Human Capital Formation in Childhood and Adolescence

Flávio Cunha

Rice University

July 20, 2016

Evolution of Inequality in USA

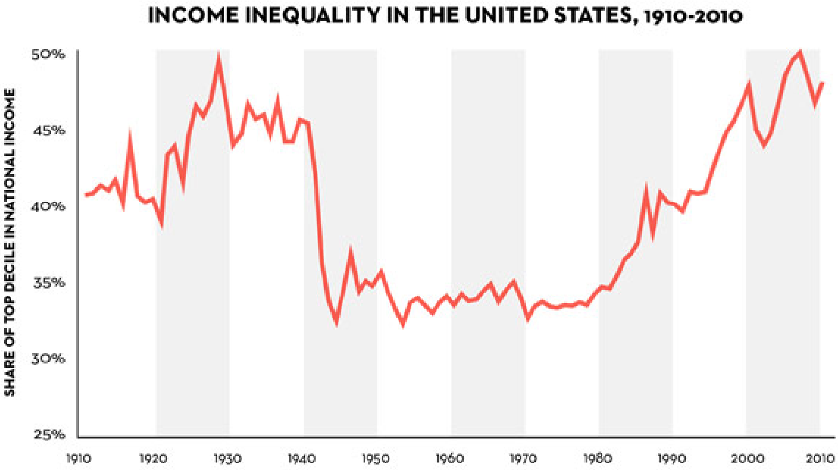


Figure
Relative Supply and Demand of Skilled Labor
Case 1: Supply and Demand Grow at Same Rate

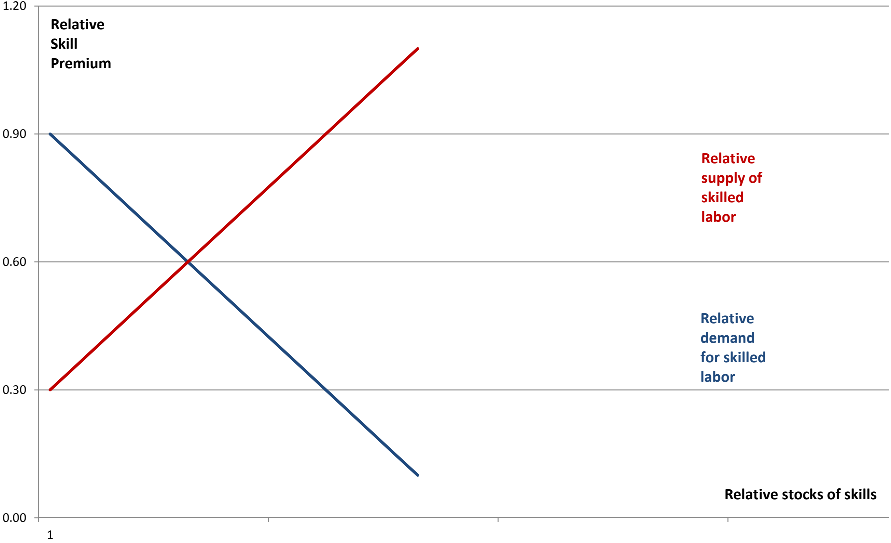


Figure
Relative Supply and Demand of Skilled Labor
Case 1: Supply and Demand Grow at Same Rate

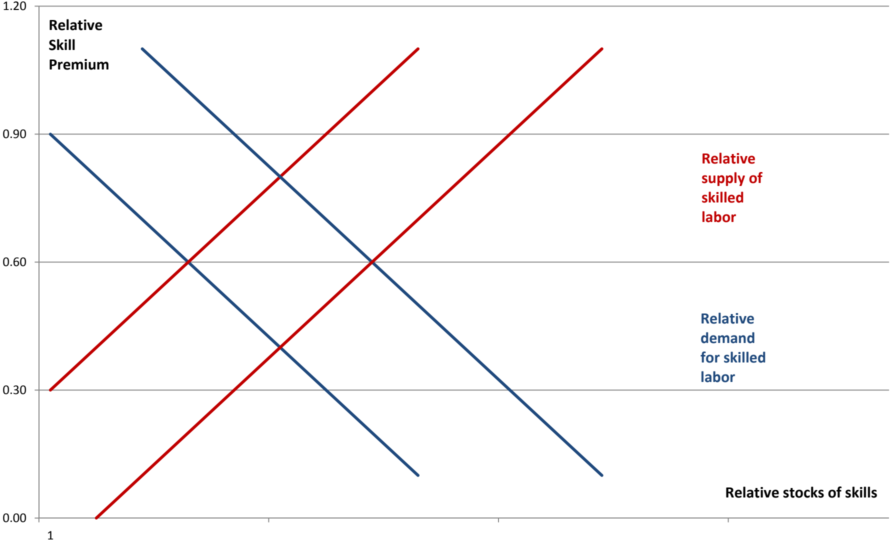


Figure
Relative Supply and Demand of Skilled Labor
Case 1: Supply and Demand Grow at Same Rate

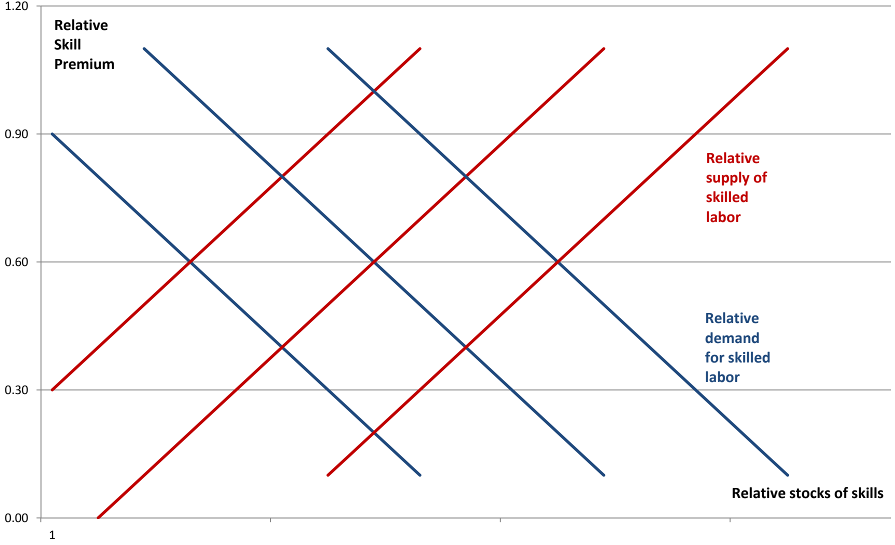


Figure
Relative Supply and Demand of Skilled Labor
Case 1: Supply and Demand Grow at Same Rate

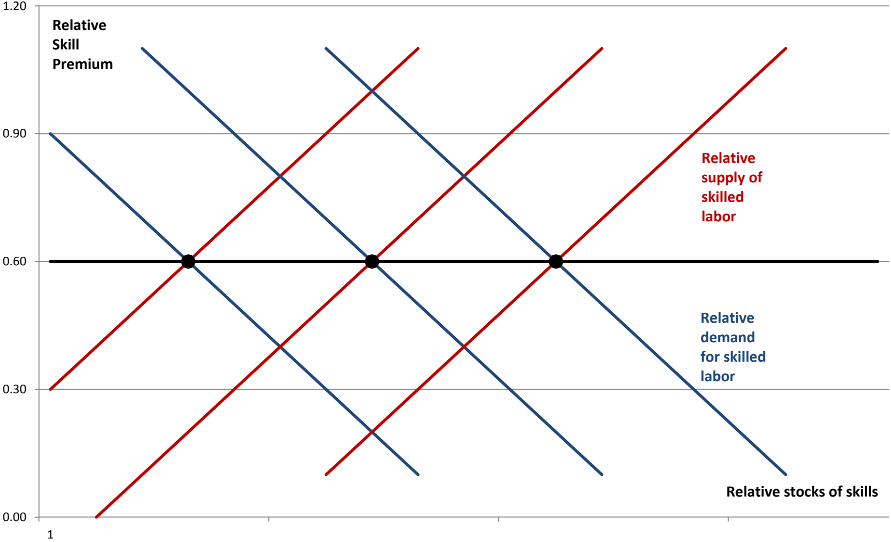


Figure
Relative Supply and Demand of Skilled Labor
Case 2: SBTC - Relative Demand Starts to Grow at a Faster Rate

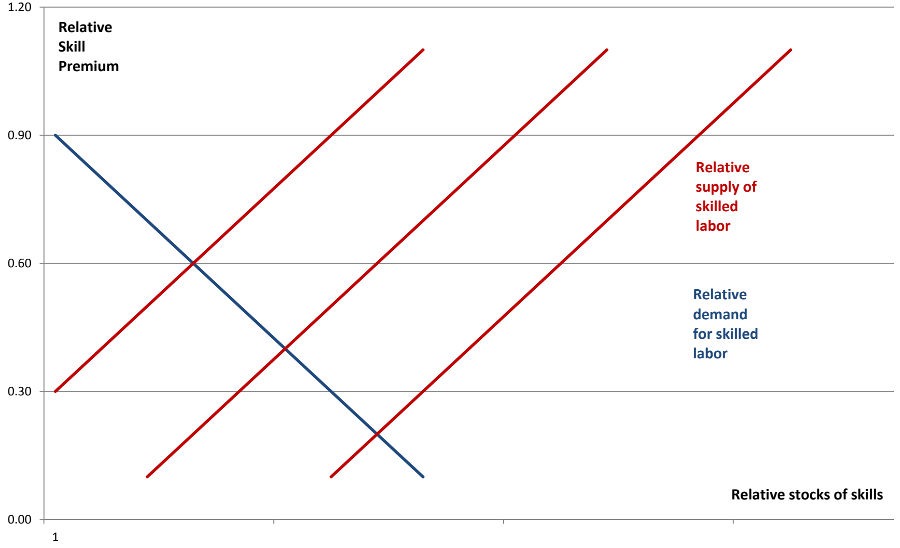


Figure
Relative Supply and Demand of Skilled Labor
Case 2: SBTC - Relative Demand Starts to Grow at a Faster Rate

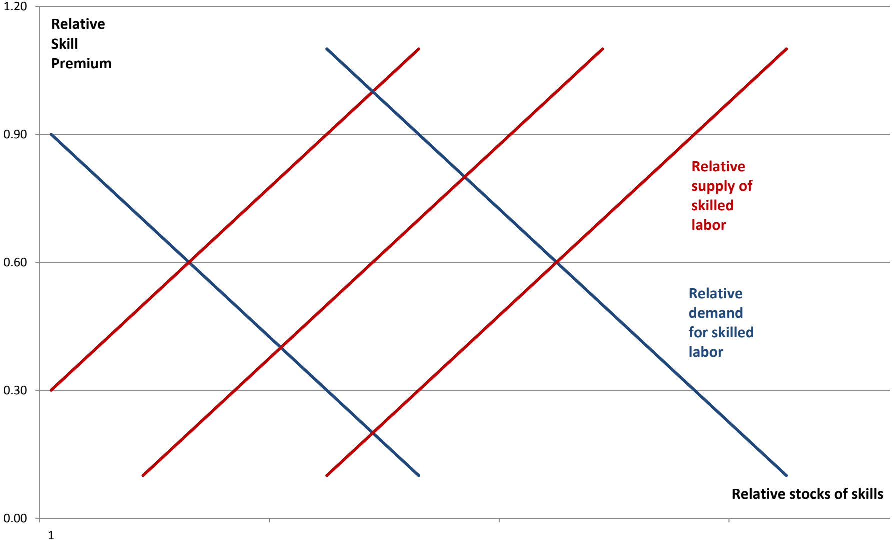
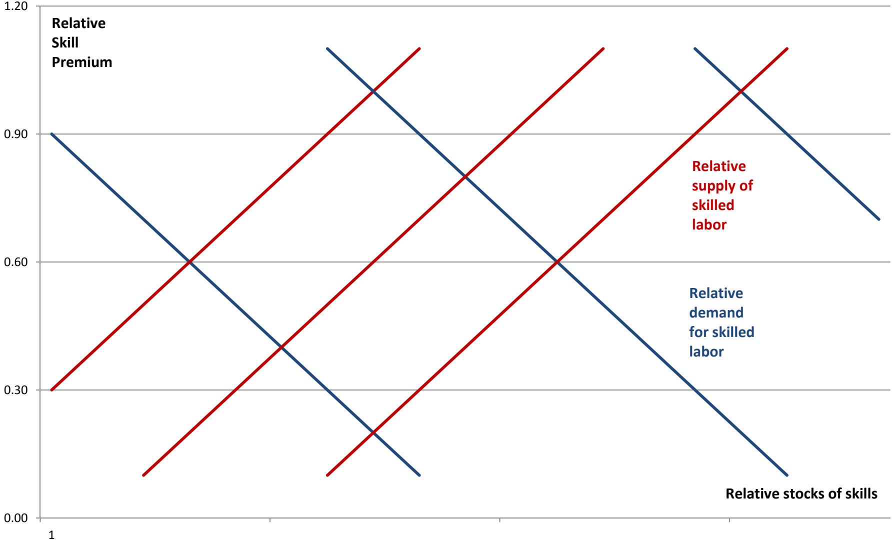


Figure
Relative Supply and Demand of Skilled Labor
Case 2: SBTC - Relative Demand Starts to Grow at a Faster Rate



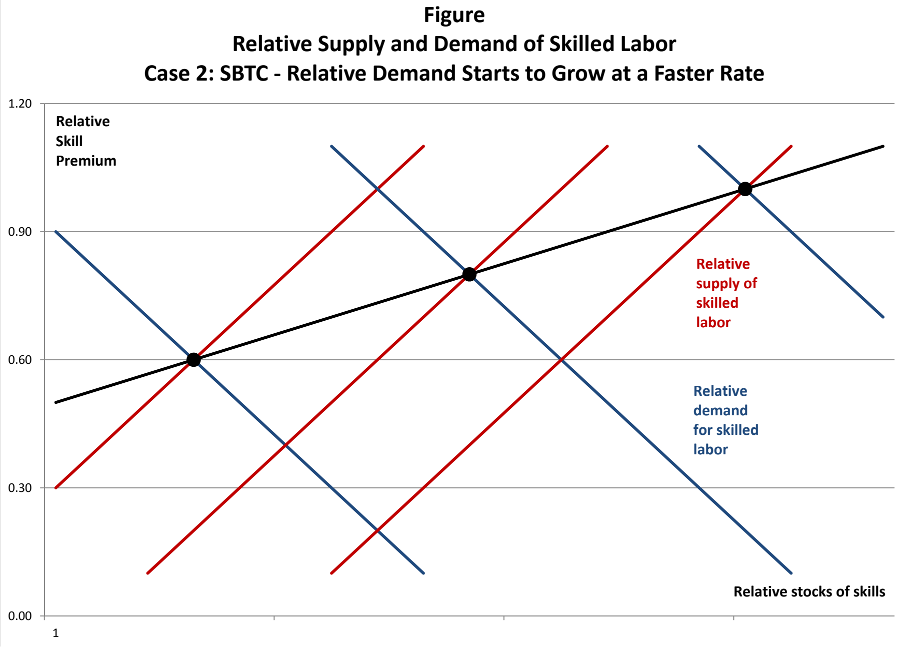


Figure
Relative Supply and Demand of Skilled Labor
Case 2: SBTC - Relative Demand Starts to Grow at a Faster Rate
1.20
0.90
0.60
0.30
0.00
1
Relative
Skill
Premium
Relative
supply of
skilled
labor
Relative
demand
for skilled
labor
Relative stocks of skills

Figure
Relative Demand and Supply of Skilled Labor
Case 3: Relative Supply Starts to Grow at Slower Rate

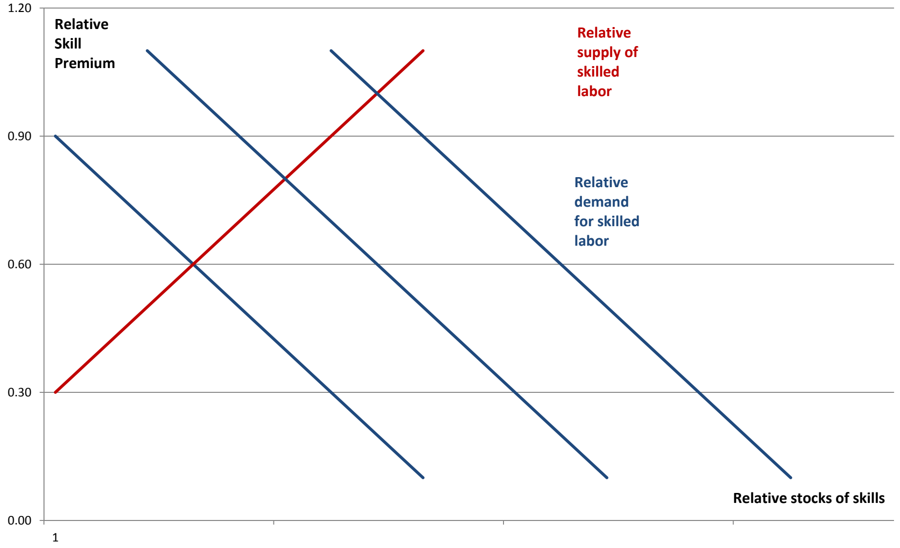


Figure
Relative Demand and Supply of Skilled Labor
Case 3: Relative Supply Starts to Grow at Slower Rate

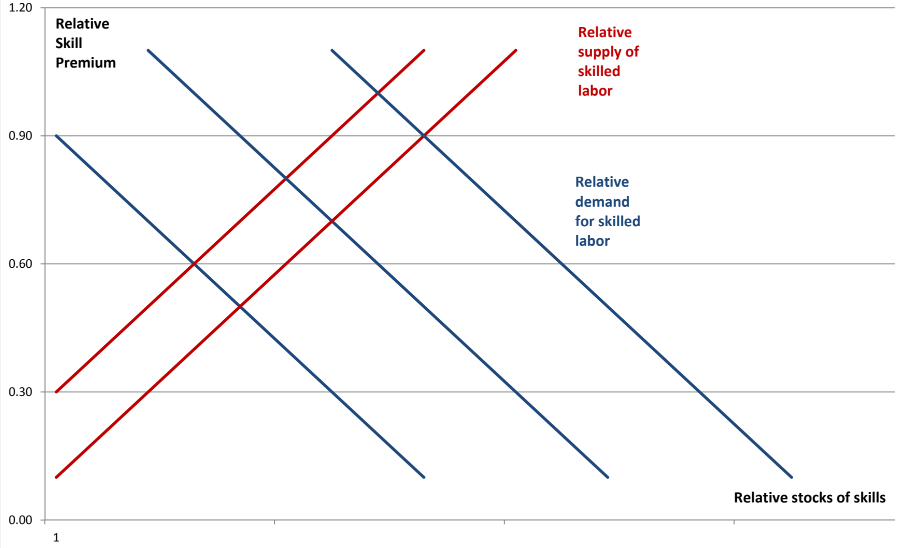


Figure
Relative Demand and Supply of Skilled Labor
Case 3: Relative Supply Starts to Grow at Slower Rate

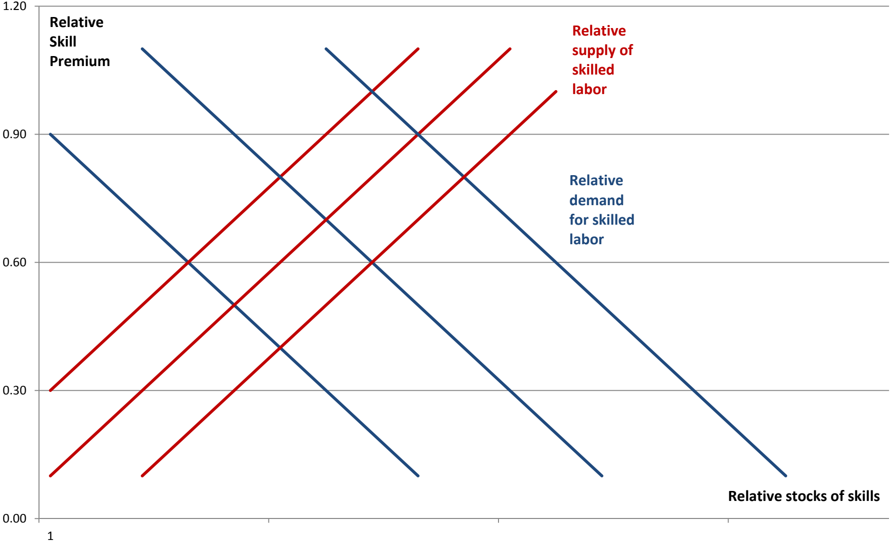
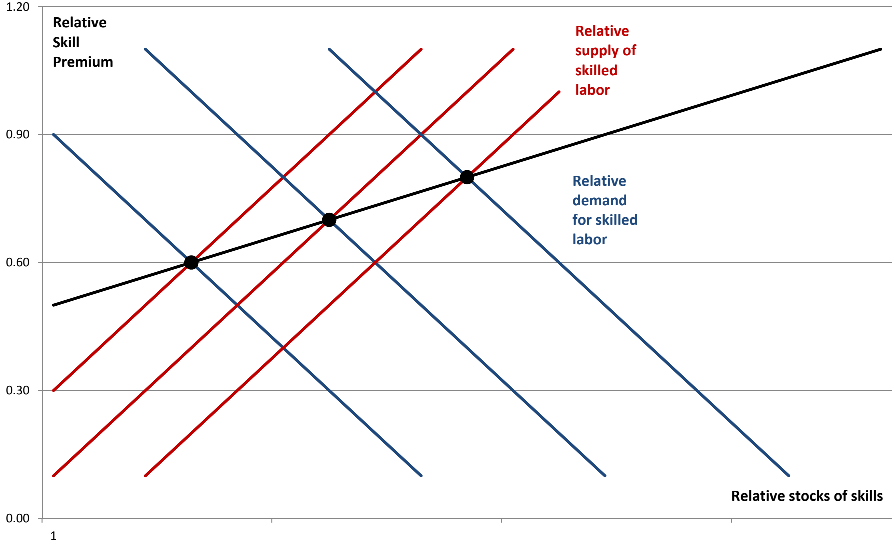


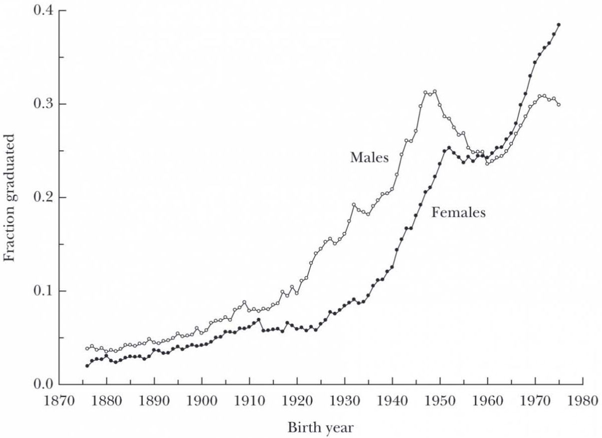
Figure
Relative Demand and Supply of Skilled Labor
Case 3: Relative Supply Starts to Grow at Slower Rate



Evolution of Inequality in USA

Figure 1

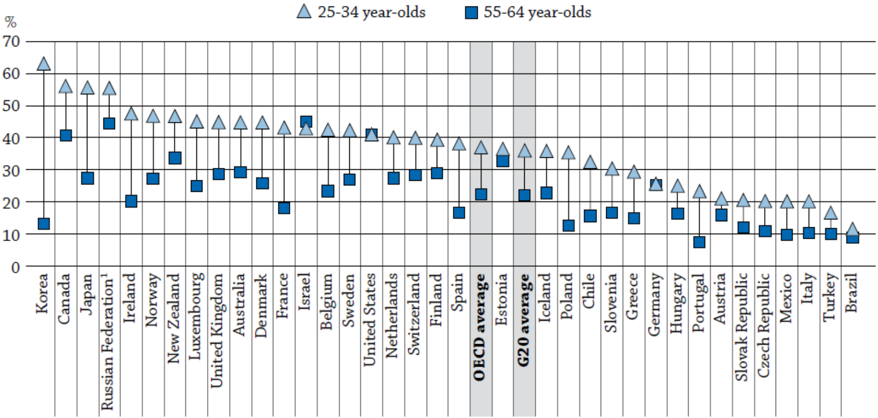
College Graduation Rates (by 35 years) for Men and Women: Cohorts Born from 1876 to 1975



Sources: 1940 to 2000 Census of Population Integrated Public Use Micro-data Samples (IPUMS).

Evolution of Inequality in USA

Chart A1.1. Percentage of population that has attained tertiary education, by age group (2009)



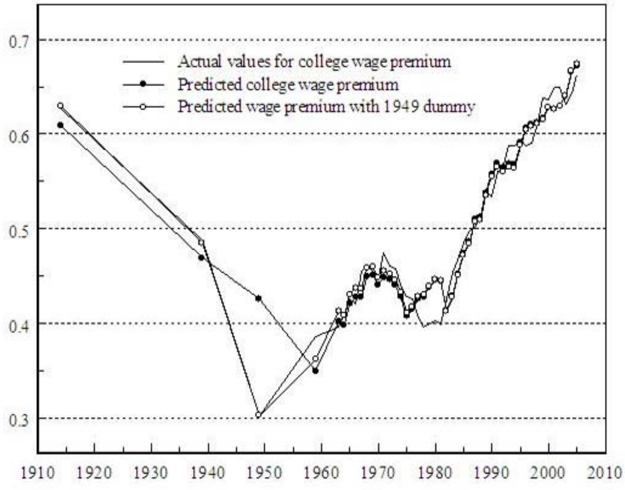
1. Year of reference 2002.

