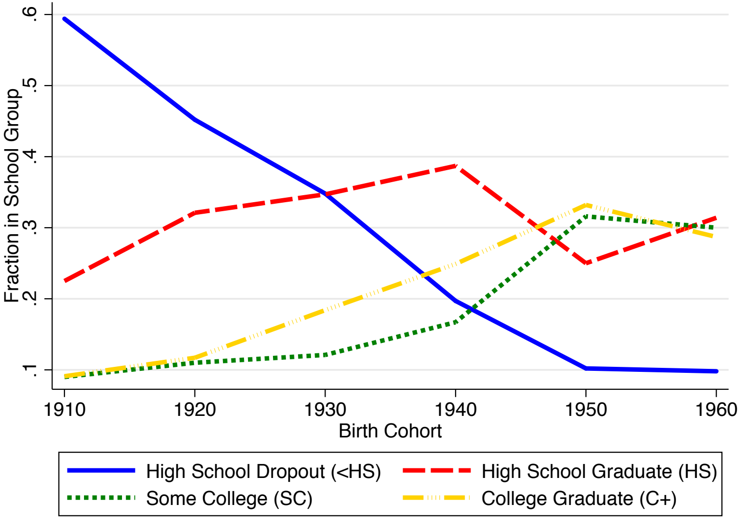
Student Abilities During the Expansion of US Education

Lutz Hendricks &Todd Schoellman

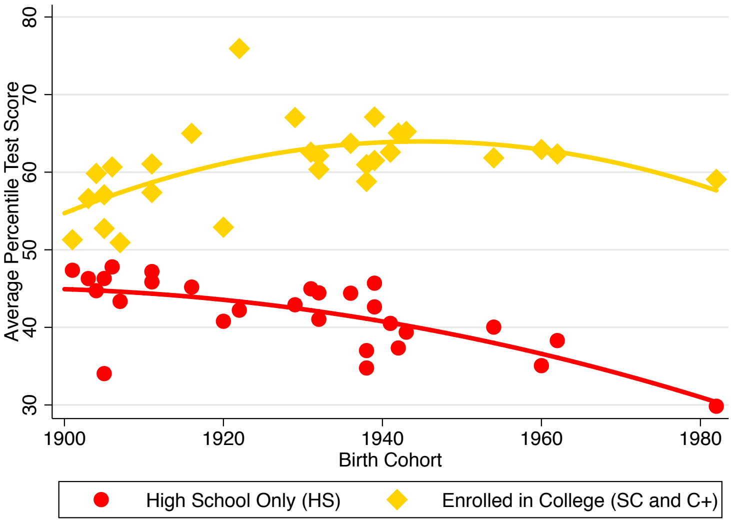
University of North Carolina, Chapel Hill & Arizona State University

September 23, 2015

Motivation: The Expansion of Education



Motivation: Changes in College Enrollment



Our Idea: Composition Effects and Wages

Our idea:
Expansion, Change $\Longrightarrow$ Composition Effects $\Longrightarrow$ Wages

Focus on college wage premium:

1. Large: 0.52 for 1960 cohort
2. Rising: 0.15 between 1910 and 1960 cohorts

Think of average wages in three parts:

1. Ability
2. Human capital
3. Skill prices

Challenge and Our Approach

- Challenge: We observe only wages
  - Not three components separately
- Our Approach: Cognitive Test Scores
  - Noisy but informative signal of ability
  - Quantify the role of ability
  - Infer remaining terms as a residual
- Results: ability explains:
  1. Around half of the college wage premium
  2. *All* of the rise in college wage premium
  3. Also, a modest slowdown in wage growth

Literature

1. Ability biases and wage premiums.
   - Card (2001); Heckman, Lochner, and Todd (2006)
2. Using cognitive test scores to learn about ability
   - Heckman, Lochner, and Taber (1998); Garriga and Keightley (2007)
3. Composition effects
   - Finch (1946); Taubman and Wales (1972); Laitner (2000)

Outline

Two Goals for the Model

1. Simple formalization of problem
   - Two key ingredients: ability heterogeneity, imperfect sorting
   - Two key parameters
   - Wages with three components
2. Tool for measurement
   - Show how cognitive test scores can be informative
   - Fit model to data, back out ability

Model: Demographics and Endowments

Discrete time, overlapping generations environment

- Cohort τ, age v, die at T

Endowments: ability and tastes

- Cognitive ability $a \sim \mathcal{N}(0, 1)$
- Taste for schooling $p \sim \mathcal{N}(0, \sigma_p)$
- Endowments are iid and uncorrelated

Summarize type by $q = (\tau, a, p)$.

