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

(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 8.3 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).
Katz and Goldin (2007): College Graduation Conditional on Enrollment in USA

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)
Application Behavior of High-Achieving Students

Panel A: High-Income Students’ Portfolios of College Applications

Nonselective | Less-selective | Safety | Match | Reach
---|---|---|---|---
40% | 30% | 20% | 10% | 0%
Relatively few high-income students apply to non-selective schools.
High-income students applications are well-distributed among reach, match, and safety schools.

Panel B: Low-Income Students’ Portfolios of College Applications

Nonselective | Less-selective | Safety | Match | Reach
---|---|---|---|---
40% | 30% | 20% | 10% | 0%
The bulk of low-income students' applications go to nonselective schools.
Low-income students are less likely than their high-income counterparts to apply to a mix of match and reach schools.

College selectivity, measured as college’s median SAT score—student’s SAT score (in percentiles)

Source: Avery and Hoxby (2012).
1. Year of reference 2011. Countries are ranked in ascending order of the percentage-point difference between the 25-34 and 55-64 year-old population with tertiary education.

Source: OECD. Table A1.3a. See Annex 3 for notes (www.oecd.org/edu/eag.htm).

### Proportion of the 25‐34 year‐old population with tertiary education (left axis)

<table>
<thead>
<tr>
<th>Country</th>
<th>Proportion</th>
<th>Proportion</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Israel</td>
<td>44.50</td>
<td>46.52</td>
<td>-2.02</td>
</tr>
<tr>
<td>United States</td>
<td>44.04</td>
<td>41.81</td>
<td>2.23</td>
</tr>
<tr>
<td>Germany</td>
<td>28.96</td>
<td>26.43</td>
<td>2.52</td>
</tr>
<tr>
<td>Brazil</td>
<td>14.46</td>
<td>10.17</td>
<td>4.28</td>
</tr>
<tr>
<td>Estonia</td>
<td>39.84</td>
<td>35.50</td>
<td>4.34</td>
</tr>
<tr>
<td>Austria</td>
<td>23.03</td>
<td>16.73</td>
<td>6.29</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>56.97</td>
<td>49.16</td>
<td>7.81</td>
</tr>
<tr>
<td>Finland</td>
<td>39.75</td>
<td>31.39</td>
<td>8.36</td>
</tr>
<tr>
<td>Chile</td>
<td>22.48</td>
<td>13.06</td>
<td>9.42</td>
</tr>
<tr>
<td>Turkey</td>
<td>21.00</td>
<td>10.32</td>
<td>10.67</td>
</tr>
<tr>
<td>Italy</td>
<td>22.25</td>
<td>11.42</td>
<td>10.83</td>
</tr>
<tr>
<td>Denmark</td>
<td>40.24</td>
<td>28.67</td>
<td>11.57</td>
</tr>
<tr>
<td>Mexico</td>
<td>24.10</td>
<td>12.51</td>
<td>11.59</td>
</tr>
<tr>
<td>Switzerland</td>
<td>40.64</td>
<td>28.74</td>
<td>11.90</td>
</tr>
<tr>
<td>New Zealand</td>
<td>46.87</td>
<td>34.59</td>
<td>12.28</td>
</tr>
<tr>
<td>Canada</td>
<td>57.26</td>
<td>44.49</td>
<td>12.77</td>
</tr>
<tr>
<td>Slovak Republic</td>
<td>26.98</td>
<td>13.68</td>
<td>13.30</td>
</tr>
<tr>
<td>Iceland</td>
<td>38.38</td>
<td>24.94</td>
<td>13.44</td>
</tr>
<tr>
<td>Australia</td>
<td>47.23</td>
<td>33.03</td>
<td>14.20</td>
</tr>
<tr>
<td>Greece</td>
<td>34.74</td>
<td>20.04</td>
<td>14.71</td>
</tr>
<tr>
<td>Sweden</td>
<td>43.47</td>
<td>28.66</td>
<td>14.82</td>
</tr>
<tr>
<td>Norway</td>
<td>45.02</td>
<td>29.91</td>
<td>15.11</td>
</tr>
<tr>
<td>Hungary</td>
<td>30.43</td>
<td>15.32</td>
<td>15.11</td>
</tr>
<tr>
<td>Netherlands</td>
<td>43.04</td>
<td>27.85</td>
<td>15.18</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>27.83</td>
<td>12.63</td>
<td>15.20</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>47.86</td>
<td>32.62</td>
<td>15.24</td>
</tr>
<tr>
<td>Latvia</td>
<td>38.72</td>
<td>22.05</td>
<td>16.68</td>
</tr>
<tr>
<td>Portugal</td>
<td>28.33</td>
<td>11.07</td>
<td>17.26</td>
</tr>
<tr>
<td>Belgium</td>
<td>42.99</td>
<td>25.34</td>
<td>17.65</td>
</tr>
<tr>
<td>Slovenia</td>
<td>35.36</td>
<td>17.14</td>
<td>18.22</td>
</tr>
<tr>
<td>Spain</td>
<td>39.26</td>
<td>19.02</td>
<td>20.24</td>
</tr>
<tr>
<td>France</td>
<td>42.91</td>
<td>19.60</td>
<td>23.31</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>49.88</td>
<td>26.44</td>
<td>23.44</td>
</tr>
<tr>
<td>Ireland</td>
<td>49.21</td>
<td>24.86</td>
<td>24.35</td>
</tr>
<tr>
<td>Japan</td>
<td>58.55</td>
<td>32.05</td>
<td>26.51</td>
</tr>
<tr>
<td>Poland</td>
<td>40.80</td>
<td>12.63</td>
<td>28.17</td>
</tr>
<tr>
<td>Korea</td>
<td>65.69</td>
<td>13.54</td>
<td>52.14</td>
</tr>
</tbody>
</table>

