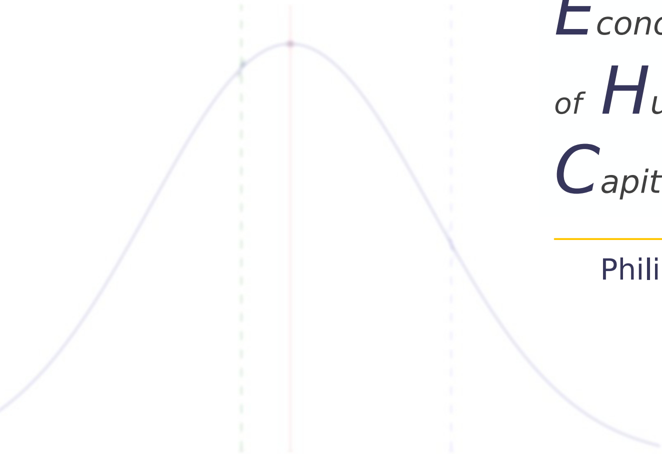


Economics of Human Capital

Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

Economics of Human Capital

Introduction

Philipp Eisenhauer

Human capital is defined as:

> *The knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being.*
>
> \- OECD (2001)

Tasks

- definition and measurement of human capital
- determining the effect of human capital on a variety of personal, social and economic outcomes
- understanding the formation of human capital

Tasks

- identifying the driving forces behind the observed heterogeneity across and within countries
- search for effective policies to ameliorate disparities
- ...

Facts

Figure: Years of schooling

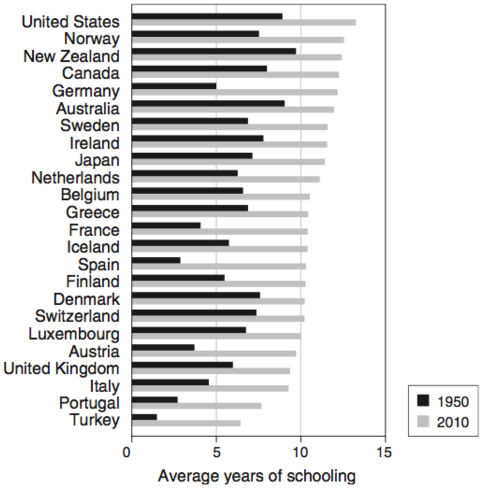


FIGURE 4.4
Years of schooling of the total population aged 25 and older.

Source: Barro and Lee (2010, education data set, available at www.barrolee.com/data).

Figure: Unemployment rates

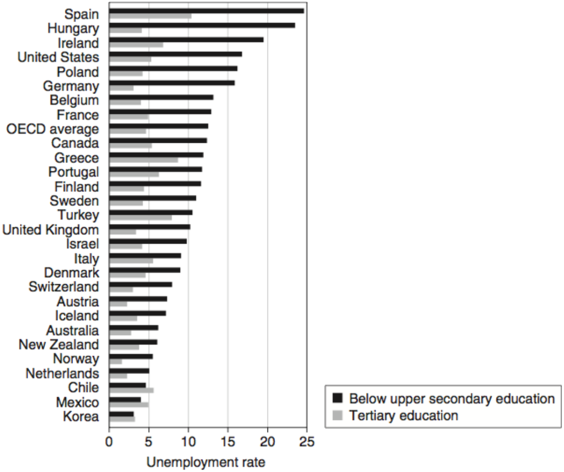


FIGURE 4.6

Unemployment rates by level of educational attainment for 25- to 64-year-olds, 2010. The OECD average is the nonweighted average of the 34 OECD countries, including those not represented on this figure. Data missing for non-OECD countries.

Source: OECD (2012, table A7.4a, p. 133).

Figure: Tertiary education

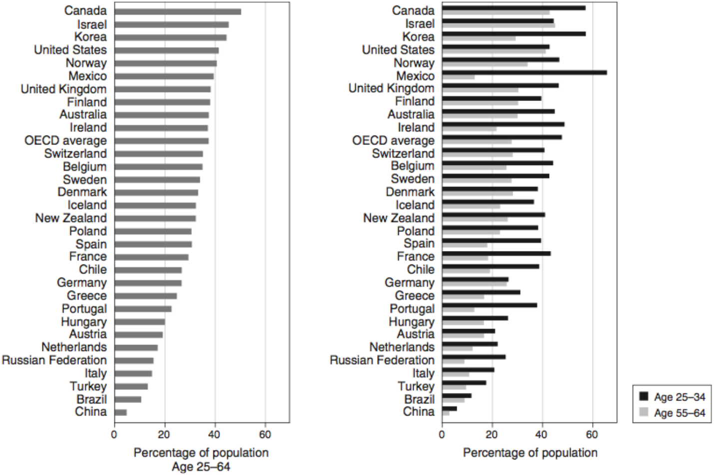


FIGURE 4.3
Percentage of the population that has attained at least tertiary education or advanced research programs, by age group, 2010. The OECD average is the nonweighted average of the 34 OECD countries, including those not represented in this figure. Brazil, China, and the Russian Federation are not part of the OECD.

Source: OECD (2012, table A1.3a, p. 36).

Figure: Secondary education

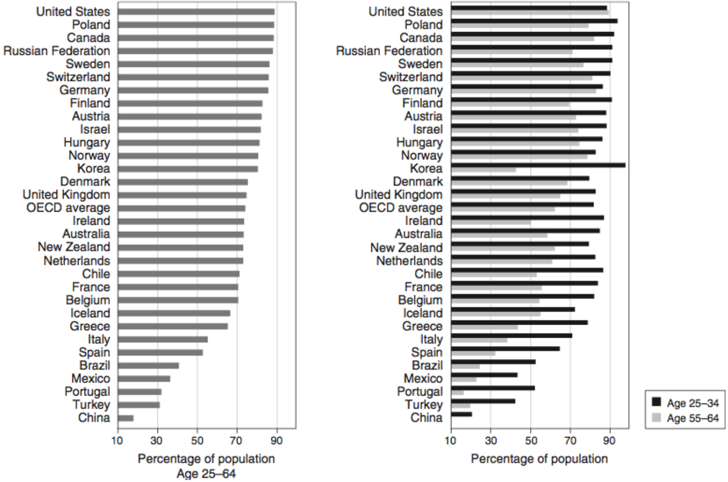


FIGURE 4.2

Percentage of the population that has attained at least upper secondary education, by age group, 2010. The OECD average is the nonweighted average of the 34 OECD countries, including those not represented in this figure. Brazil, China, and the Russian Federation are not part of the OECD.

Source: OECD (2012, table A1.2a, p. 35).

Figure: Expenditures

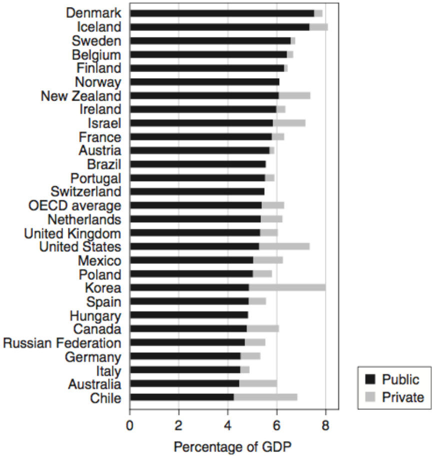


FIGURE 4.1
Expenditure on educational institutions as a percentage of GDP, 2009. The OECD average is the nonweighted average of the 34 OECD countries, including those not represented in this figure. Brazil and the Russian Federation are not part of the OECD. Private expenditure is missing for Brazil, Hungary, Norway, and Switzerland. Data are missing for China, Greece and Turkey.

Source: OECD (2012, table B2.3, p. 246).

Figure: Relative earnings

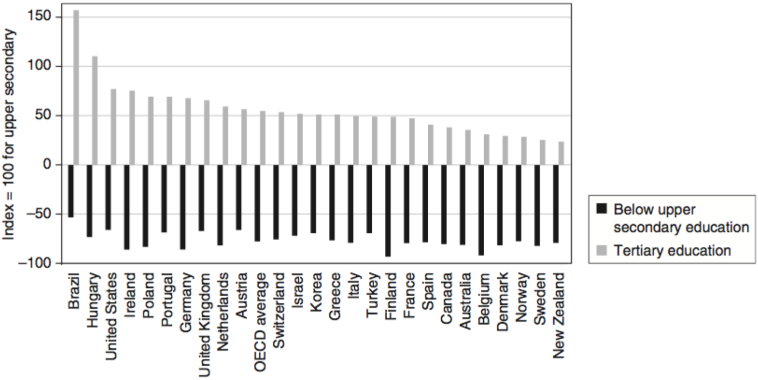


FIGURE 4.5

Relative earnings from employment among 25- to 64-year-olds, by level of educational attainment (2010 or latest available year). Upper secondary and post-secondary nontertiary education = 100. The OECD average is the nonweighted average of the 34 OECD countries, including those not represented in this figure. Brazil is not part of the OECD. Data missing for Chile, China, Iceland, Mexico, and the Russian Federation.

Source: OECD (2012, chart A8.1, p. 140).

Figure: Unemployment rates

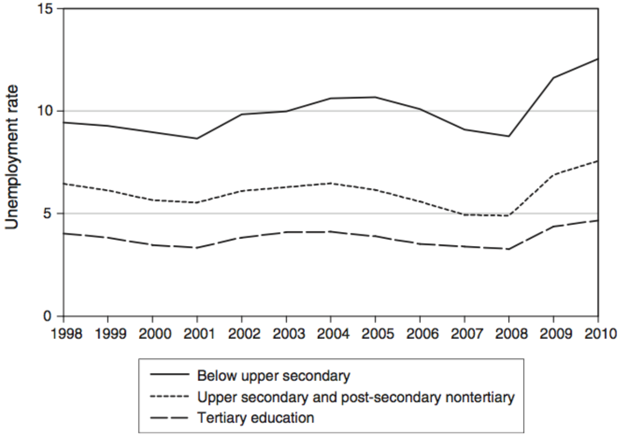


FIGURE 4.7

Unemployment rates by level of educational attainment for 25- to 64-year-olds, 2010. The OECD average is the nonweighted average of the 34 OECD countries.

Source: OECD (2012, table A7.4a, p. 133).

Life-cycle of earnings

Stylized Facts

- Life-cycle earnings are increasing at early ages and decline towards the end.
- Wages tend to increase over the life-cycle with a weak tendency to decline at the end of working life.
- Hours of work increase at early ages and decline in old age, with the peak occurring earlier than in the wage profiles.

See Weiss (1986) for comprehensive modeling framework that allows to interpret all these facts.

Figure: Wage gains

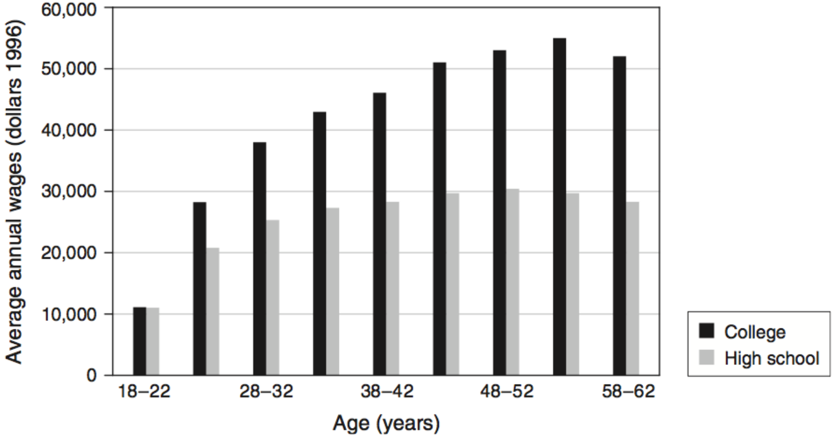


FIGURE 4.8
Average wage gains for college and high school graduates in the United States in 1996.

Source: Ashenfelter and Rouse (1999).

We study a version of the seminal Ben-Porath Model (Ben-Porath, 1967) that relates human capital accumulation to life-cycle earnings.

Why do economists use mathematical models? (slides)

This material is best studied using the following resource.

- Cahuc, P., & Zylberberg, A. (2004). *Labor economics* (1st ed.). Cambridge, MA: MIT Press.

Basic Notation

$s(t)$	fraction devoted to training
$h(t)$	stock of human capital
$w(t)$	wage
δ	depreciation of knowledge

The individual's objective is to maximize the discounted sum of wages over their life-cycle.

$$\Omega = \int_0^T w(t)\, e^{-rt} dt$$

Their economic environment is characterized by the production functions for wages and human capital.

$$
\begin{aligned}
w(t) &= A[1 - s(t)]h(t) \\
\dot{h} &= \theta g(s(t), h(t)) - \delta h(t) \qquad g' > 0, g'' < 0
\end{aligned}
$$

Notable Features

- Individuals cannot work and learn at the same time.
- There is no individual heterogeneity.
- There is no direct cost of education but there are the opportunity cost of lost wages.
- ...

Model Specification

We study the implementation in Cahuc and Zylberberg (2004).

$$g(h(t), s(t) = (h(t)s(t))^{0.71}$$

$$A = 0.75 \qquad \delta = 0.06 \qquad r = 0.05$$

$$h_0 = 5 \qquad T = 60 \qquad \theta = 0.5$$

Figure: Contour plot of human capital production function

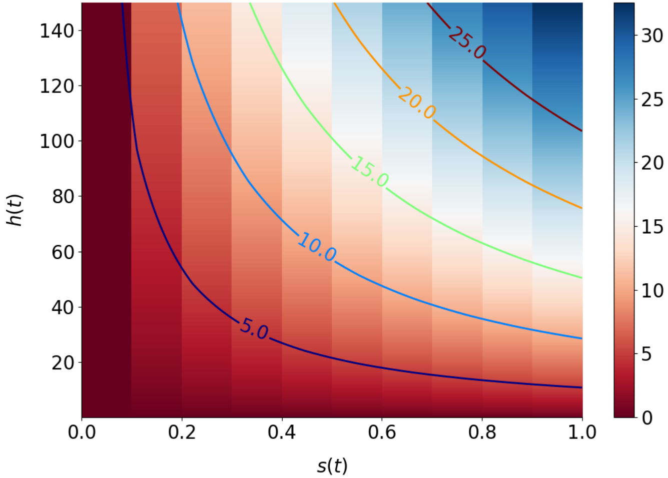


Figure: Surface plot of human capital production function

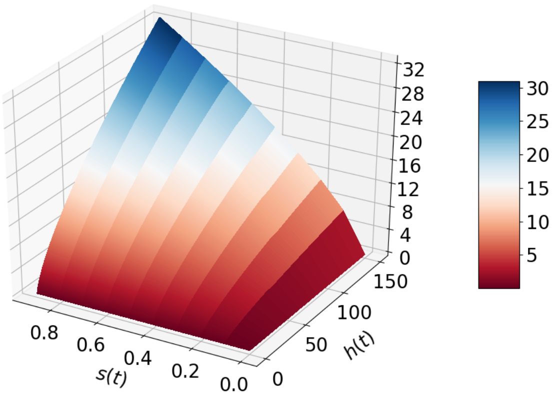


Figure: Wage production

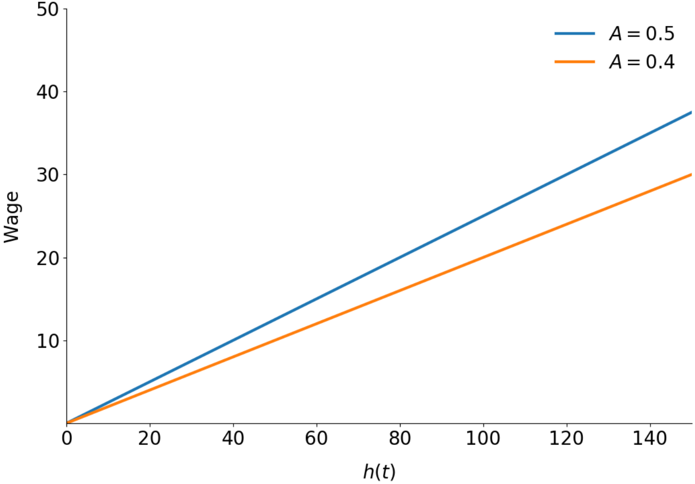


Figure: Wage $w(t)$ over the life-cycle

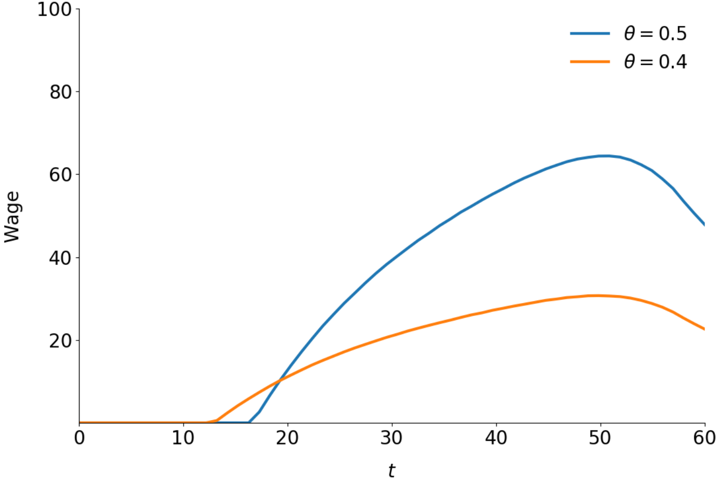


Figure: Stock of human capital $h(t)$ over the life-cycle

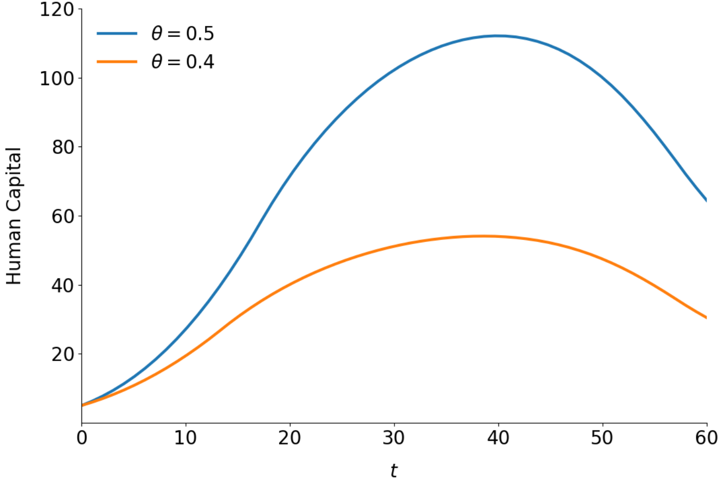


Figure: Human capital investment $s(t)$ over the life-cycle

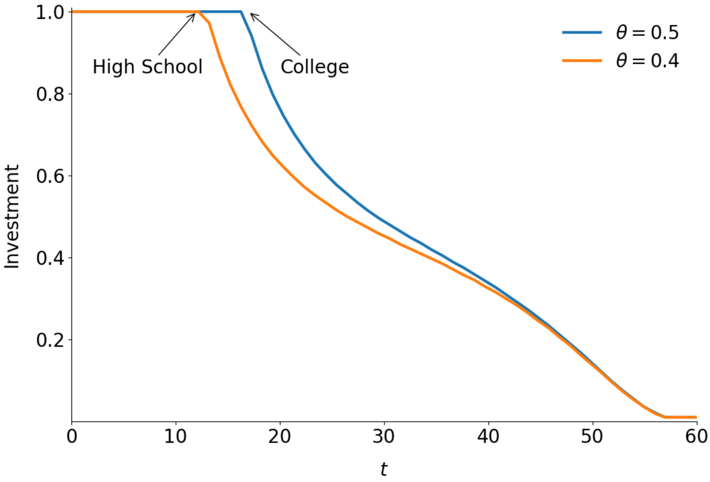
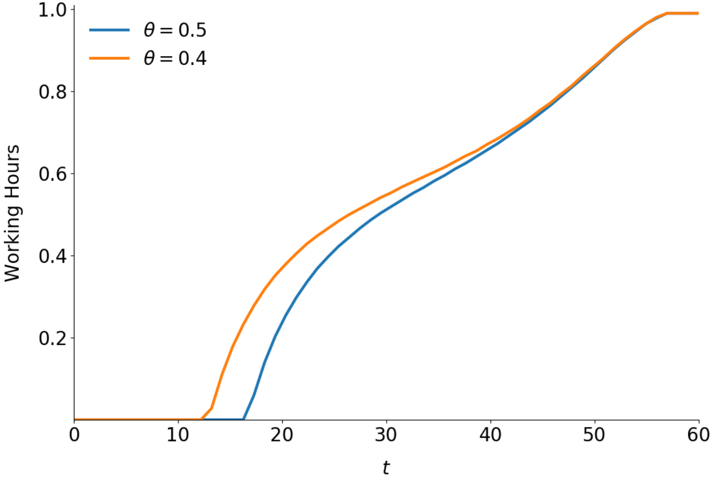


Figure: Hours worked $(1-s(t))$ over the life-cycle



How well does the model do?

