Human Capital Formation in Childhood and Adolescence

Flávio Cunha

Rice University

July 10, 2018
Evolution of Inequality in USA

INCOME INEQUALITY IN THE UNITED STATES, 1910-2010

SHARE OF TOP DECILE IN NATIONAL INCOME

25% 30% 35% 40% 45% 50%

Figure 1
College Graduation Rates (by 35 years) for Men and Women: Cohorts Born from 1876 to 1975

Sources: 1940 to 2000 Census of Population Integrated Public Use Micro-data Samples (IPUMS).
Katz and Goldin (2007): College Enrollment in USA

Figure 8.3 Educational Attainment by Birth Cohort, 1900–1980. (a) Years of Schooling. (b) College Enrollment. (c) College Degree Attainment.

Source: Data are from Goldin and Katz (2008) and tabulated from 1940 to 2000 Census of Population Integrated Public Use Microdata Samples (IPUMS). Observations are for US native-born individuals adjusted to 35 years of age. Figure 8.3a shows the fraction of each birth cohort with at least a high school degree, Fig. 8.3b shows the fraction of each cohort with some college attendance, and Fig. 8.3c shows the fraction of each cohort with a college degree. For additional details, see DeLong, Goldin, and Katz (2003).
Comparing across the panels shown in Fig. 8.3, it is clear that changes in college degree attainment have not followed changes in college enrollment consistently over the course of the last 25 years. While college enrollment rates have increased fairly consistently, college degree attainment declined before increasing among more recent cohorts. Figure 8.4 presents the trend by birth cohort in the share of enrolled college students who complete a BA degree—essentially the trend shown in Fig. 8.3c divided by the trend in Fig. 8.3b. For both men and women, the rate of college completion has been below 50% for nearly a half century, with this level appreciably below the rate of completion achieved by men in the early part of the century. A component of this stagnation has been a growing disparity in college completion rates by parental circumstances. For example, for high school students from the top quartile of the family income distribution, completion rates rose slightly from 67.4 to 71% between those starting college in the early 1980s and those starting in the early 1990s, while the college completion rates fell for students from other income groups (Bowen, Chingos, and McPherson (2009)). Indeed, for 1992 high school seniors who enrolled in college, the difference in college completion rates between the students...
Hoxby (2009): Segmented Markets in Higher Education

Mean SAT/ACT Percentile Score of Colleges, by Colleges' Selectivity in 1962

- most selective in 1962: 4-year colleges with selectivity in the 99th %ile in 1962
- 96th-98th %ile in 1962
- 91st-95th %ile in 1962
- 81st-90th %ile in 1962
- 71st-80th %ile in 1962
- 61st-70th %ile in 1962
- 51st-60th %ile in 1962
- 41st-50th %ile in 1962
- 31st-40th %ile in 1962
- 21st-30th %ile in 1962
- 11th-20th %ile in 1962
- 6th-10th %ile in 1962

least selective in 1962: 4-year colleges with selectivity in the 1st-5th %iles in 1962
- 2 year colleges (estimated)
Returns and Stocks of Skilled/Unskilled Labor

- Let $L_S$ and $L_U$ denote, respectively, skilled and unskilled labor.
- Let $w_S$ and $w_U$ denote, respectively, skilled and unskilled wage rates.
- Consider the following problem:

$$\min w_S L_S + w_U L_U$$

subject to the aggregate production function (where $\gamma \in [0, 1]$ and $\phi \leq 1$):

$$Y = \left[ \gamma L_S^\phi + (1 - \gamma) L_U^\phi \right]^{\frac{1}{\phi}}$$

- Solution satisfies:

$$\ln \frac{w_S}{w_U} = \ln \frac{\gamma}{1 - \gamma} + (\phi - 1) \ln \frac{L_S}{L_U}$$
Katz and Goldin (2007): Model vs Data

- Actual values for college wage premium
- Predicted college wage premium
- Predicted wage premium with 1949 dummy

Plan: Data on skill formation

- Inequality in skills and inequality in adult socio-economic outcomes.
- Inequality in investments and inequality in skills.
- Increasing inequality in skills.
- Increasing inequality in investments.
- Evidence from RCTs.
Figure 1: The Probability of Educational Decisions, by Endowment Levels, Dropping from Secondary School vs. Graduating

Figure 2: The Probability of Educational Decisions, by Endowment Levels, HS Graduate vs. College Enrollment
Figure 3: The Probability of Educational Decisions, by Endowment Levels, Some College vs. 4-year college degree


James Heckman
Economics and Econometrics of Human Development
Figure 4: The Effect of Cognitive and Socio-emotional endowments, (log) Wages

James Heckman Economics and Econometrics of Human Development
Inequality in Cognitive Skills as Children Age

Introduction
Simple Model
Structural Model
Data and Estimates
Conclusion and Future Work


-0.8000
-0.6000
-0.4000
-0.2000
0.0000
0.2000
0.4000
0.6000
0.8000
3 5 6 7 8 9 10 11 12 13 14
Age
Bottom Quartile
Second Quartile
Third Quartile
Top Quartile
Average percentile rank on anti-social behavior score, by income quartile

- Lowest Income Quartile
- Second Income Quartile
- Third Income Quartile
- Highest Income Quartile
Inequality in Health as Children Age


Inequality in Investments as Children Age

Figure
Unadjusted Mean Home Score
by Quartile of Permanent Income of the Family

Flávio Cunha and Anton Badev University of Pennsylvania () 10/17 6 / 44
Inequality in Investments as Children Age

Inequality in Investments as Children Age

How teacher ratings relate to a school's poverty level

Teachers who receive the state’s top value-added rating — "Most Effective" — are likely to be in schools with fewer poor students, based on value-added ratings for teachers at 1,720 public schools. Of 1,035 teachers at the wealthiest schools, 34 percent got the top rating. In contrast, of 2,411 teachers at the poorest schools, just over 9 percent were rated "Most Effective."

SOURCE: Ohio Department of Education

RICH EXNER, JAMES OWENS | THE PLAIN DEALER
Inequality in Cognitive Skills Over Time

Inequality in Noncognitive Skills Over Time

Social Trust
By parents’ education, 12th graders, 1976–2011

“Most people can be trusted” (agree)


Source: Monitoring the Future
Inequality in Health Over Time

- HS or less
- BA or more

Graph showing the percentage of inequality in health over time, with two lines representing different educational attainment levels.
Inequality in Investments Over Time
Inequality in Investments Over Time

Trends in Family Dinners
By parental education, 1978–2005

“Our whole family usually eats dinner together” (agree)

Source: DDB Lifestyle surveys, 1978–2005
Inequality in Investments Over Time

![Participation in School-Based Extracurriculars](1972–2002)

- Highest SES quartile
- Lowest SES quartile

Increasing Inequality in College Attendance

- Top quartile
- Third quartile
- Second quartile
- Bottom quartile
Evidence from RCTs in Early Childhood and Adolescence

- Early interventions:
  - Perry Preschool Program
  - Abecedarian
  - Infant Health and Development Program (IHDP)
The kernel densities reveal different patterns of the effect of the program on the distribution of skills. The cognition of females is enhanced mostly in the right tail of the distribution (panel B). In contrast, a substantial part of the improvement in externalizing behavior for females operates through enhancing low levels of the skill (panel D). Externalizing behavior in males is improved at all levels.

Academic motivation in females is improved at all levels except for the top percentiles (see panel F). There is no statistically significant difference in the distribution of cognition for males (panel A).

We also test for gender differences in skills and find that differences are not statistically significant. In other words, for each skill and for each treatment group we cannot reject the null hypothesis of equality of skills between males and females. See Figure L.1 of online Appendix L.