OECD average: 39.20 24.19 15.01

**Figure 3**

Percentage of Younger and Older Adults with Tertiary Education

- **Difference between the 25-34 and 55-64 year-old population with tertiary education (right axis)**
- **Proportion of the 25-34 year-old population with tertiary education (left axis)**
- **Proportion of the 55-64 year-old population with tertiary education (left axis)**
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 technology of skill formation:

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

where $\gamma \in [0, 1]$ and $\phi \leq 1$. 
Taking first-order conditions:

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

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

which yields:

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

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

Plan

- 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

**Figure 4:** The Effect of Cognitive and Socio-emotional endowments, (log) Wages

Average percentile rank on anti-social behavior score, by income quartile.
Inequality in Health as Children Age


James Heckman

Economics and Econometrics of Human Development
Inequality in Skills are Partially the Result of Inequality in Investments: Cunha (2007)
Inequality in Investments as Children Age

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."

---

**Teachers rated Most Effective**  **Teachers rated Least Effective**

Percent of teachers in rating category

% of students eligible for free or reduced-price lunch

SOURCE: Ohio Department of Education

RICH EXNER, JAMES OWENS | THE PLAIN DEALER
The bar chart shows the average number of team sports offered based on the Student poverty (Free/Reduced Lunch) rate. The x-axis represents the different poverty rates (0-25% FRL, 26-50% FRL, 51-75% FRL, 76-100% FRL), and the y-axis shows the average number of team sports offered.
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)

10% 15% 20% 25% 30% 35% 40%

Source: Monitoring the Future

- Upper third in parents' education
- Lower third in parents' education
Inequality in Health Over Time
Inequality in Investments Over Time

![Graph showing inequality in investments over time. The x-axis represents years from 1965-66 to 2008-13, and the y-axis represents daily minutes in developmental childcare. Two lines are shown: one for both parents having a Bachelor's degree or more, and the other for both parents having a high school degree or less.](image-url)
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

Increasing Inequality in College Attendance

- Top quartile
- Third quartile
- Second quartile
- Bottom quartile


Flávio Cunha (Rice University)
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).

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.
In PPP and ABC, and for early education programs in general, non-cognitive skills are not typically followed in the long term.

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 Female</td>
<td>Male</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&lt;sup&gt;a&lt;/sup&gt;</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&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.615)</td>
<td>(0.135)</td>
</tr>
<tr>
<td>Yearly Labor Income, 2014 USD</td>
<td>40&lt;sup&gt;a&lt;/sup&gt;</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&lt;sup&gt;a&lt;/sup&gt;</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&lt;sup&gt;a&lt;/sup&gt;</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&lt;sup&gt;d&lt;/sup&gt;</td>
<td>≤40&lt;sup&gt;a&lt;/sup&gt;</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&lt;sup&gt;a&lt;/sup&gt;</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&lt;sup&gt;a&lt;/sup&gt;</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&lt;sup&gt;a&lt;/sup&gt;</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&lt;sup&gt;a&lt;/sup&gt;</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: Flávio Cunha (Rice University)

Human Capital Formation in Childhood and Adolescence

July 9, 2018 37 / 169
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 χ² 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:**
    - Working memory;
    - Inhibitory control;
    - Cognitive flexibility.
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.
Skills Developed in Adolescence

- **Adolescent development:**
  - Fast development of the reward system potentialized by the influence of peers.
  - Slow development of emotional intelligence: self regulation:
    - Patience for reflection and thoughtfulness;
    - Comfort with ambiguity and change;
    - Ability to say no to impulsive urges.
Technology of Skill Formation

- 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 $h_i$ denote the human capital at beginning of adulthood. Assume that:

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

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

Note that:

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

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

subject to the technology of skill formation:

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

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:

$$\frac{(1 - \gamma a) \gamma e}{\gamma a} > 1 + r$$
Boundary Solution when $\phi \to -\infty$

- 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.
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></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</td>
<td>=, ≠, ≤, ≥, +, -</td>
<td>Educational attainment, dates</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>
Problem 1: Cardinality of Human Capital

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, $t = 1, \ldots, T$, 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 $\ln Y_i \sim N(\mu_h, \sigma_{h,t}^2 + \sigma_{\nu,t}^2)$.
Problem 1: Cardinality of Human Capital

- 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:

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

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

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

Solution: We need to transform at least one of the test scores at \( t \) 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:

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

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

- We conclude that:

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

- 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

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

- 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 $\varepsilon_{i,t,1}$ 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

\[
m_{i,t,1} = \frac{\bar{m}_{i,t,1}}{s_{t,1}}
\]

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

- Now we have a rescaled score that has a cardinal scale.
Applications of Anchoring: Bond & Lang (2018)

Figure 1: Raw Difference in Expected White Grade Completion conditional on Test Score

![Graph showing the raw difference in expected white grade completion conditional on test score for Math, Reading Recognition, and Reading Comprehension.](image_url)
Applications of Anchoring: Bond & Lang (2018)

Figure 2: Measurement Error Adjusted Difference in Achievement in Units of Predicted White Education

<table>
<thead>
<tr>
<th>Grade</th>
<th>Math</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade</th>
<th>Reading Recognition</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade</th>
<th>Reading Comprehension</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 $i = 1, ..., I$ and $t = 1, ..., T$.