Countries are ranked in descending order of the percentage of 25-34 year-olds who have attained tertiary education.

Source: OECD. Table A1.3a. See Annex 3 for notes (*www.oecd.org/edu/eag2011*).

StatLink http://dx.doi.org/10.1787/888932459831

Evolution of Inequality in USA



Sources of Economic Inequality

248 SEPARATING UNCERTAINTY FROM HETEROGENEITY

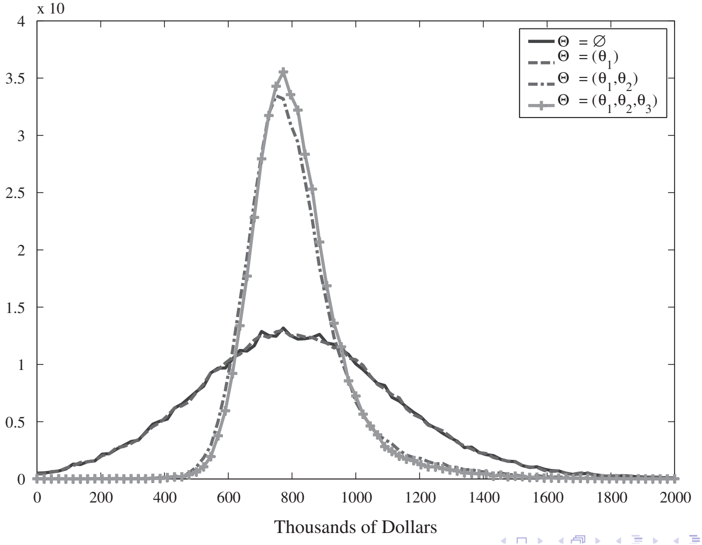


Figure 4: The Effect of Cognitive and Socio-emotional endowments, (log) Wages

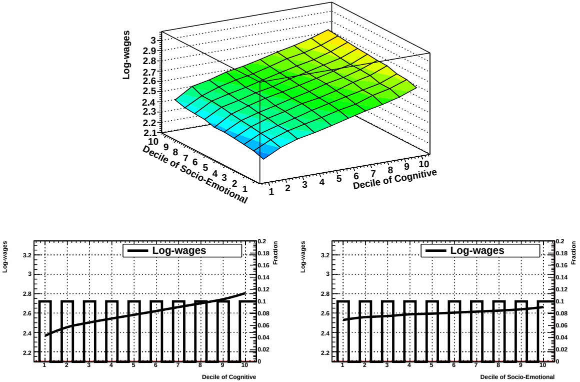


Figure 1: The Probability of Educational Decisions, by Endowment Levels, Dropping from Secondary School vs. Graduating

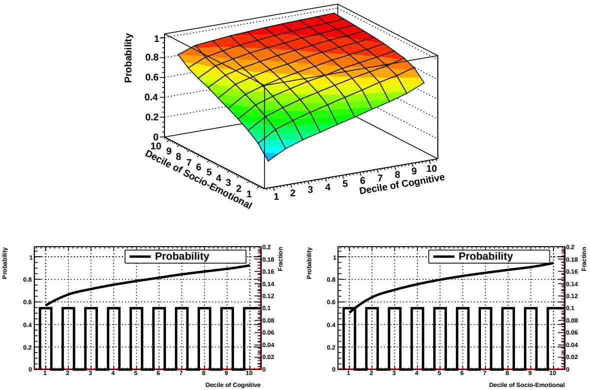


Figure 2: The Probability of Educational Decisions, by Endowment Levels, **HS Graduate** vs. College Enrollment

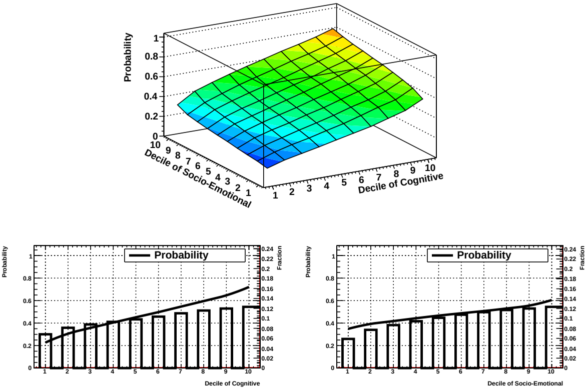


Figure 3: The Probability of Educational Decisions, by Endowment Levels, **Some College** vs. **4-year college degree**

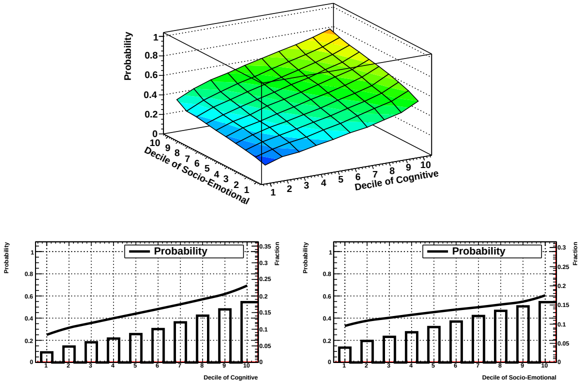


Figure 5: The Effect of Cognitive and Socio-emotional endowments, Daily Smoking

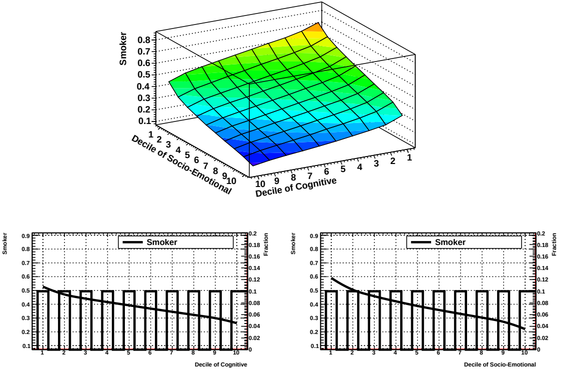
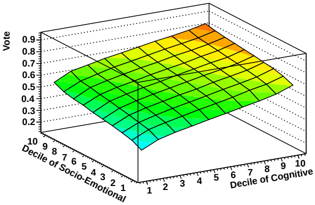
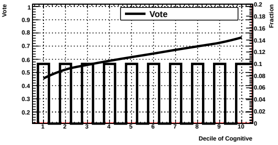


Figure 7: The Effect of Cognitive and Socio-emotional endowments, Participated in 2006 election





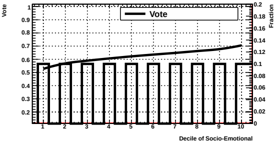
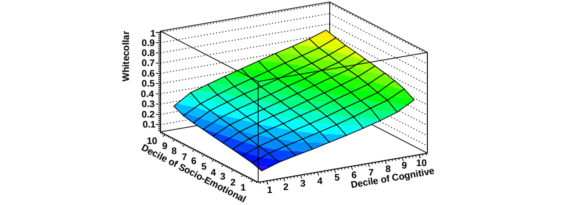
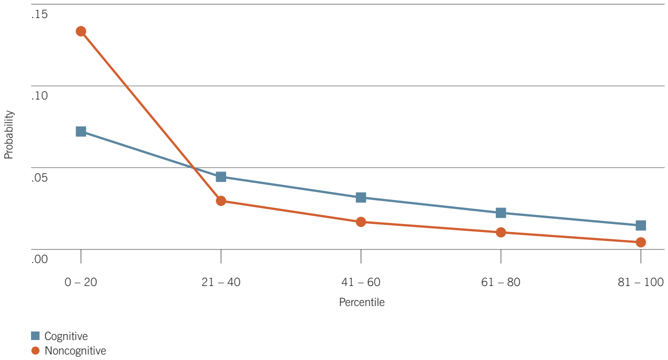


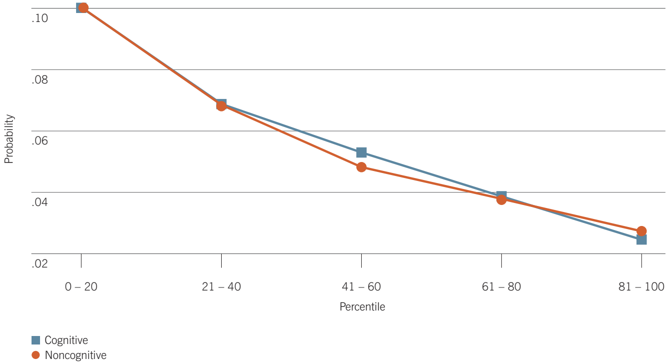
Figure 8: The Effect of Cognitive and Socio-emotional endowments on Probability of White-collar occupation (age 30)



Ever been in jail by age 30, by ability (males)

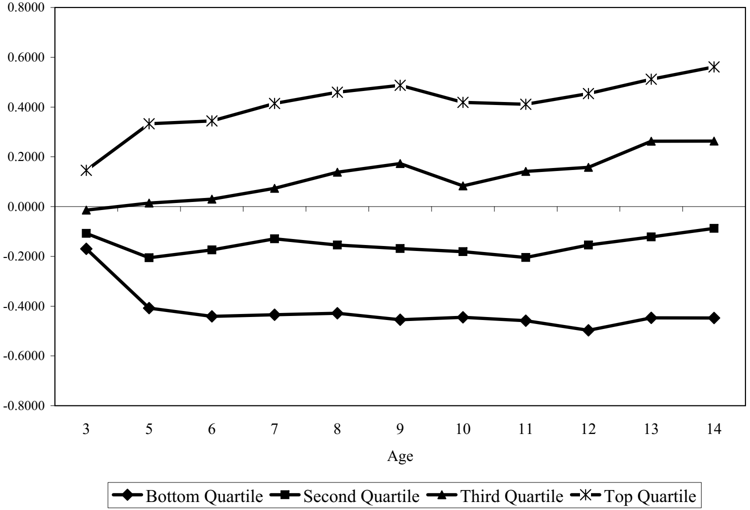


Probability of being teenage and single with children (females)



Gaps in Skills in Childhood and Adolescence

CNLSY/79 Data



Gaps in Skills in Early Childhood

Hart and Risley (1995)

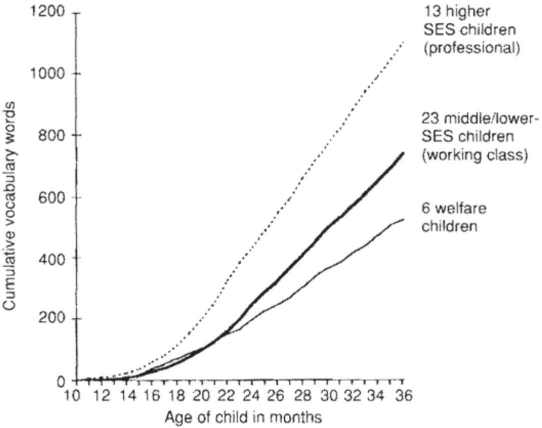
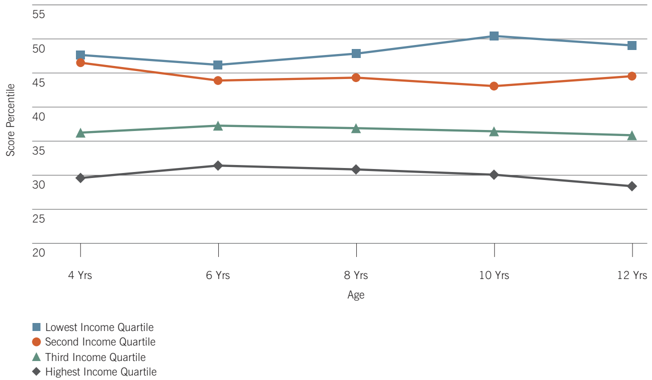


Figure 2. The widening gap we saw in the vocabulary growth of children from professional, working-class, and welfare families across their first 3 years of life. (See Appendix B for a detailed explanation of this figure.)

Gaps in Skills in Early Childhood

Carneiro and Heckman (2003)

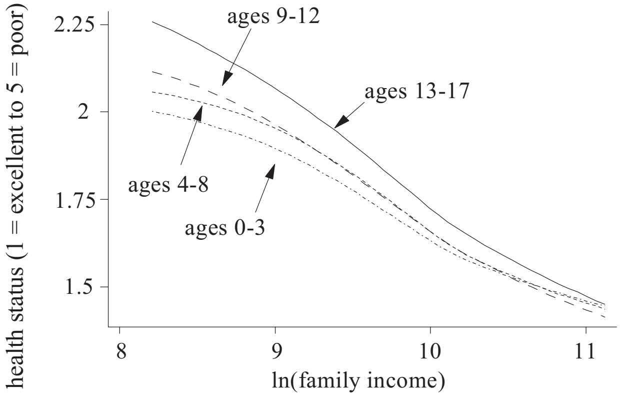
Average percentile rank on anti-social behavior score, by income quartile



Gaps in Skills in Early Childhood

Casey, Lubotsky, and Paxson (2002)

Health and income for children and adults, U.S. National Health Interview Survey 1986-1995.*

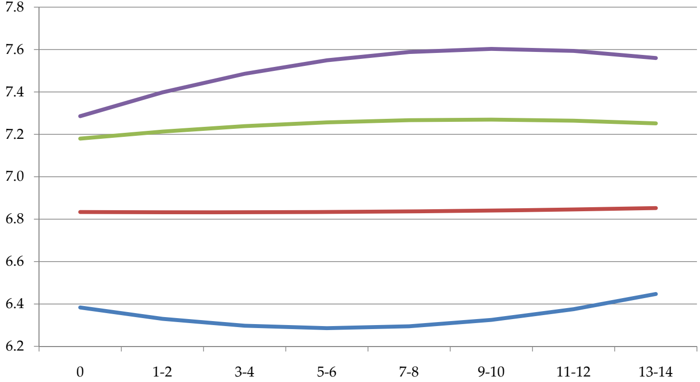


* From Case, A., Lubotsky, D. & Paxson, C. (2002), American Economic Review, Vol. 92, 1308-1334.

Gaps in Investments in Early Childhood

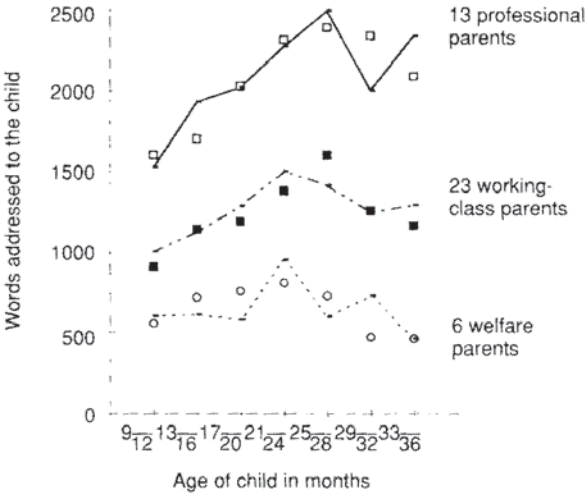
Carneiro and Heckman (2003)

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



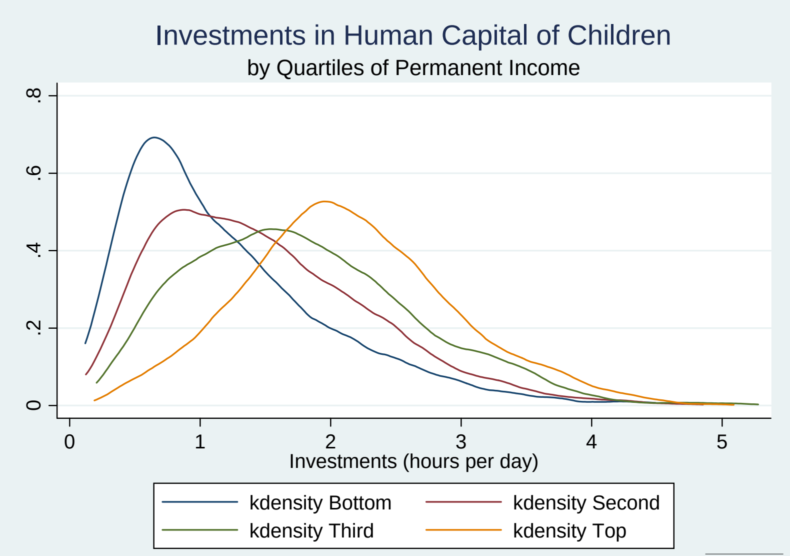
Gaps in Investments in Early Childhood

Hart and Risley (1995)



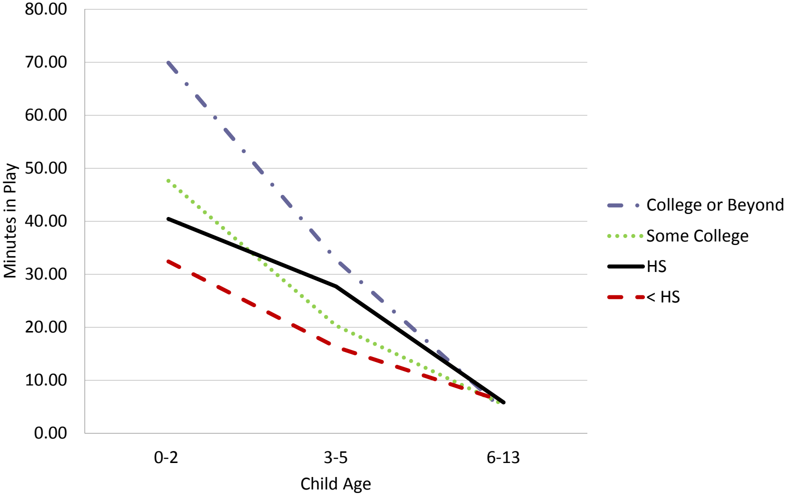
Gaps in Investments in Early Childhood

PSID, CDS



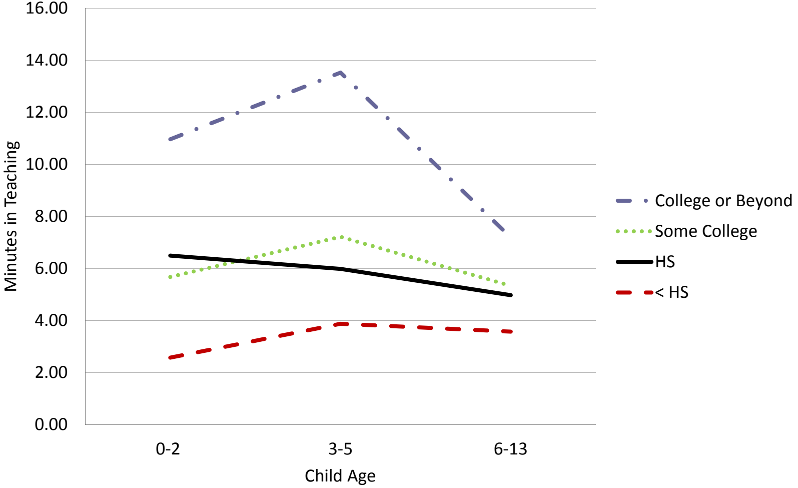
Gaps in Investments in Early Childhood

Kalil, Ryan, and Corey (2012)



Gaps in Investments in Early Childhood

Kalil, Ryan, and Corey (2012)



Gaps in Investments in Adolescence

Kalil, Ryan, and Corey (2012)

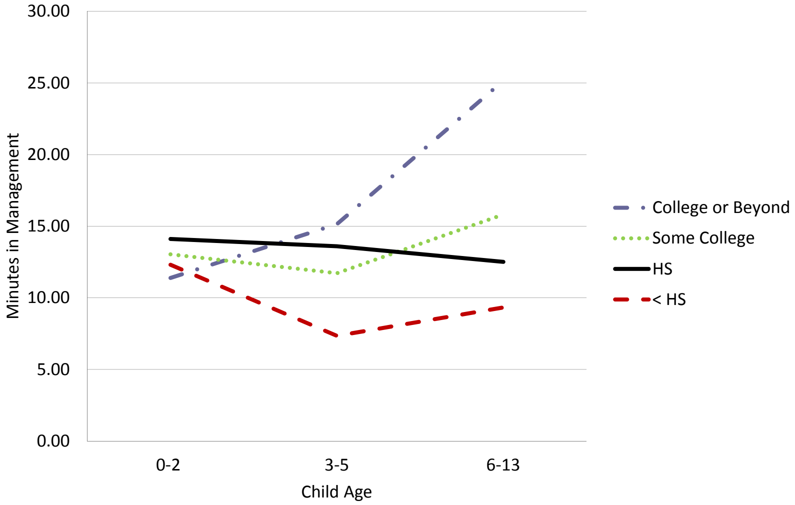
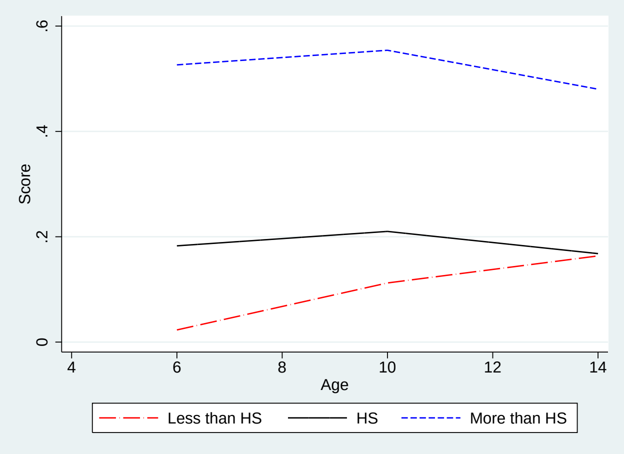


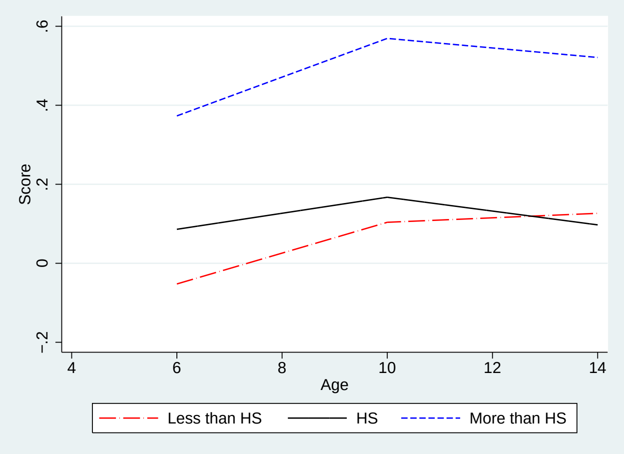
Figure 15: Parental Investment over Childhood among Whites by Mother's Education: Material Resources



Data: A balanced panel from Children of National Longitudinal Survey of Youth 1979.

Source: Moon (2012).

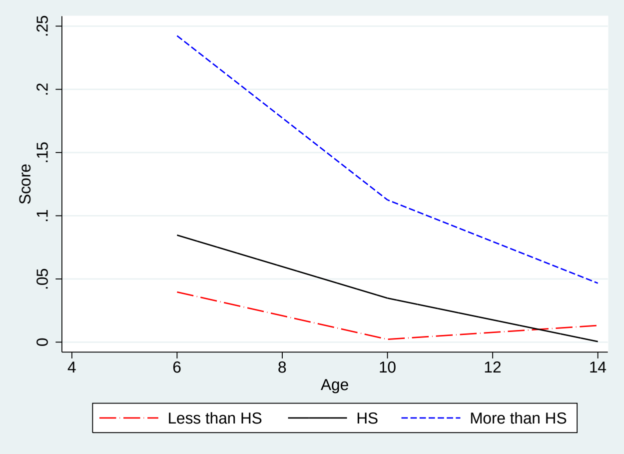
Figure 16: Parental Investment over Childhood among Whites by Mother's Education: Cognitive Stimulation



Data: A balanced panel from Children of National Longitudinal Survey of Youth 1979.

Source: Moon (2012).

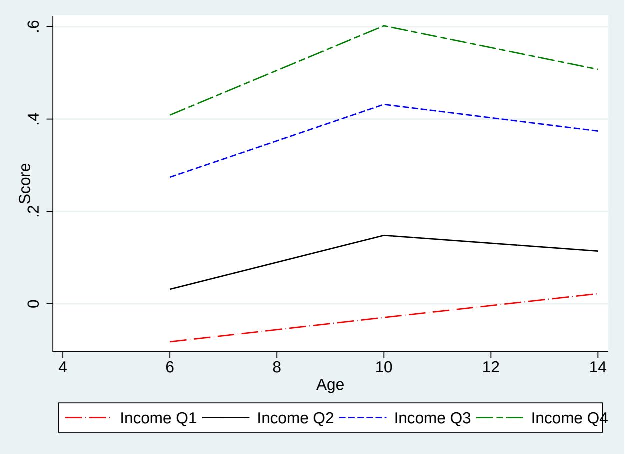
Figure 17: Parental Investment over Childhood among Whites by Mother's Education: Emotional Support



Data: A balanced panel from Children of National Longitudinal Survey of Youth 1979.

