Assumptions (2) and (3) allow for correlation in measurement error across scenarios conditional on the elicitation form f and the MSD item j.

Natural assumption because of the structure of the elicitation form.

The PHD Study

We enrolled 822 women at the second trimester of first pregnancy and measured beliefs (CEC, 2013).

When children were 8-12 months old, we measured investments (HOME for all, LENA for a random subsample 33%).

When children were 22-26 months old, we assessed development (Bayley Scales).

When children were 28-34 months old, we piloted a parenting-directed language invtervention (136 parents).

Table 1
Demographic characteristics of PHD study participants.

Characteristic	Percentage
Year of birth	
Mother born between 1968 and 1977	3.54%
Mother born between 1978 and 1987	33.21%
Mother born between 1988 and 1987	63.25%
Race and ethnicity	
Mother is Hispanic	13.02%
Mother is non-Hispanic black	53.77%
Mother is non-Hispanic white	26.64%
Other	6.57%
Educational attainment	
Less than high school diploma	42.09%
High school or some college	41.36%
Four-year college diploma or higher	16.55%
Marital status	
Single[a]	60.71%
Cohabiting	9.49%
Married	29.81%
Center for epidemiological studies depression scale	
The score is greater than or equal to 16	29.32%
Household income per year (y)	
$y <$ \$25,000	44.77%
\$25,000 $\leq y <$ \$55,0000	20.56%
\$55,000 $\leq y <$ \$105,0000	16.06%
$y \geq$ \$105,000	18.61%
Sample size	
First wave[b]	822
Second wave[c]	687

Raw Data: How Likely - One Item

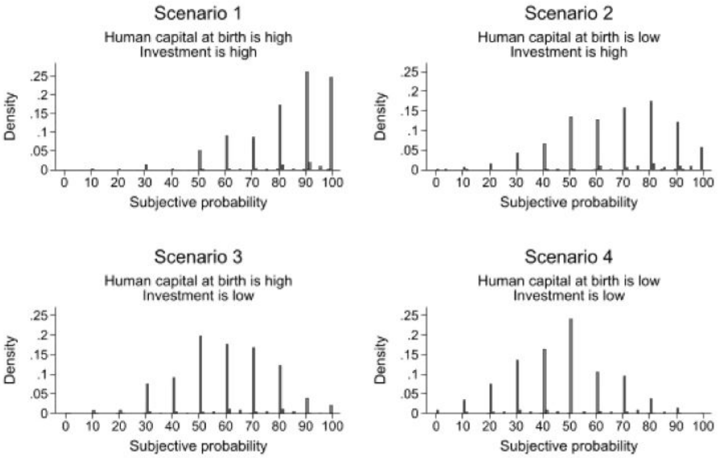


Fig. 4. This figure shows the histograms of maternal reports of subjective probability for the MSD item "child speaks a partial sentence with three words or more" for all scenarios of human capital at birth and investments. The figure shows a pattern of answers that indicates that the higher human capital at birth or investment, the higher the likelihood that the child will be able to accomplish this task by age 24 months.

Raw Data: How Likely - Many Items

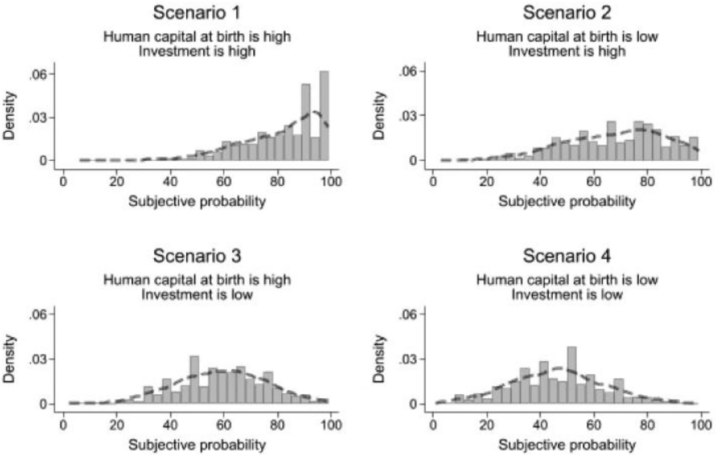


Fig. 5. This figure shows the histograms of subjective probability after we average maternal reports across the MSD items for each scenario of human capital at birth and investment. The result is that subjective probabilities are far less likely to exhibit heaping that we observed in Fig. 4.