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:

\[ p_{i,t} = \delta_t + \kappa_t \ln x_{i,t} + \zeta_{i,t} \]
Problem 2: Latent Factors

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

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

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

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

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

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

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

- Note that \( \tilde{\varepsilon}_{i,t} \sim N \left( 0, \left[ \beta_t \Sigma_{\varepsilon}^{-1} \beta_t \right]^{-1} \right) \) and
  \[ \tilde{\zeta}_{i,t} \sim N \left( 0, \left[ \kappa_t \Sigma_{\zeta}^{-1} \kappa_t \right]^{-1} \right) \] 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 \( \left[ \kappa_t \Sigma_{\zeta}^{-1} \kappa_t \right]^{-1} \) 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.
Problem 2: Latent Factors

Let \( \hat{\gamma}^B \) denote the OLS estimator that uses Bartlett scores \( h^B \) and \( w^B \) (assumed to be exogenous).

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

Note that \( w^B \) 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 \( \left[ \left( w^B_t \right)^T w^B_t \right] \).
### Problem 2: Latent Factors

- 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>$\text{plim} \ [w_t^T w_t]$</th>
<th>$\text{plim} \ [(w_t^B)^T w_t^B]$</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>$E (h_t^2) + \text{Var} (\xi_t)$</td>
</tr>
<tr>
<td>$(2,3)$</td>
<td>$E (x_t h_t^2)$</td>
<td>$E (x_t h_t^2) + E (x_t) \text{Var} (\xi_t)$</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

$$\Delta = E (x_t^2) \text{Var} (\varepsilon_t) + E (h_t^2) \text{Var} (\xi_t) + \text{Var} (\xi_t) + \text{Var} (\varepsilon_t)$$
Problem 2: Latent Factors

- 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 $i = 1, ..., I$ and $t = 1, ..., T$.

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:

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

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

- 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 \( \ln h_{i,t} \).
### Table V

*The Technology for Cognitive and Noncognitive Skill Formation*

Estimated Along With Investment Equation With Linear Anchoring on Educational Attainment (Years of Schooling); Factors Normally Distributed

**Panel A: Technology of Cognitive Skill Formation (Next Period Cognitive Skills)**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>First Stage</th>
<th>Second Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Period Cognitive Skills (Self-Productivity)</td>
<td>$\gamma_{1,C,1}$ 0.426 (0.03)</td>
<td>$\gamma_{2,C,1}$ 0.901 (0.01)</td>
</tr>
<tr>
<td>Current Period Noncognitive Skills (Cross-Productivity)</td>
<td>$\gamma_{1,C,2}$ 0.127 (0.04)</td>
<td>$\gamma_{2,C,2}$ 0.014 (0.01)</td>
</tr>
<tr>
<td>Current Period Investments</td>
<td>$\gamma_{1,C,3}$ 0.322 (0.04)</td>
<td>$\gamma_{2,C,3}$ 0.024 (0.01)</td>
</tr>
<tr>
<td>Parental Cognitive Skills</td>
<td>$\gamma_{1,C,4}$ 0.059 (0.02)</td>
<td>$\gamma_{2,C,4}$ 0.062 (0.01)</td>
</tr>
<tr>
<td>Parental Noncognitive Skills</td>
<td>$\gamma_{1,C,5}$ 0.066 (0.04)</td>
<td>$\gamma_{2,C,5}$ 0.000 (0.01)</td>
</tr>
<tr>
<td>Complementarity Parameter</td>
<td>$\phi_{1,C}$ 0.748 (0.25)</td>
<td>$\phi_{2,C}$ -1.207 (0.16)</td>
</tr>
<tr>
<td>Implied Elasticity Parameter</td>
<td>$1/(1-\phi_{1,C})$ 3.968</td>
<td>$1/(1-\phi_{2,C})$ 0.453</td>
</tr>
<tr>
<td>Variance of Shocks $\eta_{t,C}$</td>
<td>$\delta^2_{1,C}$ 0.159 (0.01)</td>
<td>$\delta^2_{2,C}$ 0.092 (0.00)</td>
</tr>
</tbody>
</table>

**Panel B: Technology of Noncognitive Skill Formation (Next Period Noncognitive Skills)**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>First Stage</th>
<th>Second Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parental Cognitive Skills</td>
<td>$\gamma_{1,N,4}$ 0.000 (0.01)</td>
<td>$\gamma_{2,N,4}$ 0.000 (0.01)</td>
</tr>
<tr>
<td>Parental Noncognitive Skills</td>
<td>$\gamma_{1,N,5}$ 0.093 (0.03)</td>
<td>$\gamma_{2,N,5}$ 0.011 (0.02)</td>
</tr>
<tr>
<td>Complementarity Parameter</td>
<td>$\phi_{1,N}$ 0.017 (0.27)</td>
<td>$\phi_{2,N}$ -0.323 (0.21)</td>
</tr>
<tr>
<td>Elasticity Parameter</td>
<td>$1/(1-\phi_{1,N})$ 1.017</td>
<td>$1/(1-\phi_{2,N})$ 0.756</td>
</tr>
<tr>
<td>Variance of Shocks $\eta_{t,N}$</td>
<td>$\delta^2_{1,N}$ 0.170</td>
<td>$\delta^2_{2,N}$ 0.104</td>
</tr>
</tbody>
</table>

Note: Standard errors in parenthesis.
Estimates of the Technology of Skill Formation