Source: Moon (2012).

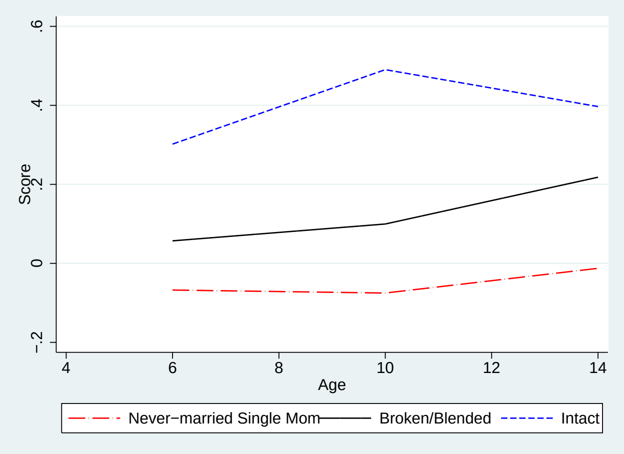
Figure 18: Parental Investment over Childhood among Whites by Family Income Quartile: Cognitive Stimulation



Data: A balanced panel from Children of National Longitudinal Survey of Youth 1979.

Source: Moon (2012).

Figure 19: Parental Investment over Childhood among Whites by Family Type: Cognitive Stimulation



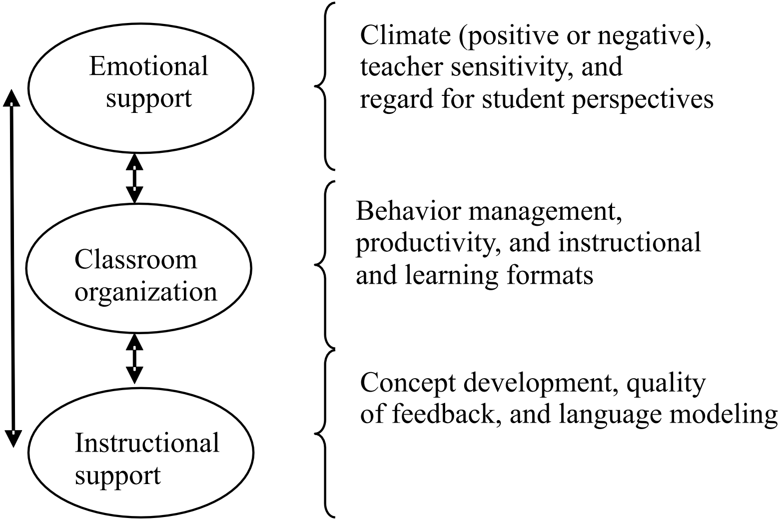
Data: A balanced panel from Children of National Longitudinal Survey of Youth 1979.

Source: Moon (2012).

The Role of Schools: Araujo et al (2015)

- How much, and in what ways, do kindergarten teachers maHer for learning outcomes?
- Two challenges:
    - Sorting of students to teachers.
        - Solution: Randomly match students to teachers.
    - Data on teachers are weakly correlated with student gain.
        - Improve the quality of data on teachers.

Classroom observation tool



Example: Teacher Behaviors and CLASS Scores for Behavior Management Dimension

Behavior Management

Encompasses the teacher's ability to provide clear behavioral expectations and use effective methods to prevent and redirect misbehavior.

	Low (1,2)	Mid (3,4,5)	High (6,7)
Clear Behavior Expectations ▪ Clear expectations ▪ Consistency ▪ Clarity of rules	Rules and expectations are absent, unclear, or inconsistently enforced.	Rules and expectations may be stated clearly, but are inconsistently enforced.	Rules and expectations for behavior are clear and are consistently enforced.
Proactive ▪ Anticipates problem behavior or escalation ▪ Rarely reactive ▪ Monitoring	Teacher is reactive and monitoring is absent or ineffective.	Teacher uses a mix of proactive and reactive responses; sometimes monitors but at other times misses early indicators of problems.	Teacher is consistently proactive and monitors effectively to prevent problems from developing.
Redirection of Misbehavior ▪ Effectively reduces misbehavior ▪ Attention to the positive ▪ Uses subtle cues to redirect ▪ Efficient	Attempts to redirect misbehavior are ineffective; teacher rarely focuses on positives or uses subtle cues. As a result, misbehavior continues/escalates and takes time away from learning.	Some attempts to redirect misbehavior are effective; teacher sometimes focuses on positives and uses subtle cues. As a result, there are few times when misbehavior continue/escalate or takes time away from learning.	Teacher effectively redirects misbehavior by focusing on positives and making use of subtle cues. Behavior management does not take time away from learning.
Student Behavior ▪ Frequent compliance ▪ Little aggression & defiance	There are frequent instances of misbehavior in the classroom.	There are periodic episodes of misbehavior in the classroom.	There are few, if any, instances of student misbehavior in the classroom.

Source: Pianta, La Paro & Hamre (2008)

The Role of Schools: Araujo et al (2015)

- Break analysis in two parts:
  - Estimate teacher effects: How much does it matter whether a child was assigned to teacher A or B in a school?
  - Estimate the associations between within-school differences in teacher characteristics or behaviors and child learning outcomes

The Role of Schools: Araujo et al (2015)

- One standard error in teacher quality leads to increases in child learning of
  - 11% of standard deviation in math.
  - 13% of standard deviation in language.
  - 7% of standard deviation in executive function.
- Same teachers have their students learn more math and more language year after year.
  - Cross-year correlation of teacher effects in math is 0.32
  - Cross-year correlation of teacher effects in language is 0.42.

The Role of Schools: Araujo et al (2015)

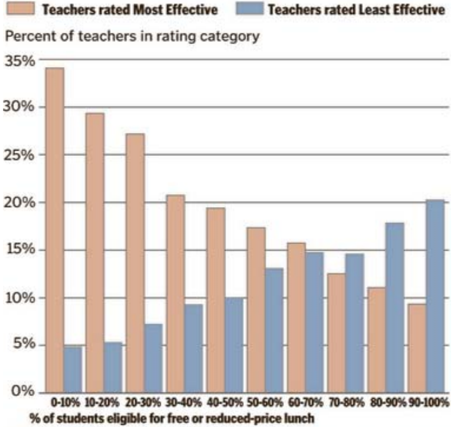
- What explains differences in teacher effectiveness?
  - One standard deviation in teacher IQ increases child's performance by 4% of a standard deviation.
  - Students randomly assigned to "rookie" teachers learn 16% of standard deviation less.
  - No correlation between teacher personality scores (Big Five) and student learning.
  - One standard deviation in CLASS explains 59% of a standard deviation in student learning.
  - Teachers with better CLASS scores get all their students to learn more: Effects are not concentrated on girls or boys, on children with high or low levels of development when they enter school, or on children of high or low socioeconomic status

The Role of Schools: Araujo et al (2015)

- Interestingly, parental reports of teacher quality correlate (very imperfectly) with teacher effectiveness:
  - Teachers who produce one standard deviation more learning are given a 0.44 higher score (on a scale from 1 to 5).
  - Rookie teachers are given 0.33 lower score by parents.
  - Teachers with higher CLASS scores also get higher scores reported by parents.
- However, parents do not adjust behaviors in response to differences in teacher quality.
  - There is no effect on the quality or quantity of parent-child interaction at home.
  - There is no effect on the child's dropping out or absenteeism.

How teacher ratings relate to a school's poverty level

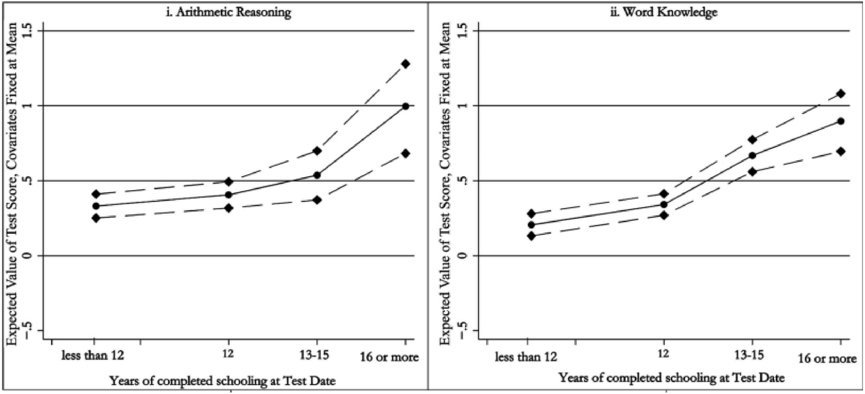
Teachers who receive the state's top value-added rating — "Most Effective" — are likely to be in schools with fewer poor students, based on value-added ratings for teachers at 1,720 public schools. Of 1,035 teachers at the wealthiest schools, 34 percent got the top rating. In contrast, of 2,411 teachers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



SOURCE: Ohio Department of Education

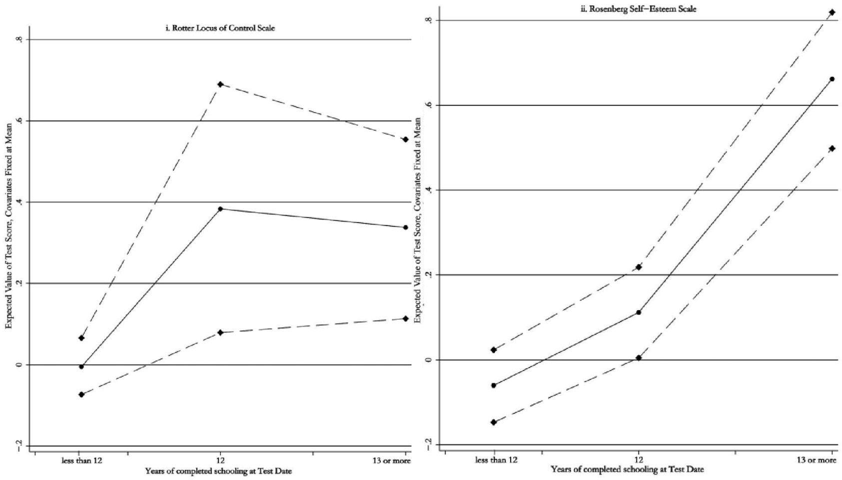
RICH EXNER, JAMES OWENS | THE PLAIN DEALER

Figure 22: Causal Effect of Schooling on ASVAB Measures of Cognition



Notes: Effect of schooling on components of the ASVAB. The first four components are averaged to create male's with average ability. We standardize the test scores to have within-sample mean zero, variance one. The model is estimated using the NLSY79 sample. Solid lines depict average test scores, and dashed lines, confidence intervals.
Source: Heckman, Stixrud and Urzua [2006, Figure 4].

Figure 23: Causal Effect of Schooling on Two Measures of Personality

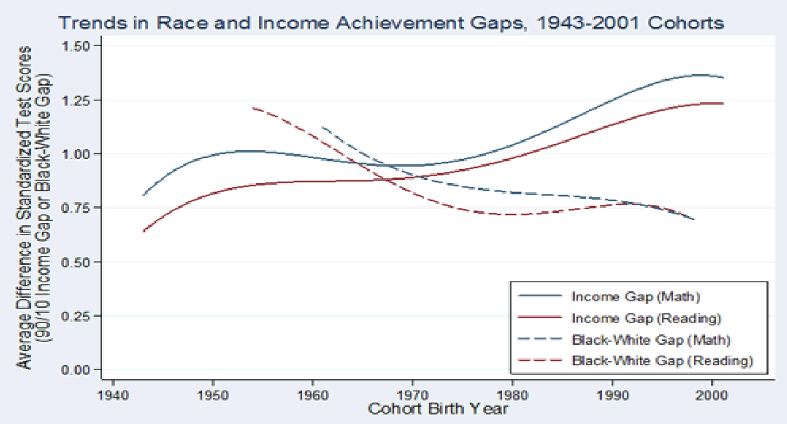


Evidence is Reinforced from Evidence from RCT

- Heckman has shown that many early childhood programs have positive impacts on long-term outcomes related to inequality.
- This is in spite of the fact that gains in cognitive skills do not tend to persist over time (more about this on "dynamic complementarity").
- Dobbie and Fryer (2009) show that Harlem Children's Zone is effective at increasing the achievement of the poorest minority children: It closes the black-white achievement gap in mathematics and reduces it by nearly half in English Language Arts.
- Heller et al (2015) show that Cognitive Behavior Therapy target to disadvantaged adolescents can reduce propensity to participate in crime and even improve educational outcomes.

Increasing Inequality in Skills

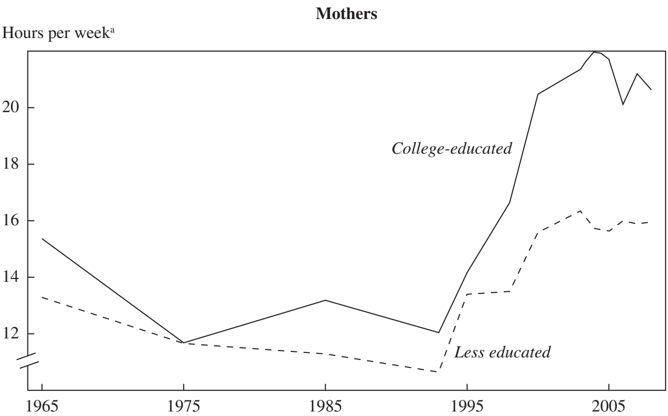
Reardon (2013)



Increasing Inequality in Investments

Ramey and Ramey (2013)

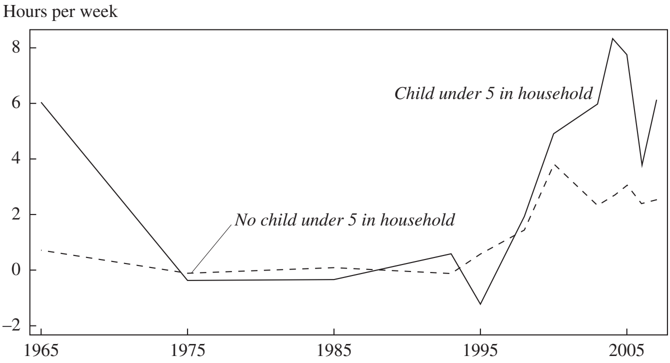
Figure 1. Time Spent on Childcare by Parents, by Educational Attainment, 1965–2008



Increasing Inequality in Investments

Saks and Stevenson (2013)

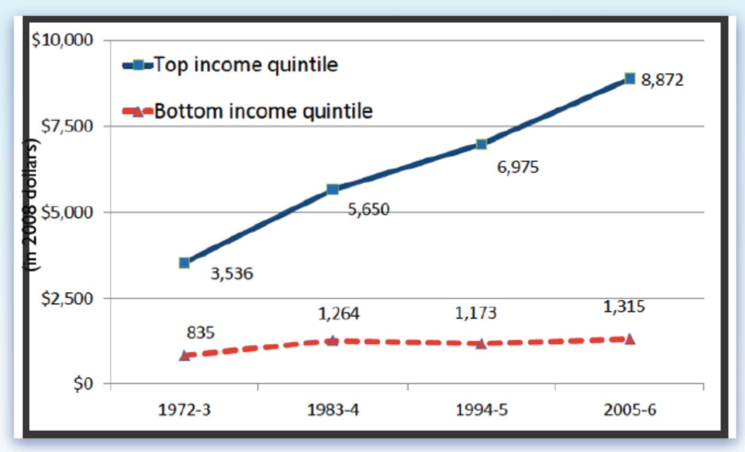
Figure 2. Difference in Childcare Time between Well-Educated and Less Educated Parents, by Presence of Young Children, 1965–2008[a]



Source: Authors' calculations.

a. Figure plots the coefficients on the interaction between the year dummy and a dummy for whether the parent completed college, using the descriptive model and data in Ramey and Ramey (this volume) and including or excluding travel time in total childcare time.

Increasing Inequality in Investments
Kornrich and Furstenberg (2011)



Theory

- Next, I will try to make sense of this data by proposing a very simple model of human capital formation.
- At the core of this model, there will be two important parameters:
  - Self-productivity of skills: I learn how to read, then I use reading to learn other skills.
  - Dynamic complementarity: The returns to the development of advanced skills are higher for the individuals who learned basic skills.

Optimal Early and Late Investments in Children

- Consider the following cost minimization problem:

$$\min x_E + \frac{1}{1+r} x_L$$

subject to the technology of skill formation:

$$h = \left[\gamma x_E^{\phi} + (1-\gamma) x_L^{\phi} \right]^{\frac{1}{\phi}}$$

where $\gamma \in [0,1]$ and $\phi \leq 1$.

- Note that:
  - The parameter γ captures self-productivity.
  - The parameter ϕ captures dynamic complementarity.

Boundary Solution when $\phi = 1$

- In this case, $h = \gamma x_E + (1-\gamma) x_L$.
- Two investment strategies: Invest early and produce γ units of human capital per unit of investment.
- Save in physical assets early and invest $1+r$ late and produce $(1+r)(1-\gamma)$ units of human capital.
- Should invest all early if, and only if:

$$\gamma > \frac{1+r}{2+r}$$

Boundary Solution when $\phi \to -\infty$

- In this case, $h = \min\{x_E, x_L\}$
- The solution to this problem is $x_E = x_L$ for whatever values of r.

Interior Solution when $-\infty < \phi < 1$

- The solution to this problem is:

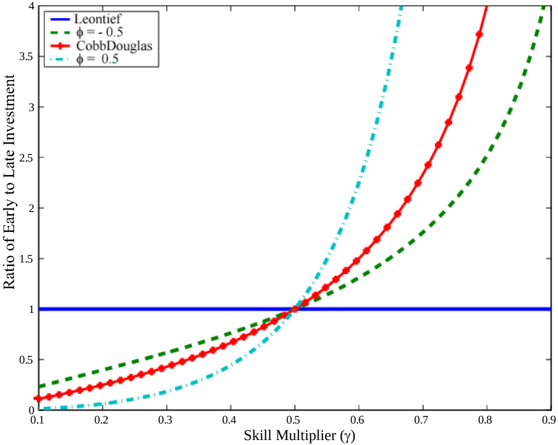
$$x_E = \frac{\gamma^{\frac{1}{1-\phi}}}{\left[\gamma^{\frac{1}{1-\phi}} + (1-\gamma)^{\frac{1}{1-\phi}} (1+r)^{\frac{\phi}{1-\phi}}\right]^{\frac{1}{\phi}}} h$$

$$x_L = \frac{(1-\gamma)^{\frac{1}{1-\phi}} (1+r)^{\frac{1}{1-\phi}}}{\left[\gamma^{\frac{1}{1-\phi}} + (1-\gamma)^{\frac{1}{1-\phi}} (1+r)^{\frac{\phi}{1-\phi}}\right]^{\frac{1}{\phi}}} h$$

- Note that we have the following ratio:

$$\ln \frac{x_E}{x_L} = \frac{1}{1-\phi} \ln \left(\frac{\gamma}{1-\gamma} \right) + \frac{1}{1-\phi} \ln \left(\frac{1}{1+r} \right)$$

Textbook Model of Investments in Children



(Assumes $r = 0$)

Dual Side of Dynamic Complementarity

- 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 Parameters of the Technology of Skill Formation: Parameterization

- $\theta_{c,t}$ denotes cognitive skills of the child at age t.
- $\theta_{n,t}$ denotes non-cognitive skills of the child at age t.
- $x_{k,t}$ is parental investment in skill k when child is t years old.
- $\theta_{c,p}$ represents parental cognitive skills.
- $\theta_{n,p}$ represents parental noncognitive skills.
- $\eta_{k,t}$ are shocks and/or unmeasurable inputs. .

Estimating Parameters of the Technology of Skill Formation: Parameterization

- There are S different developmental stages: $s = 1, ..., S$. The technology for skill k, at period t and stage s is:

$$\theta_{k,t+1} = e^{\eta_{c,t+1}} \times f_{s,k}$$

where

$$f_{s,k} = \left[\gamma_{s,k,1}\theta_{c,t}^{\phi_{s,c}} + \gamma_{s,k,2}\theta_{n,t}^{\phi_{s,c}} + \gamma_{s,k,3}x_{k,t}^{\phi_{s,c}} + \gamma_{s,k,4}\theta_{c,p}^{\phi_{s,c}} + \gamma_{s,k,5}\theta_{n,p}^{\phi_{s,c}}\right]^{\frac{1}{\phi_{s,c}}}$$

Estimating Parameters of the Technology of Skill Formation: Data

- 2207 firstborn white children (birth to age 14) from CNLSY/79.
- Large number of observations: Almost 400,000.
- Large number of parameters: Almost 1250.
- Extensive data collection on parental characteristics and child cognitive and noncognitive development.
- The data collection is every two years.
- There are eight periods: (birth, ages 1-2, ages 3-4,...,ages 13-14).

Estimating Parameters of the Technology of Skill Formation: Measures

- Child's Cognitive Skills:
    - MSD
    - Parts of Body
    - Memory for Locations
    - PPVT
    - PIAT
- Child's Noncognitive Skills:
    - Temperament and Behavior Problem Index.
- Parental Investments
    - Components of the Home Score
- Parental Cognitive Skills
    - Components of the ASVAB Tests.
- Parental Noncognitive Skills
    - Mother's Self-Esteem and Mom's Locus of Control

Estimating Parameters of the Technology of Skill Formation: Identification

- We need to deal with three problems:
  - We don't observe $\theta_{c,t}$, $\theta_{n,t}$ and $x_{k,t}$ directly.
  - We don't know which scale to use to measure $\theta_{c,t}$ and $\theta_{n,t}$.
  - Investments $x_{k,t}$ are chosen by parents based on information from $\eta_{k,t}$ that is unobserved by the econometrician.

Addressing Measurement Error

- To focus on the important ideas, suppose that we want to estimate a linear production function:

$$\theta_{k,t+1} = \beta x_{k,t} + \eta_{k,t}$$

- If we know the scale of $\theta_{k,t+1}$ and investments are exogenous, $E(\eta_{k,t} | x_{k,t}) = 0$, then, a consistent estimator for β is:

$$\beta = \frac{Cov(\theta_{k,t+1}, x_{k,t})}{Var(x_{k,t})}$$

- If we observe $\theta_{k,t+1}$ and $x_{k,t}$, that is it.

Addressing Measurement Error

- When we don't observe $\theta_{k,t+1}$ and $x_{k,t}$, we can use observed proxy variables:

$$\begin{aligned} y_{1,k,t+1} &= \theta_{k,t+1} + \varepsilon_{1,k,t+1} \\ y_{2,k,t+1} &= \theta_{k,t+1} + \varepsilon_{2,k,t+1} \\ y_{3,k,t} &= x_{k,t} + \varepsilon_{3,k,t} \\ y_{4,k,t} &= x_{k,t} + \varepsilon_{4,k,t} \end{aligned}$$

- The vector $\varepsilon_{j,k,t+1}$ captures measurement error. Assume they are uncorrelated. Then, we can estimate β by:

$$\beta = \frac{Cov(\theta_{k,t+1}, x_{k,t})}{Var(x_{k,t})} = \frac{Cov(y_{1,k,t+1}, y_{3,k,t})}{Cov(y_{3,k,t}, y_{4,k,t})}$$

- . Key to identification: covariance restrictions in measurement error (some are testable).

Addressing Measurement Error

- It is possible to relax the model so we can give each measure its own "information parameter" (also called factor loading) $\alpha_{l,k,t}$for $l = 2, ..., L$:

$$y_{1,k,t} = \theta_{k,t} + \varepsilon_{1,k,t} y_{l,k,t} = \alpha_{l,k,t}\theta_{k,t} + \varepsilon_{2,k,t}$$

- If we have at least three measurements per period, we can allow some of the measurement error components to be correlated (see Cunha and Heckman, 2008):

$$Cov(\varepsilon_{l,k,t}, \varepsilon_{l',k',t'}) \neq 0 \text{ for } l, l' = 2, ..., L \text{ and } \forall k, k', t, t'$$

- We can also identify nonlinear measurement equations:

$$\begin{aligned} y_{1,k,t} &= \theta_{k,t+1} + \varepsilon_{1,k,t} \\ y_{l,k,t} &= h_{l,k,t}(\theta_{k,t}, \varepsilon_{l,k,t}) \end{aligned}$$

Estimation algorithm

- When the transition and measurement equations are linear and the factors and measurement errors are normally distributed, we can use the Kalman Filter to compute the likelihood
- Kalman filter breaks down because of the nonlinearity.
- Two common options are:
  - Extended Kalman Filter (EKF): Linearize $f(\theta)$ around $f(E\theta)$
  - Particle Filter: Sequential Monte Carlo Method.
- We use the Mixture of Normals Unscented Kalman Filter

Illustration of Estimation Algorithm

- We have the following measurement equations:

$$\begin{aligned} y_1 &= \mu_1 + x + \varepsilon_1 \\ y_2 &= \mu_2 + x + \varepsilon_2 \\ y_3 &= \mu_3 + \theta + \varepsilon_3 \\ y_4 &= \mu_4 + \theta + \varepsilon_4 \end{aligned}$$

- As well as the following technology of skill formation (nonlinear):

$$\theta = \alpha + \beta x^2 + \eta$$

- The goal is to construct the contribution to the likelihood for each individual: $L_i = L\left(y_1, y_2, y_3, y_4\right)$

Illustration of Estimation Algorithm

- The filtering algorithm provides a set of steps that allows us to compute:$L\left(y_1, y_2, y_3, y_4\right) =$ $L\left(\left. y_4 \right| y_1, y_2, y_3\right) \times L\left(\left. y_3 \right| y_2, y_1\right) \times L\left(\left. y_2 \right| y_1\right) \times L\left(y_1\right)$
- While computation of $L\left(y_1, y_2, y_3, y_4\right)$ is time-consuming, the computation of each one of the terms $L\left(\left. y_4 \right| y_3, y_2, y_1\right)$, $L\left(\left. y_3 \right| y_2, y_1\right)$, $L\left(\left. y_2 \right| y_1\right)$, and $L\left(y_1\right)$ is fast.