Model: Preferences and Budget Constraint

- Preferences:

$$\sum_{v=1}^{T} \beta^{v} \log[c(q,v)] - \exp\left[-(p+a)\right] \chi(s,\tau)$$

  - $\chi(s,\tau) > 0$, increasing in $s \rightarrow$ complementarity

- Budget Constraint:

$$\sum_{v=1}^{T} \frac{c(q,v)}{R^{v}} = \sum_{v=T(s)+1}^{T} \frac{w(s,q,v)}{R^{v}}$$

Model: Wages

Individual wages:

$$\log[w(s,q,v)] = \theta a + z(s,\tau + v - 1) + h(s,v) + \varepsilon_w$$

Mean wages, conditional on observables:

$$E\left[\log(w)|s,\tau,v=40\right] = \underbrace{\theta E(a|s,\tau)}_{\text{Effective Ability}} + \underbrace{h(s,v) + z(s,\tau+v-1)}_{\text{Residual}}$$

Assumptions:

- ε_w, h independent of a

Model: Key Properties

Perfect sorting by $p + a$. Intuition:

- a affects benefits, opportunity cost of schooling equally
- Preferences determine school choice

Captures two key ingredients:

- Heterogeneous ability affects wages (θ)
- Imperfect sorting by ability into schooling (σ_p)

Outline

Calibration: NLSY79 and Cognitive Test Scores

Our calibration draws heavily on NLSY79

- Nationally representative
- Joint distribution of wages, school attainment, and test score (AFQT)
- 1957–1964 cohorts: pool, treat as 1960

Use AFQT as a proxy for ability

- Effect of AFQT on wages pins down θ
- Sorting of school groups by AFQT pins down σ_p
- Other parameters matter less or not at all.

Interpreting Test Scores

We observe test scores, not ability. Interpretation:

$$\hat{a} = \eta \left(a + \tilde{\varepsilon}_{\hat{a}}\right)$$
$$\tilde{\varepsilon}_{\hat{a}} \sim \mathcal{N}(0, \sigma_{\hat{a}})$$

To undo scaling effect, we standard normalize:

$$\hat{a} = \frac{a}{\sqrt{1 + \sigma_{\hat{a}}^2}} + \varepsilon_{\hat{a}}$$
$$\varepsilon_{\hat{a}} \sim \mathcal{N}\left(0, \frac{\sigma_{\hat{a}}}{\sqrt{1 + \sigma_{\hat{a}}^2}}\right)$$

How Test Scores Are Informative: A Special Case

Let $\sigma_{\hat{a}} = 0$.

- $\implies \hat{a} = a$

True wage generating process:

$$\log[w(s, q, v)] = \theta a + z(s, \tau + v - 1) + h(s, v) + \varepsilon_w$$

Our empirical regression:

$$\log(w) = \beta_{\hat{a}} \hat{a} + \sum_s \gamma_s d_s + \varepsilon_w$$

- Implementation yields $\beta_{\hat{a}} = \theta$

Returns to AFQT in the NLSY79

	Log-Wages
$\beta_{\hat{a}}$	0.104
	(0.017)
γ_{HS}	0.17
	(0.06)
γ_{SC}	0.35
	(0.06)
γ_{C+}	0.69
	(0.07)
Observations	1942
R^2	0.24

Implication: $\theta = 0.104$

Joint Distribution of AFQT & Schooling in the NLSY79

	AFQT Quartile			
School Attainment	1	2	3	4
$<$HS	86%	12%	2%	0%
HS	42%	34%	19%	5%
SC	18%	32%	31%	19%
C+	1%	11%	29%	59%

Population strongly sorted by test score ($= a$).

Results When Test Scores Are Noiseless: Direct Calculation

School Comparison	Effective Ability Gap Calculation	Wage Gap Data
<HS-HS	-0.08	-0.24
SC-HS	0.06	0.18
C+-HS	0.14	0.52

Result: one-quarter to one-third of wage gap is due to ability

- Wages measured at age 35–44, from Census
  - Census: for consistency

Results When Test Scores are Noiseless: Model

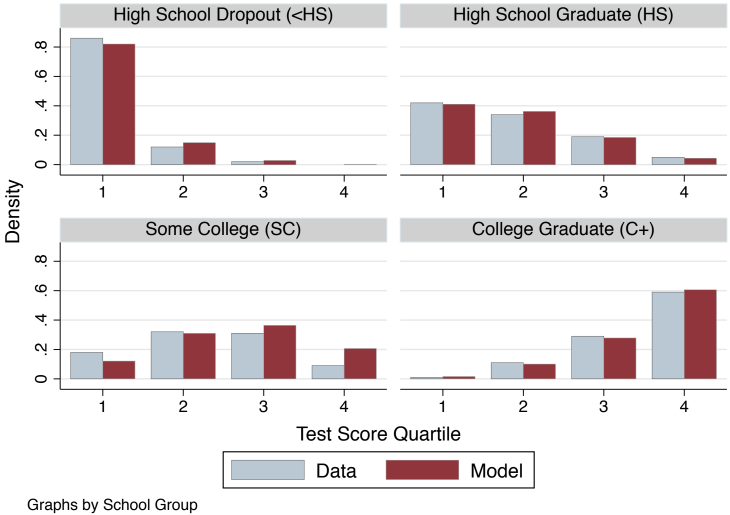
Alternative: calibrate model

- $\theta = 0.104$
- Choose σ_p to match sorting

School	Contribution of Composition Effects		
Comparison	Calculation	Model	Data
$<$HS-HS	-0.08	-0.08	-0.24
SC-HS	0.06	0.07	0.18
C+-HS	0.14	0.15	0.52