<table>
<thead>
<tr>
<th>First Stage</th>
<th>Second Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
<td>Parameters</td>
</tr>
<tr>
<td>$\gamma_{1,C,1}$</td>
<td>0.426</td>
</tr>
<tr>
<td>($0.03$)</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{1,C,2}$</td>
<td>0.901</td>
</tr>
<tr>
<td>($0.01$)</td>
<td></td>
</tr>
<tr>
<td>Current Period Noncognitive Skills (Self-Productivity)</td>
<td>Current Period Noncognitive Skills (Cross-Productivity)</td>
</tr>
<tr>
<td>$\gamma_{1,N,1}$</td>
<td>0.000</td>
</tr>
<tr>
<td>($0.02$)</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{1,N,2}$</td>
<td>0.712</td>
</tr>
<tr>
<td>($0.03$)</td>
<td></td>
</tr>
<tr>
<td>Current Period Investments</td>
<td>Current Period Investments</td>
</tr>
<tr>
<td>$\gamma_{1,N,3}$</td>
<td>0.195</td>
</tr>
<tr>
<td>($0.03$)</td>
<td></td>
</tr>
<tr>
<td>Parental Cognitive Skills</td>
<td>Parental Cognitive Skills</td>
</tr>
<tr>
<td>$\gamma_{1,N,4}$</td>
<td>0.000</td>
</tr>
<tr>
<td>($0.01$)</td>
<td></td>
</tr>
<tr>
<td>Parental Noncognitive Skills</td>
<td>Parental Noncognitive Skills</td>
</tr>
<tr>
<td>$\gamma_{1,N,5}$</td>
<td>0.093</td>
</tr>
<tr>
<td>($0.03$)</td>
<td></td>
</tr>
<tr>
<td>Complementarity Parameter</td>
<td>Complementarity Parameter</td>
</tr>
<tr>
<td>$\phi_{1,N}$</td>
<td>0.017</td>
</tr>
<tr>
<td>($0.27$)</td>
<td></td>
</tr>
<tr>
<td>Elasticity Parameter</td>
<td>Elasticity Parameter</td>
</tr>
<tr>
<td>$1/(1-\phi_{1,C})$</td>
<td>3.968</td>
</tr>
<tr>
<td></td>
<td>$1/(1-\phi_{2,C})$</td>
</tr>
<tr>
<td></td>
<td>0.453</td>
</tr>
<tr>
<td>Variance of Shocks $\eta_{C,t}$</td>
<td>Variance of Shocks $\eta_{N,t}$</td>
</tr>
<tr>
<td>$\delta^2_{1,C}$</td>
<td>0.159</td>
</tr>
<tr>
<td>($0.01$)</td>
<td></td>
</tr>
<tr>
<td>$\delta^2_{2,C}$</td>
<td>0.092</td>
</tr>
<tr>
<td>($0.00$)</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Technology of Noncognitive Skill Formation (Next Period Noncognitive Skills)

<table>
<thead>
<tr>
<th>First Stage Parameters</th>
<th>Second Stage Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{1,N,1}$</td>
<td>0.000</td>
</tr>
<tr>
<td>($0.01$)</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{1,N,2}$</td>
<td>0.868</td>
</tr>
<tr>
<td>($0.01$)</td>
<td></td>
</tr>
<tr>
<td>Current Period Investments</td>
<td>Current Period Investments</td>
</tr>
<tr>
<td>$\gamma_{1,N,3}$</td>
<td>0.121</td>
</tr>
<tr>
<td>($0.03$)</td>
<td></td>
</tr>
<tr>
<td>Parental Cognitive Skills</td>
<td>Parental Cognitive Skills</td>
</tr>
<tr>
<td>$\gamma_{1,N,4}$</td>
<td>0.000</td>
</tr>
<tr>
<td>($0.01$)</td>
<td></td>
</tr>
<tr>
<td>Parental Noncognitive Skills</td>
<td>Parental Noncognitive Skills</td>
</tr>
<tr>
<td>$\gamma_{1,N,5}$</td>
<td>0.011</td>
</tr>
<tr>
<td>($0.02$)</td>
<td></td>
</tr>
<tr>
<td>Complementarity Parameter</td>
<td>Complementarity Parameter</td>
</tr>
<tr>
<td>$\phi_{1,N}$</td>
<td>-0.323</td>
</tr>
<tr>
<td>($0.21$)</td>
<td></td>
</tr>
<tr>
<td>Elasticity Parameter</td>
<td>Elasticity Parameter</td>
</tr>
<tr>
<td>$1/(1-\phi_{1,N})$</td>
<td>1.017</td>
</tr>
<tr>
<td></td>
<td>$1/(1-\phi_{2,N})$</td>
</tr>
<tr>
<td></td>
<td>0.756</td>
</tr>
<tr>
<td>Variance of Shocks $\eta_{N,t}$</td>
<td>Variance of Shocks $\eta_{N,t}$</td>
</tr>
<tr>
<td>$\delta^2_{1,N}$</td>
<td>0.170</td>
</tr>
<tr>
<td>($0.01$)</td>
<td></td>
</tr>
<tr>
<td>$\delta^2_{2,N}$</td>
<td>0.104</td>
</tr>
<tr>
<td>($0.00$)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard errors in parenthesis.
Interpretation of Findings: Maximizing Average Education

- Suppose that \( H \) children are born, \( h = 1, \ldots, H \).
- These children represent draws from the distribution of initial conditions \( F(\theta_{c,1,h}, \theta_{n,1,h}, \theta_{c,p}, \theta_{n,p}, \pi) \).
- We want to allocate finite resources \( B \) across these children for early and late investments.
- Formally:

\[
S^* = \max \frac{1}{H} \left[ \sum_{h=1}^{H} S(\theta_{c,3}, \theta_{n,3}, \pi_h) \right]
\]

subject to the technologies for the formation of cognitive and noncognitive skills as well as:

\[
\sum_{h=1}^{H} (x_{1,h} + x_{2,h}) = B
\]
Another possibility is to minimize aggregate crime (average crime per individual).

This will lead to different optimal ratios as crime is more sensitive to changes in noncognitive skills.