Illustration of Estimation Algorithm

- Let's consider the situation in which both x and ε_l, for $l = 1, 2$, are normally distributed. This means that y_1 and y_2 are jointly normally distributed.
- More concretely, let $x \sim N\left(\mu_x, \sigma_x^2\right)$ and $\varepsilon_l \sim N\left(0, \sigma_l^2\right)$ for $l = 1, 2$.
- Note that: $y_1 \sim N\left(\mu_1 + \mu_x, \sigma_x^2 + \sigma_1^2\right)$.
- Therefore, the first term in the contribution to the likelihood is:

$$L\left(y_1\right) = \sqrt{\frac{1}{2\pi\left(\sigma_x^2 + \sigma_1^2\right)}} \exp\left\{-\frac{1}{2}\frac{\left(y_1 - \mu_1 - \mu_x\right)^2}{\left(\sigma_x^2 + \sigma_1^2\right)}\right\}$$

Illustration of Estimation Algorithm

- To derive the second term in the contribution to the likelihood, it will be very helpful to compute $E\left(x \mid y_1\right)$ and $Var\left(x \mid y_1\right)$.
- Because of the normality assumption, it follows that:

$$E\left(x \mid y_1\right) = E\left(x\right) + \frac{Cov\left(x, y_1\right)}{Var\left(y_1\right)}\left[y_1 - E\left(y_1\right)\right]$$

$$E\left(x \mid y_1\right) = \mu_x + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_1^2}\left[y_1 - \mu_1 - \mu_x\right]$$

- And:

$$Var\left(x \mid y_1\right) = Var\left(x\right) - \frac{Cov\left(x, y_1\right)^2}{Var\left(y_1\right)}$$

$$Var\left(x \mid y_1\right) = \sigma_x^2 - \frac{\sigma_x^4}{\sigma_x^2 + \sigma_1^2}$$

Illustration of Estimation Algorithm

- Note that
$$E\left(y_2 \mid y_1\right) = \mu_2 + E\left(x \mid y_1\right) + E\left(\varepsilon_2 \mid y_1\right)$$
$$E\left(y_2 \mid y_1\right) = \mu_2 + \mu_x + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_1^2}\left(y_1 - \mu_1 - \mu_x\right)$$
- And:
$$Var\left(y_2 \mid y_1\right) = Var\left(x \mid y_1\right) + Var\left(\varepsilon_2 \mid y_1\right)$$
$$Var\left(y_2 \mid y_1\right) = \sigma_2^2 + \sigma_x^2 - \frac{\sigma_x^4}{\sigma_x^2 + \sigma_1^2}$$
- Therefore:
$$L\left(y_2 \mid y_1\right) = \sqrt{\frac{1}{2\pi Var\left(y_2 \mid y_1\right)}} \exp\left\{-\frac{1}{2}\frac{\left[y_2 - E\left(y_2 \mid y_1\right)\right]^2}{Var\left(y_2 \mid y_1\right)}\right\}$$

Illustration of Estimation Algorithm

- Now, we need to compute the moments $E\left(x \mid y_1, y_2\right)$ and $Var\left(x \mid y_1, y_2\right)$

$$E\left(x \mid y_1, y_2\right) = E\left(x \mid y_1\right) + \frac{Cov\left(x, y_2 \mid y_1\right)}{Var\left(y_2 \mid y_1\right)}\left[y_2 - E\left(y_2 \mid y_1\right)\right]$$

- And:

$$Var\left(x \mid y_1, y_2\right) = Var\left(x \mid y_1\right) - \frac{Cov\left(x, y_2 \mid y_1\right)}{Var\left(y_2 \mid y_1\right)}$$

- Define $\mu_{x \mid y_1, y_2} = E\left(x \mid y_1, y_2\right)$ and $\sigma^2_{x \mid y_1, y_2} = Var\left(x \mid y_1, y_2\right)$.
- Great exercise to make sure you understand steps is to derive the expressions for

Illustration of Estimation Algorithm

- Note that y_3 is a function of θ, not x.
- For this reason, we need to compute $E\left(\theta \mid y_1, y_2\right)$ and $Var\left(\theta \mid y_1, y_2\right)$.
- One way to compute these moments is by exploring the information from the technology of skill formation which defines the relationship between θ and x.

Illustration of Estimation Algorithm

- If I use the technology of skill formation, note that:

$$E\left(\theta \mid y_1, y_2\right) = \alpha + \beta E\left(x^2 \mid y_1, y_2\right) + E\left(\eta \mid y_1, y_2\right)$$

- Endogeneity means $E\left(\eta \mid y_1, y_2\right) \neq 0$. For now, let's assume away endogeneity so that $E\left(\eta \mid y_1, y_2\right) = 0$.
- If the technology were linear in x, I could use the steps described above to derive a closed-form expression.
- This is where the Unscented Transform is helpful.

Illustration of Estimation Algorithm

- If I use the technology of skill formation, note that:

$$E\left(\theta \mid y_1, y_2\right) = \alpha + \beta E\left(x^2 \mid y_1, y_2\right) + E\left(\eta \mid y_1, y_2\right)$$

  and

$$Var\left(\theta \mid y_1, y_2\right) = \beta^2 Var\left(x^2 \mid y_1, y_2\right) + Var\left(\eta \mid y_1, y_2\right)$$

- Endogeneity means $E\left(\eta \mid y_1, y_2\right) \neq 0$. For now, let's assume away endogeneity so that $E\left(\eta \mid y_1, y_2\right) = 0$.
- Heteroskedasticity means that $Var\left(\eta \mid y_1, y_2\right)$ is a function of y_1, y_2. For now, let's assume homoskedasticity so that $Var\left(\eta \mid y_1, y_2\right) = \sigma_\eta^2$

Illustration of Estimation Algorithm

- Under these assumptions:

$$E\left(\theta \middle| y_1, y_2\right) = \alpha + \beta E\left(x^2 \middle| y_1, y_2\right)$$

  and

$$Var\left(\theta \middle| y_1, y_2\right) = \beta^2 Var\left(x^2 \middle| y_1, y_2\right) + \sigma_\eta^2$$

- If the technology were linear in x, I could use the steps described above to derive a closed-form expression.
- This is where the Unscented Transform is helpful.

Illustration of Unscented Transform

- The Unscented Transform approximates the mean and variance of a random variable that undergoes a nonlinear transformation.
- For a one-dimensional problem, it is possible to approximate both $E\left(x^2\middle|\, y_1, y_2\right)$ and $Var\left(x^2\middle|\, y_1, y_2\right)$ with the following three points:

	Points (χ_j)	Weights (ω_j)
1	$\mu_x - \sigma_x\sqrt{(1+\kappa)}$	$\frac{1}{2(1+\kappa)}$
2	μ_x	$\frac{\kappa}{(1+\kappa)}$
3	$\mu_x + \sigma_x\sqrt{(1+\kappa)}$	$\frac{1}{2(1+\kappa)}$

where κ is a parameter whose value is to be determined below.

Illustration of Unscented Transform

- We use these points and weights to compute the following moments:

$$\widetilde{E}\left(\theta \middle| y_1, y_2\right) = \sum_{j=1}^{3} \omega_j \left(\alpha + \beta\chi_j^2\right)$$

$$\widetilde{Var}\left(\theta \middle| y_1, y_2\right) = \sum_{j=1}^{3} \omega_j \left[\left(\alpha + \beta\chi_j^2\right) - \widetilde{E}\left(\theta \middle| y_1, y_2\right)\right]^2 + \sigma_\eta^2$$

Illustration of Unscented Transform

- It is easy to show that these expressions are equal to:

$$\widetilde{E}\left(\theta \middle| y_1, y_2\right) = \alpha + \beta\left(\mu^2_{x|y_1,y_2} + \sigma^2_{x|y_1,y_2}\right)$$

$$\widetilde{Var}\left(\theta \middle| y_1, y_2\right) = \beta^2\left(\kappa\sigma^4_{x|y_1,y_2} + 4\mu^2_{x|y_1,y_2}\sigma^2_{x|y_1,y_2}\right) + \sigma^2_\eta$$

- How do we pick a value for κ?

Illustration of Unscented Transform

- To understand how to pick the value for κ, note that the assumption of normality of x, together with the assumption that the technology function is quadratic in x, implies that:

$$E\left(\theta|\, y_1, y_2\right) = \alpha + \beta\left(\mu^2_{x|y_1,y_2} + \sigma^2_{x|y_1,y_2}\right)$$

$$Var\left(\theta|\, y_1, y_2\right) = \beta^2\left(2\sigma^4_{x|y_1,y_2} + 4\mu^2_{x|y_1,y_2}\sigma^2_{x|y_1,y_2}\right) + \sigma^2_\eta$$

- In contrast, the approximation via the Unscented Transform yields:

$$\widetilde{E}\left(\theta|\, y_1, y_2\right) = \alpha + \beta\left(\mu^2_{x|y_1,y_2} + \sigma^2_{x|y_1,y_2}\right)$$

$$\widetilde{Var}\left(\theta|\, y_1, y_2\right) = \beta^2\left(\kappa\sigma^4_{x|y_1,y_2} + 4\mu^2_{x|y_1,y_2}\sigma^2_{x|y_1,y_2}\right) + \sigma^2_\eta$$

- The approximation is exact when $\kappa = 2$. Kmenta (1969) showed one can approximate CES function with quadratic function (in logs)

Illustration of Unscented Transform

- This means that we can approximate the first and second conditional moments of y_3

$$\widetilde{E}\left(y_3 | y_1, y_2\right) = \mu_3 + \alpha + \beta\left(\mu^2_{x|y_1,y_2} + \sigma^2_{x|y_1,y_2}\right)$$

$$\widetilde{Var}\left(y_3 | y_1, y_2\right) = \beta^2\left(\kappa\sigma^4_{x|y_1,y_2} + 4\mu^2_{x|y_1,y_2}\sigma^2_{x|y_1,y_2}\right) + \sigma^2_\eta + \sigma^2_3$$

- We can use these moments to approximate the contribution to the likelihood term $L\left(y_3 | y_2, y_1\right)$.
- In the last step, we can follow the rules above to derive $\widetilde{E}\left(\theta | y_1, y_2, y_3\right)$ and $\widetilde{Var}\left(\theta | y_1, y_2, y_3\right)$.
- We then use these moments to obtain $\widetilde{E}\left(y_4 | y_1, y_2, y_3\right)$, $\widetilde{Var}\left(y_4 | y_1, y_{2,} y_3\right)$, and, finally, $L\left(y_4 | y_3, y_2, y_1\right)$.

Estimation of Technology of Skill Formation: Lack of Metric

- The second problem in the estimation of the technology of skill formation is the lack of metric in test scores.
- Our approach: anchor skills on adult outcomes that have clear metric. . Consider linear anchor, z, say log earnings (measured in dollars):

$$z = \mu + \alpha_c \theta_{c,T} + \alpha_n \theta_{n,T} + \nu$$

- Note that $\alpha_c \theta_{c,T}$ and $\alpha_n \theta_{n,T}$ are in dollar units. Consequently:

$$\alpha_k \theta_{k,t+1} = f(\alpha_c \theta_{c,t}, \alpha_n \theta_{n,t}, x_{c,t}, \theta_{c,p}, \eta_{k,t})$$

- Anchoring functions can be linear or nonlinear.

Estimation of Technology of Skill Formation: Endogeneity

- Suppose that Ω_t are the state variables at period t. This means $\theta_{c,T}, \theta_{n,T}, \theta_{c,p}, \theta_{c,p}, \theta_{n,p}, \pi \subset \Omega_t$, but the reverse need not be true.
- Write:

$$\alpha_c \theta_{c,T} = f\left(\alpha_c \theta_{c,t}, \alpha_n \theta_{n,T}, x_{c,t}, \theta_{c,p}, \theta_{n,p}, \pi, \nu_{c,t}\right)$$

- Suppose the policy function for investment is:

$$x_{k,t} = g\left(\Omega_t\right) + \zeta_t$$

- We need exclusion restrictions, say $z_t \in \Omega_t$, but $z_t \notin \theta_{c,T}, \theta_{n,T}, \theta_{c,p}, \theta_{c,p}, \theta_{n,p}, \pi$.
- Repeated measurements on $x_{k,t}$ allows us to identify the distribution of ζ_t.
- Intuition: Nonlinear version of 2SLS.

Estimates of the Technology of Skill Formation

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)

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

Estimates of the Technology of Skill Formation

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

		First Stage Parameters		Second Stage Parameters
Current Period Cognitive Skills (Cross-Productivity)	$\gamma_{1,N,1}$	0.000	$\gamma_{2,N,1}$	0.000
		(0.02)		(0.01)
Current Period Noncognitive Skills (Self-Productivity)	$\gamma_{1,N,2}$	0.712	$\gamma_{2,N,2}$	0.868
		(0.03)		(0.01)
Current Period Investments	$\gamma_{1,N,3}$	0.195	$\gamma_{2,N,3}$	0.121
		(0.03)		(0.03)
Parental Cognitive Skills	$\gamma_{1,N,4}$	0.000	$\gamma_{2,N,4}$	0.000
		(0.01)		(0.01)
Parental Noncognitive Skills	$\gamma_{1,N,5}$	0.093	$\gamma_{2,N,5}$	0.011
		(0.03)		(0.02)
Complementarity Parameter	$\phi_{1,N}$	0.017	$\phi_{2,N}$	-0.323
		(0.27)		(0.21)
Elasticity Parameter	$1/(1-\phi_{1,N})$	1.017	$1/(1-\phi_{2,N})$	0.756
Variance of Shocks $\eta_{N,t}$	$\delta^2_{1,N}$	0.170	$\delta^2_{2,N}$	0.104
		(0.01)		(0.00)

Note: Standard errors in parenthesis.

Interpretation of Findings: Maximizing Average Education

- Suppose that H children are born, $h = 1, ..., H$.
- 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:

$$\sum_{h=1}^{H} (x_{1,h} + x_{2,h}) = B$$

Interpretation of Findings: Minimizing Average Crime

- 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.

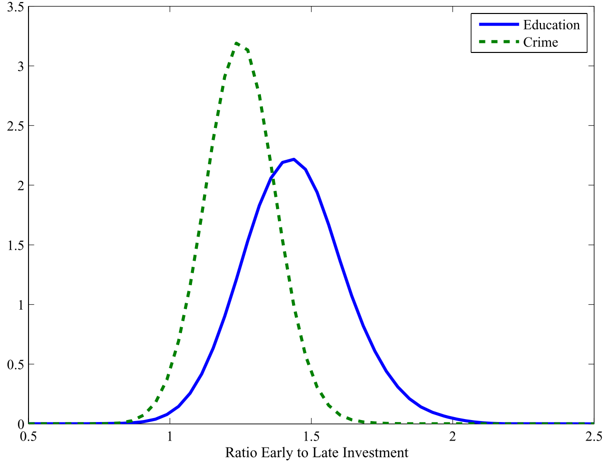


FIGURE 6.—Densities of ratio of early to late investments maximizing aggregate education versus minimizing aggregate crime.

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)

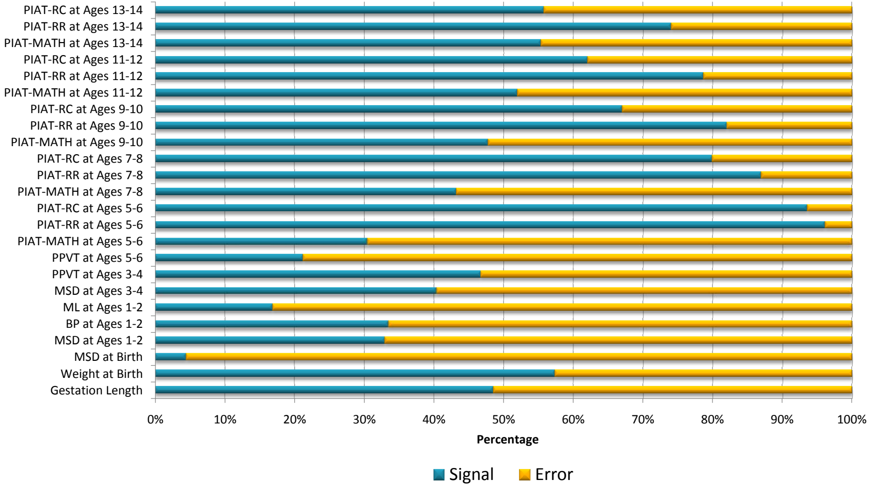


Figure 4A
Share of Residual Variance in Measurements of Noncognitive Skills Due to the Variance of Noncognitive Factor (Signal) and Due to the Variance of Measurement Error (Noise)

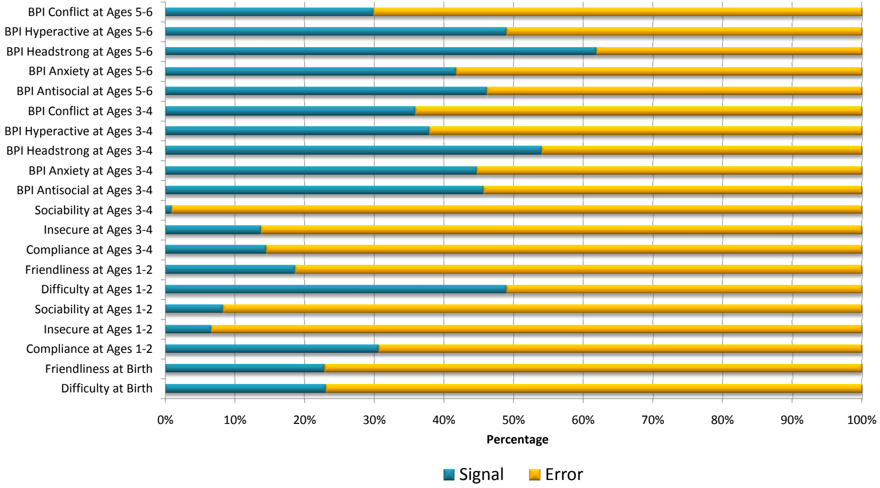


Figure 4B
Share of Residual Variance in Measurements of Noncognitive Skills Due to the Variance of Noncognitive Factor (Signal) and Due to the Variance of Measurement Error (Noise)

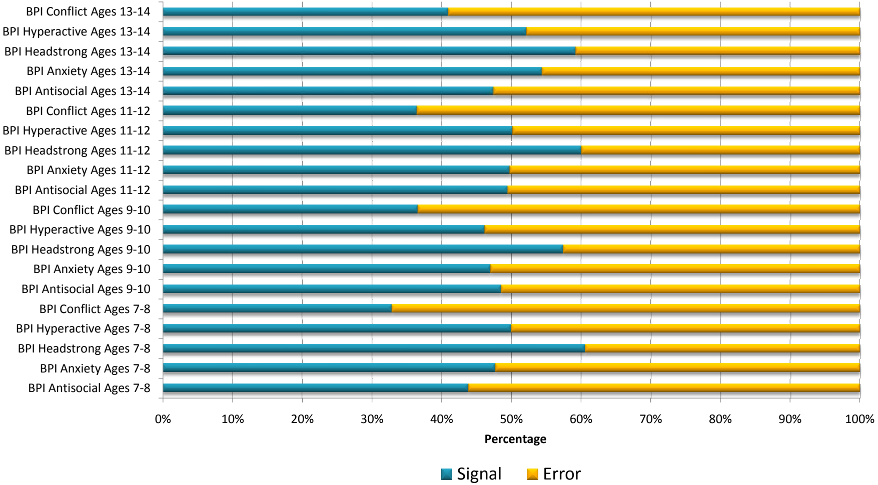


Figure 5A
Share of Residual Variance in Measurements of Investments Due to the Variance of Investment Factor (Signal) and Due to the Variance of Measurement Error (Noise)

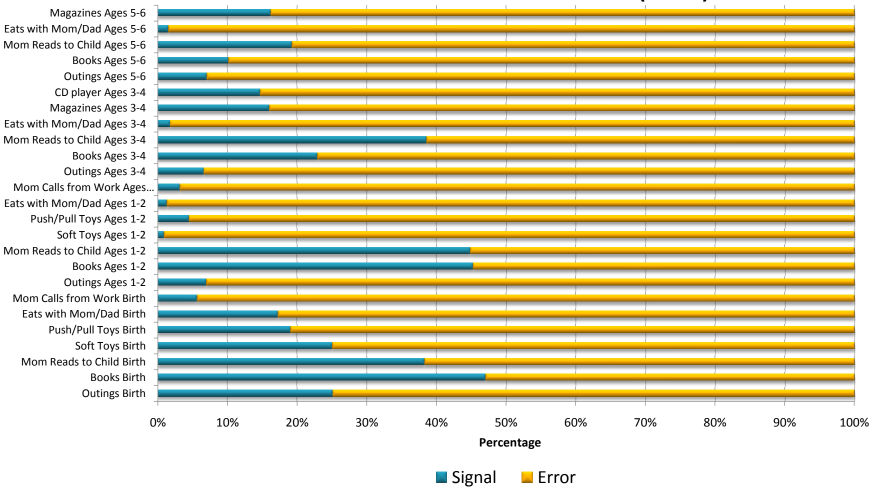


Figure 5B
Share of Residual Variance in Measurements of Investments Due to the Variance of Investment Factor (Signal) and Due to the Variance of Measurement Error (Noise)

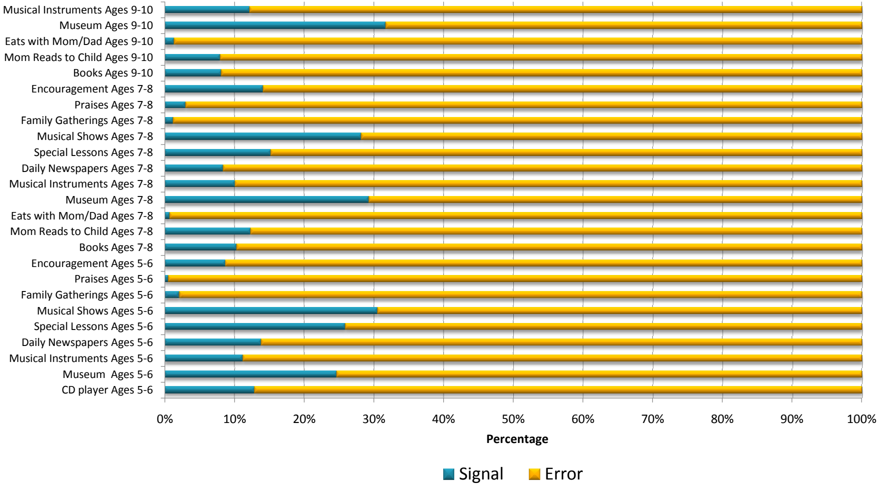


Figure 5C
Share of Residual Variance in Measurements of Investments Due to the Variance of Investment Factor (Signal) and Due to the Variance of Measurement Error (Noise)

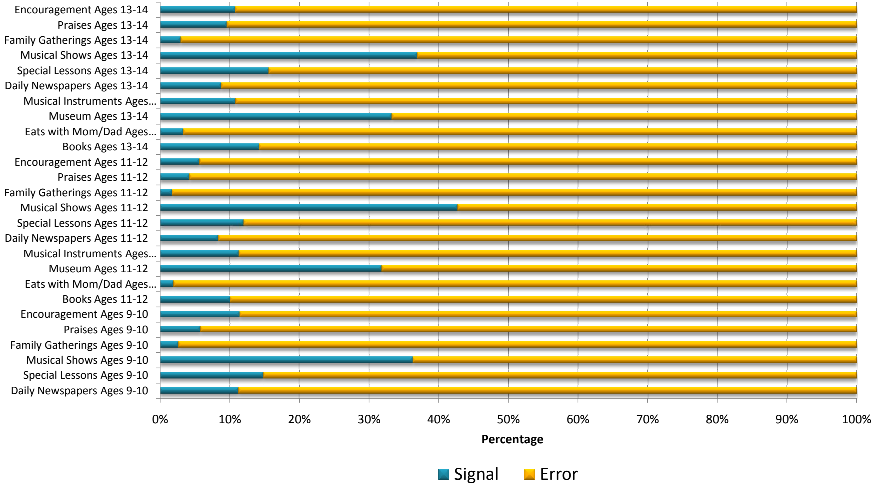
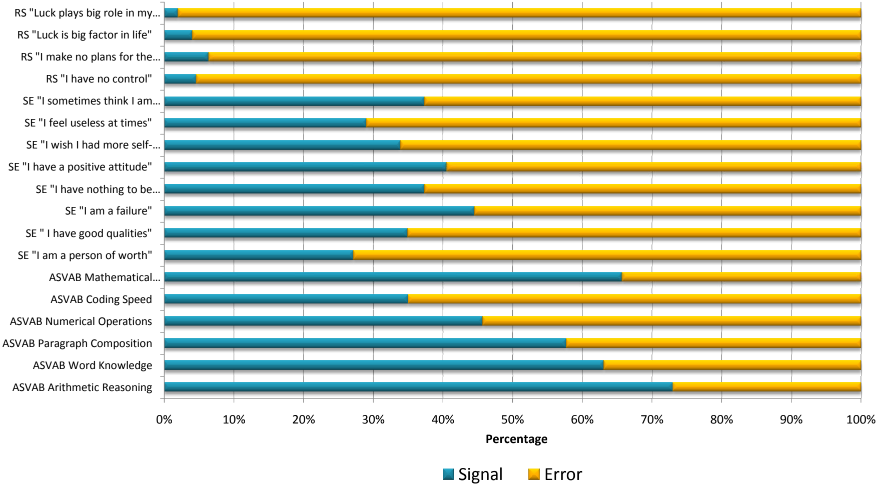


Figure 6
Share of Variance in Measurements of Maternal Cognitive and Noncognitive Skills due to the Variances of Cognitive and Noncognitive Factors (Signal) versus Variance of Measurement Error (Noise)



Home Observation for the Measurement of the Environment

- Created by Bettye Caldwell and Robert Bradley in late 1960s, early 1970s (first published in 1980s)
- Evaluates a child's home environment as well as parent-child interaction.
- Administered by trained professional at the child's home with both child and primary caregiver present.
- Semi-structured interview and observation period: 45-60 minutes.