Raw Data: How Likely, IRT Inversion, Many Items

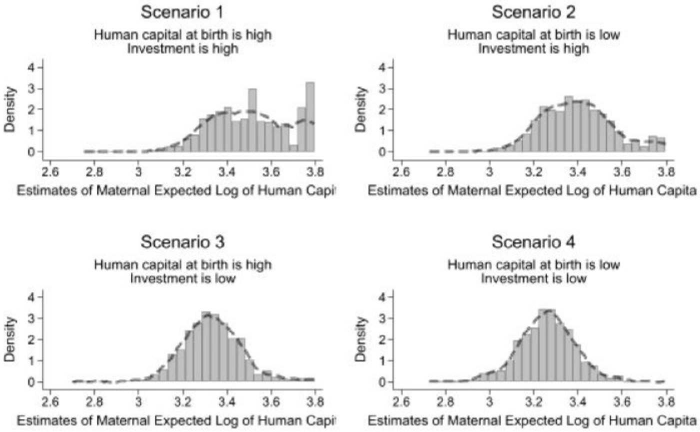


Fig. 6. This figure displays the histograms of error-ridden measures of the expectation of the natural log of human capital at age two years for each scenario of human capital at birth and investments. To produce these measures, we proceed in two steps. In the first step, we transform, for each MSD item and scenario, the subjective probability data to an error-ridden, MSD-item specific, measure of the expectation of the natural log of human capital. In the second step, we average the measures across MSD items for each scenario. We then plot the histograms of the averaged measures.

Raw Data: Age Range - One Item

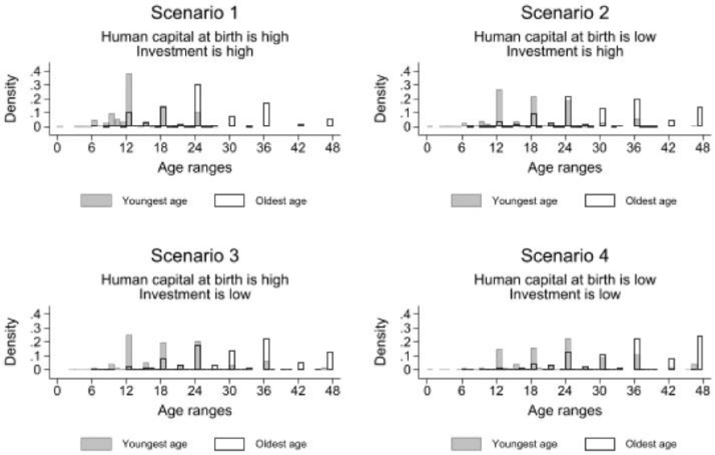


Fig. 7. This figure plots the histograms of maternal reports of age ranges for the MSD item "speak a partial sentence with three words or more". The solid gray bars show the youngest ages children will learn this MSD task for each scenario of human capital at birth and investments. The solid white bars show the oldest ages. This figure shows heaping around multiples of six months. This heaping pattern is similar to the one observed for the subjective probability data that we show in Fig. 4.