**Figure 5. Kernel Densities of Factor Scores**

Notes: Probability density functions of Bartlett (1937) factor scores are shown. Densities are computed based on a normal kernel. Numbers above the charts are one-sided p-values testing the equality of factor score means for the treatment and control groups. Higher externalizing behavior corresponds to more socially desirable behavior. See online Appendix L for the empirical CDFs of the factor scores (Figure L.5). Vertical lines locate factor score means for treatment and control groups.
Early Childhood Education

Table 7: Life-Cycle Outcomes, PPP and ABC

<table>
<thead>
<tr>
<th></th>
<th>PPP</th>
<th>ABC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age</td>
<td>Female</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Cognition and Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult IQ</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School Graduation</td>
<td>19(^a)</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.416)</td>
</tr>
<tr>
<td>Economic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>40(^a)</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.615)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Yearly Labor Income, 2014 USD</td>
<td>40(^a)</td>
<td>$6,166</td>
</tr>
<tr>
<td></td>
<td>(0.224)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>HI by Employer</td>
<td>40(^a)</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>Ever on Welfare</td>
<td>18–27(^a)</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.590)</td>
</tr>
<tr>
<td>Crime</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of Arrests(^d)</td>
<td>≤40(^a)</td>
<td>-2.77</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>No. of Non-Juv. Arrests</td>
<td>≤40(^a)</td>
<td>-2.45</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Lifestyle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-reported Drug User</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not a Daily Smoker</td>
<td>27(^a)</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.089)</td>
</tr>
<tr>
<td>Not a Daily Smoker</td>
<td>40(^a)</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td>(0.206)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Physical Activity</td>
<td>40(^a)</td>
<td>0.330</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.545)</td>
</tr>
<tr>
<td>Health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity (BMI &gt;30)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension I</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: \(a\) Heckman et al. (2010a). \(b\) Campbell et al. (2014). \(c\) Elango et al. (2015). Note: This table displays statistics for the treatment effects of PPP and ABC on important life-cycle outcome variables. Hypertension I is the first stage of high blood pressure—systolic blood pressure between 140 and 159 and diastolic pressure between 90 and 99. “HI by employer” refers to health insurance provided by the employer and is conditional on being employed.

\(d\) “No. of Arrests” includes offenses in the case of ABC, even where more than one offense was charged per arrest. For the further definitions of the outcomes, see the respective web appendices of the cited papers. Outcomes from Heckman et al. (2010a) are reported with one-sided \(p\)-value which is based on Freedman-Lane procedure, using the linear covariates of maternal employment, paternal presence and SB (Stanford-Binet) IQ, and restricting permutation orbits within strata formed by a Socio-economic Status index being above or below the sample median and permuting siblings as a block.

\(p\)-values for the outcomes from Campbell et al. (2014) are one-sided single hypothesis constrained permutation \(p\)-value’s, based on the IPW (Inverse Probability Weighting) \(t\)-statistic associated with the difference in means between treatment groups; probabilities of IPW are estimated using the variables gender, presence of father in home at entry, cultural deprivation scale, child IQ at entry (SB), number of siblings and maternal employment status.

\(p\)-values for the outcomes from Elango et al. (2015) are bootstrapped with 1000 resamples, corrected for attrition with Inverse Probability Weights, with treatment effects conditioned on treatment status, cohort, number of siblings, mothers IQ, and the ABC high risk index.
Evidence from RCTs in Adolescence

- **Becoming-A-Man (B.A.M.) Study:**
  - Student training: Learning how to “read” the context to employ the “appropriate” reaction.

- **Montreal Longitudinal Study**
  - Parent training: Improve monitoring and positive reinforcement; implement non-punitive discipline; and how to better cope with crisis.
  - Child training: Teaching social skills to reduce aggressive behavior (including how to manage anger-inducing situations).
Figure 1. Non-cognitive skills and school performance during adolescence. A, B and C show distributions for non-cognitive skills measured in early adolescence for the control, treatment and non-disruptive groups (the non-disruptive boys being those who were not disruptive in kindergarten and did not participate in the experiment as treatment or control: they serve as a normative population baseline). Kolmogorov-Smirnov test for equality of Treatment and Control distributions gives p-value of 0.003 for Trust, 0.036 for Aggression Control, and 0.023 for Attention-Impulse Control. D shows the increasing gap in the percent of subjects held back at each age. P-value from $\chi^2$ test between Treatment and Control groups is 0.60 at age 10 and 0.01 at age 17.
Figure 2. Young Adult Outcomes. As young adults, treatment subjects commit fewer crimes, are more likely to graduate from secondary school, are more likely to be active fulltime in school or work, and are more likely to belong to a social or civic group. The intervention closed part or all of the gap between boys ranked as disruptive in kindergarten but not treated (the control group) and the non-disruptive boys (who represent the normative population). Raw differences are significant for secondary diploma (p-value=0.04) and group membership (p-value=0.05), conditional differences (controlling for group imbalances) are significant for number of crimes (p-value=0.09) and percent active fulltime (p-value=0.03).
Equation that describes skill formation process.
Identification and estimation of key parameters of the equation.
Constraints: Decision maker preference and information set
Identification and estimation of subjective information set.
Skills Developed in Early Childhood

- **Early development:**
  - Development of language and cognitive skills
  - Development of **executive functions:**
    - Inhibitory control;
    - Working memory;
    - Cognitive flexibility (flexible thinking and set shifting).
Skills Developed in Adolescence

- Adolescent development:

  It seems like people accept you more if you’re, like, a dangerous driver or something. If there is a line of cars going down the road and the other lane is clear and you pass eight cars at once, everybody likes that. . . . If my friends are with me in the car, or if there are a lot of people in the line, I would do it, but if I’m by myself and I didn’t know anybody, then I wouldn’t do it. That’s no fun. — Anonymous teenager, as quoted in The Culture of Adolescent Risk-Taking (Lightfoot, 1997, p. 10)
Differential susceptibility of adolescents to peer influences on Stoplight task performance

Mean (a) percentage of risky decisions and (b) number of crashes for adolescent, young adult, and adult participants when playing the Stoplight driving game either alone or with a peer audience. Error bars indicate the standard error of the mean.
Adolescent development:

- Fast development of the reward system potentialized by the influence of peers.
- Slow development of “rational” decision making system.
We formalize the notion that human capital accumulation is one in which we produce different types of skills at different stages of the lifecycle.

This notion leads to a technology of skill formation that is described by two parameters:

- Self-productivity of skills: I learn how inhibit control early on, that helps me learn how to “read” the context before choosing an action when adolescent.

- Dynamic complementarity: The returns to the development of “reading” context are higher for the children that have learned how to inhibit control early on (and vice-versa).
Let $h_{i,0}$ and $x_{i,e}$ denote, respectively, human capital at birth and investment during early childhood.

Let $h_{i,a}$ denote the human capital at beginning of adolescence.

Assume that:

$$h_{i,a} = \left[ \gamma_e x_{i,e} + (1 - \gamma_e) h_{i,0} \right]^{\frac{1}{\phi_e}}$$
Let $x_{i,a}$ denote investment during adolescence.