Relative to cognitive skills, noncognitive skills are more malleable at later ages.
FIGURE 5.—Ratio of early to late investments by maternal cognitive and noncognitive skills maximizing aggregate education (left) and minimizing aggregate crime (right) (other endowments held at mean levels).

FIGURE 6.—Densities of ratio of early to late investments maximizing aggregate education versus minimizing aggregate crime.

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

The diagram shows the number of words addressed to children at different ages by parents in three different income groups: 13 professional parents, 23 working-class parents, and 6 welfare parents. The x-axis represents the age of the child in months, while the y-axis represents the number of words addressed to the child.
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:

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

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

- We do not impose any a priori restrictions on the parameters of these distributions.
Typical Textbook Model

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

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.

Horizontal axis = child development
Vertical axis = household consumption
Introducing Heterogeneity in Reference Points

Horizontal axis = child development
Vertical axis = household consumption

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

Horizontal axis = child development
Vertical axis = household consumption

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
Introducing Heterogeneity in Reference Points

Horizontal axis = child development
Vertical axis = household consumption

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 home visitation affect maternal subjective expectations?
- 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:

\[ E (\ln h_{i,1} | h_{0,i}, x_i, \Psi_i) = \mu_{i,0} + \mu_{i,1} \ln h_{0,i} + \mu_{i,2} \ln x_i + \mu_{i,3} \ln h_{0,i} \ln x_i \]
Model: Preferences and budget constraint

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:

\[ c_i + p x_i = y_i. \]
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

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

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 $E \left[ \ln h_{i,1} \mid h_0, x, \psi_i \right]$:
  - 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 II: (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.... ?  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?
Eliciting beliefs: Scenarios of human capital and investments

- 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 $x = \underline{x}$ – 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

Probability
Age (in months)
Logistic prediction, high
High investment

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

Probability
0.25 0.5 0.75 1
0 4 8 12 16 20 24 28 32 36 40 44 48
Age (in months)
Logistic prediction, high
High investment

- Logistic prediction, high
- High 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 $x = \bar{x}$, 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
- MKIDS

Probability to expected development
- NHANES

Speak partial sentence

Data Predicted

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

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

  $$\ln q_{i,j,k}^L = E \left[ \ln h_{i,1} \mid h_{0,k}, x_k, \psi_i \right] + \epsilon_{i,j,k}^L.$$

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

  $$\ln q_{i,j,k}^A = E \left[ \ln h_{i,1} \mid h_{0,k}, x_k, \psi_i \right] + \epsilon_{i,j,k}^A.$$

- 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 $E [\ln h_{i,1} | h_0, h, \Psi_i]$ 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 $E [\ln h_{i,1} | h_0, x, \Psi_i]$ at two different levels of investments:

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

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

- Subtracting and re-organizing terms:

  $$\mu_{i,2} = \frac{E [\ln h_{i,1} | h_0, x, \Psi_i] - E [\ln h_{i,1} | h_0, x, \Psi_i]}{\ln x - \ln \underline{x}}$$
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

Figure 5
### 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</td>
<td>0.3452</td>
<td>0.1908</td>
</tr>
<tr>
<td>$55,000 \text{ per year} \leq y &lt; $105,000 \text{ per year}$</td>
<td>-0.1701</td>
<td>0.3662</td>
<td>-0.2460</td>
</tr>
<tr>
<td>$y \geq $105,000 \text{ per year}$</td>
<td>-0.5060</td>
<td>0.4694</td>
<td>-0.5276</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.2746</td>
<td>-0.5133</td>
<td>0.0514</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 6
Correlation between the HOME Score and MSE

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Both</th>
<th>Both</th>
<th>How Likely Only</th>
<th>How Likely Only</th>
<th>Age Range Only</th>
<th>Age Range Only</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>(0.0813)</td>
<td>(0.0740)</td>
<td>(0.0799)</td>
<td>(0.0742)</td>
<td>(0.0585)</td>
<td>(0.0530)</td>
<td></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>(0.0449)</td>
<td>(0.0385)</td>
<td>(0.0435)</td>
<td>(0.0395)</td>
<td>(0.0446)</td>
<td>(0.0383)</td>
<td></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>(0.0673)</td>
<td>(0.0611)</td>
<td>(0.0662)</td>
<td>(0.0618)</td>
<td>(0.0479)</td>
<td>(0.0422)</td>
<td></td>
</tr>
</tbody>
</table>

Demographic characteristics included?*  

<table>
<thead>
<tr>
<th>Observations</th>
<th>687</th>
<th>687</th>
<th>687</th>
<th>687</th>
<th>687</th>
<th>687</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.0369</td>
<td>0.2695</td>
<td>0.0343</td>
<td>0.2706</td>
<td>0.0314</td>
<td>0.2655</td>
</tr>
</tbody>
</table>

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.
Attanasio and colleagues have adapted an influential home visitation program from the Jamaica.

Gertler et al (2014) follow up participants when they were in the early 20s and find positive impacts of the program on educational attainment and labor market outcomes.


Question: Why has investment increased?
Project Timeline

- **Baseline:**
  - Measure $h_0$: BSID (cognitive, receptive language, expressive language) and MacArthur-Bates Language Scale.
  - Intervention assignment: $d_i \in \{0, 1\}$.

- **First follow up:**
  - Measure $h_1$: BSID (cognitive, receptive language, expressive language) and MacArthur-Bates Language Scale.
  - Measure $x_1$: Family Care Indicators (materials, activities) and time diary for child.