Raw Data: Age Range - Probability, One Item

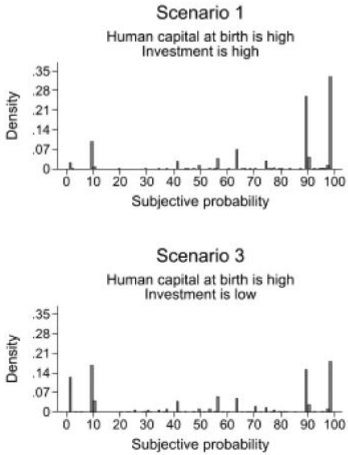


Fig. 8. This figure shows the subjective probability data after transformation from age ranges. In this figure, we show the data for the MSD item "speak a partial sentence with three words or more". To transform the data from age ranges to subjective probability, we follow the steps described in Section 2.4.2. We assume that the interpolating function is the normal CDF and that the parameters $\Delta_0 = 10\%$ and $\Delta_1 = 90\%$. We note that the subjective probability data suffers from heaping, but unlike the one in Fig. 4, the heaping in this data occurs at the extremes.

Raw Data: Age Range, IRT Inversion, One Item

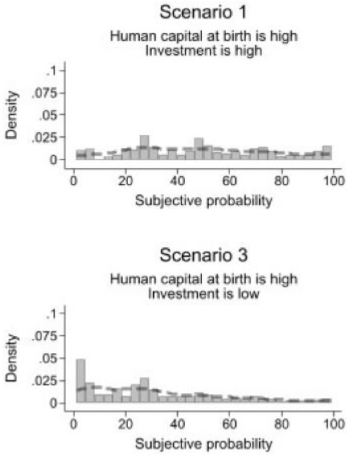


Fig. 9. This figure shows the histograms of the subjective probability average across MSD items for each scenario. In this figure, we use only the data from the age ranges elicitation form. We follow two steps to produce this figure. In the first step, we transform the age ranges data to subjective probability for each MSD item. To do so, we assume that the interpolating function is a normal CDF and the parameters $\Delta_0 = 10\%$ and $\Delta_1 = 90\%$. In the second step, we average subjective probabilities across MSD items for each scenario. The histograms show that the averaged data do not feature as much heaping.

Raw Data: Age Range, IRT Inversion, Many Items

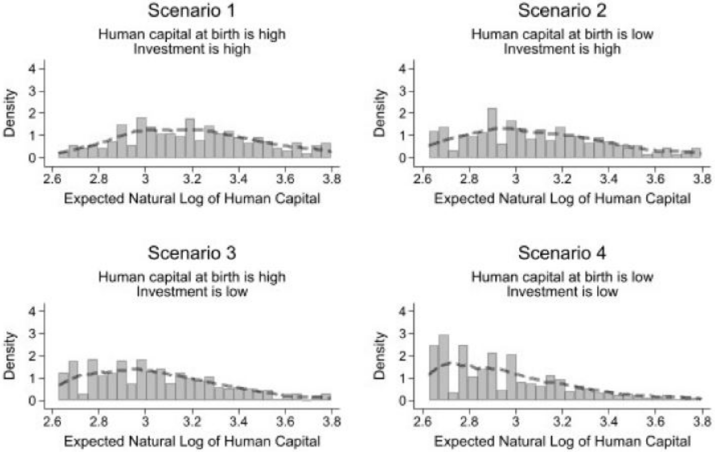


Fig. 10. This figure shows the histograms of the averaged error-ridden measures of the expectation of the natural log of human capital for each scenario when we average across MSD items for the age ranges elicitation form. To produce this figure, we follow three steps. In the first step, we use the normal CDF as the interpolating function (with parameters $\Delta_0 = 10\%$ and $\Delta_1 = 90\%$), to transform the age ranges data to subjective probability data. We do so for each MSD item and each scenario separately. Then, we use the IRT model (3), for each MSD item and scenario, to relocate and rescale the subjective probability data to the expectation of the natural log of human capital at age two years. In the last step, we average across MSD items for each scenario.