Extensions

Weiss (1986) reviews a host of alternative extensions to the basic model.

- general versus specific training
- hours worked
- uncertainty
- borrowing-constraints
- ...

Job market signaling

This material is best studied using the following resource.

- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355-374.

We study the seminal model presented in Spence (1973).

- There are two groups $j \in \{H, L\}$ in the population facing one employer, where $h_{j\in\{L,H\}}$ denotes the respective level of productivity.

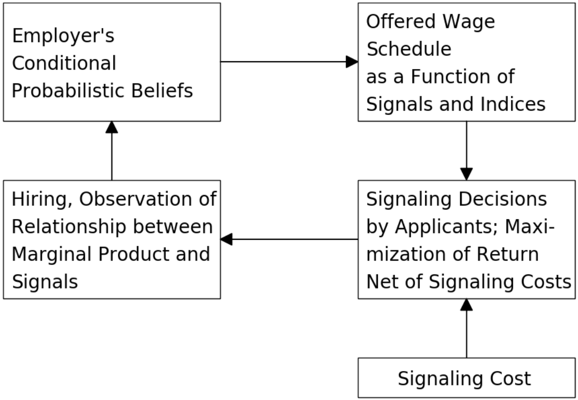
- Group L is a proportion q_L in the population.

- Education y is measured by an index y of level and achievement and is subject to individual choice.

- Education costs are both monetary and psychic and differ by group $c_{j\in\{L,H\}}$.

- The productivity type is unobservable by the employer.
- Individual decisions about y can provide a signal about the underlying productivity.
- Wage schedule is set conditional on y.

Figure: Informational feedback



We explore the following parameterized version.

$$h_L = 1 \qquad h_H = 2$$

$$c_L = y \qquad c_H = \frac{1}{2}y$$

Thus there is an inverse relationship between productivity and signaling cost.

Figure: Benefit of education

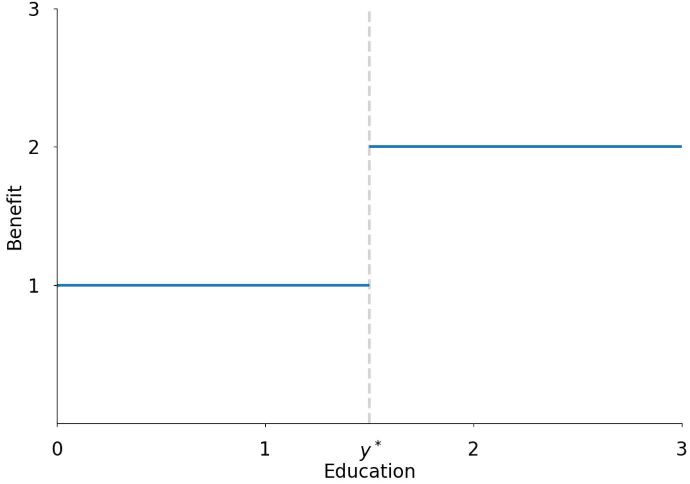


Figure: Cost of education

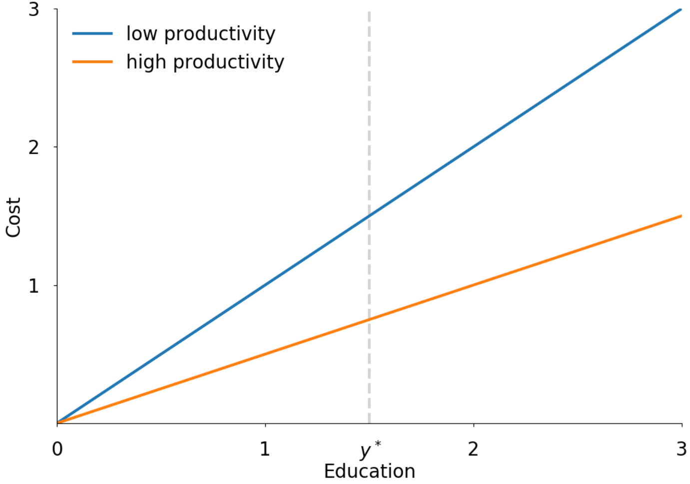
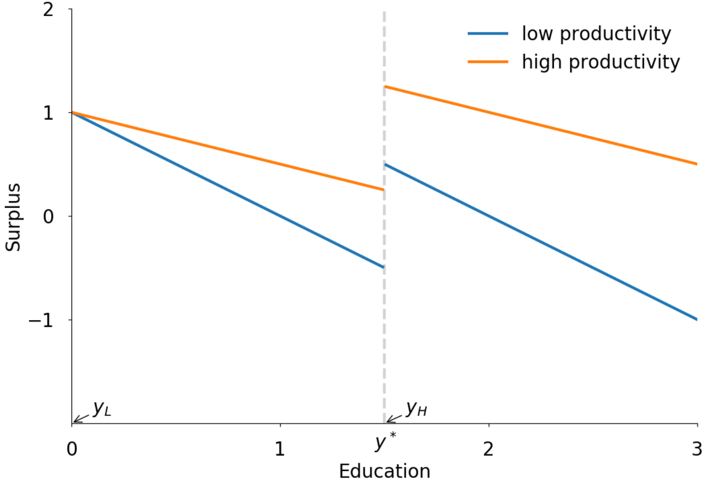


Figure: Surplus of education I



- For $y^* = 1.5$ the employer's beliefs are confirmed. More generally, L chooses $y_L = 0$ if $1 > 2 - y^*$ and H acquires $y_H = y^*$ provided that $2 - 0.5\,y^* > 1$.
- Beliefs are confirmed provided that the following holds:

$$1 < y^* < 2$$

Figure: Surplus of education II

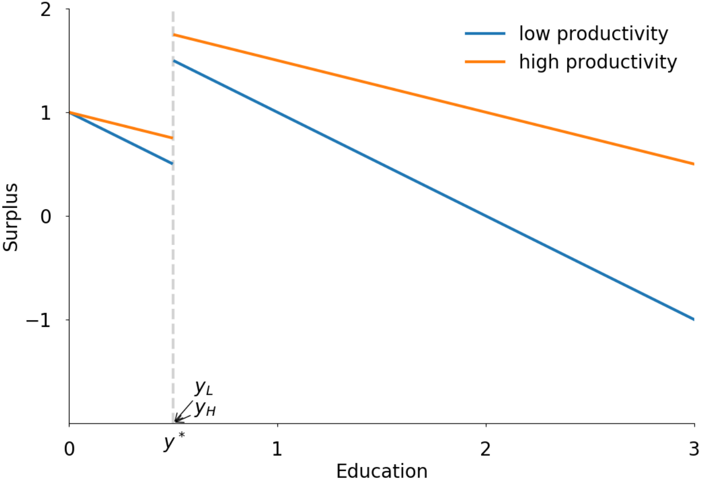
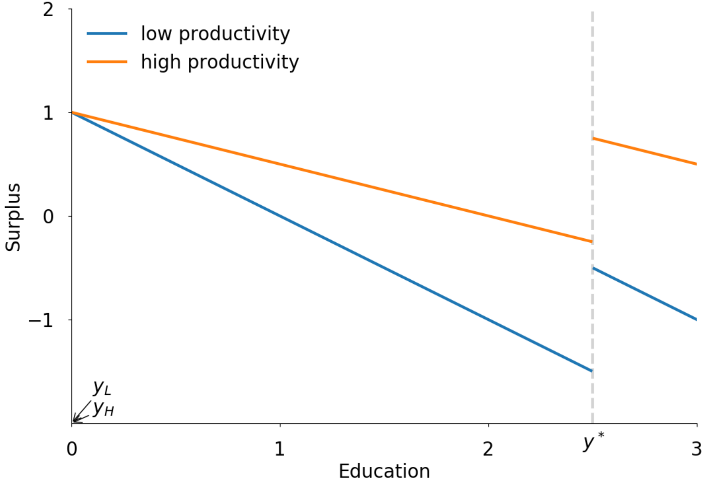


Figure: Surplus of education III



- From the outside, education appears to be productive and is for the individual. However, there is no real effect on the marginal product.

Can we distinguish between the two models based on simple information on individual education and wages?

- In the absence of signaling, both groups are paid the unconditional expected marginal product.

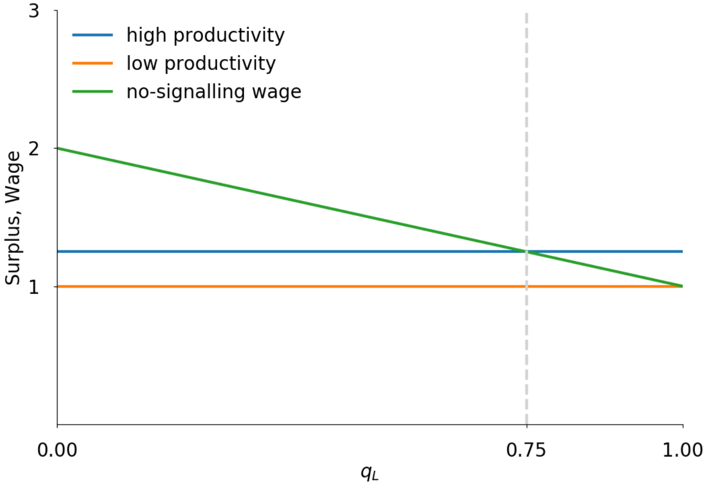
$$q_L \times 1 + (1 - q_L) \times 2$$

- It depends on the share of low productivity individuals whether high productivity individuals actually prefer a no-signaling case. Their surplus is determined as follows:

signaling	$2-\frac{1}{2}y^*$
no-signaling	$2-q_L$

- High productivity individuals prefer the signaling case as long as $y^* \leq 2q_L$.

Figure: Market structure



- The ability to signal has a detrimental effect on low productivity workers, while the consequences are ambiguous for high productivity workers.
- High productivity workers benefit from their ability to send a signal if their proportion is sufficiently small with respect to the productivity gap to low productivity individuals.

Dataset

I present a basic overview on the **National Longitudinal Survey of Youth 1979 (NLSY79)** dataset (Bureau of Labor Statistics, 2014). The slide deck is under constant development and available at the link below.

http://bit.ly/2JeEGGt

Appendix

References

Becker, G. S. (1964). *Human capital* (1st ed.). New York City, NY: Columbia University Press.

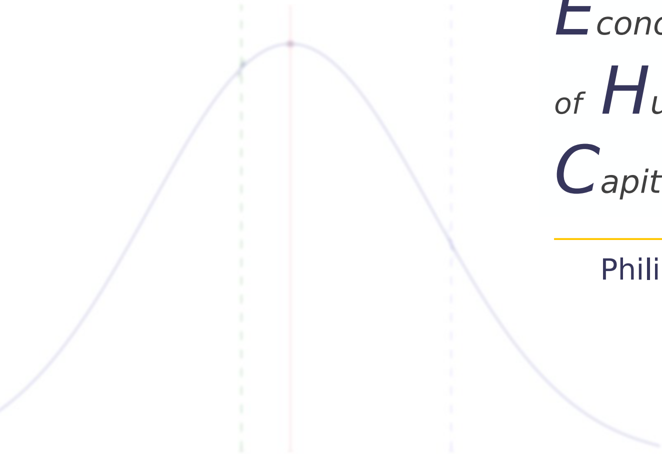
Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, *75*(4), 352–365.

Bureau of Labor Statistics. (2014). *National longitudinal survey of youth 1979 cohort, 1979-2012 (rounds 1-25)*. Columbus, OH: Produced and distributed by the Center for Human Resource Research, The Ohio State University.

Cahuc, P., & Zylberberg, A. (2004). *Labor economics* (1st ed.). Cambridge, MA: MIT Press.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355-374.

Weiss, Y. (1986). The determination of life cycle earnings: A survey. In O. C. Ashenfelter & R. Layard (Eds.), *Handbook of labor economics* (Vol. 1, pp. 603–640). Amsterdam, Netherlands: North-Holland Publishing Company.



Economics of Human Capital

Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

Economics of Human Capital

Returns to schooling

Philipp Eisenhauer

Introduction

I heavily draw on the material presented in:

- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (1st ed., Vol. 1, pp. 307–458). Amsterdam, Netherlands: North-Holland Publishing Company.

Importance of returns

- explain wage inequality within countries
- explain growth differentials across countries
- assess schooling investment on individual level
- evaluate public policies to foster educational attainment
- ...

Core parameter

The internal rate of return is the discount rate that equates the present value of two potential income streams (Becker, 1964).

Different return concepts

- Mincer rate of return
- internal rate of return
- true rate of return

⇒ We will also distinguish between ex ante and ex post returns when introducing uncertainty and the sequential revelation of uncertainty.

Mincer rate of return

Mincer Equation

$$\ln Y(s, x) = \alpha + \rho_s s + \beta_0 x + \beta_1 x^2 + \epsilon$$

$\Rightarrow$ How to interpret the *Mincer Coefficient* ρ_s?

Key features

- linear and homogeneous returns to schooling
- additive separability

Conceptual Frameworks

- compensating differences
- accounting-identity

Compensating Differences Model

Let $Y(s)$ represent the annual earnings of an individual with s years of education, assumed to be constant over his lifetime. Let r be an externally determined interest rate and T the length of working life, assumed not to depend on s. The present value of earnings associated with schooling level s is

$$V(s) = Y(s)\int_s^T e^{-rt}dt = \frac{Y(s)}{r}(e^{-rs} - e^{-rT}).$$

Note that one's working life can either be spend in school or in the labor market.

Figure: Earnings

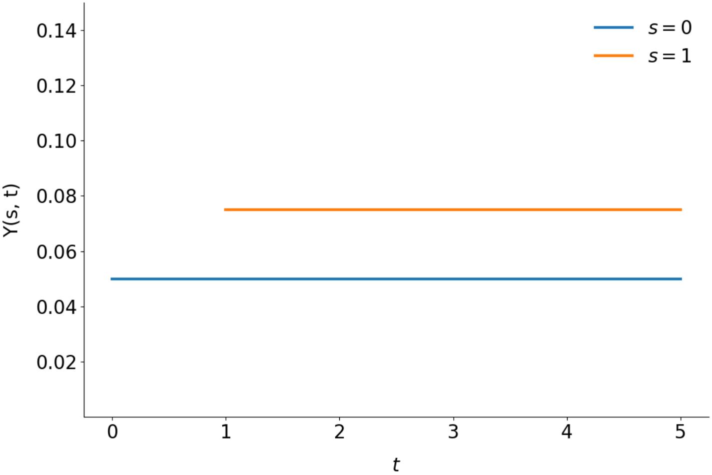
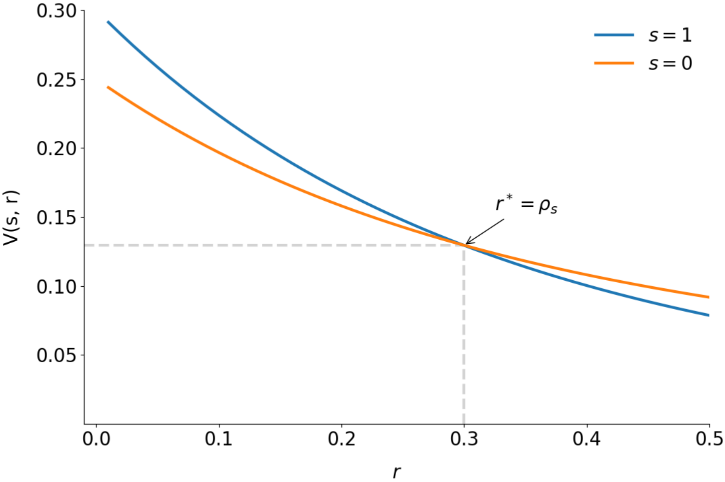


Figure: Value



Equilibrium across heterogeneous schooling levels requires that individuals be indifferent between schooling choices with $V(s) - V(0) = 0$. Equating earnings streams across schooling levels yields:

$$\frac{Y(s)}{r}\left(e^{-rs} - e^{-rT}\right) - \frac{Y(0)}{r}\left(1 - e^{-rT}\right) = 0$$

$$Y(s) = Y(0)\left(\frac{1 - e^{-rT}}{e^{-rs} - e^{-rT}}\right)$$

$$Y(s) = Y(0)\left(\frac{1}{e^{-rs}}\right)\left(\frac{1 - e^{-rT}}{1 - e^{-r(T-s)}}\right)$$

Taking the natural logarithm:

$$\ln Y(s) = \ln Y(0) + rs + \ln\left(\frac{1-e^{-rT}}{1-e^{-r(T-s)}}\right)$$

$\Rightarrow$ ρ_s equals the market interest rate and the internal rate of return to schooling by construction (as T gets large)

Model features

- identical abilities and opportunities
- no credit constraints
- perfect certainty
- no direct cost of schooling
- no nonpecuniary benefits of school and work

Accounting-Identity Model

Model ingredients

P_t	potential earnings at t
$C_t = k_t P_t$	investment cost of training at t
ρ_t	average return to investment at t
κ	decline in post-school investment with T

$$P_t \equiv P_{t-1}(1 + k_{t-1}\rho_{t-1}) \equiv \prod_{j=0}^{t-1}(1 + \rho_j k_j)P_0$$

Formal schooling is defined as years spent in full-time investment ($k_t = 1$), which is assumed to take place at the beginning of life and to yield a rate of return ρ_s that is constant across all years of schooling.

$$\begin{aligned}
\ln P_t &\equiv \ln P_0 + s \ln\left(1 + \rho_s\right) + \sum_{j=s}^{t-1} \ln\left(1 + \rho_0 k_j\right) \\
&\approx \ln P_0 + s\rho_s + \rho_0 \sum_{j=s}^{t-1} k_j
\end{aligned}$$

Note that we also assume that the return to post-school investment is constant over ages and equals ρ_0.