HOME: Strenghts and Weakenesses

- Strenghts
  - Easy to administer and score.
  - Reliability and validity.
  - Easy to adapt for specific purposes.
  - Provides objective information on home, child, and parent-child interaction.
- Weakenesses:
  - Training of administrators to follow standardized measurement.
  - Only Yes/No questions.
  - Score: Simple summation gives “too much” weight to items that do not vary a lot across households.

HOME: IRT Analysis

- Let θ_i denote the latent quality of the environment experienced by child i.
- Let $d_{i,j}^{*} = a_j \left(\theta_i - b_j\right) + \epsilon_{i,j}$ and define $d_{i,j} = 0$ if $d_{i,j}^{*} \leq 0$ and $d_{i,j} = 1$, otherwise.
- Assume $\epsilon_{i,j}$ has logistic distribution and let θ_i be normally distributed with mean zero and variance σ^2.
- Parameter a_j is item discrimination while b_j is item difficulty.

Interpretation of IRT Parameters

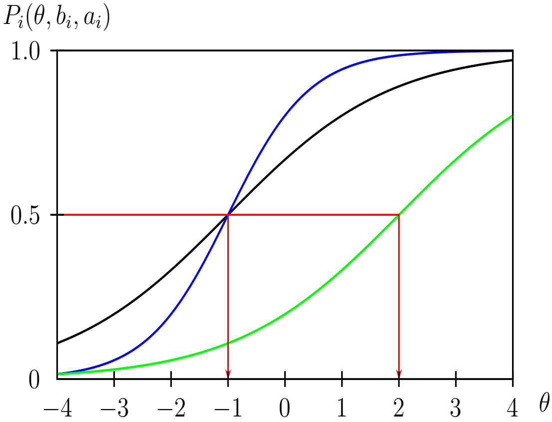
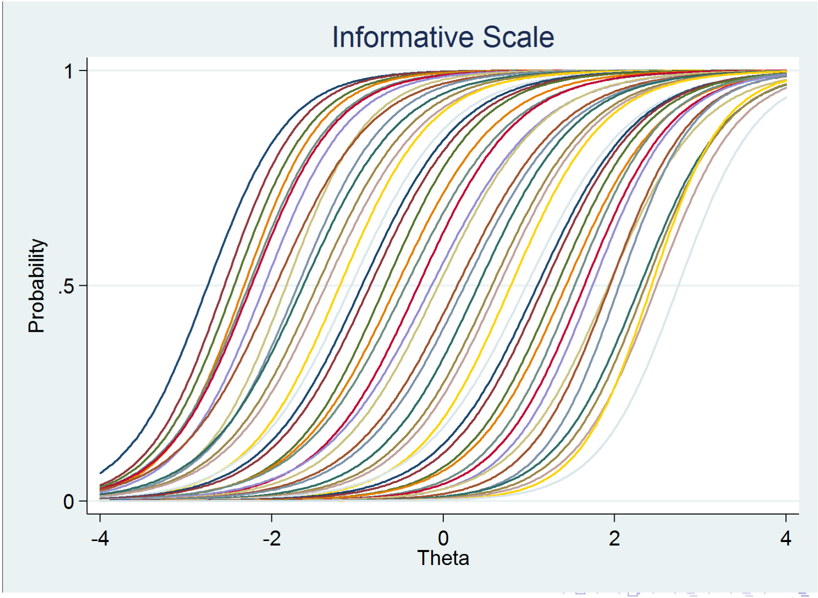
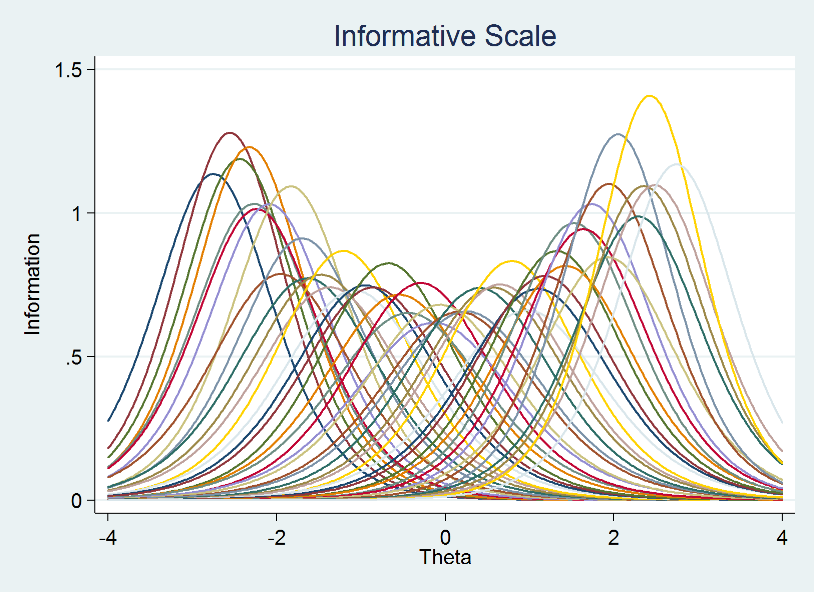


Figure 13: The item response functions of three 2PL items

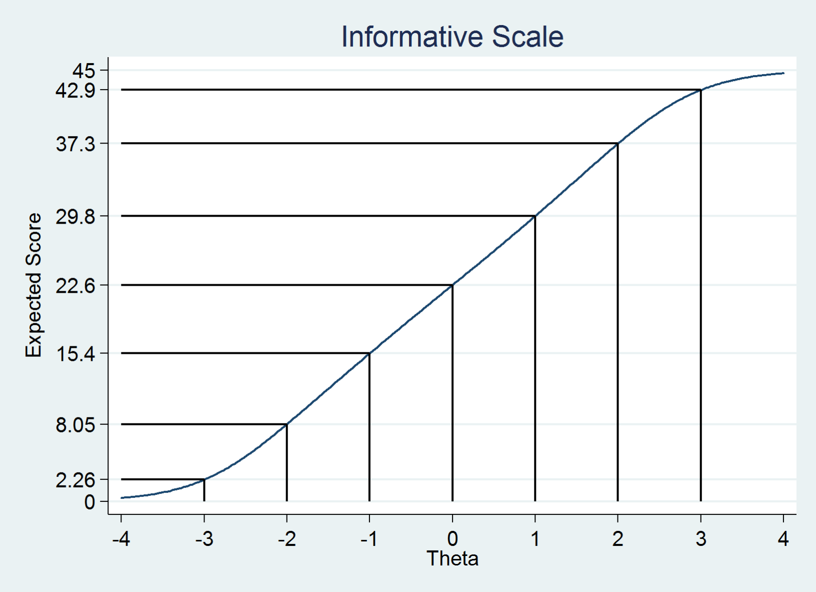
Properties of an Informative IRT Scale



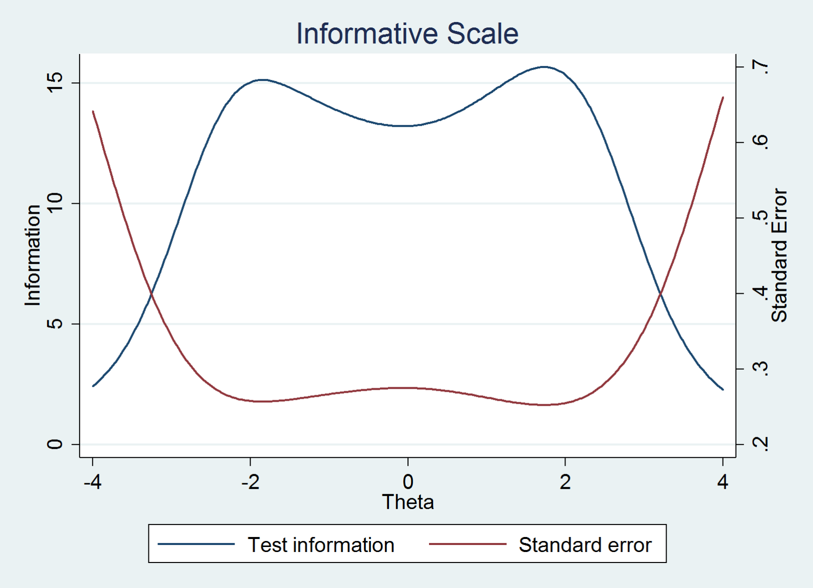
Properties of an Informative IRT Scale: IIF



Properties of an Informative IRT Scale: TCC



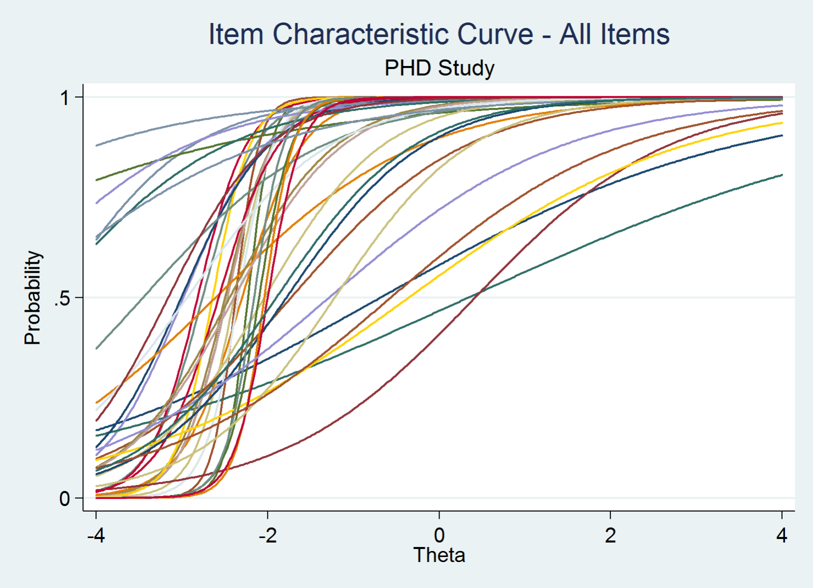
Properties of an Informative IRT Scale: IIF



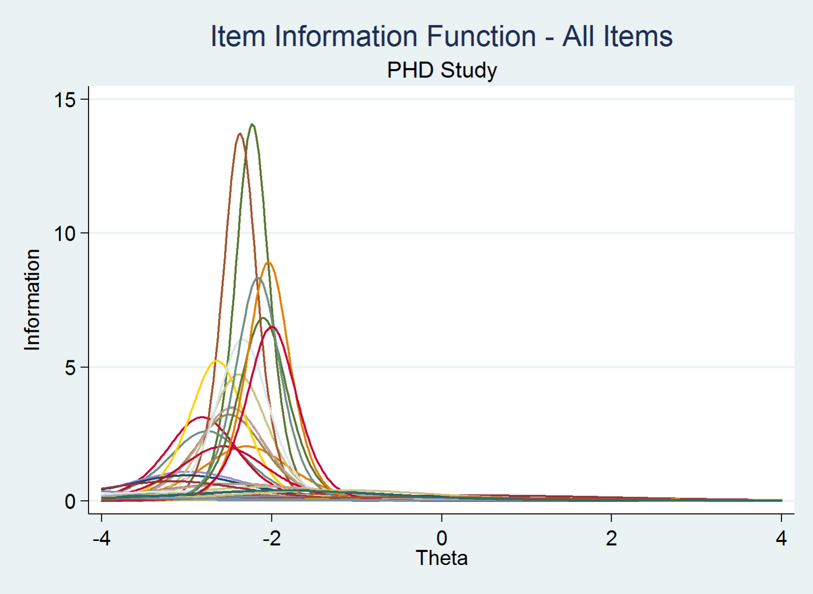
HOME: IRT Analysis

- In few words, an informative scale (as presented in the last four graphs) would have items that have good discriminatory power as well as variability in difficulty.
- This combination allows us to identify, with a lot of precision, households that have low, medium, and high quality environments.
- Unfortunately, the HOME Scale does not have this property.
- As I will show below, that are “too many” easy items and “too few” medium and difficult items.
- For this reason, the HOME Scale will be able to separate very low quality home environments from okay ones, but it will not have power to separate okay from great home environments.

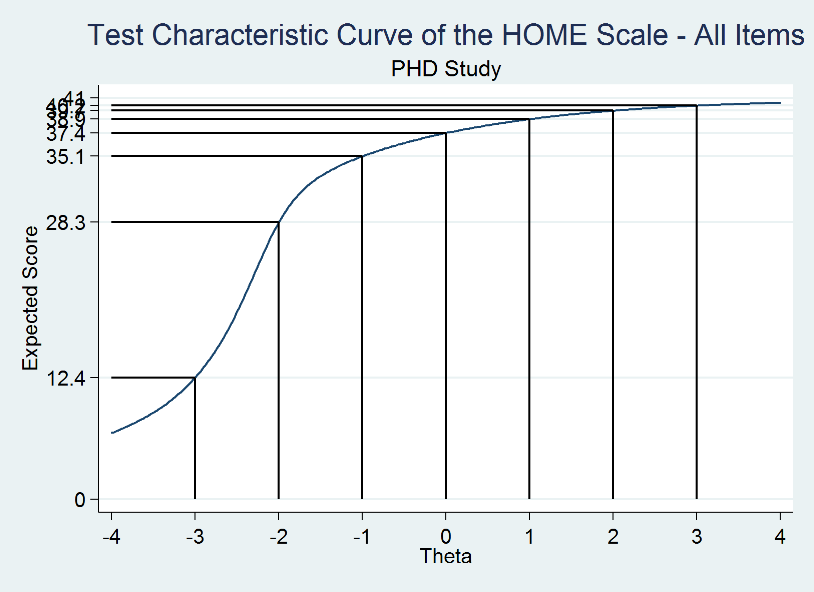
IRT Properties of Full Scale HOME



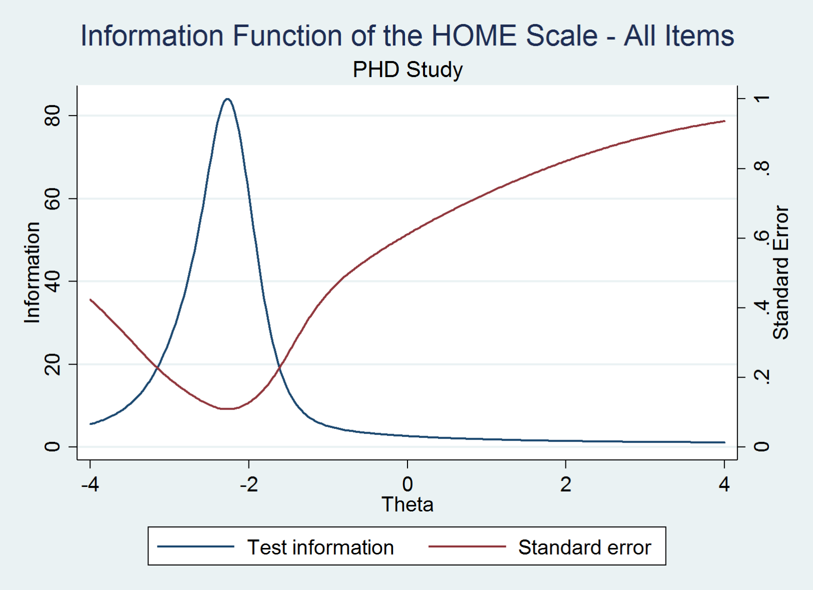
IRT Properties of Full Scale HOME



IRT Properties of Full Scale HOME



IRT Properties of Full Scale HOME

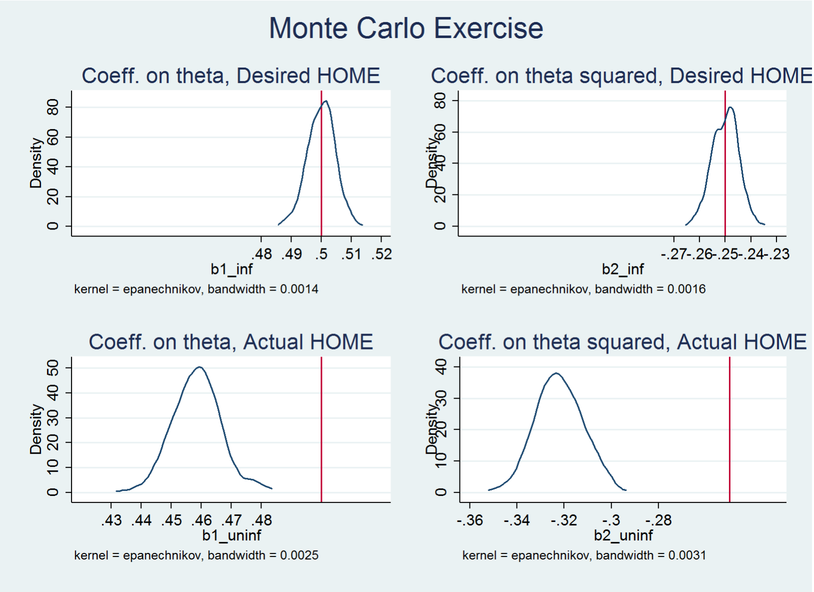


Why does the IRT Properties of the HOME Matter?

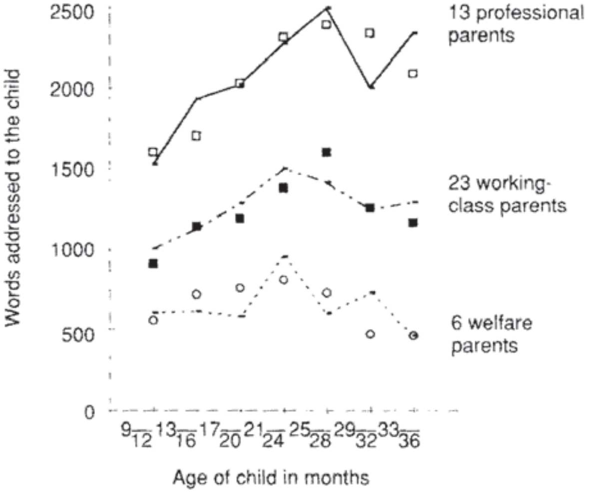
- It probably affects the estimation of the technology of skill formation.
- Why? Medium and high quality environments are difficult to separate.
- It is possible that differences between medium and high quality environments are more (or less) important for child development than differences between medium and low quality environments.
- Either case may lead to biases in the estimation of the technology of skill formation.

Monte Carlo Exercise

- Let h_1 denote human capital, θ denote investments, and ζ denote uncorrelated shocks. Consider the simple technology of skill formation:
$$h_1 = 1.0 + 0.5\theta - 0.25\theta^2 + \zeta \tag{1}$$
- To obtain an idea about potential problems of using the HOME as a measure of investment to be used in the estimation of (1):
  - Generate a HOME Scale with desirable IRT properties as the "desired" HOME Scale;
  - Generate a HOME Scale that has "flawed" IRT properties as the "actual" HOME Scale;
  - Estimate $\theta_{desired}$ from "desired" HOME Scale and θ_{flawed} from "actual" HOME Scale;
  - Regress h_1 on quadratic function of $\theta_{desired}$ and compare estimated with true coefficients;
  - Regress h_1 on quadratic function of θ_{flawed} and compare estimated with true coefficients.


Monte Carlo Exercise
Coeff. on theta, Desired HOME
Density
b1_inf
kernel = epanechnikov, bandwidth = 0.0014
Coeff. on theta squared, Desired HOME
Density
b2_inf
kernel = epanechnikov, bandwidth = 0.0016
Coeff. on theta, Actual HOME
Density
b1_uninf
kernel = epanechnikov, bandwidth = 0.0025
Coeff. on theta squared, Actual HOME
Density
b2_uninf
kernel = epanechnikov, bandwidth = 0.0031

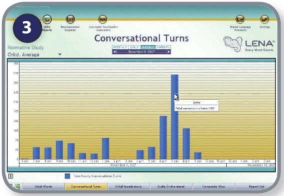
Measuring Quality and Quantity of Time: LENA Pro



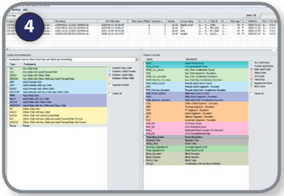
Measuring Quality and Quantity of Time: LENA Pro

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

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.

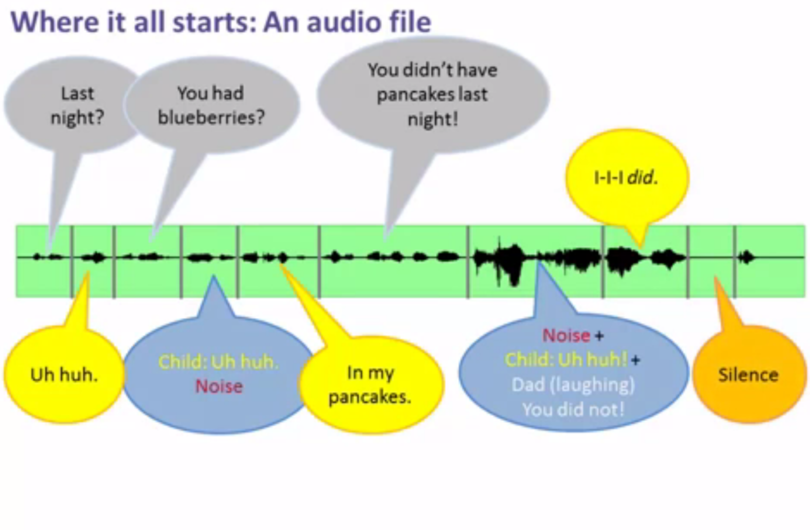


The software generates the LENA reports and other analyses.

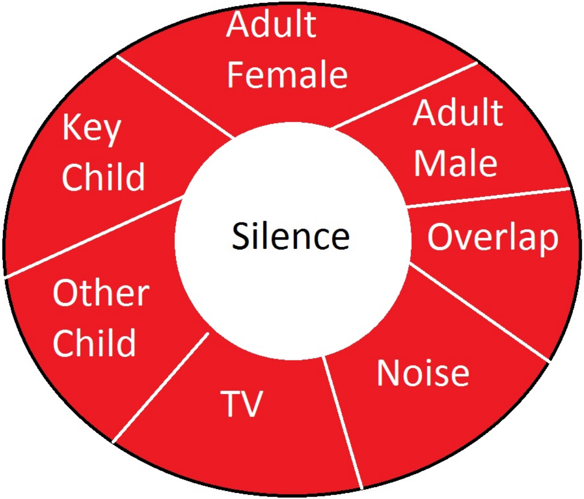


Export data from LENA Pro to mine your LENA data and perform custom in-depth analyses.