Direct calculation and model give same result.

Fit: Single Parameter (σ_p) Replicates Sorting Well



Model Contribution 1: Interpreting Noisy Test Scores

What if $\sigma_{\hat{a}} > 0$?

- Regressing wages on test scores suffers from attenuation bias
- $\beta_{\hat{a}} < \theta$
- Larger results

Noise in test scores is unknown, but can be bounded.

Bounding the Noise in Test Scores

Lower bound:

- Test scores are not perfectly reliable.
- $corr(\hat{a}_1, \hat{a}_2) \approx 0.8 < 1$
- $\implies \sigma_{\hat{a}} \geq 0.5$

Upper bound:

- Large $\sigma_{\hat{a}} \implies$ large $\theta \implies$ large results
- Results should not be implausible
- Mean ability gaps = Mean wage gaps

Iterative Calibration

Iterate over $\sigma_{\hat{a}}$

- Given $\sigma_{\hat{a}}$, we can calibrate our model
    - Calibrate θ to match $\beta_{\hat{a}} = 0.104$, given attenuation bias.
    - Calibrate σ_p to match test score-school sorting, given noise.

Increase $\sigma_{\hat{a}}$ from 0.5 until upper bound is identified

Results when Test Score is Noisy

	Data	Contribution of Composition Effects	
		Model Without Noise	Model With Noise
$<$HS	-0.24	-0.08	-0.14 – -0.22
SC	0.18	0.07	0.11 – 0.18
C+	0.52	0.15	0.25 – 0.39

Result: at least 44% of wage gaps due to ability gaps

- Helps resolve Heckman, Lochner, and Todd puzzle
- Compare to Bowlus & Robinson (2008): $\approx$ all
- Upper bound not yet precise

Outline

Model Contribution 2: Time Series Exercise

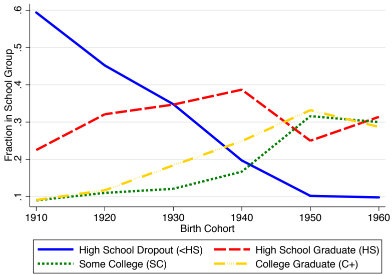
Iterate over $\sigma_{\hat{a}}$

- Given $\sigma_{\hat{a}}$, we can calibrate our model
  - Calibrate θ to match $\beta_{\hat{a}} = 0.104$, given attenuation bias.
  - Calibrate $\sigma_{p,\tau}$ to match sorting from Taubman & Wales (1972)
  - Calibrate $\chi_{s,\tau}$ to fit expansion of schooling from census

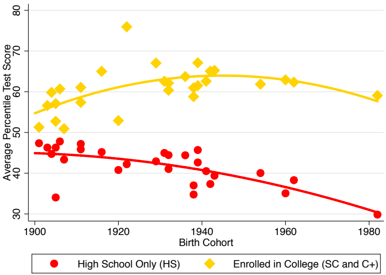
Increase $\sigma_{\hat{a}}$ from 0.5 until new upper bound is identified

Show results for $\sigma_{\hat{a}} = 0.5$ to develop intuition; then show range.

Time Series Calibration Strategy



(a) Use $\chi_{s,\tau}$ to Fit Expansion



(b) Use $\sigma_{p,\tau}$ to Fit Sorting

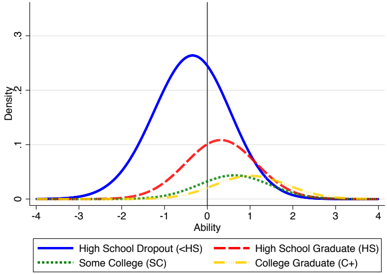
Other Calibrated Parameters if $\sigma_{\hat{a}} = 0.5$