Let $\bar{h}_i$ denote the human capital at beginning of adulthood. Assume that:

$$\bar{h}_i = \left[ \gamma_a x_{i,a}^{\phi_a} + (1 - \gamma_a) h_{i,a}^{\phi_a} \right]^{\frac{1}{\phi_a}}$$
Apply recursion and assume $\phi_e = \phi_a = \phi$:

$$\bar{h} = \left\{ \gamma_a x_{i,a} + (1 - \gamma_a) \gamma_e x_{i,e} + (1 - \gamma_a)(1 - \gamma_e) h_{i,0}^{\phi} \right\}^{\frac{1}{\phi}}$$

Note that:

- The parameter $1 - \gamma_a$ captures self-productivity.
- The parameter $\phi$ captures dynamic complementarity or substitutability.
The problem of the parent:

$$\text{min } x_{i,e} + \frac{1}{1 + r} x_{i,a}$$

subject to the technology of skill formation:

$$\bar{h} = \left\{ \gamma_a x_{i,a}^\phi + (1 - \gamma_a) \gamma_e x_{i,e}^\phi + (1 - \gamma_a) (1 - \gamma_e) h_{i,0}^\phi \right\}^{\frac{1}{\phi}}$$

where $\gamma_a \in [0, 1]$, $\gamma_e \in [0, 1]$, and $\phi \leq 1$. 
Boundary Solution when $\phi = 1$

In this case:

$$h = \gamma_a x_{i,a} + (1 - \gamma_a) \gamma_e x_{i,e} + (1 - \gamma_a) (1 - \gamma_e) h_{i,0}$$

Two investment strategies: Invest early and produce $(1 - \gamma_a) \gamma_e$ units of human capital per unit of investment.

Save in physical assets early and invest $1 + r$ late and produce $(1 + r) \gamma_a$ units of human capital.

Should invest all early if, and only if:

$$\frac{(1 - \gamma_a) \gamma_e}{\gamma_a} > 1 + r$$
In this case:

\[ \bar{h}_i = \min \{ x_{i,a}, x_{i,e}, h_{i,0} \} \]

The solution to this problem is \( x_{i,a} = x_{i,e} = h_{i,0} \) regardless of \( r \).
The solution to this problem is characterized by the following ratio:

\[
\ln \frac{x_{i,e}}{x_{i,a}} = \frac{1}{1-\phi} \ln \left[ \frac{(1-\gamma_a)\gamma_e}{\gamma_a} \right] + \frac{1}{1-\phi} \ln \left( \frac{1}{1+r} \right)
\]
Returns to late investments are higher for the individuals that have high early investments.

BUT: Returns to early investments are higher for the individuals who will also have high late investments.

In other words, if the child will not receive high late investments, then the impacts of early investments will be diminished.
Estimating the Technology of Skill Formation

- Return to the recursive formulation of the technology of skill formation:

\[ h_{i,t+1} = \left[ \gamma_t x_{i,t}^{\phi_t} + (1 - \gamma_t) h_{i,t}^{\phi_t} \right]^{\frac{\rho_t}{\phi_t}} e^{\eta_{i,t+1}} \]

- Consider (simplified version of) the Kmenta (1967) approximation:

\[ \ln h_{i,t+1} = \psi_{t,1} \ln x_{i,t} + \psi_{t,2} \ln h_{i,t} + \psi_{t,3} \ln x_{i,t} \ln h_{i,t} + \eta_{i,t+1} \]

- Where: \( \psi_{t,1} = \gamma_t \phi_t \), \( \psi_{t,2} = (1 - \gamma_t) \phi_t \), and \( \psi_{t,3} = \frac{1}{2} \rho_t \phi_t \gamma_t (1 - \gamma_t) \).

- Possible to decompose \( \eta_{i,t+1} \) into permanent and temporary shocks, but not going to do it today.
To simplify the math, I will use a simpler version of the Kmenta approximation:

$$\ln h_{i,t+1} = \psi_{t,1} \ln x_{i,t} + \psi_{t,2} \ln h_{i,t} + \psi_{t,3} \ln x_{i,t} \ln h_{i,t} + \eta_{i,t+1}$$

for $i = 1, \ldots, I$ and $t = 1, \ldots, T$.

I will illustrate three problems in the estimation of the technology of skill formation:

- Problem 1: data on measures of human capital have no cardinality: anchoring.
- Problem 2: data on measures of human capital and investment have measurement error: latent factors.
- Problem 3: data on investment is endogenous: instruments.
To simplify the math, I will use a simpler version of the Kmenta approximation:

$$\ln h_{i,t+1} = \psi_{t,1} \ln x_{i,t} + \psi_{t,2} \ln h_{i,t} + \psi_{t,3} \ln x_{i,t} \ln h_{i,t} + \eta_{i,t+1}$$

for $i = 1, \ldots, I$ and $t = 1, \ldots, T$.

I will illustrate three problems in the estimation of the technology of skill formation:

- Problem 1: data on measures of human capital have no cardinality: anchoring.
- Problem 2: data on measures of human capital and investment have measurement error: latent factors.
- Problem 3: data on investment is endogenous: instruments.
Problem 1: Cardinality of Human Capital

- The notion of a production function implies that inputs and output have a well-defined metric.
- You put $a$ units of investments and $b$ units of current-period human capital and you produce $x$ units of next-period human capital.
- Usually units of investments are time (e.g., hours per day) or money (e.g., dollars per month).
- What is the unit of human capital?
<table>
<thead>
<tr>
<th>Type of scale</th>
<th>Description</th>
<th>Possible statements</th>
<th>Allowed operators</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>Describes qualitative attributes</td>
<td>Identity, countable</td>
<td>=, ≠</td>
<td>Binary variable denoting gender</td>
</tr>
<tr>
<td>Ordinal</td>
<td>Describes objects that can be ordered in terms of &quot;greater&quot;, &quot;less&quot;, or &quot;equal&quot;</td>
<td>Identity, countable, less than/greater than relations</td>
<td>=, ≠, ≤, ≥</td>
<td>Utility levels, test scores, percentile scores</td>
</tr>
<tr>
<td>Interval (cardinal)</td>
<td>Describes objects that can be placed in equally spaced units without a true zero point.</td>
<td>Identity, countable, less than/greater than relations, equality of differences</td>
<td>=, ≠, ≤, ≥, +, -</td>
<td>Educational attainment, dates</td>
</tr>
<tr>
<td>Ratio (cardinal)</td>
<td>Describes objects that can be placed in equally spaced units that have a true zero point.</td>
<td>Identity, countable, less than/greater than relations, equality of differences, equality of ratios, true zero</td>
<td>=, ≠, ≤, ≥, +, -, ×, ÷</td>
<td>Earnings, length, age</td>
</tr>
</tbody>
</table>
Let’s approach this problem in the following way. Suppose that we have data on labor income, $Y_i$, at some point in adulthood (e.g., when the individual is 45 years old).

We can “anchor” human capital at age $t$ before adulthood, $t = 1, ..., T$, through the equation:

$$\ln Y_i = \ln h_{i,t} + \nu_{i,t}$$

Now $\ln h_{i,t}$ is cardinal. Assume that $\ln h_{i,t} \sim N (\mu_h, \sigma_{h,t}^2)$,

$$\nu_{i,t} \sim N (0, \sigma_{\nu,t}^2).$$

Note that $\ln Y_i \sim N (\mu_h, \sigma_{h,t}^2 + \sigma_{\nu,t}^2)$.
Now, we have data on scores in standardized tests $M_{i,t,j}$ for $j = 1, \ldots, J$.

Assume that the relationship between $M_{i,t,j}$ and $\ln h_{i,t}$ is:

$$M_{i,t,j} = \alpha_{t,j} + \beta_{t,j} \ln h_{i,t} + \varepsilon_{i,t,j}$$

where $\varepsilon_{i,t,j} \sim N \left(0, \sigma_{t,j}^2\right)$ is measurement error.

Therefore, we have that $M_{i,t,j} \sim N \left(\alpha_{t,j} + \beta_{t,j} \mu h, \beta_{t,j}^2 \sigma_{t,h}^2 + \sigma_{t,j}^2\right)$.