Step 1: Segmenting Audio File



Steps 2 and 3: Assign and Confirm



Reliability: Adult Word Counts

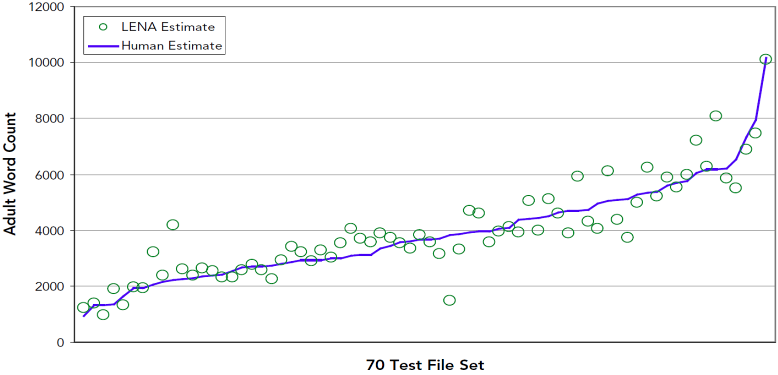
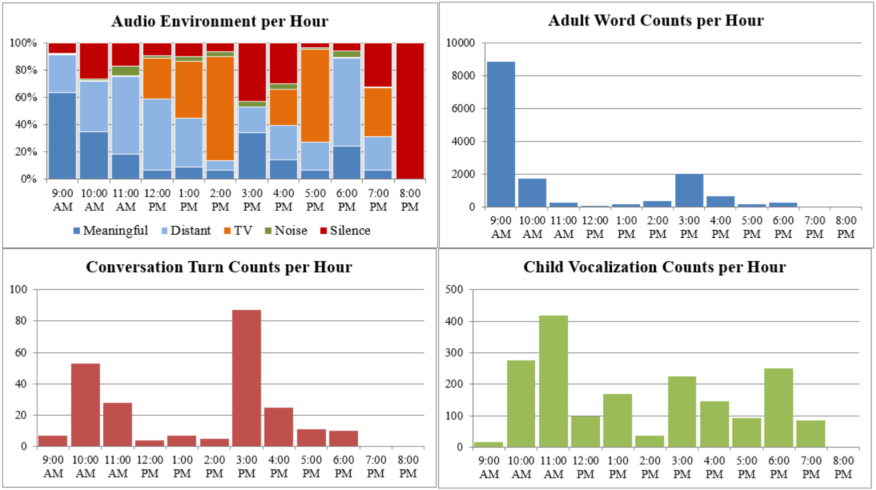


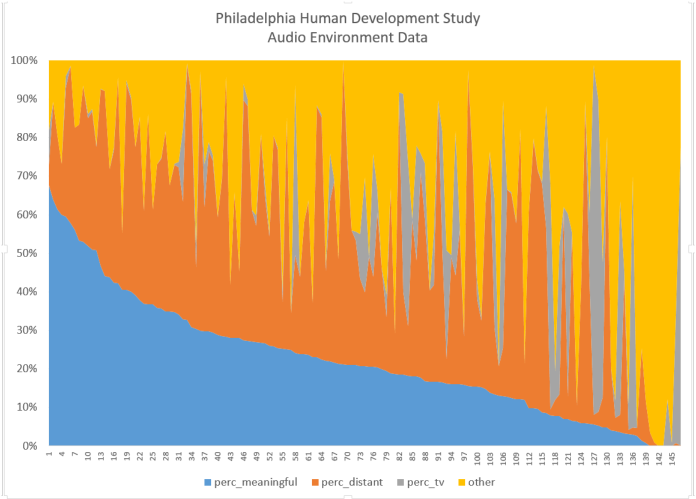
Figure 1. Human and LENA-based AWC estimates for 70 test files.

Measuring Quality and Quantity of Time: LENA Pro



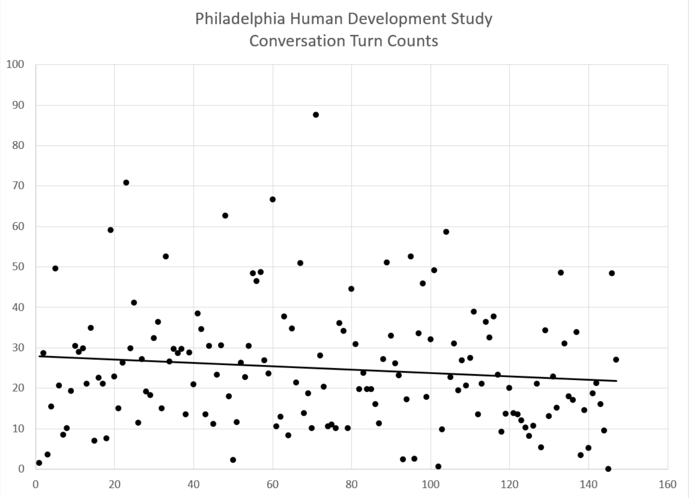
Measuring Quantity of Time: Meaningful Time

Philadelphia Human Development Study



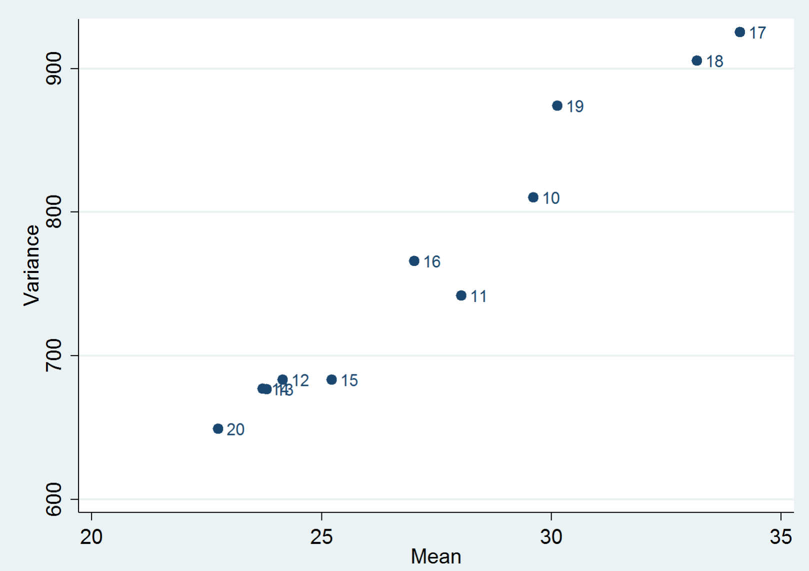
Measuring Quality of Time: Conversation Turn Counts

Philadelphia Human Development Study



Measuring Quality and Quantity of Time: LENA Pro

Philadelphia Human Development Study



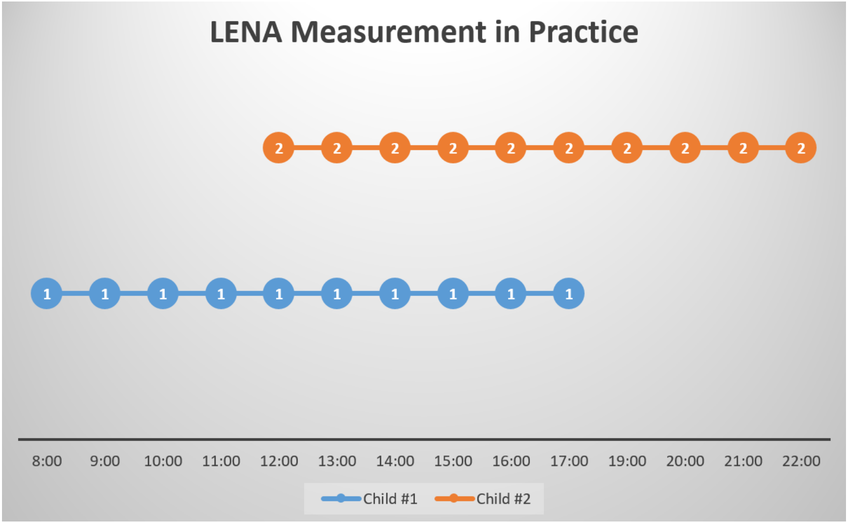
Count Data

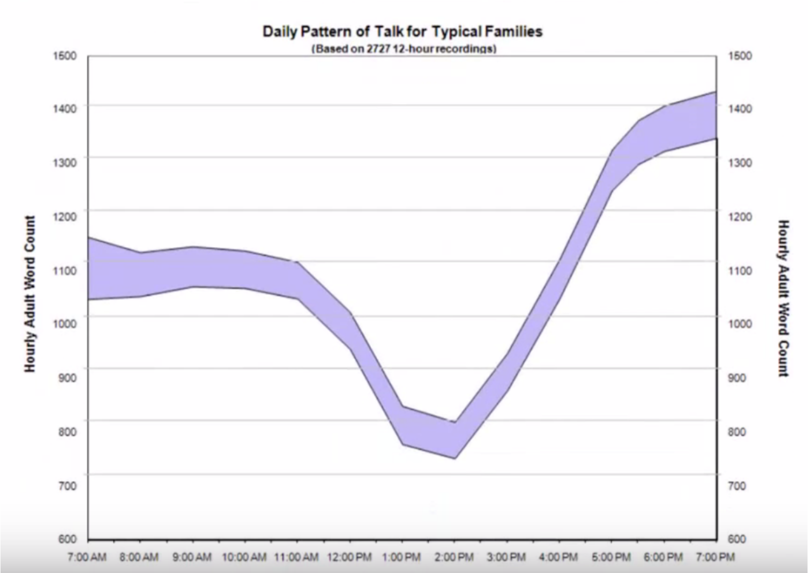
- This dependence between mean and variance (in hours when the mean is high, the variance is also high) is typical in count data.
- One may think of taking the natural log of conversation turn counts and proceed with OLS-type analysis.
- Not a good idea with count data:
  - There are many zeros; taking the logs will eliminate the zeros from the analysis, which means it reduces cases of poor language environment.
  - We want to identify households in terms of expected number of counts, not the expected log of number of counts (nonlinear transformation).

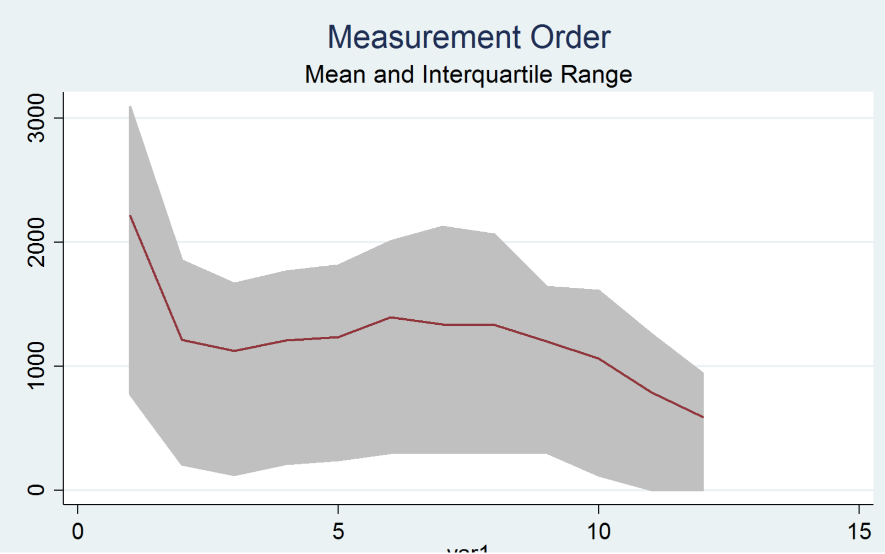
- Let $Y_{i,j}$ denote the jth observation on conversation turn counts between an adult and child i.
- Because these are counts, we model each observation as a Poisson random variable with parameter $\epsilon_i \lambda_{i,j}$ where ϵ_i is a random effect term and $\lambda_{i,j}$ is such that:

$$\ln \lambda_{i,j} = X_{i,j} \delta_j + \ln s_{i,j} \tag{2}$$

- Vector $X_{i,j}$ contains variables that describe the context of measurement and $s_{i,j}$ is "exposure" (i.e., number of seconds that the LENA device was on during the jth measurement).


LENA Measurement in Practice
2
1
8:00
9:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00
17:00
19:00
20:00
21:00
22:00
Child #1
Child #2


Daily Pattern of Talk for Typical Families
(Based on 2727 12-hour recordings)
Hourly Adult Word Count
Hourly Adult Word Count
1600
1400
1300
1200
1100
1000
900
800
700
600
1500
1400
1300
1200
1100
1000
900
800
700
600
7:00 AM
8:00 AM
9:00 AM
10:00 AM
11:00 AM
12:00 PM
1:00 PM
2:00 PM
3:00 PM
4:00 PM
5:00 PM
6:00 PM
7:00 PM


Measurement Order
Mean and Interquartile Range
3000
2000
1000
0
0
5
10
15

Analysis of LENA Conversation Turn Counts Data

- Conditional on ϵ_i, the probability of observing a count equal to:

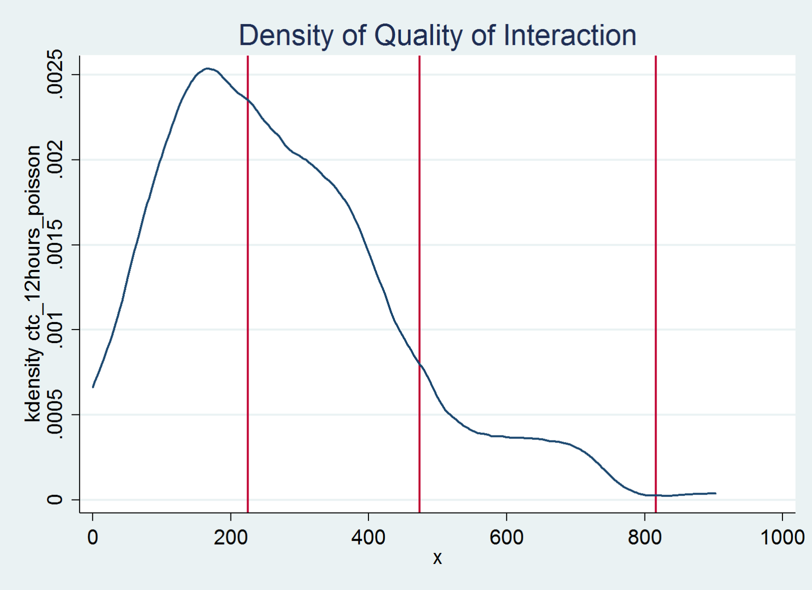
$$\Pr\left(y_{i,j} \middle| \epsilon_i \right) = \frac{\left(\epsilon_i \lambda_{i,j}\right)^{y_{i,j}}}{y_{i,j}!} e^{-\epsilon_i \lambda_{i,j}}$$

  where $\Pr\left(y_{i,j} \middle| \epsilon_i \right) = \Pr\left(Y_{i,j} = y_{i,j} \middle| \epsilon_i \right)$ is the probability that the count of variable $Y_{i,j}$ is equal to $y_{i,j}$ conditional on ϵ_i.

- Assume that, conditional on ϵ_i, the events are independent. Thus:

$$\Pr\left(y_{i,1}, \ldots, y_{i,J} \middle| \epsilon_i \right) = \left\{ \left[\prod_{j=1}^{J} \frac{\left(\lambda_{i,j}\right)^{y_{i,j}}}{y_{i,j}!} \right] \epsilon_i^{\sum_{j=1}^{J} y_{i,j}} e^{-\epsilon_i \sum_{j=1}^{J} \lambda_{i,j}} \right\} \tag{3}$$

- Because we don't observe the random effect ϵ_i, we need to integrate it out.
- We assume that ϵ_i has gamma distribution with mean one and variance $\frac{1}{\alpha}$

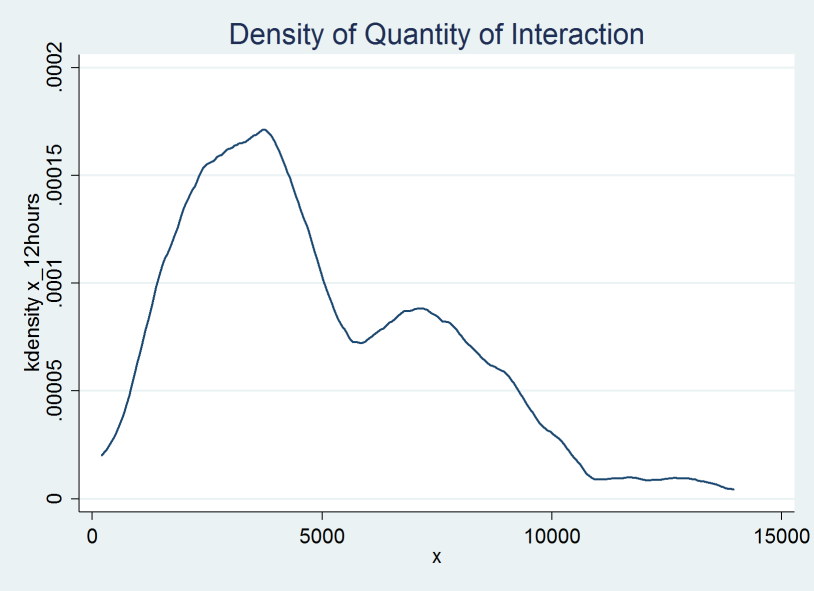

Density of Quality of Interaction
kdensity ctc_12hours_poisson
0
.0005
.001
.0015
.002
.0025
0
200
400
600
800
1000
x

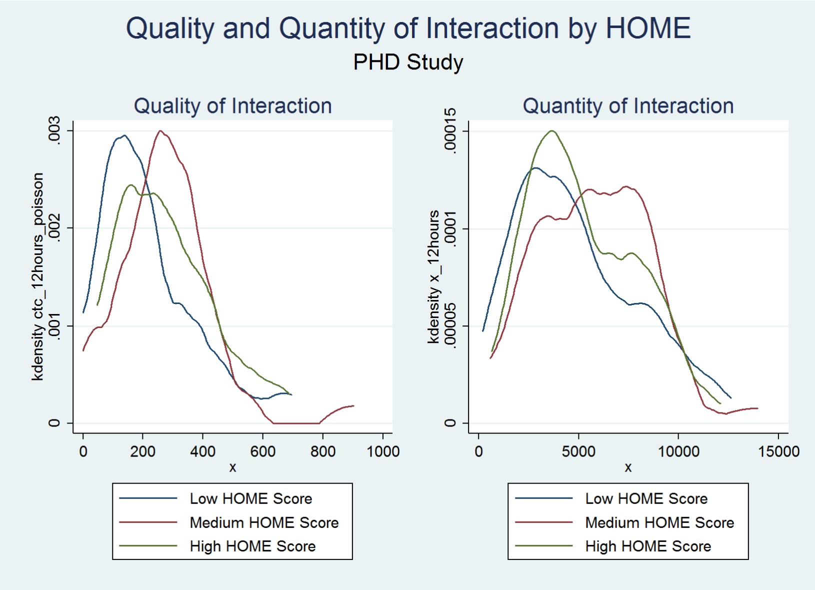
Analysis of LENA Conversation Turn Counts Data

- Let $M_{i,j}$ denote the share of meaningful time of adult-child interaction in jth observation.
- Because these are proportion data, we model each observation as the following logistic regression:

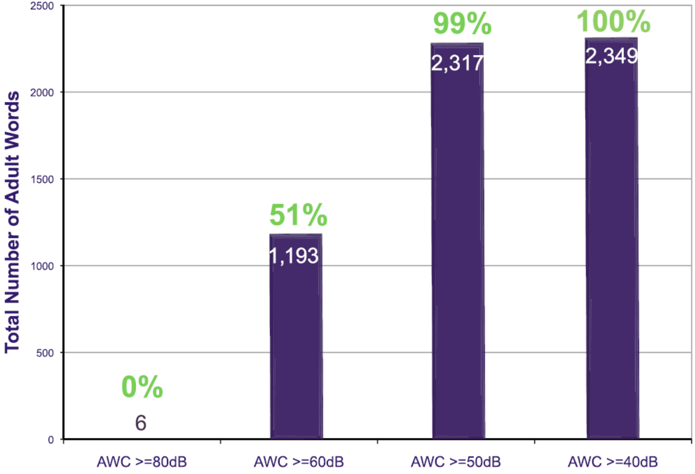
$$\ln \left\{ \frac{M_{i,j}}{1 - M_{i,j}} \right\} = X_{i,j}\rho_j + \mu_i + \nu_{i,j}$$

  where μ_i is a random effect with mean zero and variance σ_μ^2.
- We are interested in estimating the unobserved heterogeneity captured by μ_i across families.


Density of Quantity of Interaction
kdensity x_12hours
0
.00005
.0001
.00015
.0002
0
5000
10000
15000
x



Quality and Quantity of Interaction by HOME
PHD Study
Quality of Interaction
Quantity of Interaction
kdensity ctc_12hours_poisson
kdensity x_12hours
0
.001
.002
.003
0
.00005
.0001
.00015
0
200
400
600
800
1000
0
5000
10000
15000
x
x
Low HOME Score
Medium HOME Score
High HOME Score


Total Number of Adult Words
2500
2000
1500
1000
500
0
0%
6
51%
1,193
99%
2,317
100%
2,349
AWC >=80dB
AWC >=60dB
AWC >=50dB
AWC >=40dB

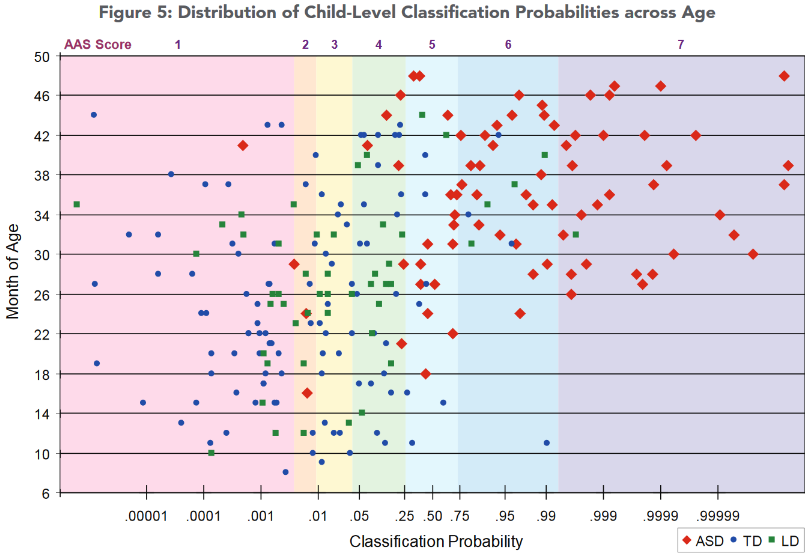
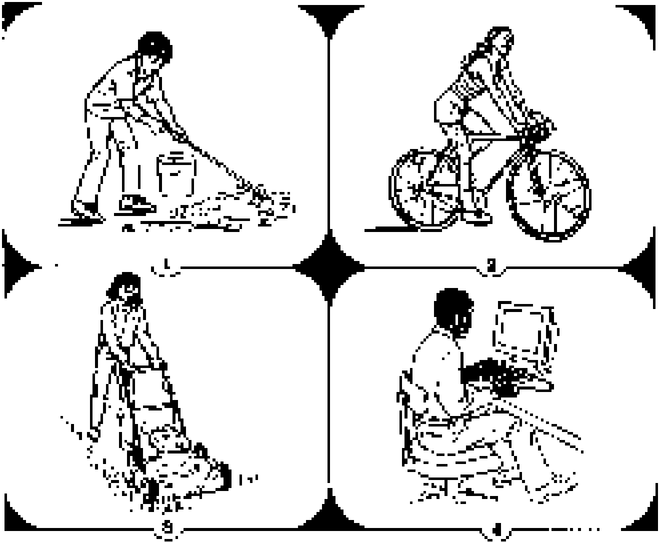


Figure 5: Distribution of Child-Level Classification Probabilities across Age
AAS Score
1
2
3
4
5
6
7
50
46
42
38
34
30
26
22
18
14
10
6
Month of Age
.00001
.0001
.001
.01
.05
.25
.50
.75
.95
.99
.999
.9999
.99999
Classification Probability
ASD
TD
LD

Innovations in the Measurement of Skills

- In the standard measurement paradigm:
  - The goal is to measure a trait or skill θ_i (could be scalar or vector).
  - Each task is a self-contained situation.
  - The task evokes a response meant to provide evidence about the construct.
  - Each response is evaluated to provide an item score.
  - A test score accumulates evidence over the many different tasks.

Say: Point to mowing.



1) At Joe's Restaurant, one-fourth of the patrons are male and one-fifth of the patrons are from out of town. What proportion would you expect to be male and out of town?

- 1/5
- 1/10
- 1/20
- 1/25
- 1/50

The Big Five Inventory (BFI)

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement.

Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree Strongly
1	2	3	4	5

<u>I see Myself as Someone Who...</u>

____1. Is talkative

____2. Tends to find fault with others

____3. Does a thorough job

____4. Is depressed, blue

____23. Tends to be lazy

____24. Is emotionally stable, not easily upset

____25. Is inventive

____26. Has an assertive personality

1. When it is mealtime, how often does your child eat what you want him/her to eat?

1	2	3	4	5
Almost Never	Less than 1/2 the time	1/2 the time	More than 1/2 the time	Almost Always

2. When your child doesn't eat what you want him/her to eat and you tell him/her to do so, how often does he/she obey and eat?