- **Second follow up:**
  - Measure $h_2$: BSID (cognitive, receptive language, expressive language) and MacArthur-Bates Language Scale.
  - Measure $x_2$: Family Care Indicators (materials, activities) and time diary for child.
  - Measure $\mu$: maternal beliefs
We would like to elicit parental beliefs about the parameters of the technology of skill formation:

\[ E [\ln h_{i,1} | h_0, x_1, \Omega] = \mu_0 + \mu_1 \ln h_0 + \mu_2 \ln x_1 \]

Approach proposed in Cunha, Elo, and Culhane (2013):
Step 3: Choose items for the elicitation instrument. Choose 9 words from MacArthur Bates

- 3 words are “easy” (α is high): \( w^e = (w^e_1, w^e_2, w^e_3) \).
- 3 words are “moderate” (α is average): \( w^m = (w^m_1, w^m_2, w^m_3) \).
- 3 words are “hard” (α is low): \( w^h = (w^h_1, w^h_2, w^h_3) \).

These 9 words are the items in the elicitation instrument:
\[ W = (w^e, w^m, w^h) \].
Step 4: Choose scenarios for human capital at baseline and investments between beginning and end of the program:

- Scenario 1: $h_0$ is low (at baseline, child understands only “easy” words) and $x_1$ is low (few materials, few activities): $s_1 = (h_0^L, x_1^L)$.
- Scenario 2: $h_0$ is low and $x_1$ is high (lots of materials, lots of activities): $s_2 = (h_0^L, x_1^H)$.
- Scenario 3: $h_0$ is high (at baseline, child understands “easy” and “difficult” words) and $x_1$ is low: $s_3 = (h_0^H, x_1^L)$.
- Scenario 4: $h_0$ is high and $x_1$ is high: $s_1 = (h_0^H, x_1^H)$.
Eliciting Beliefs - Scenarios
The estimates of the IRT model and the definitions of the scenarios can be used to establish what part of the relevant domains are spanned by the scenarios. All the chosen words had relatively high loading factors’ $\beta$’s. Easy words had low intercepts ($\alpha$’s), hard words had high $\alpha$’s and medium words medium $\alpha$s.
High’ and ’Low’ levels of maternal investment
Step 5: Elicitation instrument is based on MacArthur Bates CDI.

In order to measure $E [\ln h_{i,1} | h_0, x]$, we select words, but instead of asking: “Has your child ever spoken word X?” , we ask:

Suppose [describe scenario for $h_0$ and $x$]. At what age do you think the child will speak words $w^d$?

The index $d$ denotes the difficulty of the words.

For every combination of scenarios of $\ln h_0$ and $\ln x$, parents answer three questions.

Let $a_{i,d,s}$ denote the age reported by mother $i$ answer for word difficulty level $d$ when scenario was $s$.

For each parent, we have 12 answers. Why so many? Measurement error.
Transforming age answers into mean beliefs

- Step 6: Now we go from $a_{i,d,s}$ to $E \left[ \ln h_{i,1} \mid h_0, x_1, \Omega \right]$.
- Easier to explain with the following example.
### From maternal answers to model variables

<table>
<thead>
<tr>
<th></th>
<th>Scenario 1: $h_0^L, x_1^L$</th>
<th>Scenario 4: $h_0^H, x_1^H$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Easy words</td>
<td>Medium words</td>
</tr>
<tr>
<td>Maternal answer</td>
<td>27</td>
<td>30</td>
</tr>
<tr>
<td>Typical age children learn</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td>words</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference (developmental</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>delay)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronological age (36 months)</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>Developmental age</td>
<td>31</td>
<td>30</td>
</tr>
<tr>
<td>Error-ridden measure of $E$</td>
<td>3.43</td>
<td>3.40</td>
</tr>
<tr>
<td>$(\ln h_1</td>
<td>h_0^s, x_1^s, \Omega)$</td>
<td></td>
</tr>
</tbody>
</table>
Estimating beliefs: Intuitive explanation

- Let $E[\ln h_{i,1} | h_0, x]$ denote maternal expectation of child development at follow-up conditional on the child’s initial level of human capital, investments, and the mother’s information set.

- Let $h_{i,d,s}$ denote the error-ridden maternal report of $E[\ln h_{i,1} | h_0, x]$. Define the measurement error as $\eta_{i,d,s}$:

  $$h_{i,d,s} = E[\ln h_{i,1} | h_0, x] + \eta_{i,d,s}$$

- Now:

  $$h_{i,d,s} = \mu_0 + \mu_1 \ln h_0^s + \mu_2 \ln x_1^s + \eta_{i,d,s}$$

- Beliefs are latent factors with fixed factor loadings. We can relax assumptions on $\eta_{i,d,s}$. 
### Table: Answers about the Outcomes of Maternal Investment and Initial Conditions

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St. Dv.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low Initial Conditions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Investment</td>
<td>18.3</td>
<td>6.2</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>23.5</td>
<td>7.3</td>
<td>10</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>29.5</td>
<td>8.8</td>
<td>11</td>
<td>48</td>
</tr>
<tr>
<td>High Investment</td>
<td>15.8</td>
<td>5.7</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>20.1</td>
<td>6.8</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>25.0</td>
<td>8.2</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>14.4</td>
<td>4.8</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td>Low Investment</td>
<td>18.0</td>
<td>5.6</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>22.3</td>
<td>7.2</td>
<td>10</td>
<td>48</td>
</tr>
<tr>
<td>High Initial Conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Investment</td>
<td>13.5</td>
<td>5.3</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>16.7</td>
<td>5.9</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>20.3</td>
<td>7.2</td>
<td>9</td>
<td>48</td>
</tr>
</tbody>
</table>
Maternal Beliefs of Child Development

Density (kernel=epanechnikov)

developmental age (logs)