In particular, note that $M_{i,t,j} | \ln h_{i,t} \sim N \left(\alpha_{t,j} + \beta_{t,j} \ln h_{i,t}, \sigma_{t,j}^2\right)$.
Problem 1: Cardinality of Human Capital

Solution: We need to transform at least one of the test scores at time $t$ so that the transformed measure has cardinality.

Define $\tilde{m}_{i,t,1} = E(\ln Y_i | M_{i,t,1})$ and $s_{t,1} = \frac{\beta_{t,1}^2 \sigma_{t,h}^2}{\beta_{t,1}^2 \sigma_{t,h}^2 + \sigma_{t,j}^2}$.

Use the fact that $\ln Y_i$ and $M_{i,t,1}$ are jointly normal to conclude that:

$$\tilde{m}_{i,t,1} = (1 - s_{t,1}) \mu_h + s_{t,1} (M_{i,t,1} - \alpha_{t,1}).$$
Problem 1: Cardinality of Human Capital

Given that:

\[ \tilde{m}_{i,t,1} = (1 - s_{t,1}) \mu_h + s_{t,1} (M_{i,t,1} - \alpha_{t,1}) \]

and that:

\[ M_{i,t,j} = \alpha_{t,j} + \beta_{t,j} \ln h_{i,t} + \epsilon_{i,t,1} \]

We conclude that:

\[ \tilde{m}_{i,t,1} = (1 - s_{t,1}) \mu_h + s_{t,1} \ln h_{i,t} + \frac{s_{t,1}}{\beta_{t,1}} \epsilon_{i,t,1} \]

We need to estimate \( s_{t,1} \).
Problem 1: Cardinality of Human Capital

- We need to estimate \( s_{t,1} \), but we don’t observe \( \ln h_{i,t} \). We do observe \( \ln Y_i = \ln h_{i,t} + \nu_{i,t} \), so

\[
\tilde{m}_{i,t,1} = (1 - s_{t,1}) \mu_h + s_{t,1} \ln Y_i + \frac{s_{t,1}}{\beta_{t,1}} \epsilon_{i,t,1} - s_{t,1} \nu_{i,t}
\]

- Clearly, we can’t use OLS because \( \ln Y_i \) is correlated with \( \nu_{i,t} \).
- We need an instrument. In particular, we need something that is correlated with \( \ln Y_i \) (through \( \ln h_{i,t} \)), but not correlated with \( \epsilon_{i,t,1} \) or \( \nu_{i,t} \).
- We have a few candidates:
  - Investment at period \( t - 1 \).
  - Determinants of investment at period \( t - 1 \) (e.g., random assignment to control or treatment arms of intervention).
  - If nothing else, then \( \tilde{m}_{i,\tau,1}^* \) which is leave-one-out estimator of \( \tilde{m}_{i,\tau,1} \) where \( \tau \neq t \)
Problem 1: Cardinality of Human Capital

- Use one of these instruments to identify $s_{t,1}$ and define

$$m_{i,t,1} = \frac{\tilde{m}_{i,t,1}}{s_{t,1}}$$

$$m_{i,t,1} = \frac{(1-s_{t,1})}{s_{t,1}} \mu_h + \ln h_{i,t} + \frac{1}{\beta_{t,1}} \epsilon_{i,t,1}$$

- Now we have a rescaled score that has a cardinal scale.
Figure 1: Raw Difference in Expected White Grade Completion conditional on Test Score

- Math
  - 95% Confidence Interval

- Reading Recognition
  - 95% Confidence Interval

- Reading Comprehension
  - 95% Confidence Interval
Figure 2: Measurement Error Adjusted Difference in Achievement in Units of Predicted White Education

- **Math**: 95% Confidence Interval
- **Reading Recognition**: 95% Confidence Interval
- **Reading Comprehension**: 95% Confidence Interval
To simplify the math, I will use a simpler version of the Kmenta approximation:

\[ \ln h_{i,t+1} = \psi_{t,1} \ln x_{i,t} + \psi_{t,2} \ln h_{i,t} + \psi_{t,3} \ln x_{i,t} \ln h_{i,t} + \eta_{i,t+1} \]

for \( i = 1, \ldots, I \) and \( t = 1, \ldots, T \).

I will illustrate three problems in the estimation of the technology of skill formation:

- **Problem 1:** data on measures of human capital have no cardinality: anchoring.
- **Problem 2:** data on measures of human capital and investment have measurement error: latent factors.
- **Problem 3:** data on investment is endogenous: instruments.
Problem 2: Latent Factors

- At every age $t$ we have $J$ test scores and at least one of which (e.g., the first) is anchored:

$$m_{i,t,1} = \frac{(1-s_{t,1})}{s_{t,1}} \mu_h + \ln h_{i,t} + \frac{1}{\beta_{t,1}} \varepsilon_{i,t,1}$$

$$m_{i,t,j} = \alpha_{t,j} + \beta_{t,j} \ln h_{i,t} + \varepsilon_{i,t,j}$$

- At every age $t$ we have $J$ measures of investments:

$$p_{i,t,j} = \delta_{t,j} + \kappa_{t,j} \ln x_{i,t} + \zeta_{i,t,j}$$
Problem 2: Latent Factors

- Rewrite in vector form:

$$m_{i,t} = \alpha_t + \beta_t \ln h_{i,t} + \varepsilon_{i,t}$$

- At every age $t$ we have $J$ measures of investments:

$$p_{i,t} = \delta_t + \kappa_t \ln x_{i,t} + \zeta_{i,t}$$
Estimate $\alpha_t$, $\beta_t$, $\delta_t$, $\kappa_t$, matrix $\Sigma_\epsilon$ and matrix $\Sigma_\xi$ to predict Bartlett scores:

$$\ln h_{i,t}^B = \left[ \beta'_t \Sigma_\epsilon^{-1} \beta_t \right]^{-1} \left[ \beta'_t \Sigma_\epsilon^{-1} (m_{i,t} - \alpha_t) \right]$$

$$\ln x_{i,t}^B = \left[ \kappa'_t \Sigma_\xi^{-1} \kappa_t \right]^{-1} \left[ \kappa'_t \Sigma_\xi^{-1} (p_{i,t} - \delta_t) \right]$$
Problem 2: Latent Factors

- Estimate $\alpha_t$, $\beta_t$, $\delta_t$, $\kappa_t$, matrix $\Sigma_\epsilon$ and matrix $\Sigma_\xi$ to predict Bartlett scores:

\[
\ln h_{i,t}^B = \ln h_{i,t} + \left[ \beta_t' \Sigma_\epsilon^{-1} \beta_t \right]^{-1} \left[ \beta_t' \Sigma_\epsilon^{-1} \epsilon_{i,t} \right]
\]

\[
\ln x_{i,t}^B = \ln x_{i,t} + \left[ \kappa_t' \Sigma_\xi^{-1} \kappa_t \right]^{-1} \left[ \kappa_t' \Sigma_\xi^{-1} \xi_{i,t} \right]
\]
Problem 2: Latent Factors

- Note that:
  \[ \ln h_{i,t}^B = \ln h_{i,t} + \tilde{\varepsilon}_{i,t} \]
  \[ \ln x_{i,t}^B = \ln x_{i,t} + \tilde{\xi}_{i,t} \]

- Note that \( \tilde{\varepsilon}_{i,t} \sim N \left( 0, \left[ \beta_t \Sigma_{\varepsilon}^{-1} \beta_t \right]^{-1} \right) \) and \( \tilde{\xi}_{i,t} \sim N \left( 0, \left[ \kappa_t \Sigma_{\xi}^{-1} \kappa_t \right]^{-1} \right) \) and the variances are known.

- Using factor scores directly will not work because factor scores inherit measurement error (attenuation bias).