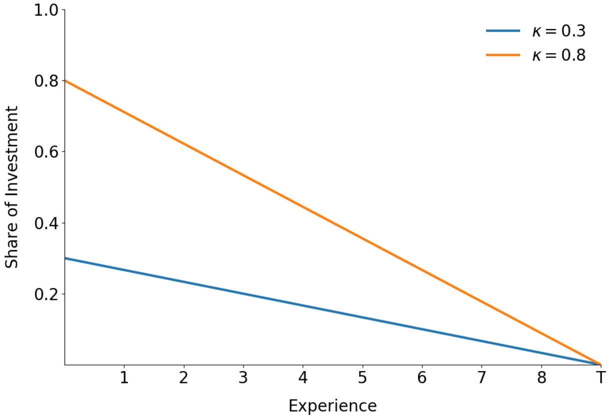
Mincer (1974) assumes a linearly declining rate of post-school investment:

$$k_{s+x} = \kappa\left(1 - x/T\right), \text{where } x = t - s$$

Thus,

$$\ln P_{x+s} = \ln P_0 + s\rho_s + \rho_0 \sum_{j=0}^{x-1} \kappa\left(1 - j/T\right).$$

Figure: Post-School Investment



The derivations draws on the following results for arithmetic series (Chapman & Hall, 2018).

$$\sum_{i=0}^{n} i = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

We need to further decompose the experience addition and use results on arithmetic series.

$$\begin{aligned}
\rho_0 \sum_{j=0}^{x-1} \kappa\left(1-j/T\right) &= \rho_0 \sum_{j=0}^{x-1} \kappa - \rho_0 \kappa \sum_{j=0}^{x-1} \left(j/T\right) \\
&= \rho_0 \kappa x - \frac{\rho_0 \kappa}{T} \sum_{j=0}^{x-1} j \\
&= \rho_0 \kappa x - \frac{\rho_0 \kappa}{T} \left(\frac{(x-1)((x-1)+1)}{2} \right) \\
&= \rho_0 \kappa x - \frac{\rho_0 \kappa}{T} \left(\frac{x^2 - x}{2} \right) \\
&= \left(\rho_0 \kappa + \frac{\rho_0 \kappa}{2T} \right) x - \frac{\rho_0 \kappa}{2T} x^2
\end{aligned}$$

Now we can substitute the experience effect back into the baseline equation of potential earnings.

$$\ln P_{x+s} = \ln P_0 + s\rho_s + \left(\rho_0\kappa + \frac{\rho_0\kappa}{2T}\right)x - \frac{\rho_0\kappa}{2T}x^2$$

Accounting for the difference in potential and observed earnings:

$$\begin{aligned}\ln Y(s,x) &\approx \ln P_{x+s} - \kappa\left(1 - x/T\right)\\ &= \left[\ln P_0 - \kappa\right] + \rho_s s + \left(\rho_0\kappa + \frac{\rho_0\kappa}{2T} + \frac{\kappa}{T}\right)x - \frac{\rho_0\kappa}{2T}x^2\end{aligned}$$

$\Rightarrow \rho_s$ is the average earnings increase with schooling

Standard Mincer Equation

$$\ln Y(s,x) = \alpha + \rho_s s + \beta_0 x + \beta_1 x^2,$$

where

$$\begin{aligned}
\alpha &= \ln P_0 - \kappa \\
\beta_0 &= \left(\rho_0 \kappa + \frac{\rho_0 \kappa}{2T} + \frac{\kappa}{T} \right) \\
\beta_1 &= -\frac{\rho_0 \kappa}{2T}
\end{aligned}$$

What about heterogeneous returns?

Random Coefficient Version

$$\ln Y(s_i, x_i) = \alpha_i + \rho_{si} s_i + \beta_{0i} x_i + \beta_{1i} x_i^2$$

and let

$$\bar{\alpha} = \mathrm{E}[\alpha_i] \qquad \bar{\rho}_s = \mathrm{E}[\rho_{si}]$$
$$\bar{\beta}_0 = \mathrm{E}[\beta_{0i}] \qquad \bar{\beta}_1 = \mathrm{E}[\beta_{1i}]$$

Dropping individual subscripts ...

$$\ln Y(s,x) = \bar{\alpha} + \bar{\rho}_s s + \bar{\beta}_0 x + \bar{\beta}_1 x^2 + \underbrace{[(\alpha - \bar{\alpha}) + (\rho_s - \bar{\rho}_s)s + (\beta_0 - \bar{\beta}_0)x + (\beta_1 - \bar{\beta}_1)x^2]}_{\epsilon}$$

⇒ If the schooling decision is determined by individual returns, then we are back in the case of a correlated random coefficient model (Heckman, Urzua, & Vytlacil, 2006).

Table 2: Estimated Coefficients from Mincer Log Earnings Regression for Men

		Whites		Blacks	
		Coefficient	Std. Error	Coefficient	Std. Error
1940	Intercept	4.4771	0.0096	4.6711	0.0298
	Education	0.1250	0.0007	0.0871	0.0022
	Experience	0.0904	0.0005	0.0646	0.0018
	Experience-Squared	-0.0013	0.0000	-0.0009	0.0000
1950	Intercept	5.3120	0.0132	5.0716	0.0409
	Education	0.1058	0.0009	0.0998	0.0030
	Experience	0.1074	0.0006	0.0933	0.0023
	Experience-Squared	-0.0017	0.0000	-0.0014	0.0000
1960	Intercept	5.6478	0.0066	5.4107	0.0220
	Education	0.1152	0.0005	0.1034	0.0016
	Experience	0.1156	0.0003	0.1035	0.0011
	Experience-Squared	-0.0018	0.0000	-0.0016	0.0000
1970	Intercept	5.9113	0.0045	5.8938	0.0155
	Education	0.1179	0.0003	0.1100	0.0012
	Experience	0.1323	0.0002	0.1074	0.0007
	Experience-Squared	-0.0022	0.0000	-0.0016	0.0000
1980	Intercept	6.8913	0.0030	6.4448	0.0120
	Education	0.1023	0.0002	0.1176	0.0009
	Experience	0.1255	0.0001	0.1075	0.0005
	Experience-Squared	-0.0022	0.0000	-0.0016	0.0000
1990	Intercept	6.8912	0.0034	6.3474	0.0144
	Education	0.1292	0.0002	0.1524	0.0011
	Experience	0.1301	0.0001	0.1109	0.0006
	Experience-Squared	-0.0023	0.0000	-0.0017	0.0000

Notes: Data taken from 1940-90 Decennial Censuses. See Appendix B for data description.

We can analyze this model in a Jupyter Noteboook. Visit

http://bit.ly/2kAtcyg

for the implementation.

Implications

- Log-earnings experience profiles are parallel across schooling levels.

$$\frac{\partial \ln Y(s,x)}{\partial s \partial x} = 0$$

- Log-earnings age profiles diverge with age across schooling levels.

$$\frac{\partial \ln Y(s,x)}{\partial s \partial t} = \frac{\rho_0 \kappa}{T} > 0$$

Figure: Experience profiles

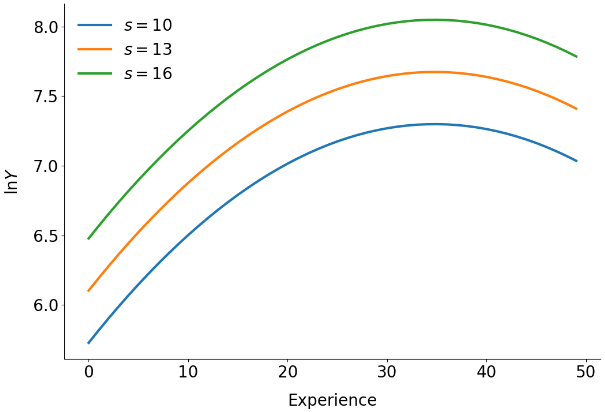
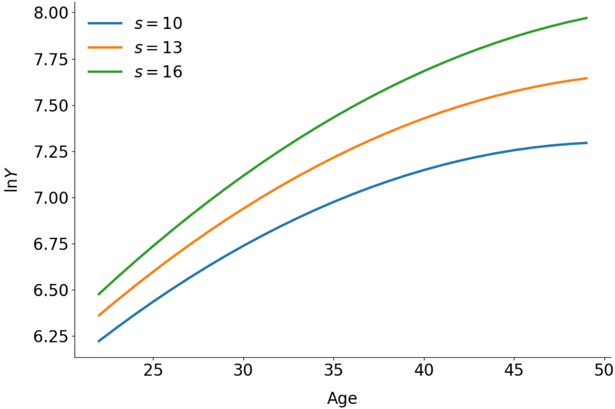


Figure: Age profiles



- The variance of earnings over the life cycle has a U-shaped pattern.

Derivation for minimizing variance

$$\begin{aligned}\ln Y(s,x) &= \ln P_{s+x} + \ln(1-k_{s+x}) \\ &\approx \ln P_s + \rho_0 \sum_{j=0}^{x-1} k_{s+j} - k_{s+x}\end{aligned}$$

Further, using the assumption of linearly declining investment yields

$$\ln Y(s,x) \approx \ln P_s + \kappa\left(\rho_0 \sum_{j=0}^{x-1}(1-j/T) - (1-x/T)\right)$$

Assuming only initial earnings potential P_s and investment levels κ vary in the population, the variance of log earnings is given by

$$
\begin{aligned}
\operatorname{var}(\ln Y(s,x)) = {} & \operatorname{var}(\ln P_s) \\
& + \left(\rho_0 \sum_{j=0}^{x-1} (1 - j/T) - (1 - x/T)\right)^2 \operatorname{var}(\kappa) \\
& + 2\left(\rho_0 \sum_{j=0}^{x-1} (1 - j/T) - (1 - x/T)\right) \operatorname{cov}(\ln P_s, \kappa).
\end{aligned}
$$

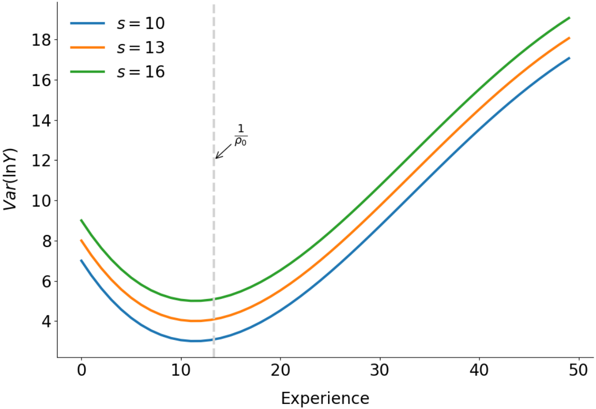
If κ and $\ln P_s$ are uncorrelated, then earnings are minimized (and equal to $Var(\ln P_s)$) when

$$\rho_0 \sum_{j=0}^{x-1} (1 - j/T) = 1 - x/T, or$$

$$\rho_0 \left(x - \frac{x(x-1)}{2T} \right) = (1 - x/T).$$

Clearly, $\lim_{T\to\infty} x^* = \frac{1}{\rho_0}$, so the variance minimizing age is $\frac{1}{\rho_0}$ when the work-life is long.

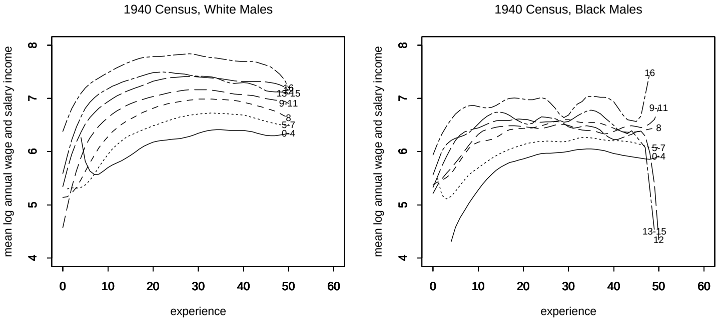
Figure: Variance profiles

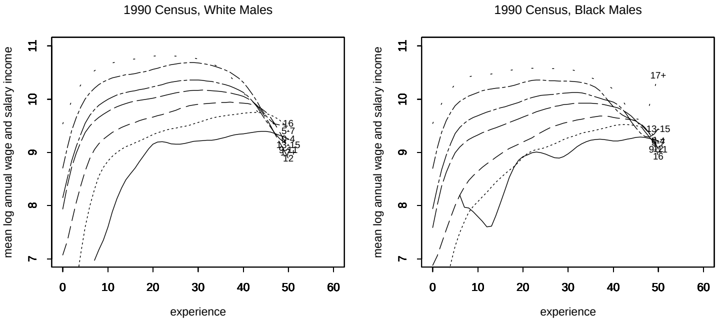


Empirical Evidence

The following results are based on the synthetic cohort approach using the United States Census.

Figure 1a: Experience-Earnings Profiles, 1940-1960





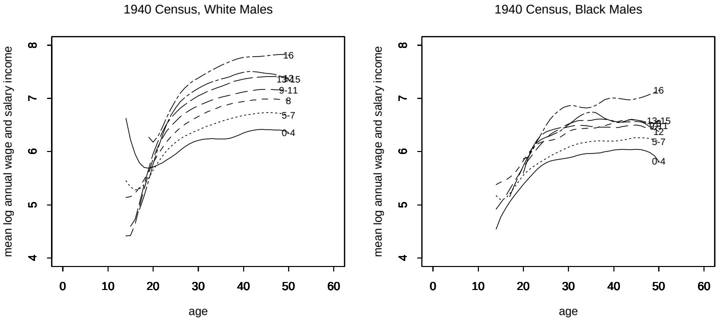
1990 Census, White Males
mean log annual wage and salary income
experience
16
5-7
0-4
13-15
9-11
12
1990 Census, Black Males
17+
13-15
0-4
9-11
16

Table 1: Tests of Parallelism in Log Earnings Experience Profiles for Men

		Estimated Difference Between College and High School Log Earnings at Different Experience Levels					
Sample	Experience Level	1940	1950	1960	1970	1980	1990
Whites	10	0.54	0.30	0.46	0.41	0.37	0.59
	20	0.40	0.40	0.43	0.49	0.45	0.54
	30	0.54	0.27	0.46	0.48	0.43	0.52
	40	0.58	0.21	0.50	0.45	0.27	0.30
	p-value	0.32	0.70	<0.001	<0.001	<0.001	<0.001
Blacks	10	0.20	0.58	0.48	0.38	0.70	0.77
	20	0.38	0.05	0.25	0.22	0.48	0.69
	30	-0.11	0.24	0.08	0.33	0.36	0.53
	40	-0.20	0.00	0.73	0.26	0.22	-0.04
	p-value	0.46	0.55	0.58	0.91	<0.001	<0.001

Notes: Data taken from 1940-90 Decennial Censuses without adjustment for inflation. Because there are very few blacks in the 1940 and 1950 samples with college degrees, especially at higher experience levels, the test results for blacks in those years refer to a test of the difference between earnings for high school graduates and persons with 8 years of education. See Appendix B for data description. See Appendix C for the formulae used for the test statistics.