Raw Data: All Forms, All Items

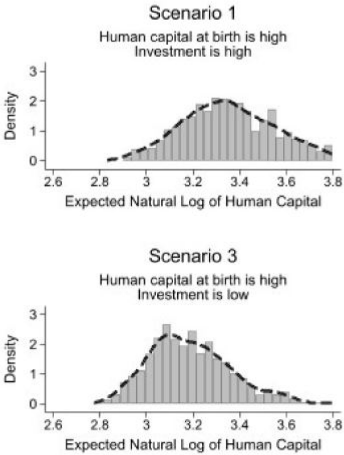


Fig. 11. This figure shows the histograms of the averaged error-ridden measures of the expectation of the natural log of human capital for each scenario when we average both across MSD items and elicitation forms. We remark that Figs. 6 and 10 present the histograms when we average across MSD items but not across elicitation forms. Once we average across both MSD items and elicitation forms, as shown in this figure, we have eliminated all forms of heaping. The estimator of the MSE also averages the data across MSD items and elicitation forms but does so in an efficient manner.

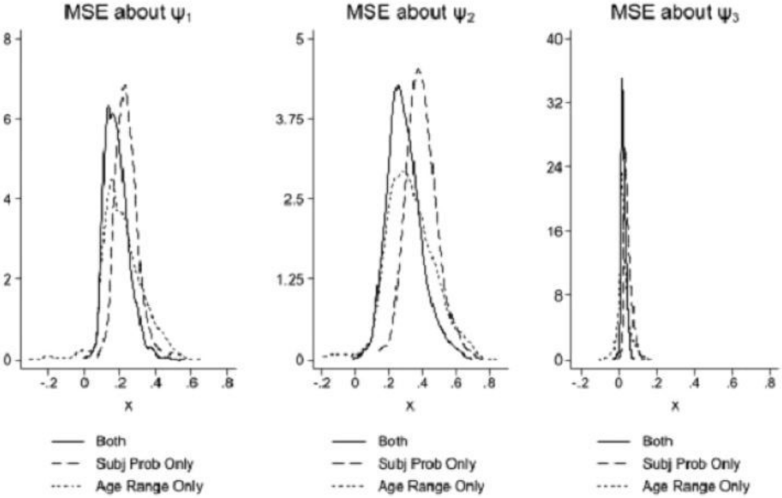
How Likely: Measurement Error

Equation Number	Latent Variable	MSD Item	1	2	3	4	5	6
1	$\mu_{\psi,3}$	1	0.008	-0.034	-0.041	**-0.886**	-0.044	**0.170**
2	$\mu_{\psi,3}$	2	-0.043	-0.053	**-0.842**	-0.062	-0.048	**0.257**
3	$\mu_{\psi,3}$	3	-0.048	-0.070	-0.033	-0.031	**-0.861**	**0.277**
4	$\mu_{\psi,3}$	4	-0.056	**-0.831**	-0.081	-0.030	-0.056	**0.249**
5	$\mu_{\psi,2}$	1	**0.693**	-0.047	-0.040	**0.489**	-0.100	-0.033
6	$\mu_{\psi,2}$	2	**0.777**	-0.054	**0.369**	-0.061	-0.047	0.008
7	$\mu_{\psi,2}$	3	**0.827**	-0.059	-0.068	-0.057	**0.381**	-0.002
8	$\mu_{\psi,2}$	4	**0.767**	**0.401**	-0.043	-0.051	-0.073	0.015
9	$\mu_{\psi,1}$	1	-0.024	0.054	0.045	**0.841**	0.056	**0.300**
10	$\mu_{\psi,1}$	2	0.029	0.076	**0.826**	0.036	0.035	**0.341**
11	$\mu_{\psi,1}$	3	0.008	0.056	0.046	0.045	**0.841**	**0.347**
12	$\mu_{\psi,1}$	4	0.011	**0.801**	0.052	0.068	0.103	**0.336**