Low Initial Conditions, Low Investment
High,Low

Low,High

High,High

Subjective Expected Returns

Expected returns
Results Expected returns
Subjective Expected Returns by initial condition
0 2 4 6 8
Density (kernel=epanechnikov)
−1 −.5 0 .5 1
developmental age (logs)
Low Initial Condition High Initial Condition
Returns to Investment
Attanasio (UCL, IFS), Cunha (Rice), Jervis (UCL, IFS)
Results Expected returns
Subjective Expected Returns by initial conditions and Intervention

Returns to Investment

Low Initial Condition

High Initial Condition

Density (kernel=epanechnikov)

developmental age (logs)

Control Stimulation

Low Initial Condition

High Initial Condition

Returns to Investment

Attanasio (UCL, IFS), Cunha (Rice), Jervis (UCL, IFS)

Parental Beliefs and Investments in Human Capital

U of Illinois - 27/3/2017 56 / 65
\[ ElnH_t^i = ElnH_t^i = \delta_0 + \delta_1 lnH_{t-1}^i + \delta_2 lnx_t^i + \delta_3 lnH_{t-1}^i lnx_t^i, \quad t = 1, 2 \]

**Table: Objective Estimation of the Production Function**

<table>
<thead>
<tr>
<th></th>
<th>First Stage</th>
<th></th>
<th>Second Stage</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scaled</td>
<td>Standardized</td>
<td>Scaled</td>
<td>Standardized</td>
</tr>
<tr>
<td>Log of Investment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.0877</td>
<td>0.0000</td>
<td>1.7260</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>(0.5422)</td>
<td>(0.1541)</td>
<td>(0.1142)</td>
<td>(0.5967)</td>
</tr>
<tr>
<td>Log of Human Capital at Baseline</td>
<td>1.1626</td>
<td>0.1896</td>
<td>0.4630</td>
<td>0.3776</td>
</tr>
<tr>
<td></td>
<td>(0.1879)</td>
<td>(0.0306)</td>
<td>(0.0910)</td>
<td>(0.0742)</td>
</tr>
<tr>
<td>Log of Investments at Follow Up</td>
<td></td>
<td></td>
<td>0.1461</td>
<td>0.7309</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0707)</td>
<td>(0.1533)</td>
</tr>
<tr>
<td>Treatment dummy</td>
<td>0.1740</td>
<td>0.0909</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0587)</td>
<td>(0.0307)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\[ E_i \ln(h_1) = \mu_{0,i} + \mu_{1,i} \ln(h_0) + \mu_{2,i} \ln(X) + \mu_{3,i} [\ln(h_0) \ln(X)] \]

Our procedure yields estimates of the coefficients for each mother:

**Table: Estimation of Perceived Production Function**

<table>
<thead>
<tr>
<th>Dependent Variable: Expected Human Capital at Follow Up</th>
<th>Cobb Douglas</th>
<th>Translog</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scaled</td>
<td>Scaled</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.7410</td>
<td>-0.0450</td>
</tr>
<tr>
<td></td>
<td>(0.0401)</td>
<td>(0.1313)</td>
</tr>
<tr>
<td>Human Capital at Baseline</td>
<td>0.5703</td>
<td>1.1922</td>
</tr>
<tr>
<td></td>
<td>(0.0132)</td>
<td>(0.0450)</td>
</tr>
<tr>
<td>Investment</td>
<td>0.0542</td>
<td>0.5638</td>
</tr>
<tr>
<td></td>
<td>(0.0023)</td>
<td>(0.0357)</td>
</tr>
<tr>
<td>Investment x Human Capital at B.</td>
<td>-0.1775</td>
<td>-0.1385</td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td>(0.0096)</td>
</tr>
</tbody>
</table>
## Table: Beliefs and Demographic Characteristics

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Standardized $\mu_0$</th>
<th>Standardized $\mu_1$</th>
<th>Standardized $\mu_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment dummy</td>
<td>0.0018</td>
<td>-0.0118</td>
<td>0.0410</td>
</tr>
<tr>
<td></td>
<td>(0.0607)</td>
<td>(0.0611)</td>
<td>(0.0611)</td>
</tr>
<tr>
<td>Child is male</td>
<td>0.0132</td>
<td>-0.0042</td>
<td>-0.0448</td>
</tr>
<tr>
<td></td>
<td>(0.0613)</td>
<td>(0.0616)</td>
<td>(0.0613)</td>
</tr>
<tr>
<td>Standardized Human Capital at Baseline</td>
<td>-0.0350</td>
<td>0.0283</td>
<td>0.0464</td>
</tr>
<tr>
<td></td>
<td>(0.0296)</td>
<td>(0.0298)</td>
<td>(0.0298)</td>
</tr>
<tr>
<td>Standardized Household Wealth</td>
<td>-0.0456</td>
<td>0.0387</td>
<td>0.0518*</td>
</tr>
<tr>
<td></td>
<td>(0.0311)</td>
<td>(0.0312)</td>
<td>(0.0313)</td>
</tr>
<tr>
<td>Mother’s Standardized Raven Score</td>
<td>-0.1960***</td>
<td>0.1661***</td>
<td>0.2230***</td>
</tr>
<tr>
<td></td>
<td>(0.0309)</td>
<td>(0.0311)</td>
<td>(0.0326)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0179</td>
<td>0.0150</td>
<td>0.0212</td>
</tr>
<tr>
<td></td>
<td>(0.0524)</td>
<td>(0.0522)</td>
<td>(0.0552)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,017</td>
<td>1,017</td>
<td>1,017</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.0487</td>
<td>0.0350</td>
<td>0.0634</td>
</tr>
</tbody>
</table>