Figure 2: Age-Earnings Profiles, 1940,1960,1980



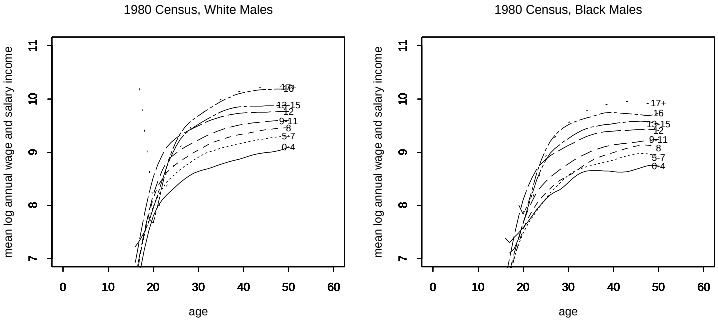
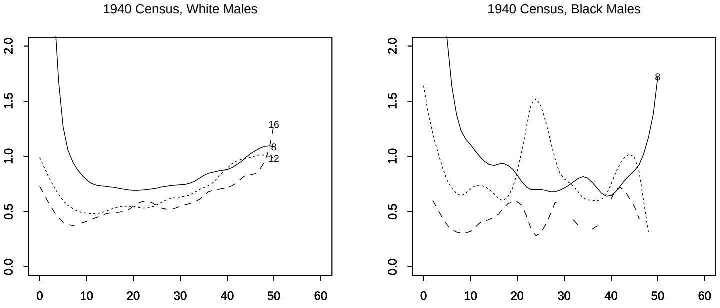
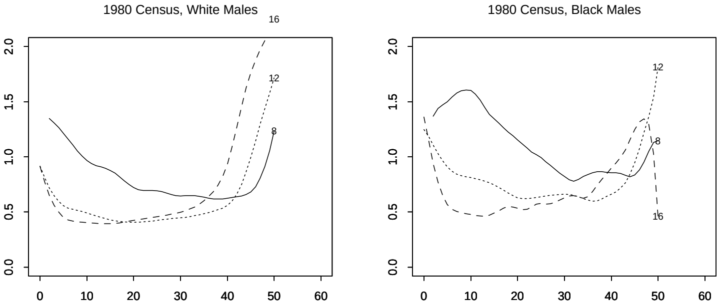

1980 Census, White Males
mean log annual wage and salary income
age
17+
16
13-15
12
9-11
8
5-7
0-4
1980 Census, Black Males
mean log annual wage and salary income
age
17+
16
13-15
12
9-11
8
5-7
0-4

Figure 3: Experience-Variance Log Earnings




1980 Census, White Males
16
12
8
2.0
1.5
1.0
0.5
0.0
0
10
20
30
40
50
60
1980 Census, Black Males
12
8
16
2.0
1.5
1.0
0.5
0.0
0
10
20
30
40
50
60

In the end, Heckman, Lochner, and Todd (2006) conclude:

> *In common usage, the coefficient on schooling in a regression of log earnings on years of schooling is often called a rate of return. In fact, it is a price of schooling from a hedonic market wage equation. It is a growth rate of market earnings with years of schooling and not an internal rate of return measure, except under stringent conditions which we specify, test and reject in this chapter.*

Internal rate of return

We study two different concepts of the rate of return in schooling:

- marginal differences
- non-marginal differences

We treat schooling as a continuous choice initially but then account for its discrete nature.

Income Maximization under Perfect Certainty

s	schooling level
x	experience level
$Y(s,x)$	wage income
$T(s)$	last age of earnings
v	tuition and psychic cost of schooling
τ	proportional tax rate
r	before-tax interest rate

Present Discounted Value of Lifetime Earnings

$$V(s) = \int_0^{T(s)-s} (1-\tau)e^{-(1-\tau)r(x+s)}Y(s,x)dx - \int_0^{s} ve^{-(1-\tau)rz}dz$$

First-Order Condition

$$
\begin{aligned}
&[T'(s)-1]e^{-(1-\tau)r(T(s)-s)}Y(s,T(s)-s) \\
&-(1-\tau)r\int_0^{T(s)-s} e^{-(1-\tau)rx}Y(s,x)dx \\
&+\int_0^{T(s)-s} e^{-(1-\tau)rx}\frac{\partial Y(s,x)}{\partial s}dx \\
&-\frac{v}{1-\tau}=0
\end{aligned}
$$

Rearranging and defining $\tilde{r} = (1-\tau)r$ ...

$$\tilde{r} = \frac{[T'(s)-1]e^{-\tilde{r}(T(s)-s)}Y(s,T(s)-s)}{\int_0^{T(s)-s} e^{-\tilde{r}x}Y(s,x)dx} \tag{1}$$

$$+\frac{\int_0^{T(s)-s} e^{-\tilde{r}x}\left[\frac{\partial Y(s,x)}{\partial s}\right]dx}{\int_0^{T(s)-s} e^{-\tilde{r}x}Y(s,x)dx} \tag{2}$$

$$-\frac{\frac{v}{1-\tau}}{\int_0^{T(s)-s} e^{-\tilde{r}x}Y(s,x)dx} \tag{3}$$

Interpretation

- (1) ... the change in the present value of earnings due to a change in working-life with additional schooling
- (2) ... weighted average effect of schooling on log earnings by experience
- (3) ... tuition and psychic costs

All components are expressed as a fraction of the present value of earnings measured at age s

Getting back to Mincer

- no tuition and psychic costs of schooling

  $\Rightarrow v = 0$

- no loss of working life from schooling

  $\Rightarrow T'(s) = 1$

- multiplicative separability between schooling and experience component of earnings

  $\Rightarrow Y(s,x) = \mu(s)\psi(x)$

$$\tilde{r} = \frac{\mu'(s)}{\mu(s)} \quad \forall \quad s$$

Thus, wage growth must be log linear in schooling and $\mu(s) = \mu(0)e^{\rho_s s}$ and $\tilde{r} = \rho_s$.

Structural Approach for the IRR

The internal rate of return for schooling level s_1 versus s_2, $r_I(s_1, s_2)$ solves ...

$$\int_0^{T(s_1)-s_1} (1-\tau)e^{-r_I(x+s_1)}Y(s_1,x)dx - \int_0^{s_1} ve^{-rz}dz$$
$$= \int_0^{T(s_2)-s_2} (1-\tau)e^{-r_I(x+s_2)}Y(s_2,x)dx - \int_0^{s_2} ve^{-rz}dz$$

Back to Mincer ....

- no taxes and no direct or psychic costs of schooling

  $\Rightarrow v = 0$ and $\tau = 0$

$$\int_0^{T(s_1)-s_1} e^{-r_I(x+s_1)} Y(s_1, x) dx = \int_0^{T(s_2)-s_2} e^{-r_I(x+s_2)} Y(s_2, x) dx$$

- equal work-lives irrespective of years of schooling

$$\Rightarrow T = T(s_1) - s_1 = T(s_2) - s_2$$

$$\int_0^T e^{-r_I(x+s_1)} Y(s_1, x) dx = \int_0^T e^{-r_I(x+s_2)} Y(s_2, x) dx$$

- parallelism in experience across schooling categories

$$\Rightarrow Y(s,x) = \mu(s)\psi(x)$$

$$\int_0^T e^{-r_I(x+s_1)}\mu(s)\psi(x)dx = \int_0^T e^{-r_I(x+s_2)}\mu(s)\psi(x)dx$$

- linearity of log earnings in schooling

$$\Rightarrow \quad \mu(s) = \mu(0)e^{\rho_s s}$$

$$\int_0^T e^{-r_I(x+s_1)}\mu(0)e^{\rho_s s_1}\psi(x)dx = \int_0^T e^{-r_I(x+s_2)}\mu(0)e^{\rho_s s_2}\psi(x)dx$$

After some further rearranging ...

$$e^{(\rho_s - r_I)s_1} = e^{(\rho_s - r_I)s_2}$$

$$\Rightarrow \rho_s = r_I$$

Heckman, Lochner, and Todd (2006) thus establish ...

> *After allowing for taxes, tuition, variable length of working life, and a flexible relationship between earnings, schooling and experience, the coefficient on years of schooling in a log earnings regression need no longer equal the internal rate of return.*

Empirical Evidence

Specifications

- **relax linearity in** S, including indicator variables for each level of schooling
- **relax linearity and parallelism**, nonparametrically estimated functions of experience, separately within each schooling level

Table 3a: Internal Rates of Return for White Men: Earnings Function Assumptions
(Specifications Assume Work Lives of 47 Years)

	Schooling Comparisons					
	6-8	8-10	10-12	12-14	12-16	14-16
1940						
Mincer Specification	13	13	13	13	13	13
Relax Linearity in S	16	14	15	10	15	21
Relax Linearity in S & Quad. in Exp.	16	14	17	10	15	20
Relax Lin. in S & Parallelism	12	14	24	11	18	26
1950						
Mincer Specification	11	11	11	11	11	11
Relax Linearity in S	13	13	18	0	8	16
Relax Linearity in S & Quad. in Exp.	14	12	16	3	8	14
Relax Linearity in S & Parallelism	26	28	28	3	8	19
1960						
Mincer Specification	12	12	12	12	12	12
Relax Linearity in S	9	7	22	6	13	21
Relax Linearity in S & Quad. in Exp.	10	9	17	8	12	17
Relax Linearity in S & Parallelism	23	29	33	7	13	25
1970						
Mincer Specification	13	13	13	13	13	13
Relax Linearity in S	2	3	30	6	13	20
Relax Linearity in S & Quad. in Exp.	5	7	20	10	13	17
Relax Linearity in S & Parallelism	17	29	33	7	13	24
1980						
Mincer Specification	11	11	11	11	11	11
Relax Linearity in S	3	-11	36	5	11	18
Relax Linearity in S & Quad. in Exp.	4	-4	28	6	11	16
Relax Linearity in S & Parallelism	16	66	45	5	11	21
1990						
Mincer Specification	14	14	14	14	14	14
Relax Linearity in S	-7	-7	39	7	15	24
Relax Linearity in S & Quad. in Exp.	-3	-3	30	10	15	20

Table 3b: Internal Rates of Return for Black Men: Earnings Function Assumptions
(Specifications Assume Work Lives of 47 Years)

	Schooling Comparisons					
	6-8	8-10	10-12	12-14	12-16	14-16
1940						
Mincer Specification	9	9	9	9	9	9
Relax Linearity in S	18	7	5	3	11	18
Relax Linearity in S & Quad. in Exp.	18	8	6	2	10	19
Relax Linearity in S & Parallelism	11	0	10	5	12	20
1950						
Mincer Specification	10	10	10	10	10	10
Relax Linearity in S	16	14	18	-2	4	9
Relax Linearity in S & Quad. in Exp.	16	14	18	0	3	6
Relax Linearity in S & Parallelism	35	15	48	-3	6	34
1960						
Mincer Specification	11	11	11	11	11	11
Relax Linearity in S	13	12	18	5	8	11
Relax Linearity in S & Quad. in Exp.	13	11	18	5	7	10
Relax Linearity in S & Parallelism	22	15	38	5	11	25
1970						
Mincer Specification	12	12	12	12	12	12
Relax Linearity in S	5	11	30	7	10	14
Relax Linearity in S & Quad. in Exp.	6	11	24	10	11	12
Relax Linearity in S & Parallelism	15	27	44	9	14	23
1980						
Mincer Specification	12	12	12	12	12	12
Relax Linearity in S	-4	1	35	10	15	19
Relax Linearity in S & Quad. in Exp.	-4	6	29	11	14	17
Relax Linearity in S & Parallelism	10	44	48	8	16	31
1990						
Mincer Specification	16	16	16	16	16	16
Relax Linearity in S	-5	-5	41	15	20	25
Relax Linearity in S & Quad. in Exp.	-3	-3	35	17	19	22

Table 4: Internal Rates of Return for White & Black Men: Accounting for Taxes and Tuition (General Non-Parametric Specification Assuming Work Lives of 47 Years)

		Schooling Comparisons					
		Whites			Blacks		
		12-14	12-16	14-16	12-14	12-16	14-16
1940	No Taxes or Tuition	11	18	26	5	12	20
	Including Tuition Costs	9	15	21	4	10	16
	Including Tuition & Flat Taxes	8	15	21	4	9	16
	Including Tuition & Prog. Taxes	8	15	21	4	10	16
1950	No Taxes or Tuition	3	8	19	-3	6	34
	Including Tuition Costs	3	8	16	-3	5	25
	Including Tuition & Flat Taxes	3	8	16	-3	5	24
	Including Tuition & Prog. Taxes	3	7	15	-3	5	21
1960	No Taxes or Tuition	7	13	25	5	11	25
	Including Tuition Costs	6	11	21	5	9	18
	Including Tuition & Flat Taxes	6	11	20	4	8	17
	Including Tuition & Prog. Taxes	6	10	19	4	8	15
1970	No Taxes or Tuition	7	13	24	9	14	23
	Including Tuition Costs	6	12	20	7	12	18
	Including Tuition & Flat Taxes	6	11	20	7	11	17
	Including Tuition & Prog. Taxes	5	10	18	7	10	16
1980	No Taxes or Tuition	5	11	21	8	16	31
	Including Tuition Costs	4	10	18	7	13	24
	Including Tuition & Flat Taxes	4	9	17	6	12	21
	Including Tuition & Prog. Taxes	4	8	15	6	11	20
1990	No Taxes or Tuition	10	16	26	18	25	35
	Including Tuition Costs	9	14	20	14	18	25
	Including Tuition & Flat Taxes	8	13	19	13	17	22
	Including Tuition & Prog. Taxes	8	12	18	13	17	22

Notes: Data taken from 1940-90 Decennial Censuses. See discussion in text and Appendix B for a description of tuition and tax amounts.

Figure 4a: Average College Tuition Paid (in 2000 dollars)

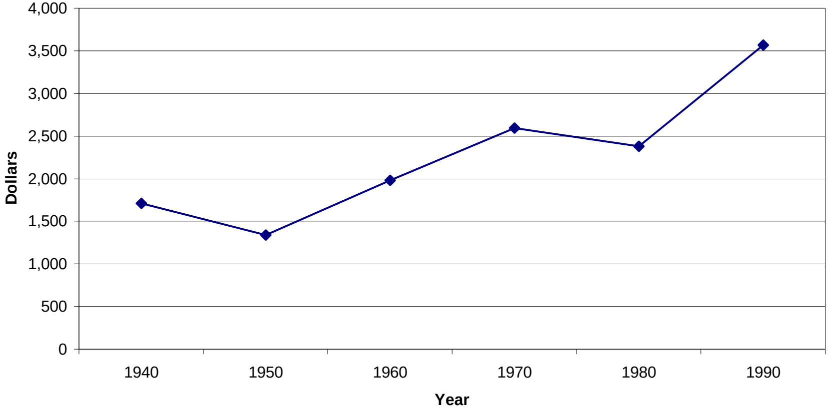


Figure 4b: Marginal Tax Rates
(from Barro & Sahasakul, 1983, Mulligan & Marion, 2000)

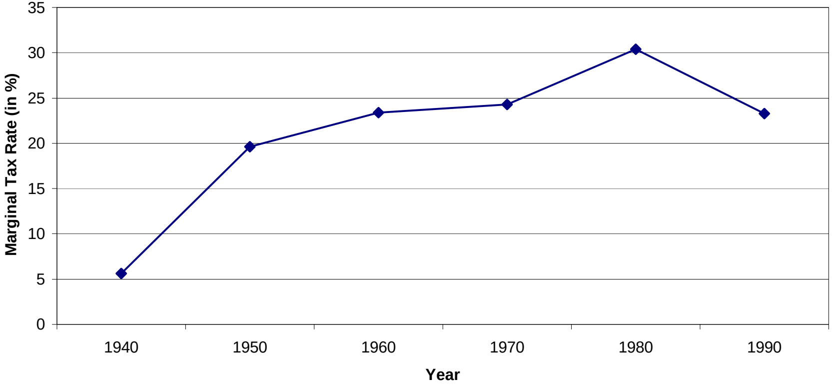


Figure 5: IRR for High School Completion (White and Black Men)

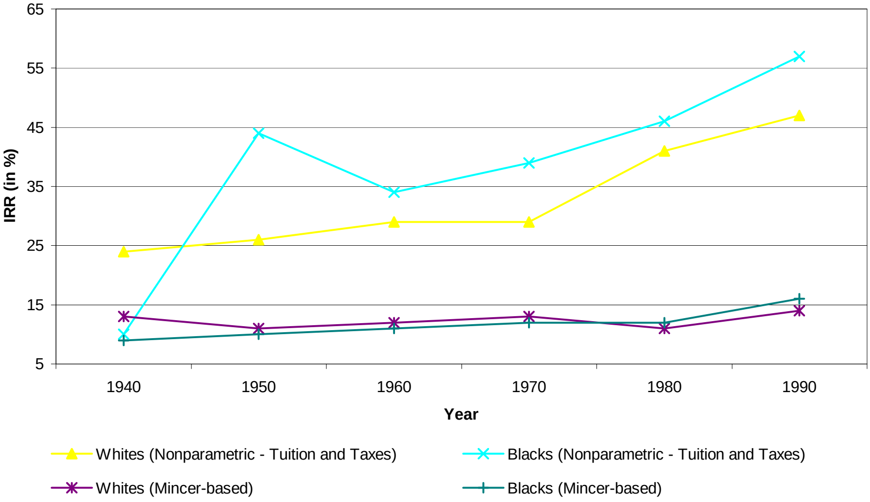
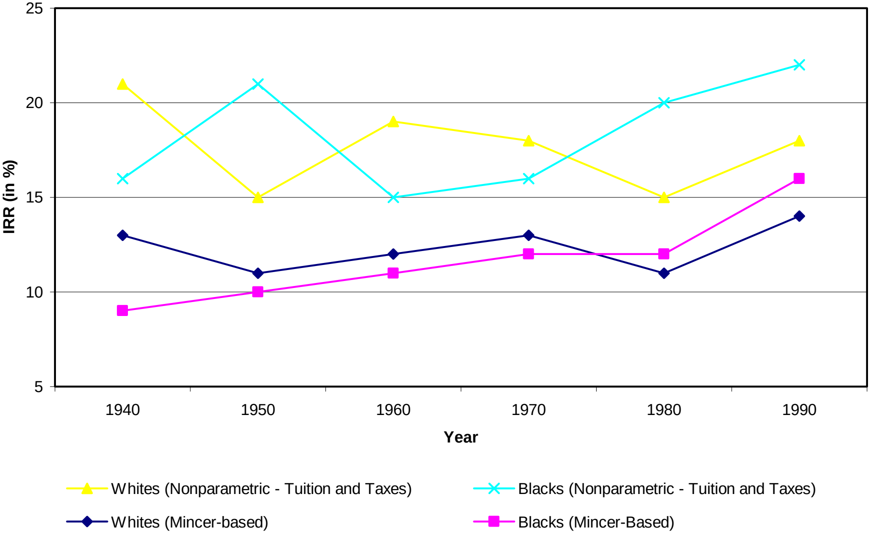


Figure 6: IRR for College Completion (White and Black Men)

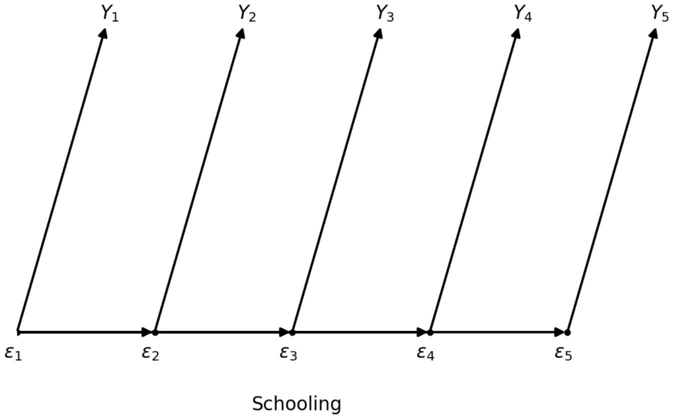


True rate of return

Suppose their is uncertainty about net earnings conditional on s and actual lifetime earnings for someone with s years of schooling are:

$$Y_s = \underbrace{\left[\sum_{x=0}^{T}(1+r)^{-x}Y(s,x)\right]}_{\bar{Y}_s}\epsilon_s$$

Figure: Model structure



The decision problem for a person with s years of schooling given the sequential revelation of information is to complete another year of schooling if

$$Y_s \leq \frac{E_s(V_{s+1})}{1+r}.$$

So the value of schooling level s, V_s, is

$$V_s = \max\left\{Y_s, \frac{E_s(V_{s+1})}{1+r}\right\}$$

for $s \leq \bar{S}$. At the maximum schooling level, $\bar{S}$, after all information is revealed, we obtain $V_{\bar{s}} = Y_{\bar{s}} = \bar{Y}_{\bar{s}}\epsilon_{\bar{s}}$.