Parameter	Value
Noise in Test Scores ($\sigma_{\hat{a}}$)	0.50
Effect of Ability on Wages (θ)	0.155
Dispersion of Preferences, 1960 Cohort ($\sigma_{p,1960}$)	0.62
Dispersion of Preferences, 1950 Cohort ($\sigma_{p,1950}$)	0.80
Dispersion of Preferences, 1940 Cohort ($\sigma_{p,1940}$)	1.12
Dispersion of Preferences, 1930 Cohort ($\sigma_{p,1930}$)	1.10
Dispersion of Preferences, 1920 Cohort ($\sigma_{p,1920}$)	1.28
Dispersion of Preferences, 1910 Cohort ($\sigma_{p,1910}$)	1.44

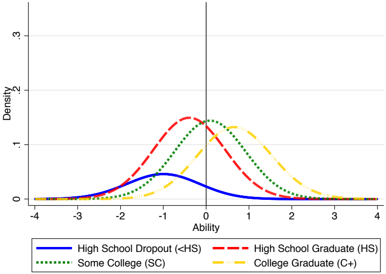
Variation in *a* explains 32% of school choice in 1910 cohort

- $\rightarrow$72% of school choice in 1960 cohort

The Evolution of Abilities: Expansion Effect

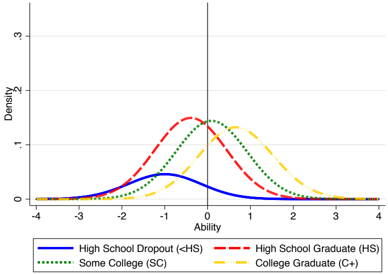


(c) 1910 Cohort (Baseline)

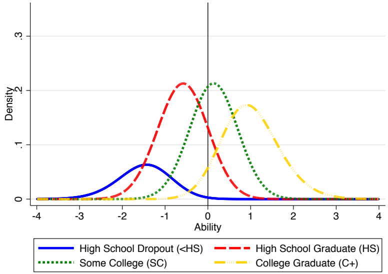


(d) 1960 Cohort, Constant Sorting

The Evolution of Abilities: Sorting Effect



(e) 1960 Cohort, Constant Sorting



(f) 1960 Cohort, Increased Sorting (Baseline)

Wage Growth Results, 1910–1960 Cohorts

		Model-Implied Decomposition	
	Data	Composition Effects	$h + z$
$<$HS	0.22	-0.17	0.40
HS	0.29	-0.14	0.42
SC	0.30	-0.08	0.38
$C+$	0.43	0.00	0.44

Generally: declining mean ability $\implies$ depressed wages

- Except college graduates

Wage Premium Growth Results, 1910–1960 Cohorts

	Data	Model-Implied Decomposition	
		Composition Effects	Relative $z + h$ Growth
$<$HS	-0.06	-0.03	-0.03
SC	0.02	0.06	-0.04
$C+$	0.15	0.14	0.01

College ability rises relative to high school

- Bowlus & Robinson (2008): 72% of rise comes from quantity, 1980–1995

Range of Results for Composition Effects

Statistic	Data	Contribution of Composition Effects	
		No Noise	Noise (Baseline)
1960 C+ Premium	0.52	0.15	0.25 – 0.37
Δ HS Wage	0.29	-0.08	-0.14 – -0.21
Δ C+ Prem.	0.15	0.08	0.14 – 0.20

Composition effects explain:

1. Half of college wage premium;
2. A wage slowdown of one-third;
3. The entire rise in the college wage premium

Outline

Robustness and Decomposition Exercises

What drives results:

1. Improvement in sorting: decomposition/robustness
   - Perhaps early evidence is untrustworthy
2. Dispersion of abilities: robustness
   - Perhaps $\beta_{\hat{a}}$ is lower $\implies$ lower θ.
3. Strong sorting in recent cohorts

Additional:

- Changes in ability distribution?

Range of Results when Sorting is Constant

Statistic	Data	Contribution of Composition Effects	
		Baseline	Constant Sorting
1960 C+ Premium	0.52	0.25 – 0.37	0.25 – 0.27
Δ HS Wage	0.29	-0.14 – -0.21	-0.16 – -0.18
Δ C+ Prem.	0.15	0.14 – 0.20	0.08 – 0.09

Range of Results when $\beta_{\hat{a}} = 0.07$

Statistic	Data	Contribution of Composition Effects	
		Baseline	Lower Return
1960 $C+$ Premium	0.52	0.25 – 0.37	0.17 – 0.37
Δ HS Wage	0.29	-0.14 – -0.21	-0.09 – -0.21
Δ C+ Prem.	0.15	0.14 – 0.20	0.09 – 0.20

WIth lower return to schooling:

- Lower bound allows for results one-third lower, across the board

Results with Flynn Effect

Flynn (1984,2009): general rise in IQs

- US: 1.2–1.8 standard deviations during our time frame
- Weak consensus: symmetric
- No consensus: why

Thought experiment: what if Flynn Effect is real ability gain?

- No role for relative ability, wage premiums
- Offsets much of ability level, wage declines

Outline

Conclusion

I won’t make it to this slide, anyway.