1	2	3	4	5
Almost Never	Less than 1/2 the time	1/2 the time	More than 1/2 the time	Almost Always

Concerns about Standardized Tests

23 **Within a substance, atoms that collide frequently and move independently of one another are *most* likely in a**

A liquid.

B solid.

C gas.

D crystal.

CSZ20827

Concerns about Standardized Tests

Concerns about Standardized Tests

- To be able to do science, we need to have:
  - Clear understanding of goals;
  - Motivation;
  - Collaboration skills;
  - Understanding of steps and beliefs;
  - Science process skills;
  - Science content knowledge.
- Traditional tests measure only the science knowledge.
- Simulation tests aim to measure all of the factors of interest.

Concerns about Self Reports from Surveys

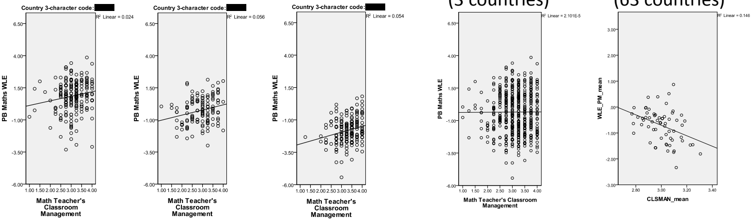
- Motivated responding (e.g., wanting to look good, or bad)
- Reference group bias (to whom you compare yourself)
- Response style bias (e.g., extreme responses, modesty)
- Lack of differentiation in others' ratings (halo, horn)
- Cross-cultural comparability (to compare countries x and y)

Concerns about Self Reports from Surveys

Within-country

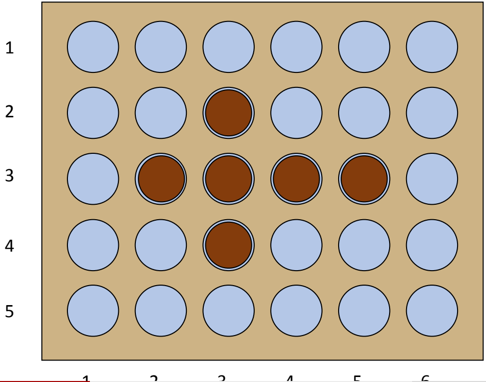
Pooled sample (3 countries)

Country means (63 countries)



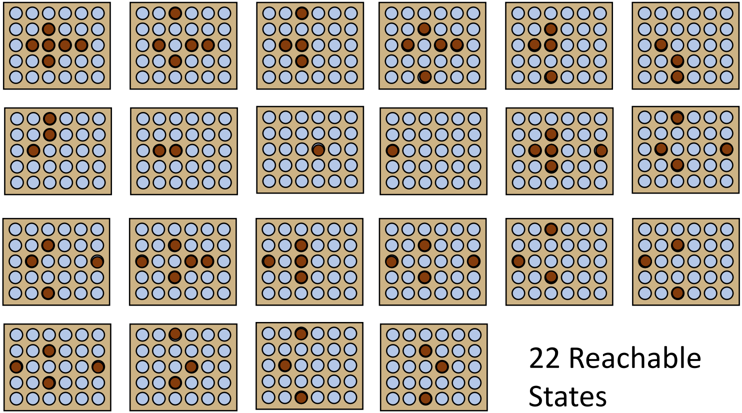
Peg Solitaire Game

- Very interesting work by Rafferty, LaMar, and Griffiths (2015)
- The goal is to leave as few pegs on the board as possible.
- Jump pegs to remove them.
- Way to evaluate an individual's decision-making process, strategic thinking, motivation.



States

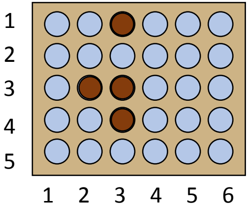
State Space



22 Reachable States

States

Each state presents a choice:



Available Actions A_s

$(3,3) \longrightarrow (1,3)$

$(3,3) \longrightarrow (5,3)$

$(3,2) \longrightarrow (3,4)$

$(4,3) \longrightarrow (2,3)$

Score

Reset

Statistics: MDP / Economics: DDCM

State Space	$S = \{s_1, s_2, \ldots s_{\bar{S}}\}$
Action Set	$A = \{a_1, a_2, \ldots a_{\bar{A}}\}$
Transition Function	$T(s, a, s') = p(s' \| s, a)$
Reward Structure	$R(s, a, s')$

$\Downarrow$

Policy: $p(a|s, \xi)$

Statistics: MDP / Economics: DDCM

- Let $V(s|a)$ denote the value of choosing action a when state is s:

  $$V(s|a) = R(a|s) + p(s'|s,a)\,V(s'|a)$$

  where $R(a|s)$ is the reward of choosing action a and $p(s'|s,a)$ is the state transition function.

Statistics: MDP / Economics: DDCM

- Let $\theta_i \in [0, \infty)$ denote the individual "capability". Then, define

$$\Pr\left(\, chooses \quad a \middle|\, s\right) = \frac{e^{\theta_i V(\,s|a)}}{\sum_b e^{\theta_i V(\,s|b)}}$$

- If $\theta_i = 0$,

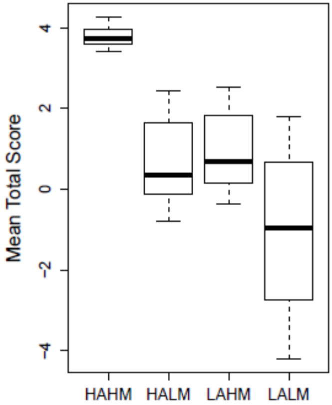
$$\Pr\left(\, chooses \quad a \middle|\, s\right) = \frac{1}{\sum_b}$$

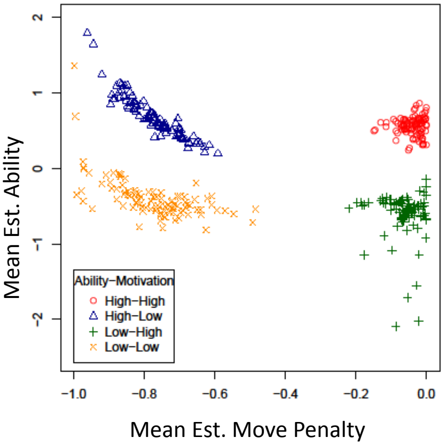
- As θ_i approaches infinity:

$$\Pr\left(\, chooses \quad a \middle|\, s\right) = \begin{cases} 0, & \text{if } a \text{is not optimal} \\ 1, & \text{if } a \text{is optimal} \end{cases}$$

Statistics: MDP / Economics: DDCM

Rafferty, LaMar, and Griffiths (2015)



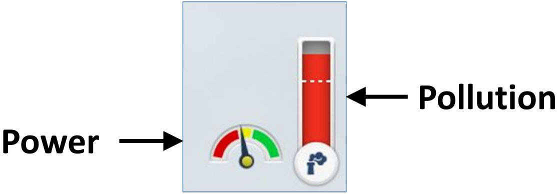


http://vimeo.com/69007945

DDCM Applied to SimCity

Rafferty, LaMar, and Griffiths (2015)

- Designed to assess systems thinking.
- There are 17 actions and 25,420 reachable states.
- Students must optimize two variables simultaneously.



Technological Progress: Text

A New Class Pet

Have you ever considered getting a new class pet?
Well if you happen to be searching for one now I've got a few
suggestions for you. If your looking for a pet thats really easy to take
care of try a chameleon. Just get a jar or tank and put in
a few twigs, a bit of grass, some leaves, and plenty of live pray!
And if thats not what you had in mind and wanted
a more active animal try a rabbit. They are cute, cuddly,
and active. You can buy rabbit mix at Wallmart or at
a pet shop.
But If You want to Traumatize the children
(which I Hope you don't) get an untamed King cobra!
If you wanted something more like a dog you
could get a golden retriever in my oppinion the are
the best Kind of ~~dog~~. They are gentle and calm
and love children.
But I will allways say that
a rat is the best pet. Not a greasy, slimy, Bloodsuckin

Technological Progress: Text

Class	Frequency Spectrum **V(m,N)**	
m	V(m,N)	mV(m,n)
1	94	94
2	18	36
3	10	30
4	2	8
6	5	30
7	1	7
8	2	16
10	1	10
13	1	13
21	1	21
Total	135	265

Estimating Vocabulary Size

- Think of writing text as selecting words from an urn.
- Let π_i denote the probability of selecting word ω_i from the urn.
- The frequency of word ω_i in a sample of N tokens is binomially distributed (N, π_i) distributed.

Estimating Vocabulary Size

- Expected Frequency Spectrum (expected number of words that will be used m times in a text of N words):

$$E\left[V\left(m,N\right)\right]=\sum_{i=1}^{S}\left(\begin{array}{c}N\\ m\end{array}\right)\pi_i^m\left(1-\pi_i\right)^{N-m}$$

- Expected Vocabulary Size (given a text of N words):

$$E\left[V\left(N\right)\right]=\sum_{m=1}^{N}\sum_{i=1}^{S}\left(\begin{array}{c}N\\ m\end{array}\right)\pi_i^m\left(1-\pi_i\right)^{N-m}$$

Estimating Vocabulary Size

- Formulas are not too helpful because we can't use observed data to estimate π_i (too many words in the urn, regardless of size N, the sample will always be too small).
- Make distributional assumptions (requires lots and lots of text data - thousands of words, which is unlike to get from children in Grade 1).

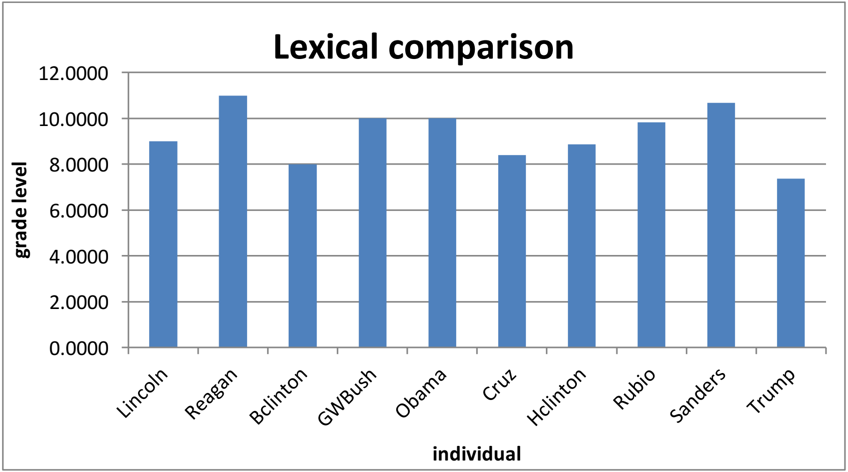
Estimating Vocabulary Size

- Use outside data to estimate π_i.
  - Treebank 3 (Wall Street Journal): 1 million tokens.
  - The New York Times Annotated Corpus: 100 million tokens.
  - Web 1T 5-gram Version 1: 1 trillion tokens from the web.
- Text "New Class Pet" is from 3rd grade student, but vocabulary is more like someone in 5th grade.
- Possible to use tools to analyze grammar.

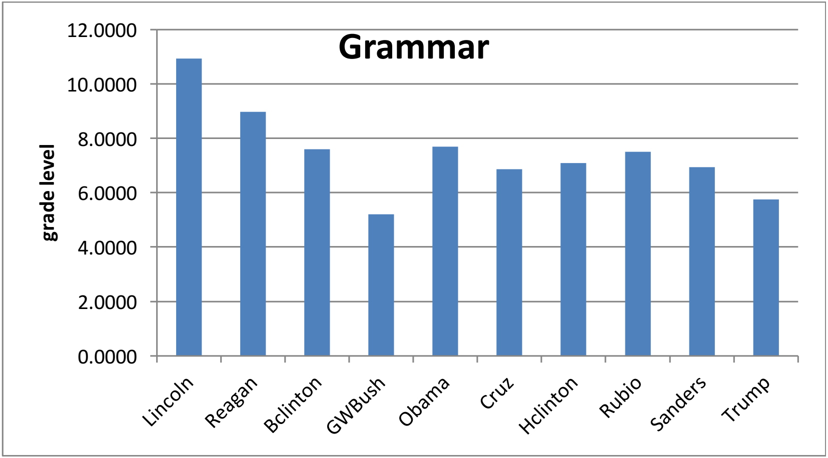
Estimating Vocabulary Size

- Use outside data to estimate π_i.
  - The British National Corpus (BNC) is a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of British English, both spoken and written, from the late twentieth century.
  - Use Adam Kilgarriff estimates of π_i (http://www.kilgarriff.co.uk/bnc-readme.html).
- Text "New Class Pet" is from 3rd grade student, but vocabulary is more like someone in 5th grade.
- Possible to use tools to analyze grammar.

Presidential Campaign: Lexicon



Presidential Campaign: Grammar



Benefits of Text Data

- Better measure of communication.
- Usually not used to evaluate teachers and schools, so not so sensitive to "teaching to test."

Advances in Self Report Data: Anchoring Vignettes

Kyllonen and Bertling (2013)

ST61

02 Thinking about the mathematics teacher who taught your last mathematics class. To what extent do you agree with the following statements?

(Please check only one box on each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) My teacher lets students know they need to work hard.	☐1	☐2	☐3	☐4

Advances in Self Report Data: Anchoring Vignettes

Kyllonen and Bertling (2013)

ST61

01 Below you will find descriptions of three mathematics teachers. Read each of the descriptions of these teachers. Then let us know to what extent you agree with the final statement.

(Please check only one box on each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) Ms. Anderson assigns mathematics homework every other day. She always gets the answers back to students before examinations. **Ms. Anderson is concerned about her students' learning.**	☐1	☐2	☐3	☐4
b) Mr. Crawford assigns mathematics homework once a week. He always gets the answers back to students before examinations. **Mr. Crawford is concerned about his students' learning.**	☐1	☐2	☐3	☐4
c) Ms. Dalton assigns mathematics homework once a week. She never gets the answers back to students before examinations. **Ms. Dalton is concerned about her students' learning.**	☐1	☐2	☐3	☐4

Teacher Support Questions

Advances in Self Report Data: Anchoring Vignettes

Kyllonen and Bertling (2013)

ST61

01 Below you will find descriptions of three mathematics teachers. Read each of the descriptions of these teachers. Then let us know to what extent you agree with the final statement.

Student "A's" responses

(Please check only one box on each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) Ms. Anderson assigns mathematics homework every other day. She always gets the answers back to students before examinations. **Ms. Anderson is concerned about her students' learning.**	☒ 1	☐ 2	☐ 3	☐ 4
b) Mr. Crawford assigns mathematics homework once a week. He always gets the answers back to students before examinations. **Mr. Crawford is concerned about his students' learning.**	☐ 1	☒ 2	☐ 3	☐ 4
c) Ms. Dalton assigns mathematics homework once a week. She never gets the answers back to students before examinations. **Ms. Dalton is concerned about her students' learning.**	☐ 1	☐ 2	☐ 3	☒ 4

Advances in Self Report Data: Anchoring Vignettes

Kyllonen and Bertling (2013)

01 Below you will find descriptions of three mathematics teachers. Read each of the descriptions of these teachers. Then let us know to what extent you agree with the final statement.

(Please check only one box on each row.)

Student "A's" responses

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) Ms. Anderson assigns mathematics homework every other day. She always gets the answers back to students before examinations. **Ms. Anderson is concerned about her students' learning.**	☒	☐	☐	☐
b) Mr. Crawford assigns mathematics homework once a week. He always gets the answers back to students before examinations. **Mr. Crawford is concerned about his students' learning.**	☐	☒	☐	☐
c) Ms. Dalton assigns mathematics homework once a week. She never gets the answers back to students before examinations. **Ms. Dalton is concerned about her students' learning.**	☐	☐	☐	☒
02 My teacher lets students know they need to work hard.	☒	☐	☐	☐

For Student "A" this can be interpreted as "at the same level as the best hypothetical teacher"

Advances in Self Report Data: Anchoring Vignettes

Kyllonen and Bertling (2013)

01 Below you will find descriptions of three mathematics teachers. Read each of the descriptions of these teachers. Then let us know to what extent you agree with the final statement.

Student "A's" responses

(Please check only one box on each row.)

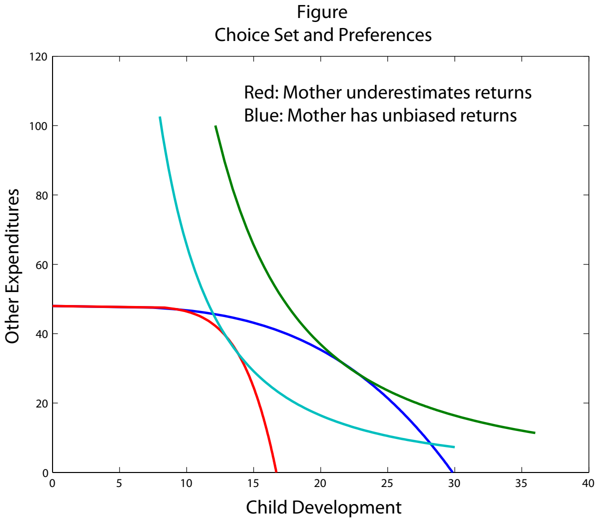
	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) Ms. Anderson assigns mathematics homework every other day. She always gets the answers back to students before examinations. **Ms. Anderson is concerned about her students' learning.**	☒ 1	☐ 2	☐ 3	☐ 4
b) Mr. Crawford assigns mathematics homework once a week. He always gets the answers back to students before examinations. **Mr. Crawford is concerned about his students' learning.**	☐ 1	☒ 2	☐ 3	☐ 4
c) Ms. Dalton assigns mathematics homework once a week. She never gets the answers back to students before examinations. **Ms. Dalton is concerned about her students' learning.**	☐ 1	☐ 2	☐ 3	☒ 4
02 My teacher lets students know they need to work hard.	☐ 1	☒ 2	☐ 3	☐ 4

For Student "A" this can be interpreted as "like the middle hypothetical teacher"

Other Measurements Using Technology

- Altruism
- Reciprocity
- Risk Preference
- Patience
- Trust
- Decision Making Capacity (Rationality)

Introducing Heterogeneity in Beliefs



Why Heterogeneity in Beliefs?

- Language acquisition: Hart and Risley (1995); Rowe (2008).
- Time spent in activities that are appropriate for the child's age (Kalil et al, 2012).
- Home visitation programs on parenting:
  - Nurse-Family Partnership (Olds et al, 2012).
  - Jamaican Nutrition Supplementation and Cognitive Stimulation Program (Gertler et al, 2014; Attanasio et al, 2014).
  - HIPPY Program (Baker et al, 2002).
  - Parent as Teachers (PAT, Wagner et al, 1998)
  - Play and Learning Strategies (PALS, Landry et al, 1996).
  - Thirty Million Words Program (Suskind and Lefler, 2013).
  - Many others (Healthy Families, Healthy Start, CHIP of Virginia, MOM of Philadelphia, etc.)

Current Research

- My current research aims to answer the following questions:
    - Can we measure parental beliefs about the technology of skill formation?
    - If so:
        - How do parental beliefs compare with objective estimates of the technology of skill formation?
        - Is there heterogeneity in parental beliefs?
    - If so, does the heterogeneity in beliefs predict heterogeneity in investments?
    - If so, can we change parental investments by affecting parental beliefs?

Model: The technology of skill formation

- The technology of skill formation is:

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

Model: The mother's information set

- Let Ψ_i denote the mother's information set.
- Let $E\left(\psi_j \middle| h_{0,i}, x_i, \Psi_i\right) = \mu_{i,j}$ and assume that $E\left(\nu_i \middle| \Psi_i\right) = 0$.
- From the point of view of the mother:

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

Model: Preferences and budget constraint

- Consider a simple static model. Parent's utility is:

$$u\left(c_i, h_{i,1}; \alpha_{i,1}, \alpha_{i,2}\right) = \ln c_i + \alpha_{i,1} \ln h_{i,1} + \alpha_{i,2} \ln x_i$$

- 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 α_i from $\mu_{i,\gamma}$ if we only observe x_i, y_i, and p.

Identification

- Elicit maternal beliefs.
- Elicit maternal preferences.
- Estimate the technology of skill formation.

Eliciting beliefs: 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 H: (22 MONTHS - 3 YEARS, 11 MONTHS)

MOTHER/GUARDIAN:
If ____________________ (Child's Name) is **at least 22 months** old, but **not yet 4 years** old, please answer these 15 questions.

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, θ_i is deviation from typical development for age.

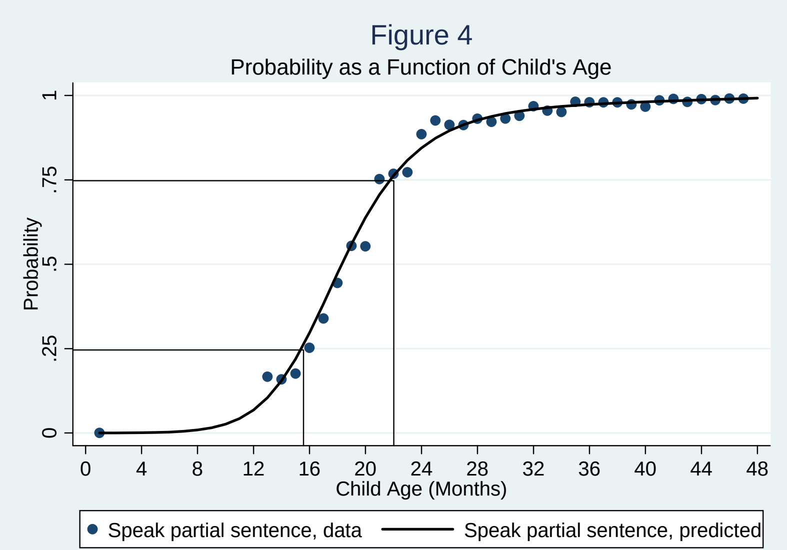


Figure 4
Probability as a Function of Child's Age
Probability
0
.25
.5
.75
1
0
4
8
12
16
20
24
28
32
36
40
44
48
Child Age (Months)
Speak partial sentence, data
Speak partial sentence, predicted

Eliciting beliefs: Changing wording of the MSD Instrument

- In order to measure $E\left[\ln h_{i,1} \middle| h_0, x, \psi_i\right]$, 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 $x = \overline{x}$ – 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.