The endogenously determined probability of going on from school level s to $s+1$ is

$$p_{s+1,s} = \Pr\left(\epsilon_s \leq \frac{E_s(V_{s+1})}{(1+r)\bar{Y}_s}\right),$$

where $E_s(V_{s+1})$ may depend on ϵ_s because it enters the agent's information set.

Thus, the expected value of schooling level s as perceived at current schooling $s-1$ is:

$$E_{s-1}(V_s) = \underbrace{(1-p_{s+1,s})\bar{Y}_s E_{s-1}\left(\epsilon \mid \epsilon \geq \frac{E_s(V_{s+1})}{(1+r)\bar{Y}_s}\right)}_{\text{direct return}}$$

$$+ \underbrace{p_{s+1,s}\left(\frac{E_{s-1}(V_{s+1})}{1+r}\right)}_{\text{option value}}.$$

Objects of interest

- Option value

$$O_{s,s-1} = \mathrm{E}_{s-1}\left[V_s - Y_s\right]$$

- True rate of return

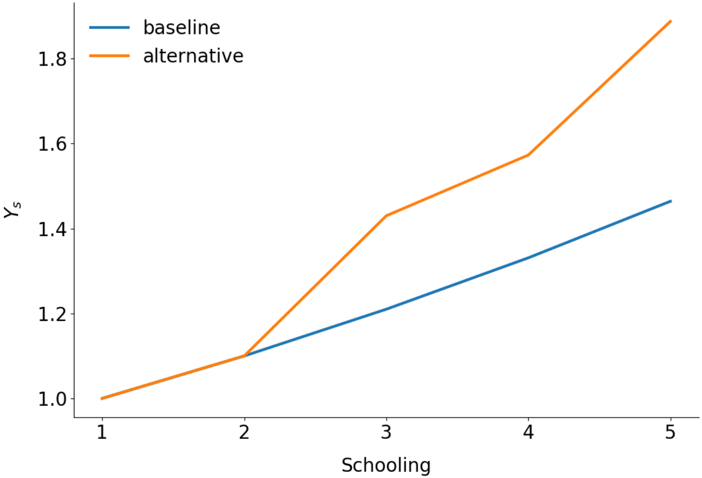
$$R_{s,s-1} = \frac{\mathrm{E}_{s-1}\left[V_s\right] - Y_{s-1}}{Y_{s-1}}$$

Model specification

$$\ln(\epsilon_s) \sim \mathbb{N}(0, \sigma) \qquad r = 0.1$$

$$Y_{s+1} = (1 + \rho_{s+1}) Y_s \qquad \sigma = 0.1$$

Figure: Scenarios



We can analyze this model in a Jupyter Noteboook. Visit

`http://bit.ly/2skwwli`

for the implementation.

Figure: Option values and uncertainty

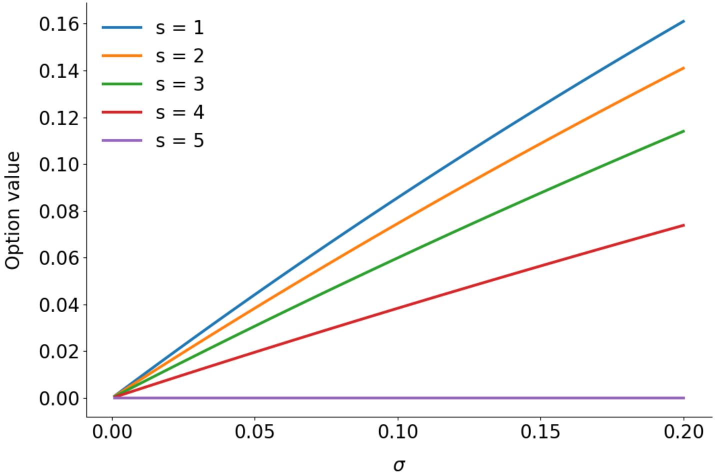
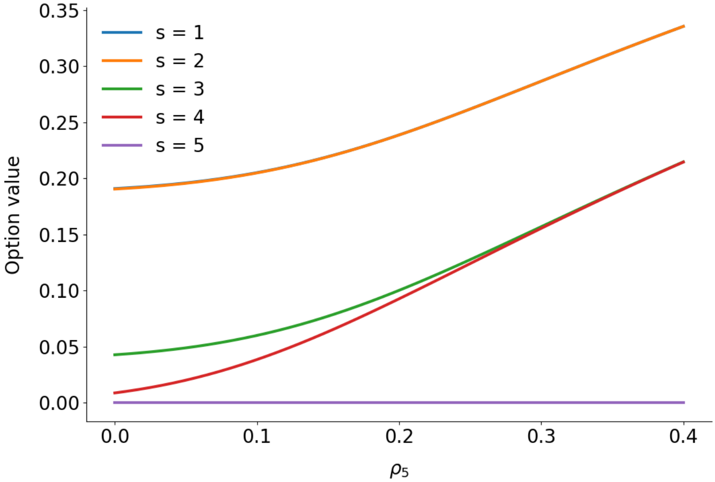


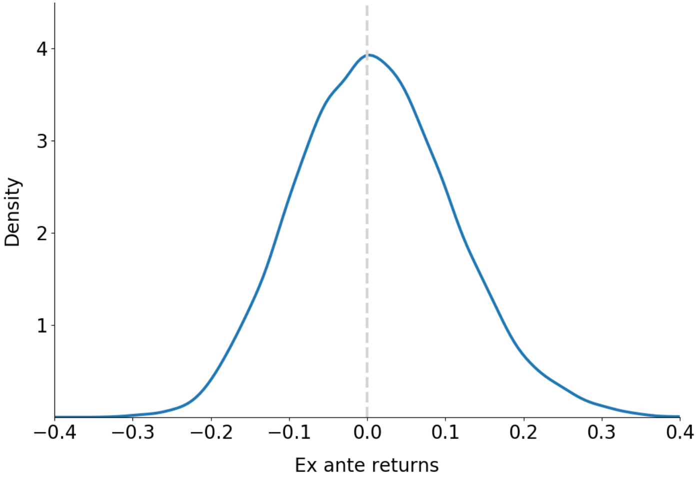
Figure: Option values and sheepskin effects



Key features

- The convergence of $O_{5,4}$ and $O_{4,3}$ follows from the additional benefits of obtaining 5 years of schooling as we increase the growth rate ρ_5. It simply dominates the increase going from three to four years of schooling parameterized by ρ_4.

Figure: Returns



Conclusion

Appendix

References

Becker, G. S. (1964). *Human capital* (1st ed.). New York City, NY: Columbia University Press.

Chapman, & Hall. (2018). *Crc standard mathematical tables and formulas* (33rd Edition ed.; D. Zwillinger, Ed.). Florida, FL: CRC Press.

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (1st ed., Vol. 1, pp. 307–458). Amsterdam, Netherlands: North-Holland Publishing Company.

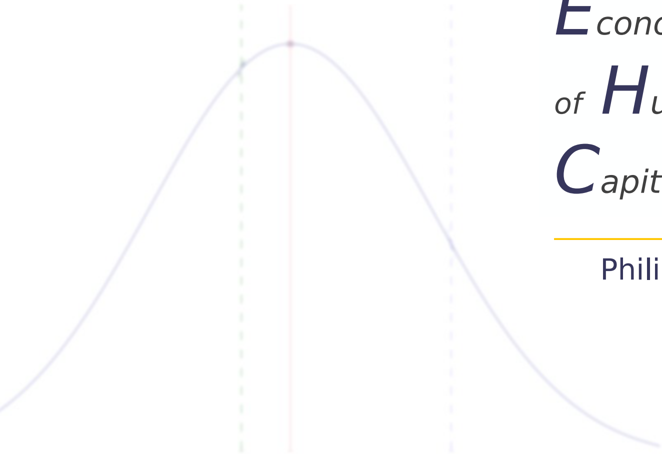
Heckman, J. J., Urzua, S., & Vytlacil, E. J. (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics*, *88*(3), 389–432.

Mincer, J. (1974). *Schooling, experience and earnings*. New York, NY: National Bureau of Economic Research.

Rosen, S. (1977). Human capital: A survey of empirical research. In R. G. Ehrenberg (Ed.), *Research in labor economics* (Vol. 1, pp. 3–40). Greenwich, CT: JAI Press.

Willis, R., & Rosen, S. (1979). Education and self-selection. *Journal of Political Economy*, *87*(5, Part 2), S7–S36.

Economics of Human Capital



Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

Economics of Human Capital

Multidimensionality of skills

Philipp Eisenhauer

Introduction

Multidimensionality of skills

- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*(3), 411–482.
- Eisenhauer, P., Heckman, J. J., & Mosso, S. (2015). Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. *International Economic Review*, *56*(2), 331–357.

The Effects of Cognitive and Noncognitive Abilities on Labor Outcomes and Social Behavior

James Heckman
University of Chicago and American Bar Foundation

Jora Stixrud and Sergio Urzua
University of Chicago

The Chicago Workshop on Black-White Inequality
Chicago
April 21, 2006

Introduction

- Although the importance of cognitive skills for success in a variety of dimensions of social and economic life is well established, the importance of noncognitive skills has largely been overlooked.

- In the 1970's, Marxist economists documented the importance of noncognitive traits in the workplace (Bowles; Gintis; Edwards).

- Early work by Peter Muesser reported in Jencks (1979) found that skills such as industriousness, perseverance, and leadership have significant influences on wages – comparable to estimated effects of schooling, IQ, and parental socioeconomic status – even after controlling for standard human capital variables.

- Osborne (2000) studies the effect of personality and behavioral traits on wages of females.

- Bowles, Gintis, and Osborne (2001) present a model in which non-cognitive skills are rewarded by employers, in the form of increased wages. In their model, employee preferences that allow their employer to induce greater effort at a lower cost are termed incentive enhancing. If, for example, costlessly enforceable contracts for labor services are unavailable, then incentive enhancing preferences will be valued by the employer, and may be rewarded as such.

- Examples of incentive enhancing preferences are: a low time discount rate (i.e., greater future orientation), high degree of self-directedness and personal efficacy, a predisposition to truth-telling, a low disutility of effort, a high marginal utility of income, identification with the objectives of a firm's owners and managers, a tendency of helpful (non-disruptive) behavior toward other employees, and a high sense of shame at being without a job or receiving handouts.

- Heckman and Rubinstein (2001) use evidence from the General Education Development (GED) testing program (an exam-certified alternative high school degree) to demonstrate the quantitative importance of noncognitive skills. GED recipients have the same cognitive ability as high school graduates that do not go onto college, as measured by scores on the Armed Forces Qualifying Test (AFQT). However, once cognitive ability is controlled for, GED recipients earn the same, have lower hourly wages, and obtain lower levels of schooling than high school dropouts. Some other factor must account for this striking difference, and the authors identify this as noncognitive skill.

- Heckman, LaFontaine and Urzua (2004) show that GEDs have higher turnover rates, are more likely to drop out of the army and post secondary schooling, and are less likely to persevere in many tasks than HS dropouts.

- Carneiro and Heckman (2002), and Heckman and Masterov (2004) argue that parents play an important role in producing both the cognitive and non-cognitive skills of their children, and more able and engaged parents have greater success in doing so. Because cognitive and non-cognitive abilities are shaped early in the lifecycle, differences in these abilities are persistent, and both are crucial to the social and economic success of an individual, gaps among income and racial groups begin early and persist.

- **Early interventions**, such as enriched childcare centers coupled with home visitations, have been successful in alleviating some of the initial disadvantages of children born into adverse conditions. The success of these interventions has primarily been due not to their success in improving the cognitive skills (IQ) of these children, but rather to their success in boosting non-cognitive skills and increasing child motivation.

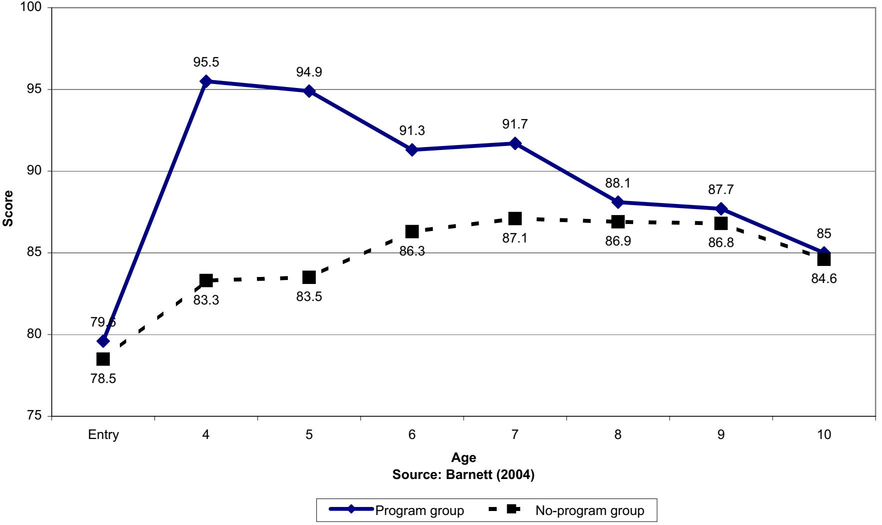
- The Perry Preschool Program, an enriched early childhood intervention evaluated by random assignment where treatments and controls are followed to age 40, did not boost IQ but raised a achievement test scores, schooling and social skills.

- Raised noncognitive skills but not cognitive skills, at least as measured by IQ.

- Effects were not uniform across gender groups (Heckman, 2004).

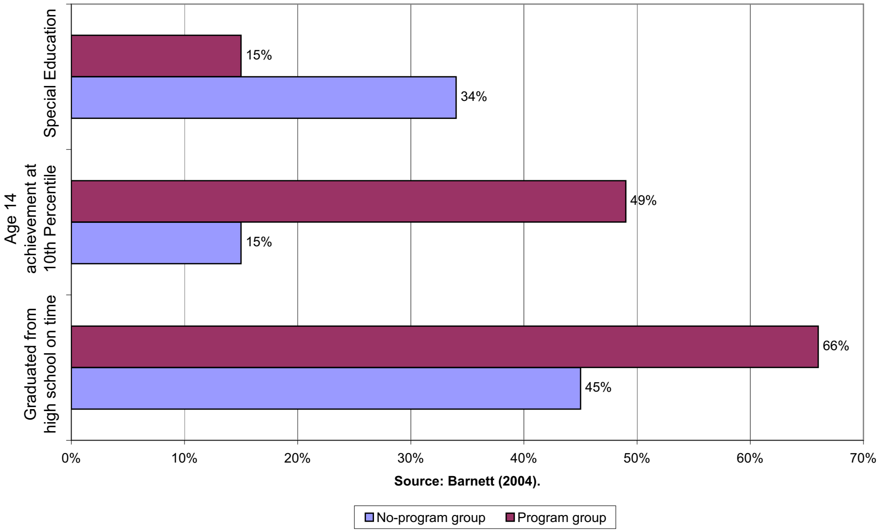
- See the evidence in Cunha, Heckman, Lochner and Masterov (2005).

Figure 1A
Perry Preschool IQ Over Time



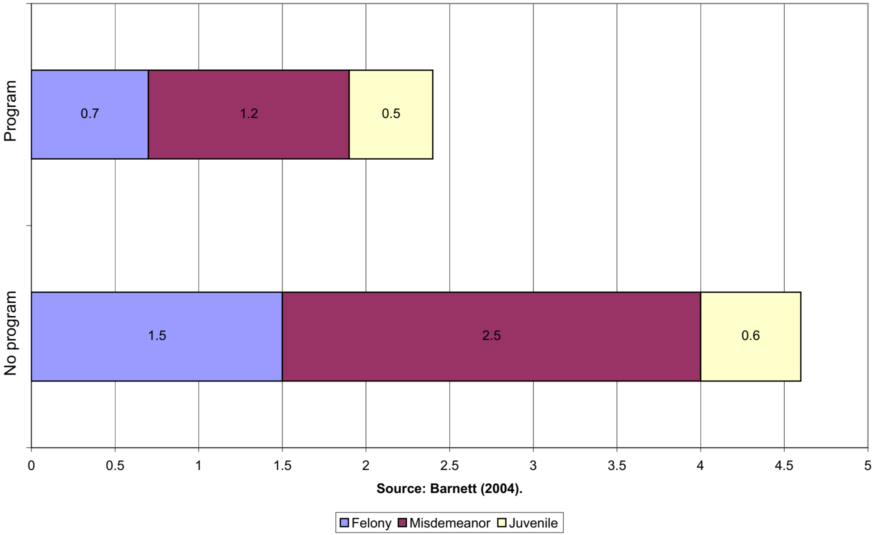
Source: Barnett (2004)

Figure 1B
Perry Preschool: Educational Effects



Source: Barnett (2004).

Figure 1C
Perry Preschool: Arrests Per Person by Age 27



Source: Barnett (2004).

Problems with the Recent Literature and Our Solution

- Naive regressions of earnings on test scores (cognitive and noncognitive) are problematic.
- Problem is reverse causality: Schooling may cause both earnings and test scores.
- The recent literature notes that schooling and age may influence cognitive measures and corrects for the effect of schooling on ability (Hansen, Heckman and Mullen, 2004).
- Age at test and schooling levels differ among individuals in our samples. It is necessary to account for the fact that our test measures are not directly comparable across people as they reflect different input levels.
- The test scores we use are corrected for the fact that different individuals have different amounts of schooling at the time they take the test.
- Our analysis generalizes Hansen, Heckman and Mullen (2004).