- However, bias is a function of \( \left[ \beta_t \Sigma_{\varepsilon}^{-1} \beta_t \right]^{-1} \) and \( \left[ \kappa_t \Sigma_{\xi}^{-1} \kappa_t \right]^{-1} \) which are known. Therefore, we can account for the bias.
Problem 2: Latent Factors

- Define
  \[
  h_t = \{\ln h_{i,t}\}_{i=1}^I \\
  w_t = \{(\ln h_{i,t}, \ln x_{i,t}, \ln h_{i,t} \times \ln x_{i,t})\}_{i=1}^I \\
  \gamma_t = (\gamma_{t,1}, \gamma_{t,2}, \gamma_{t,3})
  \]

- Rewrite:
  \[
  h_{t+1} = w_t \gamma_t + \eta_{t+1}
  \]

- Let \( \hat{\gamma}_t \) denote the infeasible OLS estimator that uses \( h \) and \( w \) (assumed to be exogenous).
  \[
  \hat{\gamma}_t = \left( w_t^T w_t \right)^{-1} \left( w_t^T h_{t+1} \right)
  \]

- Easy to show that \( \hat{\gamma}_t \) is consistent.
Let $\hat{\gamma}^B_t$ denote the OLS estimator that uses Bartlett scores $h^B_t$ and $w^B_t$ (assumed to be exogenous).

$$\hat{\gamma}^B_t = \left[ (w^B_t)^T w^B_t \right]^{-1} \left[ (w^B_t)^T h^B_{t+1} \right]$$

Note that $w^B_t$ is error-ridden measure of $w$, so standard attenuation bias arises.

Difference: attenuation bias is a function of variance of measurement error.

The bias arises because of matrix $\left[ (w^B_t)^T w^B_t \right]$. 
The matrices \([w_t^T w_t]\) and \([ (w_t^B)^T w_t^B ]\) are symmetric with the following elements:

<table>
<thead>
<tr>
<th>Element</th>
<th>plim ([w_t^T w_t])</th>
<th>plim ([(w_t^B)^T w_t^B])</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 1)</td>
<td>(E (x_t^2))</td>
<td>(E (x_t^2) + \text{Var} (\xi_t))</td>
</tr>
<tr>
<td>(1, 2)</td>
<td>(E (x_t h_t))</td>
<td>(E (x_t h_t))</td>
</tr>
<tr>
<td>(1, 3)</td>
<td>(E (x_t^2 h_t))</td>
<td>(E (x_t^2 h_t) + E (h_t) \text{Var} (\xi_t))</td>
</tr>
<tr>
<td>(2, 2)</td>
<td>(E (h_t^2))</td>
<td>(E (h_t^2) + \text{Var} (\varepsilon_t))</td>
</tr>
<tr>
<td>(2, 3)</td>
<td>(E (x_t h_t^2))</td>
<td>(E (x_t h_t^2) + E (x_t) \text{Var} (\varepsilon_t))</td>
</tr>
<tr>
<td>(3, 3)</td>
<td>(E (x_t^2 h_t^2))</td>
<td>(E (x_t^2 h_t^2) + \Delta)</td>
</tr>
</tbody>
</table>

where

\[
\Delta = E (x_t^2) \text{Var} (\varepsilon_t) + E (h_t^2) \text{Var} (\xi_t) + \text{Var} (\xi_t) + \text{Var} (\varepsilon_t)
\]
Define matrix $A = (w_t^B)^T w_t^B - B$ where

$$B = \begin{bmatrix}
\text{Var} (\xi_t) & 0 & E (h_t) & \text{Var} (\xi_t) \\
\text{Var} (\varepsilon_t) & E (x_t) & \text{Var} (\varepsilon_t) & \Delta
\end{bmatrix}$$

Feasible estimator $\hat{\gamma}^A$ is consistent:

$$\hat{\gamma}^A = \left( (w_t^B)^T w_t^B - B \right)^{-1} \left( (w_t^B)^T h_{t+1}^B \right)$$
Figure 3
Share of Residual Variance in Measurements of Cognitive Skills
Due to the Variance of Cognitive Factor (Signal)
and Due to the Variance of Measurement Error (Noise)
Figure 4A
Share of Residual Variance in Measurements of Noncognitive Skills
Due to the Variance of Noncognitive Factor (Signal)
and Due to the Variance of Measurement Error (Noise)
Figure 4B
Share of Residual Variance in Measurements of Noncognitive Skills
Due to the Variance of Noncognitive Factor (Signal)
and Due to the Variance of Measurement Error (Noise)
Figure 5A
Share of Residual Variance in Measurements of Investments Due to the Variance of Investment Factor (Signal) and Due to the Variance of Measurement Error (Noise)
Figure 5B
Share of Residual Variance in Measurements of Investments Due to the Variance of Investment Factor (Signal) and Due to the Variance of Measurement Error (Noise)
Figure 5C
Share of Residual Variance in Measurements of Investments
Due to the Variance of Investment Factor (Signal)
and Due to the Variance of Measurement Error (Noise)
To simplify the math, I will use a simpler version of the Kmenta approximation:

\[
\ln h_{i,t+1} = \psi_{t,1} \ln x_{i,t} + \psi_{t,2} \ln h_{i,t} + \psi_{t,3} \ln x_{i,t} \ln h_{i,t} + \eta_{i,t+1}
\]

for \( i = 1, \ldots, I \) and \( t = 1, \ldots, T \).

I will illustrate three problems in the estimation of the technology of skill formation:

- Problem 1: data on measures of human capital have no cardinality: anchoring.
- Problem 2: data on measures of human capital and investment have measurement error: latent factors.
- Problem 3: data on investment is endogenous: instruments.
Problem 3: Instruments

Note that:

\[ \ln h_{i,t+1} = \psi_{t,1} \ln x_{i,t} + \psi_{t,2} \ln h_{i,t} + \psi_{t,3} \ln x_{i,t} \ln h_{i,t} + \eta_{i,t+1} \]

\[ \ln x_{i,t} = z_{i,t} + v_{i,t} \]

Here \( z_{i,t} \) is the instrument.

Valid instruments address not only endogeneity (\( \ln x_{i,t} \) correlated with \( \eta_{t+1} \)) but also problems created by measurement error in \( \ln x_{i,t}^B \).

Instrument does not address bias due to measurement error in \( \ln h_{i,t}^B \) unless we have a specific instrument for \( \ln h_{i,t} \).
FIGURE 6.—Densities of ratio of early to late investments maximizing aggregate education versus minimizing aggregate crime.

Figure 2. The widening gap we saw in the vocabulary growth of children from professional, working-class, and welfare families across their first 3 years.
Hart and Risley (1995): Adult Words per Hour
Preferences are represented by the following utility function:

\[ U(c, h_1, h_R^1) = \ln c + \alpha \ln h_1 + \beta 1 (\ln h_1 \leq \ln h_R) \]

Where:

- \( c \) is consumption;
- \( h_1 \) is the child’s human capital at the end of the early childhood period;
- \( h_R \) is the parent’s reference point for the child’s human capital level at the end of the early childhood period.
- From the point of view of the parent, \( \ln h_R \sim N(\mu_R, \sigma_R^2) \).
I assume that parents cannot borrow or save:

\[ c + px = y \]

Where:

- \( p \) is the relative price of the investment good;
- \( x \) is the investment good;
- \( y \) is the family income during the early childhood period.
I assume that the child’s human capital at the end of the early childhood period is determined according to:

$$\ln h_1 = \gamma_0 + \gamma_1 \ln h_0 + \gamma_2 \ln x + \nu$$

Where:

- $h_0$ is the child’s human capital at birth;
- $\nu$ is a shock that is unanticipated by the parent and unobserved by the economist.
- From the point of view of the parent, $\gamma_k \sim N (\mu_k, \sigma_k^2)$. 
The parent’s information set:

\[ \Omega = \left\{ p, y, h_0, \epsilon, \Phi \left( \mu_R, \sigma_R^2 \right), \left[ \Phi \left( \mu_k, \sigma_k^2 \right) \right]_{k=0}^3 \right\} \]

Note that from the point of view of the parent:
- \( \Phi \left( \mu_R, \sigma_R^2 \right) \) is the parent’s perceived distribution of \( \ln h_R \).
- \( \Phi \left( \mu_k, \sigma_k^2 \right) \) is the parent’s perceived distribution of \( \gamma_k \).