Standard errors (in parentheses) are clustered at municipality level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
## Table: Beliefs and Investment

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Standardized Log Investment</th>
<th>Standardized Time</th>
<th>Standardized Activities</th>
<th>Standardized Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardized $\mu_2$</td>
<td>0.1084***</td>
<td>0.0912***</td>
<td>0.0617*</td>
<td>0.0652**</td>
</tr>
<tr>
<td></td>
<td>(0.0294)</td>
<td>(0.0309)</td>
<td>(0.0317)</td>
<td>(0.0286)</td>
</tr>
<tr>
<td>Dummy for Treatment</td>
<td>0.1762***</td>
<td>0.0266</td>
<td>0.0299</td>
<td>0.2745***</td>
</tr>
<tr>
<td></td>
<td>(0.0622)</td>
<td>(0.0625)</td>
<td>(0.0620)</td>
<td>(0.0620)</td>
</tr>
<tr>
<td>Dummy for Male</td>
<td>-0.0081</td>
<td>-0.0031</td>
<td>-0.0034</td>
<td>-0.0115</td>
</tr>
<tr>
<td></td>
<td>(0.0597)</td>
<td>(0.0620)</td>
<td>(0.0620)</td>
<td>(0.0607)</td>
</tr>
<tr>
<td>Standardized Human Capital at Birth</td>
<td>0.1469***</td>
<td>0.0830***</td>
<td>0.1042***</td>
<td>0.1220***</td>
</tr>
<tr>
<td></td>
<td>(0.0303)</td>
<td>(0.0302)</td>
<td>(0.0340)</td>
<td>(0.0315)</td>
</tr>
<tr>
<td>Standardized Household Wealth</td>
<td>0.1196***</td>
<td>0.0287</td>
<td>0.0791***</td>
<td>0.1473***</td>
</tr>
<tr>
<td></td>
<td>(0.0315)</td>
<td>(0.0328)</td>
<td>(0.0303)</td>
<td>(0.0336)</td>
</tr>
<tr>
<td>Mother's Standardized Raven Score</td>
<td>0.1862***</td>
<td>0.0819**</td>
<td>0.1567***</td>
<td>0.1446***</td>
</tr>
<tr>
<td></td>
<td>(0.0325)</td>
<td>(0.0330)</td>
<td>(0.0319)</td>
<td>(0.0347)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0911**</td>
<td>-0.0877</td>
<td>-0.0147</td>
<td>-0.1400***</td>
</tr>
<tr>
<td></td>
<td>(0.0458)</td>
<td>(0.0544)</td>
<td>(0.0543)</td>
<td>(0.0543)</td>
</tr>
<tr>
<td>Observations</td>
<td>1.017</td>
<td>1.017</td>
<td>1.017</td>
<td>1.017</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.0199</td>
<td>0.0172</td>
<td>0.0086</td>
<td>0.0256</td>
</tr>
</tbody>
</table>

Standard errors (in parentheses) are clustered at municipality level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 

---

**Flávio Cunha (Rice University)**

**Parental Beliefs and Investments in Human Capital**

**U of Illinois - 27/3/2017**

---

**Attanasio (UCL, IFS), Cunha (Rice), Jervis (UCL, IFS)**

**Human Capital Formation in Childhood and Adolescence**

**July 9, 2018**
Infant stunting dropped from 29% to 14% in Peru between 2007 and 2014
Data

- 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.
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 $\mu_{R_{yyv}}$
Estimation Results: Fit of the Model III

Fit general Height Change pattern.
Model Fit: Protein Consumption

Estimation Results: Fit of the Model IV
Fit Protein Trend over time

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.
**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.
Group sessions of approximately 12-15 parents.

- Importance of language environment for language development.
- Lasts 13 weeks.
- Each week parent provides one 16-hour recording of the child’s language environment.
- Team analyzes data and provides feedback to parents.
LENA: How it works

Measuring Quality and Quantity of Time: LENA Pro

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

12/10/2017

Name: 1493 1493
ID: 1493
Age: 31 months as of 09/10/17

**LENA Online™**

**CONFIDENTIAL**

**Daily Adult Words**
Star goal was: 7103 words

**Hourly Adult Words on 09/10/17**

**Daily Conversational Turns**
Star goal was: 392 turns

**Hourly Conversational Turns on 09/10/17**

**Daily Child Vocalizations**

**Hourly Child Vocalizations on 09/10/17**

---

Flávio Cunha (Rice University)
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 adulthood.

Inequality in stocks of human capital is increasing in the last 20 years.

Inequality in investments in human capital is also increasing in the last 20 years.
At different stages of the lifecycle, investments produce different dimensions of human capital.

To estimate production functions of human capital:
- Address lack of cardinality of measures of human capital.
- Address measurement error in measures of human capital.
- Address endogeneity of investments.

Previous work shows that some of the inequality in investments is due to inequality in family resources: family income, parents’ education, etc.

We don’t know much about parental preferences, parental information sets, and other constraints that families face when choosing how much to invest in their children.

This lack of knowledge limits our ability to think of new public policies that can foster human capital formation.
Conclusion: Skill Formation

- Previous work shows that some of the inequality in investments is due to inequality in family resources: family income, parents’ education, etc.
- We don’t know much about the nature and importance of heterogeneity in parental preferences (that can be manipulated by policy); how this heterogeneity predicts investments; and whether public policy can affect these preferences; and if so, the quantitative importance of this mechanism.
- Same is true for parental beliefs.
- Even more problematic is that parents may face many other constraints that are so far unidentified by theoretical or empirical work.
- Lots of work for young, talented researchers with interest in theory, in empirical work, or in any convex combination of the two.
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.