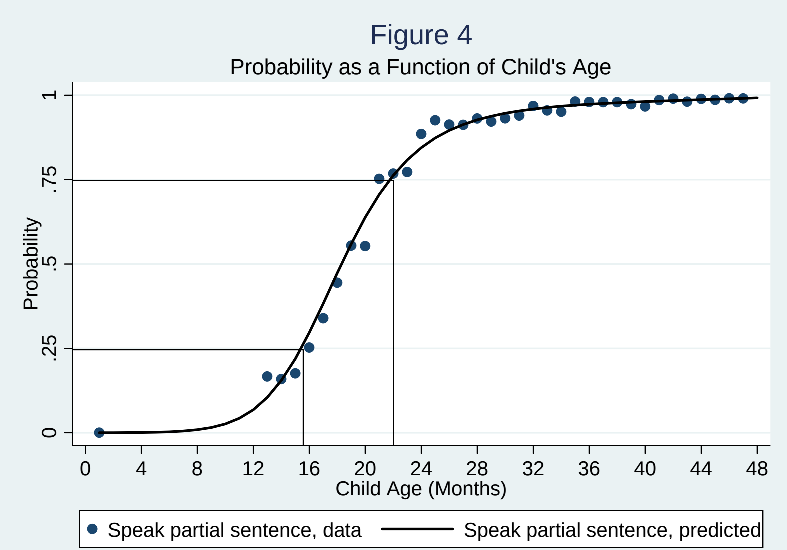
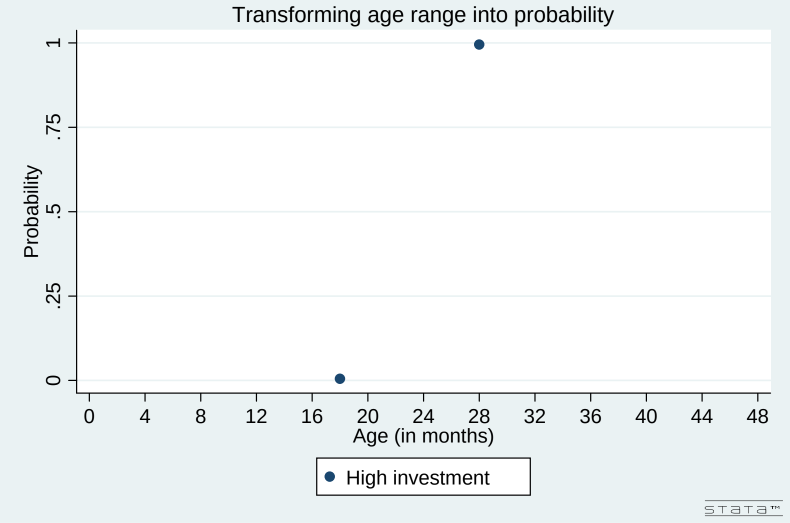


Figure 4
Probability as a Function of Child's Age
Probability
0
.25
.5
.75
1
0
4
8
12
16
20
24
28
32
36
40
44
48
Child Age (Months)
Speak partial sentence, data
Speak partial sentence, predicted

Method 2: Transforming age ranges into probabilies

- **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 $x = \overline{x}$, 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
Probability
0
.25
.5
.75
1
0
4
8
12
16
20
24
28
32
36
40
44
48
Age (in months)
High investment

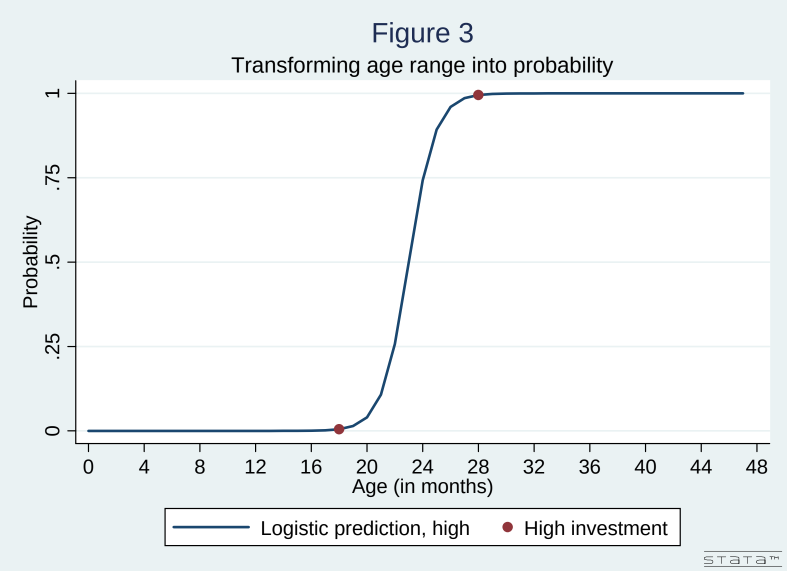


Figure 3
Transforming age range into probability
Probability
0
.25
.5
.75
1
0
4
8
12
16
20
24
28
32
36
40
44
48
Age (in months)
Logistic prediction, high
High investment

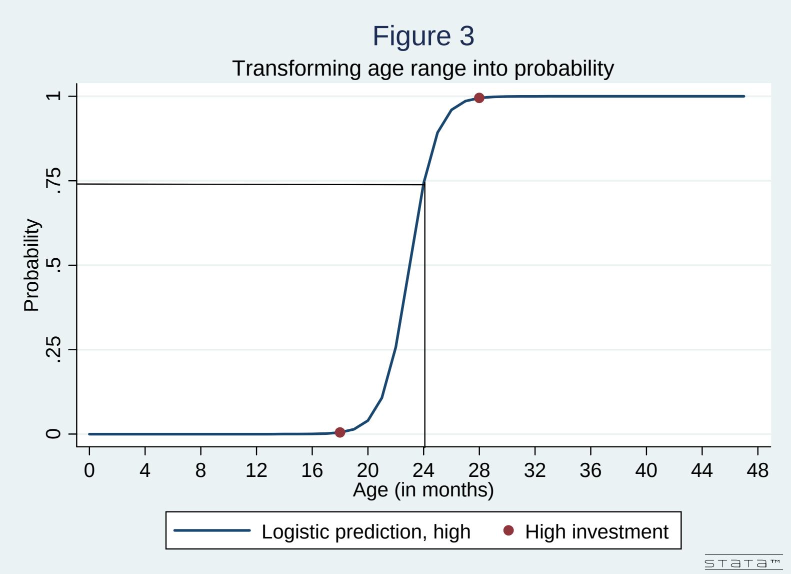


Figure 3
Transforming age range into probability
Probability
0
.25
.5
.75
1
0
4
8
12
16
20
24
28
32
36
40
44
48
Age (in months)
Logistic prediction, high
High investment

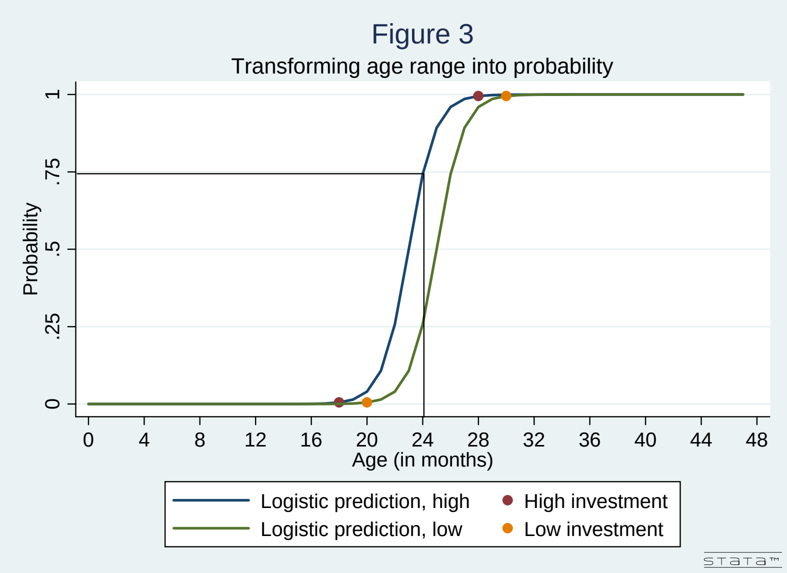


Figure 3
Transforming age range into probability
Probability
0
.25
.5
.75
1
0
4
8
12
16
20
24
28
32
36
40
44
48
Age (in months)
Logistic prediction, high
Logistic prediction, low
High investment
Low investment

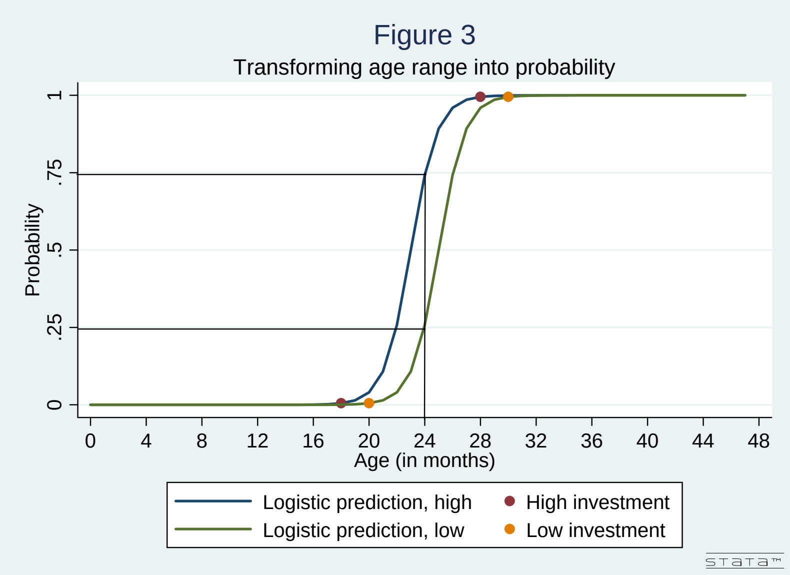


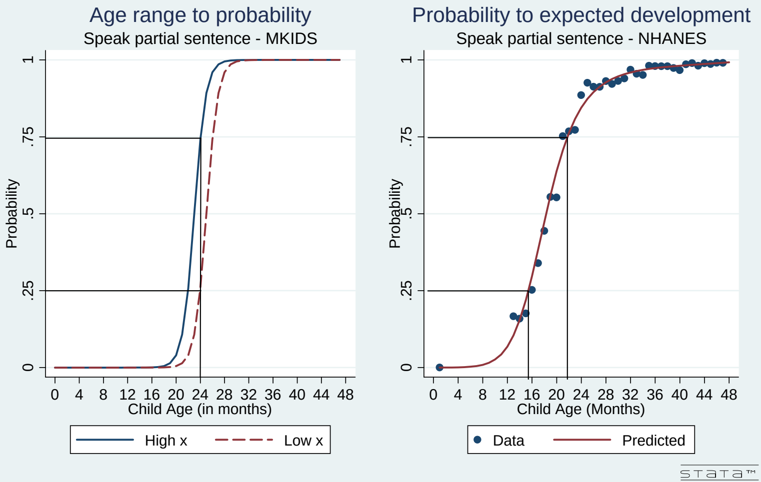
Figure 3
Transforming age range into probability
Probability
0
.25
.5
.75
1
0 4 8 12 16 20 24 28 32 36 40 44 48
Age (in months)
Logistic prediction, high
Logistic prediction, low
High investment
Low investment
stata™

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 $x = \overline{x}$, 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 $x = \underline{x}$, 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)



Recovering mean beliefs: Measurement error model

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

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

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

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

- 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:

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

.

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

- 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;
  - $\epsilon_{i,j,k} = \left(\epsilon^L_{i,j,k}, \epsilon^A_{i,j,k}\right)$ are the uniquenesses.

Eliciting beliefs: Intuitive explanation

- Let $E\left[\ln h_{i,1} \middle| h_0, h, \Psi_i\right]$ denote maternal expectation of child development at age 24 months conditional on the child's intial level of human capital, investments, and the mother's information set.
- Assume, for now, technology is Cobb-Douglas.
- Suppose we measure $E\left[\ln h_{i,1} \middle| h_0, x, \Psi_i\right]$ at two different levels of investments:

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

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

- Subtracting and re-organizing terms:

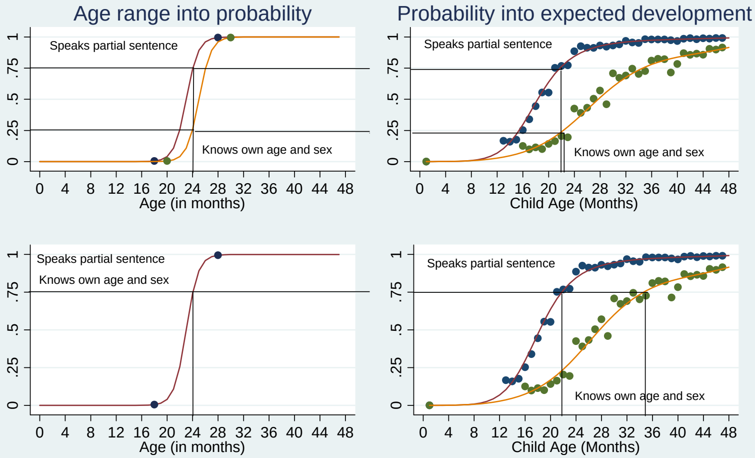
$$\mu_{i,2} = \frac{E\left[\ln h_{i,1} \middle| h_0, \overline{x}, \Psi_i\right] - E\left[\ln h_{i,1} \middle| h_0, \underline{x}, \Psi_i\right]}{\ln \overline{x} - \ln \underline{x}}$$

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



Estimation of Preferences

- The investment policy function is:

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

where $\alpha_{i,1}$ and $\alpha_{i,2}$ captures heterogeneity in preferences.
- The usual procedure is to work with observed investment data.
- We are in the field collecting these investment data.

Estimation of Preferences

- Today, we elicit the preference parameters by stated-choice data (as it is commonly applied in Marketing).
- We tell the respondent to assume that the child's initial level of human capital is high.
- Then, we create nine hypothetical scenarios of monthly income and prices:

		Price		
		\$30	\$45	\$60
Income	\$1500	Scenario 1	Scenario 2	Scenario 3
	\$2000	Scenario 4	Scenario 5	Scenario 6
	\$2500	Scenario 7	Scenario 8	Scenario 9

Estimation of Preferences

- In order to link investment to time, we prepared a three-minute video in which we explain to the respondent that the more time that the mother interacts with the child, the more money she has to spend every month buying educational goods such as child books and educational toys.
- Our goal is to pass on to the respondent the idea that investment is costly.
- Respondents are not familiar with the concept of "opportunity cost."

Estimation of Preferences

- For each combination of prices and income, we ask the respondents the following question: *Suppose that your household income is \$$y$ per month and that for each hour per day that the mother spends interacting with the child she has to spend \$$p$ per month on educational goods. Consider the following four options:*
- The four options correspond to two, three, four, and five hours of investments per day.
- Thus, if the respondent reports $x_{i,m,n}$ hours of investment per day when price is p_m and income is y_n, then share of income allocated to investments, $s_{m,n}$ is:

$$s_{i,m,n} = \frac{p_m x_{i,m,n}}{y_n}$$

Estimation of Preferences

- Note that the ratio, $r_{i,m,n}$ is:

$$r_{i,m,n} = \frac{s_{i,m,n}}{1 - s_{i,m,n}} = \alpha_{i,1} \left(\mu_{i,2} + \mu_{i,3} \ln h_{0,i} \right) + \alpha_{i,2} + \xi_{i,m,n}$$

.

- The parameters $\alpha_{i,1}$ and $\alpha_{i,2}$ can be estimated as a simple random-effects model.

Descriptive Information about Participants: MKIDS and PHD

- 777 participants, all African-American.
- MKIDS: 60% are primiparous; PHD: 100% are primiparous.
- 80% are single (not cohabiting or married).
- 80% are at most 25 years-old.
- Median household income is below the second decile of U.S. distribution.
- Low education sample: only 12% of respondents have a two-year college degree or more.

Table 1
Comparison of Datasets

	MKIDS		PHD		Total	
Number of observations	323		454		777	
	N	%	N	%	N	%
Type of Elicitation Method						
Only probability	20	6.2	0	0.0	20	2.6
Only age ranges	233	72.1	0	0.0	233	30.0
Both methods	70	21.7	454	100.0	524	67.4
MSD Items						
Wearing wet pants bothers child	323	100.0	0	0.0	323	41.6
Speak partial sentence	323	100.0	454	100.0	777	100.0
Say first and last name	323	100.0	454	100.0	777	100.0
Count 3 objects correctly	323	100.0	454	100.0	777	100.0
Know own age and sex	323	100.0	454	100.0	777	100.0
Says the names of 4 colors	323	100.0	0	0.0	323	41.6
Count out loud up to 10	323	100.0	0	0.0	323	41.6
Draw picture of man/woman	323	100.0	0	0.0	323	41.6
Hypothetical scenarios						
Baseline	158	48.9	454	100.0	612	78.8
Alternative scenario #1	42	13.0	0	0.0	42	5.4
Alternative scenario #2	91	28.2	0	0.0	91	11.7
Alternative scenario #3	32	9.9	0	0.0	32	4.1
Stated choice data						
Hypothetical scenarios for prices of investment and income	158	48.9	0	0.0	158	20.3

Table 2

Basic Features of Raw Data

MSD Items ranked in ascending order of difficulty		NHANES	Probability					Age ranges				
				Scenarios					Scenarios			
Rank	Item Description		Obs.	1	2	3	4	Obs.	1	2	3	4
1	Chils lets someone know that wearing wet pants bothers him/her?	0.99	90	0.78 (0.24)	0.55 (0.27)	0.70 (0.27)	0.51 (0.26)	303	0.64 (0.33)	0.55 (0.36)	0.50 (0.36)	0.43 (0.37)
2	Child speaks a partial sentence of 3 words or more	0.72	544	0.81 (0.18)	0.63 (0.22)	0.61 (0.20)	0.45 (0.20)	757	0.60 (0.36)	0.44 (0.38)	0.42 (0.38)	0.31 (0.36)
3	Child counts 3 objects correctly?	0.39	544	0.84 (0.18)	0.67 (0.22)	0.62 (0.20)	0.47 (0.20)	757	0.41 (0.38)	0.32 (0.36)	0.26 (0.33)	0.19 (0.30)
4	Child knows own age and sex	0.31	544	0.83 (0.19)	0.66 (0.23)	0.62 (0.21)	0.47 (0.21)	757	0.33 (0.36)	0.26 (0.33)	0.23 (0.31)	0.17 (0.29)
5	Child says first and last name together without someone's help	0.26	544	0.80 (0.20)	0.64 (0.22)	0.60 (0.21)	0.46 (0.21)	757	0.31 (0.36)	0.24 (0.33)	0.22 (0.31)	0.17 (0.29)
6	Child says the names of at least 4 colors	0.20	90	0.81 (0.23)	0.59 (0.28)	0.74 (0.22)	0.56 (0.27)	303	0.26 (0.31)	0.22 (0.29)	0.19 (0.28)	0.16 (0.26)
7	Child counts out loud up to 10?	0.07	90	0.80 (0.20)	0.58 (0.27)	0.75 (0.19)	0.53 (0.27)	303	0.24 (0.30)	0.20 (0.28)	0.19 (0.28)	0.16 (0.27)
8	Child draws a picture of a man/woman, 2 parts besides head	0.02	90	0.71 (0.25)	0.51 (0.28)	0.67 (0.21)	0.48 (0.26)	303	0.15 (0.26)	0.15 (0.26)	0.14 (0.26)	0.13 (0.25)

Note: Standard errors in parenthesis.

Beliefs about the technology of skill formation

Table 3

Maternal Beliefs about the Technology of Skill Formation

	25th percentile	Median	75th percentile	Mean	Variance
$\mu_{\psi,0}$	-0.015	0.101	0.236	0.115	0.035
	(0.009)	(0.008)	(0.009)	(0.007)	(0.002)
$\mu_{\psi,1}$	0.077	0.296	0.554	0.365	0.204
	(0.011)	(0.016)	(0.022)	(0.016)	(0.026)
$\mu_{\psi,2}$	0.065	0.166	0.285	0.192	0.046
	(0.006)	(0.007)	(0.010)	(0.008)	(0.005)
$\mu_{\psi,3}$	-0.008	0.094	0.335	0.190	0.320
	(0.007)	(0.010)	(0.024)	(0.020)	(0.051)

Note: Standard errors in parenthesis.

Sensitivity analysis

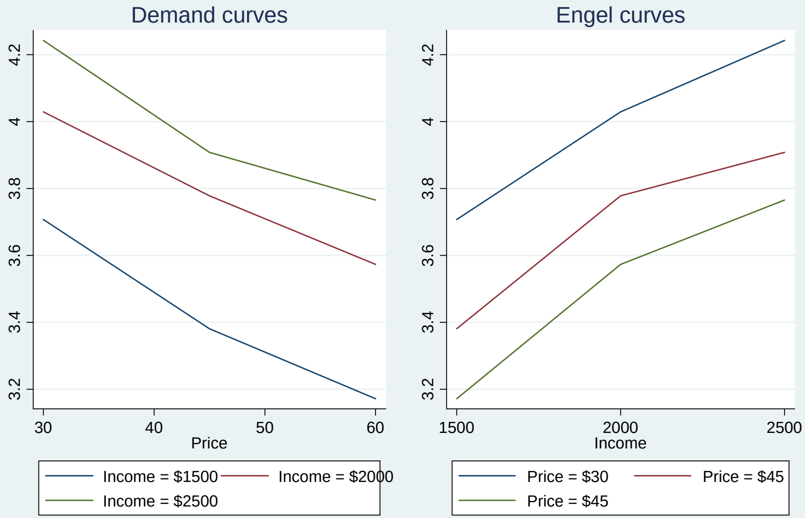
Table 4

Alternative Definition of Scenarios and Maternal Beliefs

Regressors	Dependent variables				F Test
	$\mu_{\psi,0}$	$\mu_{\psi,1}$	$\mu_{\psi,2}$	$\mu_{\psi,3}$	(p-value)
Intercept (baseline)	0.018	0.147	0.112	0.070	-
	(0.017)	(0.043)	(0.022)	(0.062)	-
Dummy for alternative scenario #1	0.067	-0.027	-0.032	-0.081	1.080
	(0.037)	(0.094)	(0.048)	(0.136)	(0.364)
Dummy for alternative scenario #2	0.280	0.469	0.175	0.424	33.910
	(0.028)	(0.071)	(0.037)	(0.103)	(0.000)
Dummy for alternative scenario #3	0.206	0.027	0.051	0.091	6.750
	(0.041)	(0.104)	(0.054)	(0.152)	(0.000)

Note: Standard errors in parenthesis, except in the F-test column where we report p-values.

Figure 9


Demand curves
Engel curves
4.2
4
3.8
3.6
3.4
3.2
30
40
50
60
Price
1500
2000
2500
Income
Income = $1500
Income = $2000
Income = $2500
Price = $30
Price = $45
Price = $45

Preferences

Table 5

Maternal Beliefs about the Technology of Skill Formation

	25th percentile	Median	75th percentile	Mean	Variance
$\alpha_{i,1}$	0.0261	0.0312	0.0400	0.0313	0.0002
	(0.0004)	(0.0002)	(0.0007)	(0.0004)	(0.0000)
$\alpha_{i,2}$	0.0669	0.0777	0.0942	0.0795	0.0003
	(0.0005)	(0.0008)	(0.0007)	(0.0005)	(0.0000)

Note: Standard errors in parenthesis.

Table 7

Comparative Statics of Investments

	Median	75th percentile	% Change in investments	% Change in parameter	Elasticity
α_1	1.70	1.73	1.6%	28.0%	5.8%
α_2	1.70	2.01	18.3%	21.4%	85.2%
$\mu_{\psi,2}$	1.70	1.77	4.1%	72.0%	5.8%
$\mu_{\psi,3}$	1.70	1.70	0.2%	257.1%	0.1%
$\mu_{\psi,3}$	1.70	1.86	9.3%	257.1%	3.6%

This table shows the comparative statics of optimal investments in relation to preference and belief parameters. Each row shows what happens to investments as we move one parameter and fix the other parameters at the median value. In the last row, we replace the human capital at birth from the mean value to the value at the first percentile.

Table 8

Maternal Beliefs and Technology

Cases	Factual investment	Counterfactual investment	% Change	Effect size
$\mu_{\psi,2} = 0.267$ $\mu_{\psi,3} = 0.000$	1.84	1.92	4.4%	10.3%
$\mu_{\psi,2} = 0.454$ $\mu_{\psi,3} = 0.000$	1.84	2.05	11.7%	26.9%

Beliefs and Investments: Anthropology

- !Kung San in the Kalahari desert in Botswana and Namibia (e.g., Lee, 1979) vs. Ache in Paraguay (see Kaplan and Dove, 1987; Hill and Hurtado, 1996).
    - Both groups believe that the development of motor skills by children depends on parental encouragement and teaching.
    - Different environmnents lead both groups to behave in very different ways.
- Gusii in Kenya (see LeVine et al, 1994).

Expectations and Investments: Psychology

- The argument that low subjective expectations about returns may affect investments has been recognized in developmental psychology for over 50 years (Hunt, 1961; Vygostky, 1978).
- Huge empirical literature attempting to estimate what parents know about child developmental milestones (Epstein, 1979; Hess et al., 1980; Ninio, 1988; Mansbach and Greenbaum, 1999).

Epstein (1979)

Expecting too little, too late however is not characteristic of teenagers' knowledge in all areas of development. In fact, when we look at items about basic care, health and nutrition, and perceptual and motor development, we discover that their expectations are quite accurate. By contrast, when we look at how they view infant needs and abilities in the areas of mental development – cognitive, social, and language, it is here that we find teenagers attributing skills to babies many months too late. And, not surprisingly, our analyses show that it is the younger infant who is most likely viewed as a creature of physical needs and growth without corresponding mental activity.

This view of the infant is also evident in teenagers' responses to the videotape measure. Mean ratings indicate that they can neither observe the signs of learning in babies nor recognize the appropriate activities by which adults support this learning.

Beliefs and Investments: Sociology

- Lynd and Lynd (1929, 1937) reported that working-class mothers ranked "strict obedience" as their most important childrearing goal more frequently than higher-SES mothers did. Many studies, conducted in the US in the 1990s or in other developed countries, replicate these findings.
- Kohn (1963) argues that the stronger preferences towards socio-emotional skills by lower-SES mothers reflect those mothers' forecasts for their children choosing occupations in which obedience and conformity have relatively higher returns.
- This finding is also reported in Lareau's ethnographic study *Unequal Childhood*:"Natural Accomplishment of Growth" and "Concerted Cultivation."

Beliefs and Investments: Economics

- Aizer and Stroud (2010) track the smoking habits of educated and non-educated pregnant women before and after the release of the 1964 Surgeon General Report on Smoking and Health.
- Before the release of the report, educated and non-educated pregnant women smoked at roughly the same rates.
- After the report, the smoking habits of educated women decreased immediately, and there was suddenly a ten-percentage point gap between pregnant women who were educated and non-educated in smoking.
- Could the divergence of early investments in the last 20 years be the result of divergence in expectations? We don't know, but it is possible that this is the case.

Discussion

- I presented research in which we aim to formulate a model of human development in which mothers have subjective expectations about a parameter of the technology of skill formation.
- The model is useful to understand how maternal knowledge about the importance of investments in children affect investment choices.
- Large body of literature in many fields suggest that beliefs may play an important role in determining familial investments in children.

Discussion

- At the same time, the literature suggests that these beliefs are endogenous.
- Parents expectations about future occupations of children, or the skills that will be most important for their survival, determine parental beliefs about what skills children should learn, and what skills they believe are malleable.
- So, if correct, this framework suggests that it may be difficult to change parental beliefs.
- At the same time, research in economics shows that most educated parents react to information that improves children's health.
- And some home visitation programs have been very successful in positively affecting children's health (but not all).
- So, future research should aim to understand the process of belief formation.