We do not impose any a priori restrictions on the parameters of these distributions.
Optimal choice when the parent knows true gamma or chooses the technology with the high gamma.

Horizontal axis = child development
Vertical axis = household consumption
Introducing Heterogeneity in Beliefs

Horizontal axis = child development
Vertical axis = household consumption

Choice of household consumption and child development when the parent knows the true gamma or adopts the technology with high gamma

Choice when the parent does not know the true gamma or adopts a technology with low gamma
Introducing Heterogeneity in Reference Points

Preferences of a parent that values child development relatively more than household consumption

Preferences of a parent that values child development relatively less than household consumption
Introducing Heterogeneity in Reference Points

Preferences with reference-point dependence

Parent is over-sensitive to the perceived under-development of the child

Parent is under-sensitive to the perceived over-development of the child

Horizontal axis = child development
Vertical axis = household consumption

Flávio Cunha (Rice University)  Human Capital Formation in Childhood and Adolescence  July 10, 2018  93 / 143
Introducing Heterogeneity in Reference Points

Choice when parents have preferences that are reference-point dependent
Introducing Heterogeneity in Reference Points

Horizontal axis = child development
Vertical axis = household consumption

Choice with "low" reference point and the parent knows the true gamma
Choice with "high" reference point and parent knows the true gamma
Introducing Heterogeneity in Reference Points

Horizontal axis = child development
Vertical axis = household consumption

Choice with "low" reference point and the parent knows the true gamma
Choice with "high" reference point and the parent knows the true gamma
Choice with "high" reference point and the parent does not know the true gamma
Three Papers

- Can we elicit maternal subjective expectations? 

- Does intervention that provide information and objective feedback affect parental beliefs, investments, and development? 

- Do reference points affect parental investments in children? 
  Wang, Puentes, Behrman, and Cunha (2018): Use RCT to see if reference points affect children’s height by age 2 years.
Philadelphia Human Development (PHD) Study.

Round 1: Elicit maternal subjective expectations during 2nd trimester of 1st pregnancy.

Round 2: Measure maternal investments when child is 9-12 months old.

Round 3: Measure child development when child is 22-26 months old.

Round 4: RCT about language development when child is 28-32 months old.
Defining Subjective Expectation

- The technology of skill formation is:
  \[ \ln h_{i,1} = \psi_0 + \psi_1 \ln h_{0,i} + \psi_2 \ln x_i + \psi_3 \ln h_{0,i} \ln x_i + \nu_i \]

- Let \( \Psi_i \) denote the mother’s information set.
- Let \( E(\psi_j| h_{0,i}, x_i, \Psi_i) = \mu_{i,j} \) and assume that \( E(\nu_i| \Psi_i) = 0 \).
- From the point of view of the mother:
  \[ E(\ln h_{i,1}| h_{0,i}, x_i, \Psi_i) = \mu_{i,0} + \mu_{i,1} \ln h_{0,i} + \mu_{i,2} \ln x_i + \mu_{i,3} \ln h_{0,i} \ln x_i \]
Consider a simple static model. Parent’s utility is:

\[ u(c_i, h_{i,1}; \alpha_{i,1}, \alpha_{i,2}) = \ln c_i + \alpha_{i,1} \ln h_{i,1} + \alpha_{i,2} \ln x_i \]

Budget constraint is:

\[ c_i + p x_i = y_i. \]
Model

- The problem of the mother is to maximize expected utility subject to the mother’s information set, the budget constraint, and the technology of skill formation.

- The solution is

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

- Clearly, we cannot separately identify \( \alpha_i \) from \( \mu_{i,\gamma} \) if we only observe \( x_i, y_i, \) and \( p \).
Eliciting subjective expectations: Steps

- Measure actual child development: MSD and Item Response Theory (IRT).
- Develop the survey instrument to elicit beliefs $E [\ln h_{i,1} \mid h_0, x, \psi_i]$:
  - Reword MSD items.
  - Create hypothetical scenarios of $h_0$ and $x$.
- Estimate beliefs from answers allowing for error in responses.
SECTION 3: MOTOR AND SOCIAL DEVELOPMENT

PART H: (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 without any help except for turning the water on and off? YES..... 1 NO..... 0 75/

5. Has your child ever counted 3 objects correctly? YES..... 1 NO..... 0 76/
Eliciting beliefs: Item response theory

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

- We observe $d_{i,j} = 1$ if $d_{i,j}^* \geq 0$ and $d_{i,j} = 0$, otherwise.

- Measure of (log of) human capital: $\ln h_i = \ln a_i + \frac{b_{2,j}}{b_{1,j}} \theta_i$.

- In this sense, $\theta_i$ is deviation from typical development for age.
Figure 4

Probability as a Function of Child's Age

- **Speak partial sentence, data**
- **Speak partial sentence, predicted**
Eliciting beliefs: Changing wording of the MSD Instrument

- In order to measure $E \left[ \ln h_{i,1} \mid h_0, x, \psi_i \right]$, we take the tasks from the MSD Scale, but instead of asking: “Has your child ever spoken a partial sentence with three words or more?”, we ask:
- **Method 1:** How likely is it that a baby will speak a partial sentence with three words or more by age 24 months?
- **Method 2:** What is the youngest and oldest age a baby learns to speak a partial sentence with three words or more?
We consider four scenarios:

- Scenario 1: Child is healthy at birth (e.g., normal gestation, birth weight, and birth length) and investment is high (e.g., six hours per day).
- Scenario 2: Child is healthy at birth and investment is low (e.g., two hours per day).
- Scenario 3: Child is not healthy at birth (e.g., premature, low birth weight, and small at birth) and investment is high.
- Scenario 4: Child is not healthy at birth and investment is low.

Scenarios are described to survey respondents through a video.
Method 1: Transforming probabilities into mean beliefs

- Method 1: How likely is it that a baby will speak a partial sentence with three words or more by age 24 months?
- Let’s say that when investment is high – that is, when $x = \bar{x}$ – the mother states that there is a 75% chance that the child will learn how to speak a partial sentence with three words or more.
- And when investment is low – that is, when $x = x$ – the mother states that there is a 25% chance that the child will learn how to speak a partial sentence with three words or more.
- We convert this probability statement into an age-equivalent statement using the NHANES data.
Figure 4
Probability as a Function of Child's Age

- **Speak partial sentence, data**
- **Speak partial sentence, predicted**
Method 2: Transforming age ranges into probabilities

- Method 2: What is the youngest and oldest age a baby learns to speak a partial sentence with three words or more?
- Let’s say that when investment is high, so that \( x = \bar{x} \), the mother states that the youngest and oldest ages a baby will learn how to speak a sentence with three words or more are, respectively, 18 and 28 months.
- And when investment is low, so that \( x = \underline{x} \), the mother states that the ages are 20 and 30 months.
- We need to transform the age ranges into probabilities. We use the age ranges to estimate a mother-specific IRT model.
Figure 3

Transforming age range into probability

- High investment
Figure 3
Transforming age range into probability

- Logistic prediction, high
- High investment

Graph showing the transformation of age range into probability with a logistic prediction curve.
Figure 3
Transforming age range into probability

Logistic prediction, high
High investment
Figure 3
Transforming age range into probability

- Logistic prediction, high
- Logistic prediction, low
- High investment
- Low investment
Figure 3
Transforming age range into probability

- Logistic prediction, high
- Logistic prediction, low
- High investment
- Low investment
Method 2: Transforming probabilities into mean beliefs

- Method 2: Given scenario for $h_0$ and $x$, how likely is it that a baby will speak a partial sentence with three words or more by age 24 months?