- We allow for common factors explaining wages and schooling to account for correlated risky behavior across youth (Biglan, 2004).
- We broaden previous analyses and explain wages, schooling and a variety of social behaviors from a low-dimensional set of factors.
- Start with cognitive tests to get the main idea of our procedure, and provide an intuition for how we secure identification.

1 Relationship Between Ability and Schooling: An Introduction

- $T(s_T)$ is the test score of a person with s_T years of schooling at the date of the test. S is final schooling level.
- Test is taken at schooling level s_T. X are other determinants. We suppress them to simplify the notation.

Latent ability f (IQ)

Test score $T(s_T)$:

$$T(s_T) = \mu(s_T) + \lambda(s_T)f + \varepsilon(s_T) \tag{1.1}$$

Assume that $\varepsilon(s_T)$ is independent of f.
f and $\varepsilon(s_T)$ are assumed to have zero means.

$\mu(s_T)$ in equation (1.1) is the effect of schooling that is uniform across latent ability levels

$\lambda(s_T)$ is the effect of schooling on revealing or transforming latent ability f.

The marginal causal effects of changing schooling from s'_T to s_T on levels and slopes are

$$\mu(s_T) - \mu(s'_T) \text{ and } \left[\lambda(s_T) - \lambda(s'_T)\right] f$$

The empirical literature on cognitive ability recognizes the problem of reverse causality. (Herrnstein and Murray, 1994, Neal and Johnson, 1996, and Winship and Korenman, 1997)

$$\text{Assumes } \mu(s_T) = s_T\beta \text{ (linearity).}$$

Uses instruments (Neal and Johnson, 1996) or proxies (Herrnstein and Murray, 1994; Winship and Korenman, 1997) to solve for endogeneity problems. Both methods are controversial.

2 Simple Idea Motivating How to Control for Endogeneity of Schooling on Test Scores

(Hansen, Heckman and Mullen, 2004)

- We observe Test at the date of the test (S_T).
- The agent has S_T years of schooling at the date of the test ($S_T = s_T$).
- The agent completes schooling and has final schooling $S = s$.
- We have panel data and follow the person after taking the test.
- Suppose we observe test score for a person with $S = s$, $S_T = s_T$.

- Then if

$$\begin{aligned} T(S,S_T) &= \mu(s_T)+U(S,S_T) \\ E\left(T(S,S_T)\mid S_T=s_T,S=s\right) &= \mu(s_T)+E\left(U\left(S,S_T\right)\mid S=s,S_T=s_T\right) \end{aligned}$$

where it is assumed

$$E\left(U\left(S,S_T\right)\mid S=s,S_T=s_T\right)=E\left(U\left(S,S_T\right)\mid S=s\right)$$

(All selection controlled by conditioning on final schooling)

- Form Contrasts

$$\begin{aligned} & E\left(T(S,S_T) \mid S=s, S_T=s_T\right) - E\left(T\left(S,S_T\right) \mid S=s, S_T=s'_T\right) \\ = \; & \mu(s_T) - \mu(s'_T) \end{aligned}$$

Get "Effects," e.g.

$$\mu(s_T) = s_T\beta$$

- We identify β.
- This is a "Matching" type of assumption.
- We can relax it using a semiparametric model (Hansen, Heckman and Mullen, 2004).
- They show that both approaches produce very similar estimates.

3 Extensions of the Model

- We extend the model to have two factors:

$$\begin{array}{ll} f^C & \text{Cognitive} \\ f^N & \text{Noncognitive} \end{array}$$

  Using these factors, we can explain a variety of outcomes:

  1. Schooling attainment
  2. Wages given schooling
  3. Wages overall
  4. Work experience
  5. Occupational choice
  6. Social behaviors and risky correlated behaviors:
     (a) Crime and incarceration
     (b) Teenage pregnancy
     (c) Drug use
     (d) Smoking

- We use test scores on both cognitive and noncognitive skills to proxy the latent factors.

- We account for measurement error. (Produces a downward bias in OLS).

- We adjust for effect of schooling on test scores. (Produces an upward bias in OLS).

4 Our Model Approximates an Explicit Economic Model of Preferences and Behavior

Our model is an approximation to a simple life cycle model of youth and adult decision making over horizon $\bar{T}$.

- Let consumption and labor supply at period t be $c(t)$ and $l(t)$, respectively. $c(t)$ can be a vector of choices by agent.

- Utility $U\left(c(t), l(t), \eta\right)$, where the η are preference parameters.

- Time preference rate ρ.

- Human Capital in period t is $h(t)$. Its time rate of change is $\dot{h}(t)$.

$$\dot{h}(t) = \varphi\left(h(t), I(t), \tau\right)$$

  τ are productivity parameters, $I(t)$ is investment at t, and $h(t)$ denotes the rate of change of the human capital stock.

- The initial condition is $h(0)$.

Wages at period $t(Y(t))$ are given by human capital and productivity traits θ :

$$Y(t) = R\left(h(t);\theta\right).$$

- Perfect credit markets at interest rate r
- Law of motion for assets at period $t(A(t))$, given initial condition $A(0)$ and ignoring taxes, is

$$\dot{A}(t) = Y(t)h(t)l(t) - P(t)'c(t) + rA(t)$$

- Agent maximizes

$$\int_0^{\overline{T}} \exp(-\rho t)U\left(c(t), l(t), \eta\right) dt$$

subject to initial conditions and dynamic constraints

- Cognitive and noncognitive skills can affect:

$$\begin{aligned}
\text{preferences } \eta &= \left(\eta\left(f^C, f^N\right), \rho = \rho\left(f^C, f^N\right)\right), \\
\text{human capital productivity } \tau &= \tau\left(f^C, f^N\right), \\
\text{and direct market productivity } \theta &= \left(\theta\left(f^C, f^N\right)\right)
\end{aligned}$$

- We examine how factors are priced out in different schooling markets.

- The factors also affect initial endowments:

$$\begin{aligned}
h(0) &= h_0\left(f^C, f^N\right) \\
A(0) &= A_0\left(f^C, f^N\right)
\end{aligned}$$

- Our econometric model is a linear-in-the-parameters approximation to the more general model.

- Underway is a more explicit structural model.

- Will talk about this at the end.

5 Data

National Longitudinal Survey of Youth (NLSY79). The NLSY is a representative sample of young Americans between the ages of 14 and 21 at the time of the first interview in 1979. We use the random sample of 6111 noninstitutionalized civilian youths.

The NLSY collects information on parental background, schooling decisions, labor market experiences, cognitive and noncognitive test scores and other behavioral measures of these individuals on an annual basis.

5.1 Cognitive Test Scores: (ASVAB)

The following tests are used in our analysis: (i) arithmetic reasoning, (ii) word knowledge, (iii) paragraph comprehension, (iv) numerical operations, and (v) coding speed.

5.2 Non-Cognitive Measures

5.2.1 Rotter Internal-External Locus of Control Scale

The Rotter Internal-External Locus of Control Scale, collected as part of the 1979 interviews, is a four-item abbreviated version of a 23-item forced choice questionnaire adapted from the 60-item Rotter scale developed by Rotter (1966). The scale is designed to measure the extent to which individuals believe they have control over their lives, i.e., self-motivation and self-determination, (internal control) as opposed to the extent that the environment (i.e., chance, fate, luck) controls their lives (external control).

5.2.2 Rosenberg Self-Esteem Scale

The Rosenberg Self-Esteem Scale, measures an individual's degree of approval or disapproval toward himself.

6 Traditional OLS Results

- To benchmark our analysis, we present traditional reduced form results of the effects of cognitive and non-cognitive skills on educational attainment, wages, smoking, going to jail, and teenage pregnancy.

- They assume that test scores are perfect proxies and they ignore problems arising from reverse causality.

Table 1 - Estimated Coefficients from Log Wage Regressions

NLSY79 - Males and Females at Age 30 [(a)]

	Males		Females	
Variables [(b)]	(A)	(B)	(A)	(B)
GED	0.017		-0.002	
	(0.048)		(0.056)	
High School Graduate	0.087		0.059	
	(0.035)		(0.044)	
Some College	0.146		0.117	
	(0.044)		(0.052)	
2yr College Graduate	0.215		0.233	
	(0.058)		(0.058)	
4yr College Graduate	0.292		0.354	
	(0.046)		(0.054)	
AFQT [(c)]	0.121	0.1900	0.169	0.251
	(0.016)	(0.013)	(0.017)	(0.014)
ATTITUDES [(d)]	0.042	0.052	0.028	0.041
	(0.011)	(0.012)	(0.013)	(0.013)
Constant	2.558	2.690	2.178	2.288
	(0.057)	(0.050)	(0.063)	(0.052)

Notes: (a) We exclude the oversample of blacks, hispanics and poor whites, the military sample, and those currently enrolled in college; (b) The model includes includes a set of cohort dummies, local labor market conditions (unemployment rate), the region of residence, and race. The column A presents the estimates obtained from OLS. Column B presents the results from an OLS model in which the schooling dummies are excluded; (c) the cognitive measure represents the standardized average over the ASVAB scores (arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations and coding speed); (d) The Non-cognitive measure is computed as a (standardized) average of the Rosenberg self-esteem scale and Rotter internal-external locus of control. Standard errors in parentheses.

- Both cognitive and non-cognitive measures are found to significantly affect wages, educational attainment, work experience and behavioral outcomes. Interesting gender differences also emerge.

- Such gender differences are a major finding of the intervention response literature (Heckman, 2004).

- Reduced form results are problematic because of measurement error and reverse causality (simultaneity).

7 A Model of Schooling and Wages

- We posit the existence of two underlying factors representing latent cognitive and non-cognitive ability. Let f^C and f^N denote the levels of latent cognitive and non-cognitive abilities.

- The levels of an individual's factors may result from some combination of inherited ability, the quality of the environment provided by his parents, early effort on his part, and the effects of any early interventions.

- Our sample starts at age 14 so we cannot investigate the effects of early environments in this study. We take f^C and f^N as initial conditions.

- We show some results from a project with Flavio Cunha at the end of this paper. This analysis starts at early ages and shows the determinants of skill formation over the life cycle.

- We assume that levels of both abilities are known by each individual but not by the researcher, and that they are fixed by the time the individual makes his schooling choice.

- We assume that latent abilities are mutually independent, and both determine the individual's wage and schooling decision.

- This does not mean that the manifest abilities are independent.

7.1 A Hedonic Model for Wages

- We assume that different schooling levels are priced differently in the labor market (Hedonic model).

- Both latent abilities (jointly with observable variables) determine log wages $\ln Y_s$

$$\ln Y_s = \beta_{Y,s} X_Y + \alpha^C_{Y,s} f^C + \alpha^N_{Y,s} f^N + e_{Y,s} \qquad \text{for} \qquad s = 1, \ldots, \bar{S},$$

where

$$e_{Y,s} \perp\!\!\!\perp (f^N, f^C, X_Y).$$

- "$\perp\!\!\!\perp$" denotes independence.

- We determine how factors are priced out in different schooling markets, $s = 1, \ldots, \bar{S}$.

7.2 The Schooling Model

Let s^* denote this optimal schooling level as a choice among utilities in different states: I_j, $j = 1, ..., \bar{S}$.

$$s^* = \arg \max_{s=\{1,...,\bar{S}\}} \{I_1, \ldots, I_{\bar{S}}\}.$$

where

$$I_s = \beta_s X_s + \alpha_s^C f^C + \alpha_s^N f^N + e_{S,s} \quad \text{for} \quad s = 1, \ldots, \bar{S} \tag{7.1}$$

is a reduced form net utility, where

$$e_{S,s} \perp\!\!\!\perp (f^N, f^C, X_s).$$

From (7.1)

$$D = \begin{cases} 1 \text{ if } I_1 = \max\{I_1, \ldots, I_{\overline{S}}\} \\ \vdots \\ \bar{S} \text{ if } I_{\bar{S}} = \max\{I_1, \ldots, I_{\overline{S}}\}. \end{cases}$$

- D indicates the schooling decision of the individual.

7.3 The Measurement System and Identification of the Model

- Identification of the above model can be directly established using the strategy developed in Carneiro, Hansen, and Heckman (2003).

- Our identification strategy assumes the existence of a set of cognitive and noncognitive measures (test scores). It assumes the existence of two sets of variables (each with at least two elements) measuring cognitive and non-cognitive skills. Each set is exclusively devoted to its respective latent ability. Latent cognitive ability is only allowed to affect scores on cognitive measures, and latent non-cognitive is only allowed to affect scores on non-cognitive measures.

- The specification pursued here makes the interpretation of f^C and f^N as cognitive and non-cognitive abilities more transparent.

- We address the potential problem of reverse causality between schooling and test scores and schooling and attitude scales.
- The observed measures may not be fully informative about the actual skills of the individuals, since they may be influenced by the schooling level at the moment of the test.
- They may also depend on school quality and family environment.

- Denote by s_T the schooling level at the moment of the test ($s_T = 1, \ldots, \bar{S}_T$), the model for the cognitive measure C_i ($i = 1, \ldots, n_C$) is

$$C_i(s_T) = \beta_{C_i}(s_T) X_C + \alpha_{C_i}(s_T) f^C + e_{C_i}(s_T)$$

with $i = 1, \ldots, n_C$, $s_T = 1, \ldots, \bar{S}_T$ and

$$e_{C_i}(s_T) \perp\!\!\!\perp \left(f^C, X_C\right) \text{ and } e_{C_i}(s_T) \perp\!\!\!\perp e_{C_j}(s_T')$$

for all C_i and C_j in C and schooling levels s_T and s_T' such that $C_i \neq C_j$ and $s_T \neq s_T'$.

- $\alpha_{C_i}(s_T)$ and $\beta_{C_i}(s_T)$ can also depend on many other determinants of family and environment.

- Likewise, if we denote by s_T the schooling level at the moment of the test ($s_T = 1, \ldots, \bar{S}_T$), the model for the non-cognitive measure N_i ($i = 1, \ldots, n_N$) is

$$N_i(s_T) = \beta_{N_i}(s_T) X_N + \alpha_{N_i}(s_T) f^C + e_{N_i}(s_T)$$

with $i = 1, \ldots, n_N$, $s_T = 1, \ldots, \bar{S}_T$ and

$$e_{N_i}(s_T) \perp\!\!\!\perp \left(f^N, X_N\right) \text{ and } e_{N_i}(s_T) \perp\!\!\!\perp e_{N_j}(s_T')$$

for all N_i and N_j in N and schooling levels s_T and s_T' such that $N_i \neq N_j$ and $s_T \neq s_T'$.

- $\alpha_{N_i}(s_T)$ and $\beta_{N_i}(s_T)$ can also depend on background features and schooling.

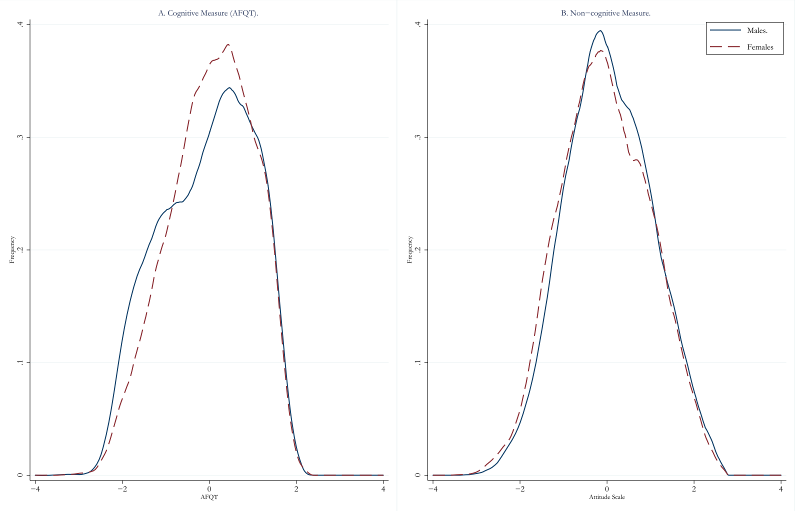
- There are no natural units for latent ability. Therefore, for some C_i ($i = 1, \ldots, n_C$) and N_j ($j = 1, \ldots, n_N$) we set $\alpha_{C_i} = \alpha_{N_i} = 1$.

- This extends traditional factor analysis by having endogenous loadings $\left(\alpha_{C_i}(s_T), \alpha_{N_i}(s_T)\right)$

8 Evidence on the Importance of Cognitive and Noncognitive Skills

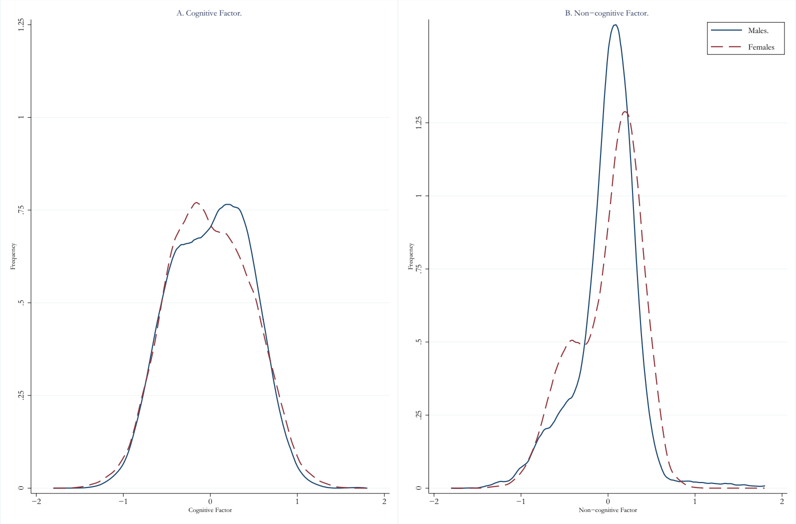
- We use a robust semiparametric approach to estimation.
- We make no distributional assumptions.
- Our evidence argues strongly against normality.
- Male distributions more variable; higher right tail in male cognitive distributions.
- Thicker lower tail for male noncognitive distributions.