Age Ranges: Measurement Error

Equation Number	Latent Variable	MSD Item	1	2	3	4	5	6
1	$\mu_{\psi,3}$	1	**-0.930**	-0.020	-0.026	-0.024	0.046	**0.056**
2	$\mu_{\psi,3}$	2	0.024	-0.004	**-0.896**	-0.001	0.021	**0.129**
3	$\mu_{\psi,3}$	3	-0.001	**-0.908**	-0.026	-0.027	-0.014	**0.147**
4	$\mu_{\psi,3}$	4	-0.001	-0.025	0.011	**-0.888**	0.003	**0.180**
5	$\mu_{\psi,2}$	1	**0.755**	-0.010	-0.019	-0.020	**0.188**	-0.054
6	$\mu_{\psi,2}$	2	0.010	-0.013	**0.638**	0.004	**0.411**	-0.024
7	$\mu_{\psi,2}$	3	0.016	**0.605**	-0.051	-0.049	**0.447**	0.005
8	$\mu_{\psi,2}$	4	-0.026	-0.005	-0.026	**0.617**	**0.422**	0.020
9	$\mu_{\psi,1}$	1	**0.858**	0.007	-0.014	0.002	-0.023	**0.203**
10	$\mu_{\psi,1}$	2	0.017	0.036	**0.886**	-0.002	0.011	**0.183**
11	$\mu_{\psi,1}$	3	0.019	**0.865**	0.034	0.021	0.010	**0.188**
12	$\mu_{\psi,1}$	4	0.032	0.001	0.027	**0.863**	-0.003	**0.164**

Heterogeneity in Beliefs



Mean Beliefs

	Both Forms	Subjective-Probability Form	Age-Range Form
$\hat{\mu}_{\psi,1}$	0.182	0.238	0.216
	(0.008)	(0.010)	(0.011)
$\hat{\mu}_{\psi,2}$	0.293	0.396	0.330
	(0.009)	(0.011)	(0.012)
$\hat{\mu}_{\psi,3}$	0.028	0.050	0.029
	(0.006)	(0.009)	(0.009)
	Test of Parameter Constancy		
H_0: $\mu_{i,\psi,j} = \mu_{\psi,j}\ \forall\ i, j$ H_1: $\mu_{i,\psi,j} \neq \mu_{\psi,j}$ *for at least one* i, j	Reject H_0	Reject H_0	Reject H_0
	Hypotheses Tests (% Reject H_0)		
H_0: $\mu_{i,\psi,1} = 0$; H_1: $\mu_{i,\psi,1} \neq 0$	68.0%	92.0%	44.3%
H_0: $\mu_{i,\psi,2} = 0$; H_1: $\mu_{i,\psi,2} \neq 0$	88.2%	98.7%	71.3%
H_0: $\mu_{i,\psi,3} = 0$; H_1: $\mu_{i,\psi,3} \neq 0$	5.7%	14.5%	2.6%

Note: Generalized least squares standard error in parentheses.

The Parent's Problem

Given that we cannot reject that $\mu_{i,3}$ is equal to zero for approximately 95% of the sample, the solution is:

$$x_i = \left[\frac{\alpha_{i,1}\mu_{i,2} + \alpha_{i,2}}{1 + \alpha_{i,1}\mu_{i,2} + \alpha_{i,2}} \right] \frac{y_i}{p_i} \tag{9}$$

Note that neither $\mu_{i,1}$ nor $\mu_{i,3}$ should predict investments.

Correlation with Household Income

	Both Forms		
Dummies for household income per year (y)	$\mu_{i,\psi,1}$	$\mu_{i,\psi,2}$	$\mu_{i,\psi,3}$
1($25,000 ≤ y < $55,0000)	0.22**	0.35***	0.19*
	(0.10)	(0.09)	(0.10)
1($55,000 ≤ y < $105,0000)	-0.17	0.37***	-0.25**
	(0.13)	(0.12)	(0.11)
1(y ≥ $105,000)	-0.51***	0.47***	-0.53***
	(0.13)	(0.14)	(0.12)
Observations	822	822	822
R^2	0.071	0.064	0.090

Correlation: MSE and Structural Factors

Table 9
Relationship between MSE parameters and structural factors.