- Given maternal supplied age range and the logistic assumption, we conclude that when $x = \bar{x}$, the mother believes that there is a 75% chance that the child will learn how to speak a partial sentence with three words or more.

- Analogously, when $x = x$, the mother believes that there is a 25% chance that the child will learn how to speak a partial sentence with three words or more.

- We convert this probability statement into an age-equivalent statement using the NHANES data.
Figure 3

Expected development for two levels of investments ($x$)

**Age range to probability**
Speak partial sentence - MKIDS

**Probability to expected development**
Speak partial sentence - NHANES

Expected development for two levels of investments ($x$)
Recovering mean beliefs: Measurement error model

- Let $\ln q^L_{i,j,k}$ denote an error-ridden measure of $E[\ln h_{i,1} | h_{0,k}, x_k, \psi_i]$ generated by “how likely” questions:

  $$\ln q^L_{i,j,k} = E[\ln h_{i,1} | h_{0,k}, x_k, \psi_i] + \epsilon^L_{i,j,k}.$$ 

- Let $\ln q^A_{i,j,k}$ denote an error-ridden measure of $E[\ln h_{i,1} | h_{0,k}, x_k, \psi_i]$ generated by “age range” questions:

  $$\ln q^A_{i,j,k} = E[\ln h_{i,1} | h_{0,k}, x_k, \psi_i] + \epsilon^A_{i,j,k}.$$ 

- For each scenario, we have multiple measures of the same underlying latent variable.
Recovering mean beliefs:

- Use technology of skill formation, and the mother’s information set, to obtain:

\[
\ln q_{i,j,k}^L = \mu_{i,0} + \mu_{i,1} \ln h_{0,k} + \mu_{i,2} \ln x_k + \mu_{i,3} \ln h_{0,k} \ln x_k + \epsilon_{i,j,k}^L.
\]

\[
\ln q_{i,j,k}^A = \mu_{i,0} + \mu_{i,1} \ln h_{0,k} + \mu_{i,2} \ln x_k + \mu_{i,3} \ln h_{0,k} \ln x_k + \epsilon_{i,j,k}^A.
\]

- We have a factor model where:
  - \( \mu_i = (\mu_{i,0}, \mu_{i,1}, \mu_{i,2}, \mu_{i,3}) \) are the latent factors;
  - \( \lambda_k = (1, h_{0,k}, \ln x_k, \ln h_{0,k} \ln x_k) \) are the factor loadings;
  - \( \epsilon_{i,j,k} = (\epsilon_{i,j,k}^L, \epsilon_{i,j,k}^A) \) are the uniquenesses.
Eliciting beliefs: Intuitive explanation

- Let $E \left[ \ln h_{i,1} \mid h_0, h, \Psi_i \right]$ denote maternal expectation of child development at age 24 months conditional on the child’s initial level of human capital, investments, and the mother’s information set.

- Assume, for now, technology is Cobb-Douglas.

- Suppose we measure $E \left[ \ln h_{i,1} \mid h_0, x, \Psi_i \right]$ at two different levels of investments:

$$E \left[ \ln h_{i,1} \mid h_0, \bar{x}, \Psi_i \right] = \mu_{i,0} + \mu_{i,1} \ln h_0 + \mu_{i,2} \ln \bar{x}$$

$$E \left[ \ln h_{i,1} \mid h_0, x, \Psi_i \right] = \mu_{i,0} + \mu_{i,1} \ln h_0 + \mu_{i,2} \ln x$$

- Subtracting and re-organizing terms:

$$\mu_{i,2} = \frac{E \left[ \ln h_{i,1} \mid h_0, \bar{x}, \Psi_i \right] - E \left[ \ln h_{i,1} \mid h_0, x, \Psi_i \right]}{\ln \bar{x} - \ln \bar{x}}$$
Important issue

- We could use only one MSD item to elicit beliefs.
- But, if we use more items, we can relax assumptions about measurement error.
- And, we can check for consistency in answers.
Figure 5
Comparing answers across scenarios

Age range into probability

Probability into expected development

Comparing answers across scenarios
Table 5
Correlation between MSE and demographic characteristics of PHD Study Participants

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Standardized ( \mu_{i,\psi,1} )</th>
<th>Standardized ( \mu_{i,\psi,2} )</th>
<th>Standardized ( \mu_{i,\psi,3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummies for household income (( y ))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$25,000 \text{ per year} \leq y &lt; $55,000 \text{ per year}</td>
<td>0.2243 (0.1003)</td>
<td>0.3452 (0.0928)</td>
<td>0.1908 (0.1027)</td>
</tr>
<tr>
<td>$55,000 \text{ per year} \leq y &lt; $105,000 \text{ per year}</td>
<td>-0.1701 (0.1265)</td>
<td>0.3662 (0.1209)</td>
<td>-0.2460 (0.1135)</td>
</tr>
<tr>
<td>( y \geq $105,000 \text{ per year} )</td>
<td>-0.5060 (0.1278)</td>
<td>0.4694 (0.1405)</td>
<td>-0.5276 (0.1203)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.2746 (0.1581)</td>
<td>-0.5133 (0.1758)</td>
<td>0.0514 (0.1664)</td>
</tr>
<tr>
<td>Observations</td>
<td>822</td>
<td>822</td>
<td>822</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.0709</td>
<td>0.0641</td>
<td>0.0900</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses.
<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Both</th>
<th>Both</th>
<th>How Likely</th>
<th>How Likely</th>
<th>Age Range</th>
<th>Age Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardized $\mu_1$</td>
<td>-0.0237</td>
<td>-0.0015</td>
<td>-0.0946</td>
<td>-0.0577</td>
<td>-0.0136</td>
<td>0.0306</td>
</tr>
<tr>
<td></td>
<td>(0.0813)</td>
<td>(0.0740)</td>
<td>(0.0799)</td>
<td>(0.0742)</td>
<td>(0.0585)</td>
<td>(0.0530)</td>
</tr>
<tr>
<td>Standardized $\mu_2$</td>
<td>0.1667</td>
<td>0.1141</td>
<td>0.1185</td>
<td>0.0980</td>
<td>0.1699</td>
<td>0.0834</td>
</tr>
<tr>
<td></td>
<td>(0.0449)</td>
<td>(0.0385)</td>
<td>(0.0435)</td>
<td>(0.0395)</td>
<td>(0.0446)</td>
<td>(0.0383)</td>
</tr>
<tr>
<td>Standardized $\mu_3$</td>
<td>-0.0856</td>
<td>0.0096</td>
<td>-0.0401</td>
<td>0.0344</td>
<td>-0.0581</td>
<td>-0.0137</td>
</tr>
<tr>
<td></td>
<td>(0.0673)</td>
<td>(0.0611)</td>
<td>(0.0662)</td>
<td>(0.0618)</td>
<td>(0.0479)</td>
<td>(0.0422)</td>
</tr>
</tbody>
</table>

Demographic characteristics included?*  
<table>
<thead>
<tr>
<th>No</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observations 687 687 687 687 687 687  
R-squared 0.0369 0.2695 0.0343 0.2706 0.0314 0.2655

Table 6

Correlation between the HOME Score and MSE
Dependent variable: Standardized HOME Score

Robust standard errors in parentheses.