Figure 2A. Distribution of Test Scores by Gender



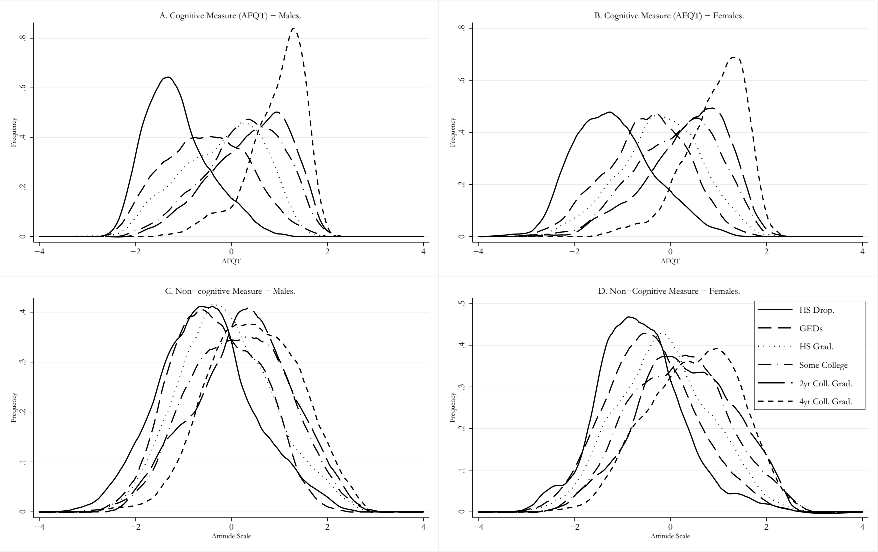
Notes: The AFQT is the mean raw score computed using ASVAB tests.The Attitude Scale is the average raw score between the Rosenberg scale of Self−Steem and the Rotter scale of internal−external locus of control.

Figure 2B. Distribution of Factors by Gender



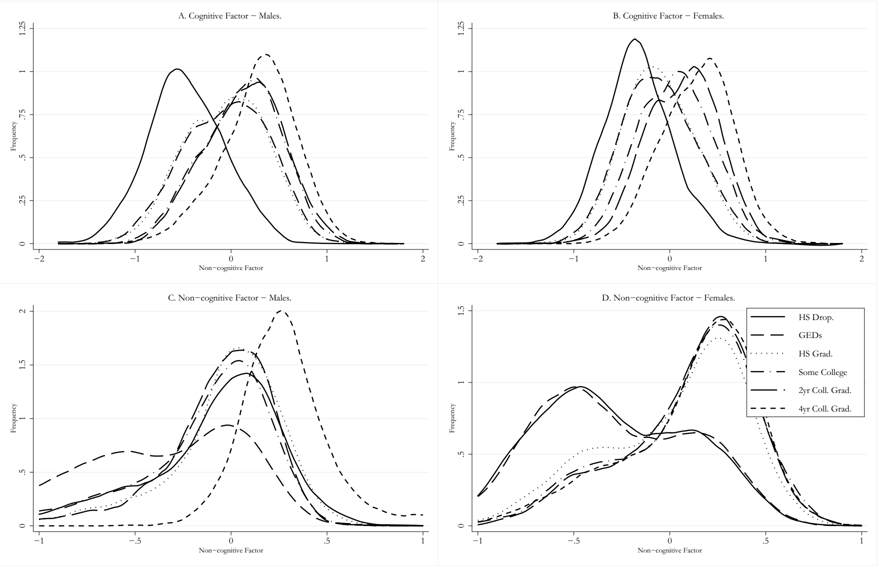
Notes: The factors are simulated from the estimates of the model. The simulated data contain 19,600 observations.

Figure 3A. Distribution of Test Scores by Gender and Schooling Level



Notes: The cognitive measure represents the standardized average over the ASVAB scores (arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations and coding speed). The Noncognitive measure is computed as a (standardized) average of the Rosenberg self−esteem scale and Rotter internal−external locus of control. The schooling levels represent the observed schooling level by age 30 in the NLSY79 sample (See Appendix A for details).

Figure 3B. Distribution of Factors by Gender and Schooling Level

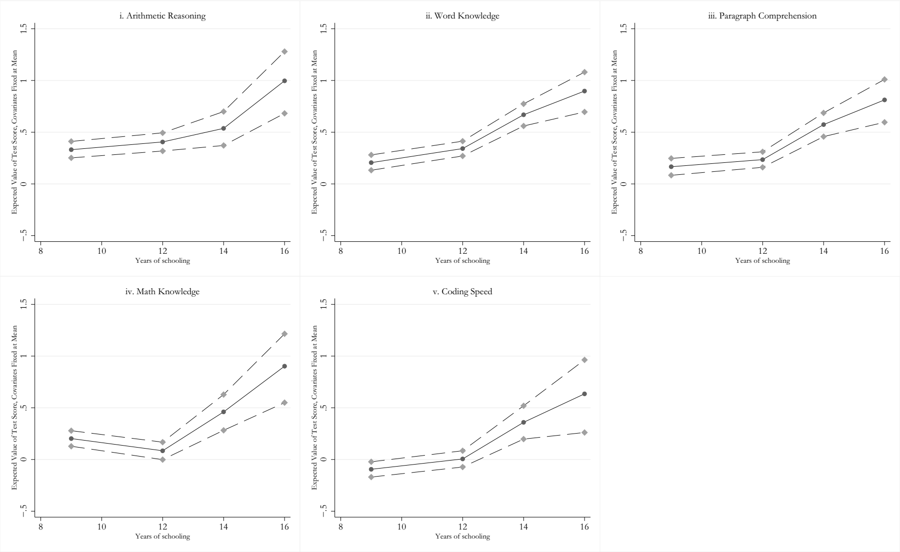


Notes: The factors are simulated from the estimates of the model. The schooling levels represent the predicted schooling level by age 30. These schooling levels are obtained. from the structure and estimates of the model and our sample of the NLSY79 (See Appendix A for details). The simulated data contain 19,600 observations.

The Effect of Schooling on Test Scores

Figure 4A. Effect of schooling on ASVAB Components for person with average ability

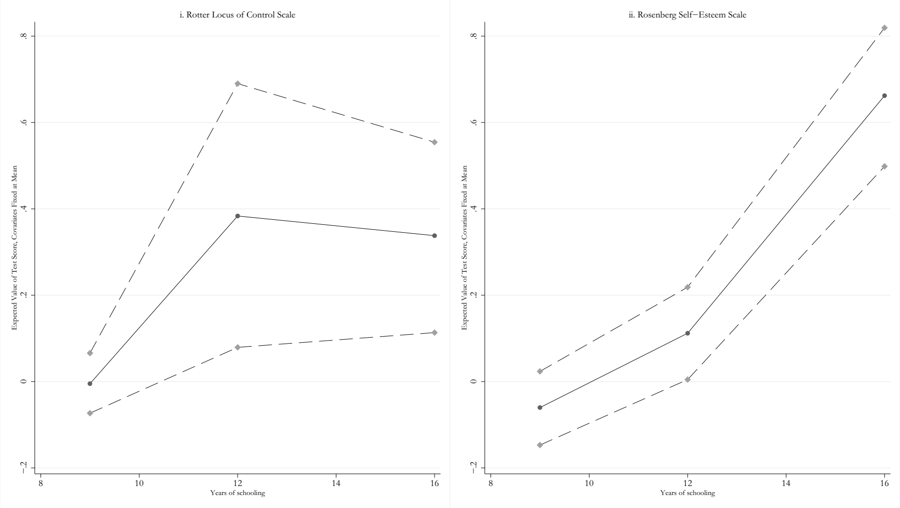
with 95% confidence bands−−Males



Notes: We standardize the test scores to have within−sample mean 0, variance 1. The model is estimated using the Age 30 NLSY79 Sample (See Appendix A for details).

Figure 4B. Effect of schooling on Noncognitive scales for person with average ability

with 95% confidence bands−−Males



Notes: The locus of control scale is based on the four−item abbreviated version of the Rotter Internal−External Locus of Control Scale. This scale is designed to measure the extent to which individuals believe they have control over their lives through self−motivation or self−determination (internal control) as opposed to the extent that the enviroment controls their lives (external control). The Self−Esteem Scale is based on the 10−item Rosenberg Self−Esteem scale. This scale describes a degree of approval or disapproval toward oneself. In both cases, we standardize the test scores to have within−sample mean 0 and variance 1, after taking averages over the respective sets of scales. The model is estimated using the Age 30 NLSY79 Sample (See Appendix A for details).

Evidence From
The Semiparametric Model

Results for Wages

Figure 5A. Mean Log Wages by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

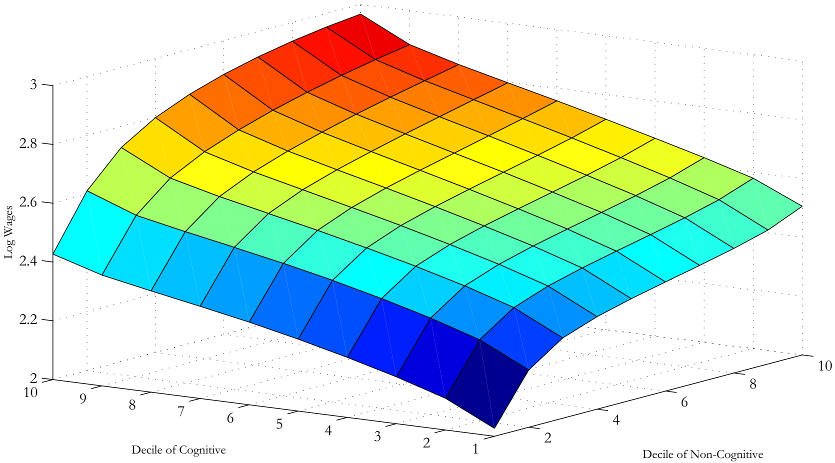
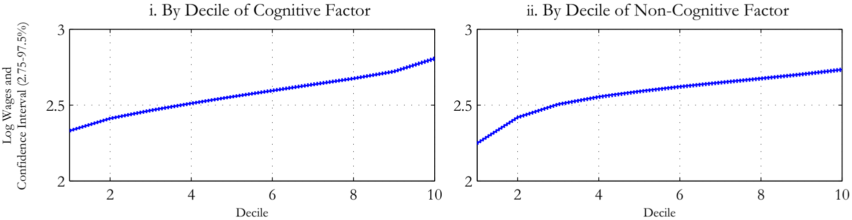


Figure 5B. Mean Log Wages by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 6A. Mean Log Wages by Age 30 - Females
i. By Decile of Cognitive and Non-Cognitive Factors

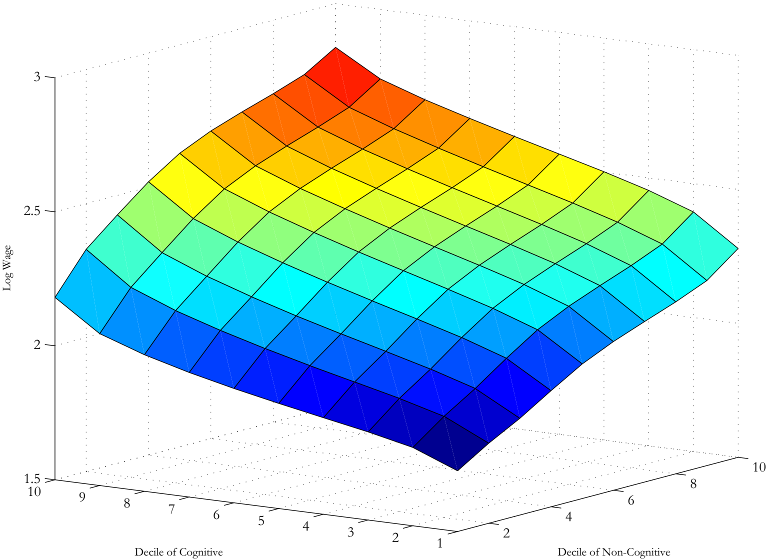
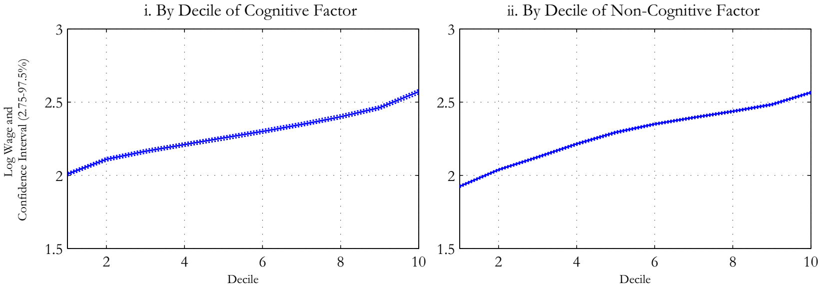


Figure 6B. Mean Log Wages by Age 30 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Results for Wages

By Schooling Level

(Hedonic Markets)

Figure 7A. Mean Log Wages of High School Dropouts by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

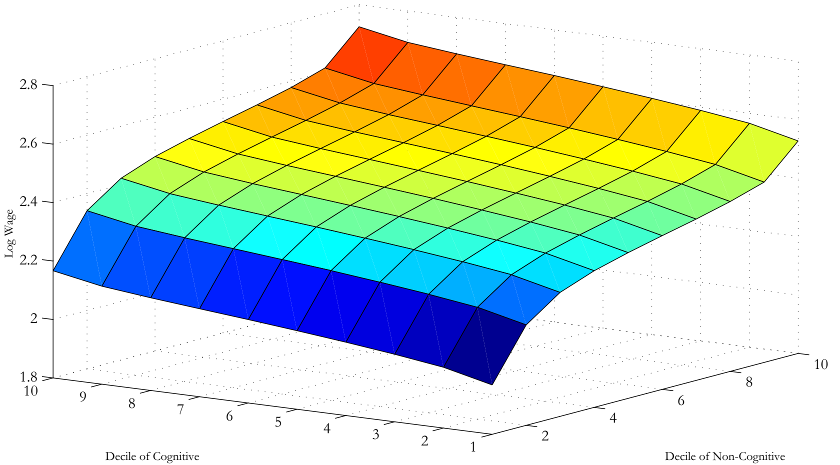
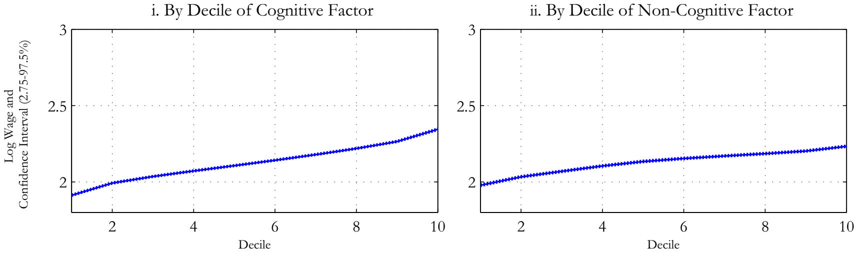


Figure 7B. Mean Log Wages of High School Dropouts by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 8A. Mean Log Wages of GEDs by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

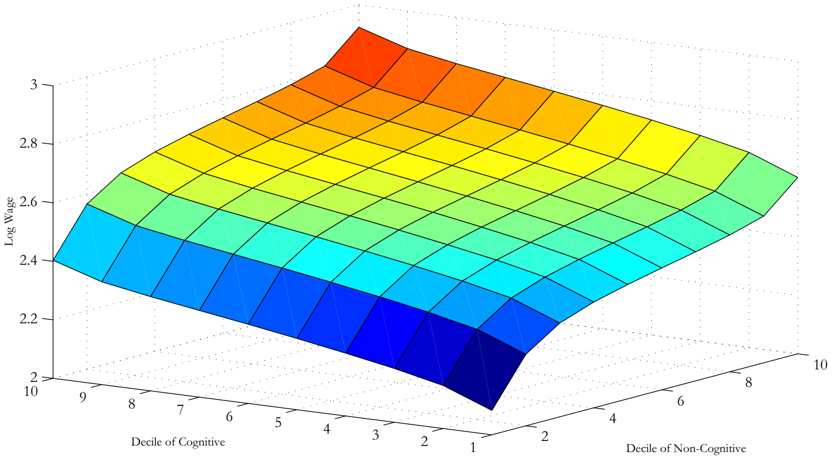
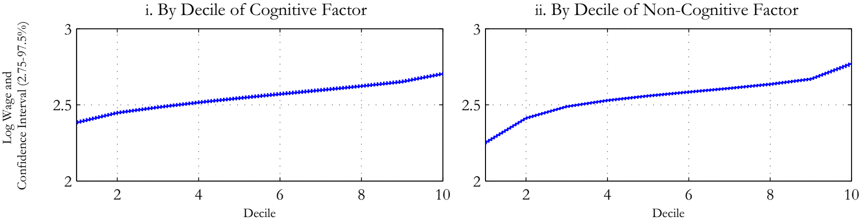


Figure 8B. Mean Log Wages of GEDs by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 9A. Mean Log Wages of GEDs by Age 30 - Females
i. By Decile of Cognitive and Non-Cognitive Factors

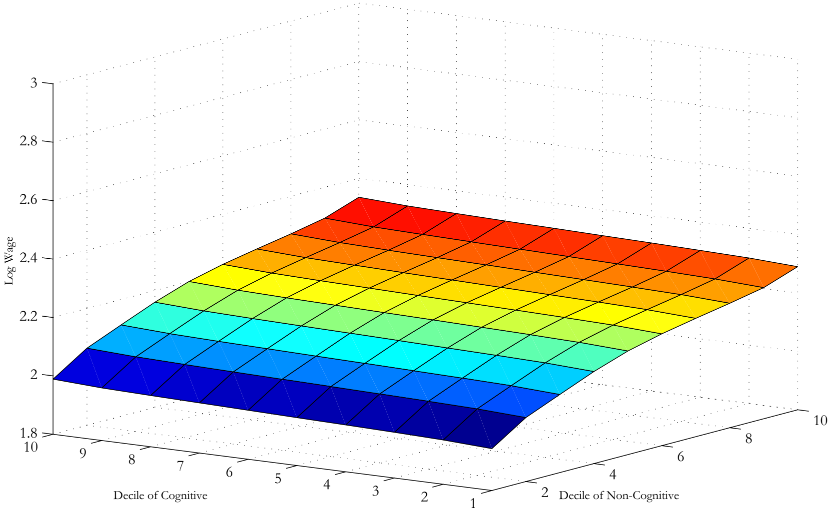
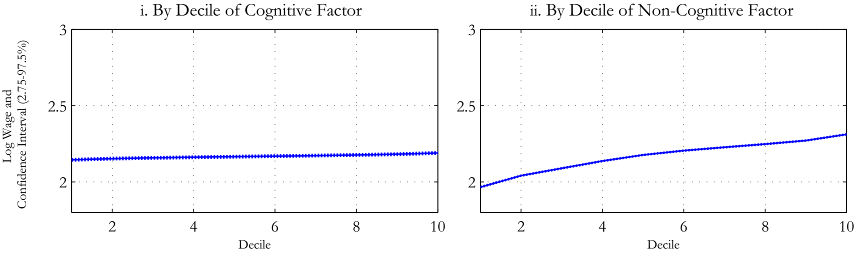