Variable	Subjective-probability form		
	$\mu_{i,\psi,1}$	$\mu_{i,\psi,2}$	$\mu_{i,\psi,3}$
Structural Factor 1	−0.294***	0.966***	−0.052
	(0.010)	(0.023)	(0.033)
Structural Factor 6	0.892***	0.031*	0.677***
	(0.017)	(0.017)	(0.043)
Observations	822	822	822
R^2	0.935	0.871	0.468
Variable	**Age-range form**		
	$\mu_{i,\psi,1}$	$\mu_{i,\psi,2}$	$\mu_{i,\psi,3}$
Structural Factor 5	−0.038***	0.855***	−0.015
	(0.014)	(0.022)	(0.035)
Structural Factor 6	0.923***	0.147***	0.506***
	(0.019)	(0.022)	(0.042)
Observations	822	822	822
R^2	0.849	0.727	0.255

MSE Predicts HOME One Year Later

Variable	Both		Subjective Probability		Age Range	
Standardized $\mu_{i,\psi,1}$	-0.024	-0.002	-0.095	-0.058	-0.014	0.031
	(0.081)	(0.074)	(0.080)	(0.074)	(0.059)	(0.053)
Standardized $\mu_{i,\psi,2}$	0.167***	0.114***	0.119***	0.098**	0.170***	0.083**
	(0.045)	(0.039)	(0.043)	(0.040)	(0.045)	(0.038)
Standardized $\mu_{i,\psi,3}$	-0.086	0.010	-0.040	0.034	-0.058	-0.014
	(0.067)	(0.061)	(0.066)	(0.062)	(0.048)	(0.042)
Demographic variables included	No	Yes	No	Yes	No	Yes
Observations	687	687	687	687	687	687
R^2	0.037	0.270	0.034	0.271	0.031	0.265

MSE Predicts HOME One Year Later: LASSO

Table 11
LASSO regression.

Dependent variable: Standardized HOME scores						
Variable	Both forms		Subjective probability		Age range	
	CV	Adaptive	CV	Adaptive	CV	Adaptive
$\mu_{i,\psi,1}$			x		x	
$\mu_{i,\psi,2}$	x	x	x	x	x	x
$\mu_{i,\psi,3}$						

Conclusion CEC, 2020

Simple model linking beliefs to investments.

Elicitation of beliefs at pregnancy.

Measurement of investment one year later.

Raw belief data: Heaping, measurement error that follows elicitation instrument structure.

Beliefs vary across sample, correlate with observable characteristics, and predict investments.

Are beliefs malleable?

Introduction CGN, 2020

Pilot evaluation of the LENA Start Program.

Invited 136 families from PHD Study.

Randomly allocated to control and treatment.

Imperfect compliance, but very strong prediction of assignment dummy.

LENA Start Program

Center- and group-based.

Lasts 13 (then, now 10) weeks.

Each session:

1. Education.
2. Coaching.
3. Objective feedback.

Education and coaching: Joint attention and speech casting/recasting.

Objective Feedback

Figure 5

Impacts on Conversational Turns: Effect sizes

ITT (OLS): 31.4% (11.3 %) of a standard deviation.

ATE (Roy Model): 48.7% (19.4 %) of a standard deviation.

TT (Roy Model): 61.4% (23.4%) of a standard deviation.

LATE (2SLS): 56.3% (22.2%) of a standard deviation.

Other noteworthy findings

We find that the increase in conversational turns is driven by the audio segments the child initiates.

We also find impact on conversational turns in audio segments initiated by a male adult.

What explains change?

Impact on parent self efficacy: No.

Impact on parent’s sense of social support: No.

Maternal knowledge: maybe (borderline).

Maternal beliefs (ATE): 44% (18%) of a standard deviation.

Next steps

Partnership with a school district in Houston (N = 600) and the Harris County Public Libraries (N = 400).

Long-term follow up through primary data collection and administrative data sharing agreements.

Soon (new paper): How are parental beliefs formed? How do parents update beliefs?