*Note: The following variables describe demographic characteristics: A dummy variable that takes the value one if the mother's year of birth is between 1978 and 1987 and zero otherwise; a dummy variable that takes the value one if the mother's year of birth is between 1988 and 1997 and zero otherwise; a dummy variable that takes the value one if the mother is Hispanic and zero otherwise; a dummy variable that takes the value one if the mother is non-Hispanic black and zero otherwise; a dummy variable that takes the value one if the mother has at least a college degree and zero otherwise; a dummy variable that takes the value one if the mother is married and zero otherwise; a dummy variable that takes the value one if the maternal score in CESD is greater than or equal to 16 and zero otherwise; three dummy variables indicating the level of household income (see Table 2). Appendix Table B5 displays all of the estimated coefficients and their robust standard errors.
Cunha, Gerdes, and Nihtianova (2018): Language Environment

- Group sessions of approximately 12-15 parents.
- Lasts 13 weeks.
- Each week there is a one-hour session:
  - Importance of language environment for language development.
  - Tips on how to improve language environment.
  - Objective feedback of the language environment based on recording provided by the parent.
Recording of the Parental Environment

1. Turn on the DLP and place it in the pocket of the child’s LENA clothing.

2. After completing recording, plug the DLP into a PC running LENA Pro. The sophisticated language environment analysis software automatically uploads and processes the audio file.

3. The software generates the LENA reports and other analyses.

4. Export data from LENA Pro to mine your LENA data and perform custom in-depth analyses.
Baseline: AWC and CTC

Flávio Cunha (Rice University)  Human Capital Formation in Childhood and Adolescence  July 10, 2018  127 / 143
During Intervention: AWC and CTC

Flávio Cunha (Rice University)
## Impact of the Intervention on the Language Environment

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random assignment</td>
<td>Random Assignment</td>
<td>Dummy for Attendance</td>
</tr>
<tr>
<td>Adult Word Counts</td>
<td>1,216.687</td>
<td>1,210.724</td>
<td>2,361.115</td>
</tr>
<tr>
<td></td>
<td>(1,549.894)</td>
<td>(1,346.493)</td>
<td>(2,615.559)</td>
</tr>
<tr>
<td>Conversation Turns</td>
<td>116.021*</td>
<td>117.170**</td>
<td>225.151**</td>
</tr>
<tr>
<td></td>
<td>(64.473)</td>
<td>(57.078)</td>
<td>(113.421)</td>
</tr>
<tr>
<td>Child Vocalizations</td>
<td>444.771</td>
<td>465.820*</td>
<td>863.127*</td>
</tr>
<tr>
<td></td>
<td>(281.505)</td>
<td>(259.943)</td>
<td>(496.136)</td>
</tr>
<tr>
<td>AVA (Standardized Score)</td>
<td>0.255</td>
<td>0.269*</td>
<td>0.466*</td>
</tr>
<tr>
<td></td>
<td>(0.155)</td>
<td>(0.141)</td>
<td>(0.256)</td>
</tr>
<tr>
<td>Observations</td>
<td>91</td>
<td>128</td>
<td>91</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
Impact of the Intervention on Parental Beliefs

Table 7

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Baseline</th>
<th>Endline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random assignment</td>
<td>0.2581</td>
<td>0.3310**</td>
</tr>
<tr>
<td></td>
<td>(0.1453)</td>
<td>(0.1143)</td>
</tr>
<tr>
<td>Dummy for LENA Start attendance</td>
<td>0.6301***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2086)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>134</td>
<td>128</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.0840</td>
<td>0.2156</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>First Stage</th>
<th>Second Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random assignment</td>
<td>0.4331*** (0.1234)</td>
<td>289.5987** (118.4140)</td>
</tr>
<tr>
<td>Maternal Beliefs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observations: 91
R-squared: 0.6804 (1) 0.4874 (2)

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
Infant stunting dropped from 29% to 14% in Peru between 2007 and 2014

Percentage of children under five affected by stunting
Nutritional supplementation trial from 1969 until 1977:
- A high-protein nutritional supplement was delivered in the two treatment villages (Atole).
- A non-protein supplement was delivered in two control villages (Fresco).
- Initial Height, Height at Month 24 Protein (and Calorie) intakes every 3 months in first 2 years (24-hour and 72-hour recall)
- Prices of eggs, chicken, pork, beef, dry beans, corn, and rice.
Adaptive expectations: Reference points for age two height at year $t$ were determined by the height of children born in year $t - 2$.

We show this implies two exclusion restrictions:

- Random assignment to treatment or control: Identifies coefficients on investment in production function.
- Interaction between random assignment and calendar time: Identifies preference parameter on reference point.
Consumption of Protein in Treatment vs Control Villages

Estimation: Identification III

- We find that there is a gap in protein choice up to month 24 between Atole and Fresco villages.
- Income is constant over time, and price is the same across locations.
- Only the increasing reference point gap, through $\lambda$, can explain the choice gap’s widening.
Age Two Height in Treatment vs. Control Villages

Estimation: Identification II

- We find that there is a gap in height at month 24 between Atole and Fresco villages.
- The gap is increasing over time.
- These curves are our $\mu_{R_2}$
Model Fit: Height

**Estimation Results: Fit of the Model III**

Fit general Height Change pattern.

![Graph showing height change over different birth cohorts between 1970 and 1975.](image)

We assume that parents, when making nutritional choice for children born in 1971, considers as reference height the current 2 year olds who were born in 1969. Reference point is identified from the increasing gap between atole and fresco villages.
Model Fit: Protein Consumption

Estimation Results: Fit of the Model IV

Fit Protein Trend over time

![Graph showing protein consumption trends](image)

We assume that parents, when making nutritional choice for children born in 1971 consider as reference height the current 2 year olds who were born in 1969. Reference point is identified from the increasing gap between atole and fresco villages.
Counterfactual 1a

For children in fresco villages:

- **Yellow**: Increase in height with just 42% price discount, fixing reference point.
- **Green**: Remaining contribution from reference point change.
Inequality in socio-economic outcomes is partly caused by inequality in human capital.

Inequality in human capital is partly caused by inequality in investments in human capital during early childhood, adolescence, and early adulthood.

Inequality in stocks of human capital has been increasing.

Inequality in investments in human capital has also been increasing.

Correlational studies. It is not determinate if we can make causal links from this data.
Conclusion: Skill Formation

- At different stages of the lifecycle, investments produce different dimensions of human capital.
- The skills acquired in one stage of the lifecycle promote the emergence of other skills in later stages (self-productivity).
- The skills acquired in different stages of the lifecycle complement each other (dynamic complementarity, “success begets success”).
- The evidence is built on estimation of technology of skill formation.
- To do so, we showed how we can
  - Address lack of cardinality of measures of human capital.
  - Address measurement error in measures of human capital.
  - Address endogeneity of investments.
Conclusion: Skill Formation

- Inequality in early investments in human capital is partially determined by:
  - Parental beliefs about the technology of skill formation.
  - Parental beliefs about what constitutes “normal” development.
- Intervention that provides information with feedback based on objective information positively affects parental beliefs, investments, and development.
- This is common in the intervention in Philadelphia, but also the intervention in Peru regarding stunting.
Conclusion: Skill Formation

- Lots of work for young researchers:
  - Theory: How to model within family decision making processes? How to model these processes when parents are not
  - Theory: How to model parent-child interaction (child is a “player”).
  - Data: How to measure investments? How to measure human capital in cardinal ways?
  - Data: Implement and evaluate pilot programs that can foster human capital formation.
  - Data and Theory: Identify mechanisms to validate or reject theories and to identify new opportunities for interventions.