Figure 9B. Mean Log Wages of GEDs by Age 30 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 10A. Mean Log Wages of High School Graduates by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

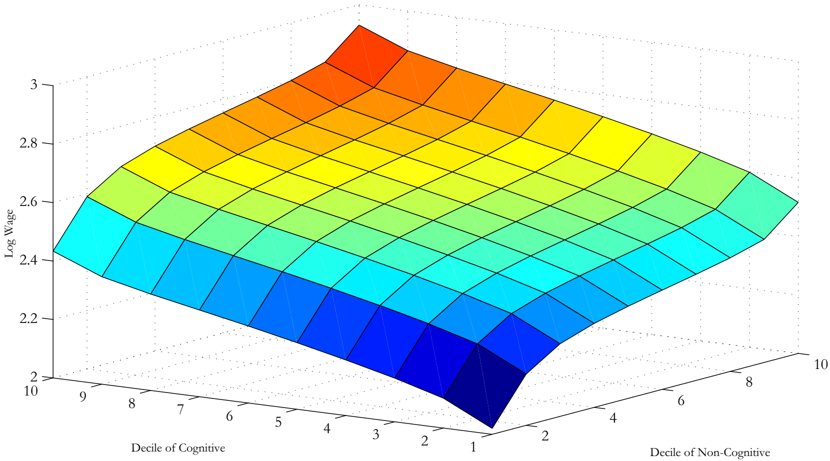
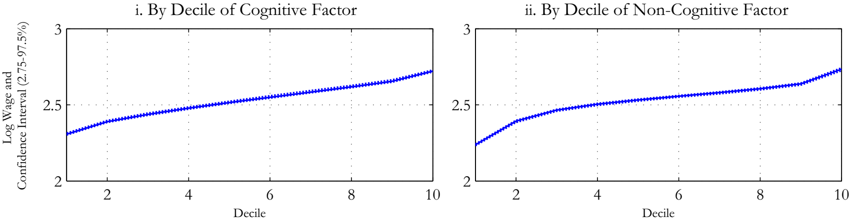


Figure 10B. Mean Log Wages of High School Graduates by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 11A. Mean Log Wages of 2-yr College Graduates by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

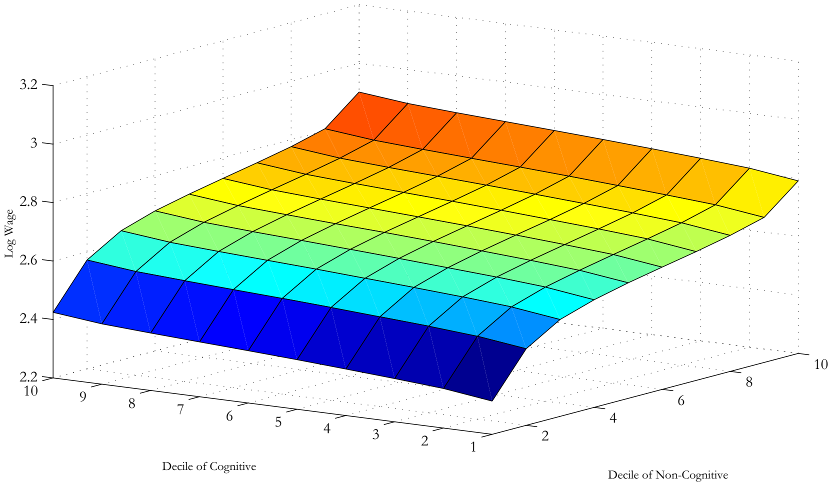
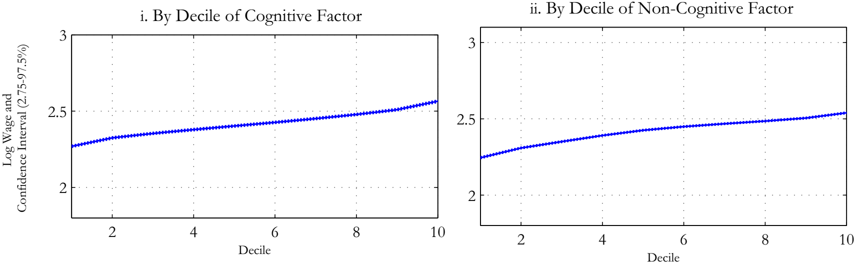


Figure 11B. Mean Log Wages of 2-yr College Graduates by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 12A. Mean Log Wages of 4-yr College Graduates by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

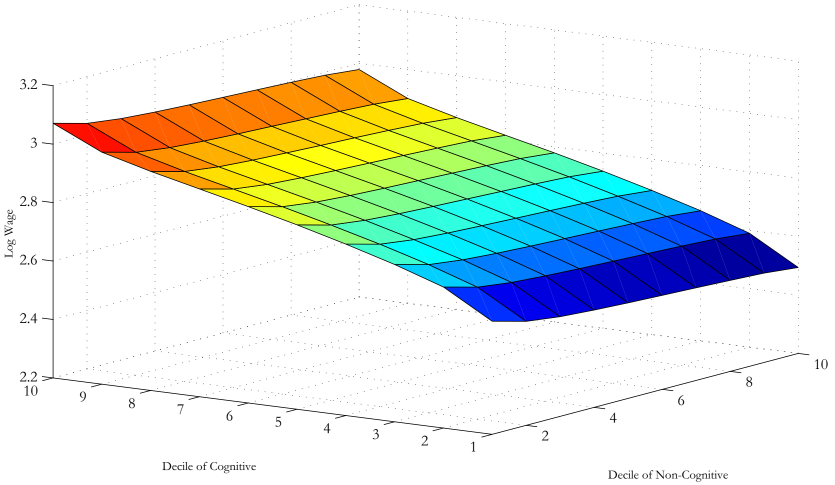
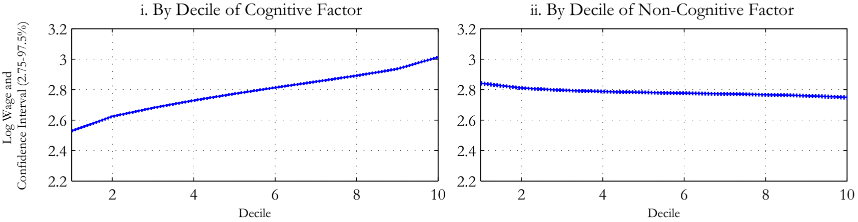


Figure 12 B. Mean Log Wages of 4-yr College Graduates by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 13 A. Mean Log Wages of 4-yr College Graduates by Age 30 - Females
i. By Decile of Cognitive and Non-Cognitive Factors

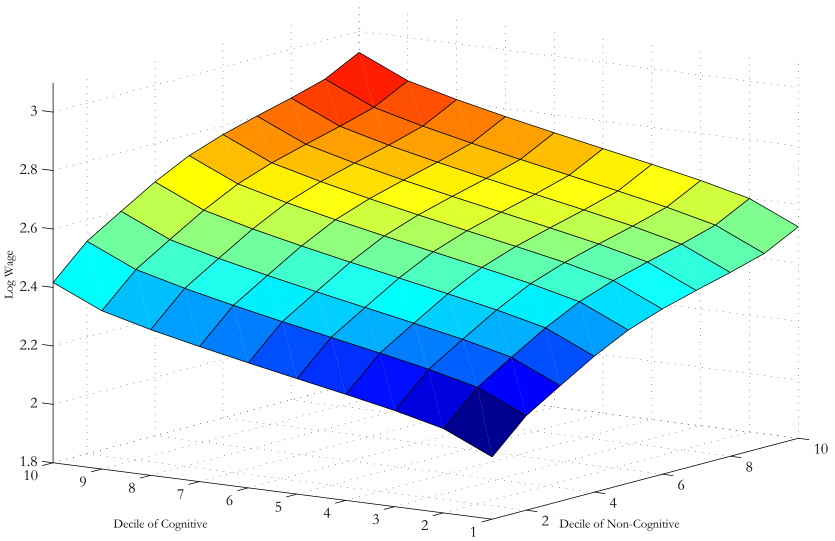
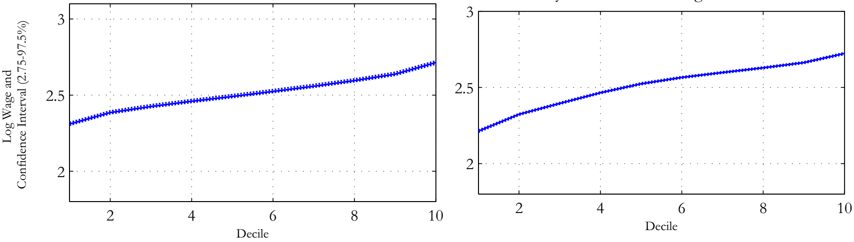


Figure 13 B. Mean Log Wages of 4-yr College Graduates by Age 30 - Females

i. By Decile of Cognitive Factor

ii. By Decile of Non-Cognitive Factor



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Results for Other Outcomes

Figure 14A. Probability of Employment by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factor

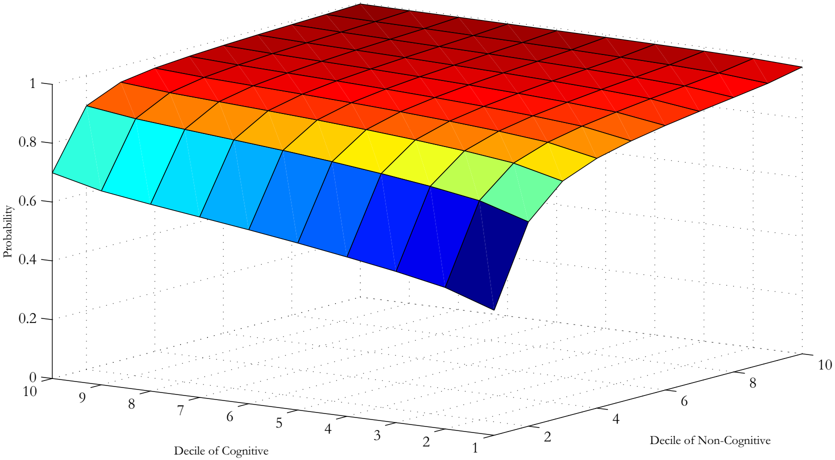
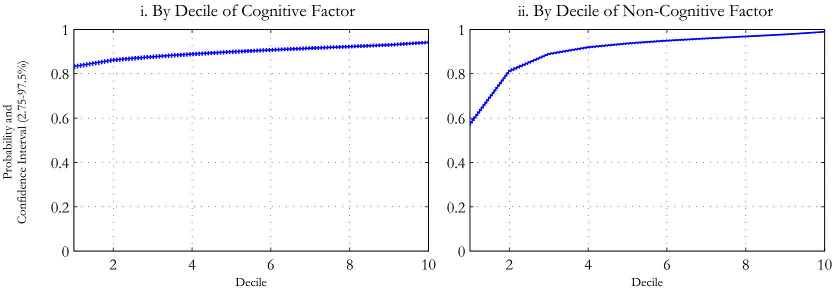


Figure 14B. Probability of Employment by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 15A. Probability of Employment by Age 30 - Females
i. By Decile of Cognitive and Non-Cognitive Factor

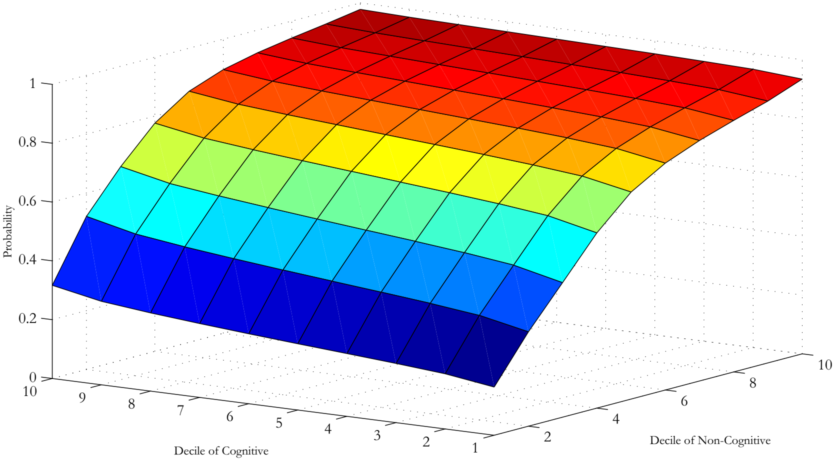
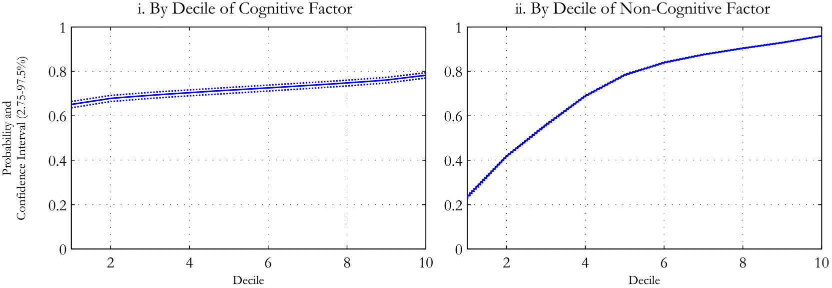


Figure 15B. Probability of Employment by Age 30 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 16A. Mean Work Experience of High School Dropouts by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

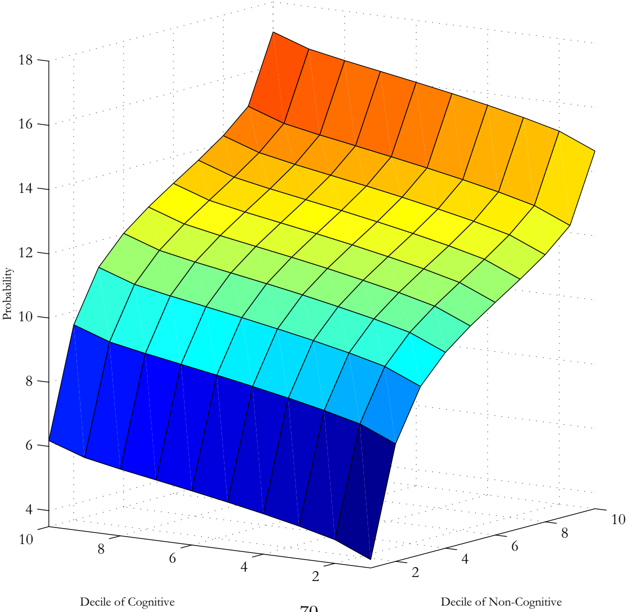
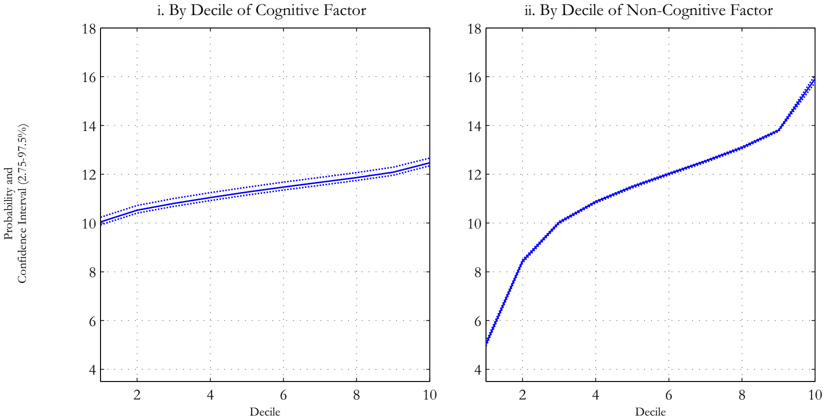


Figure 16B. Mean Work Experience of High School Dropouts by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 17A. Mean Work Experience of High School Graduates by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

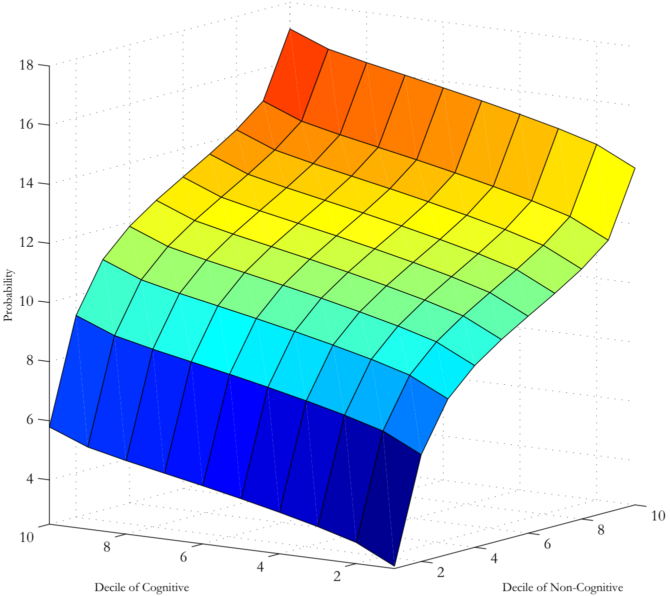
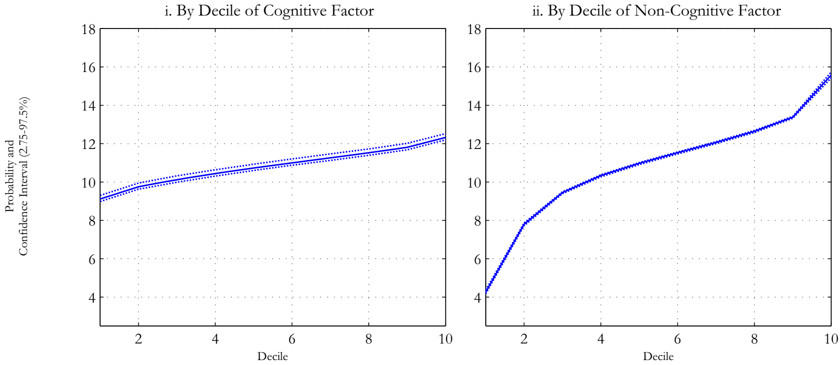


Figure 17B Mean Work Experience of High School Graduates by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 18A. Mean Work Experience of 4-yr College Graduates by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

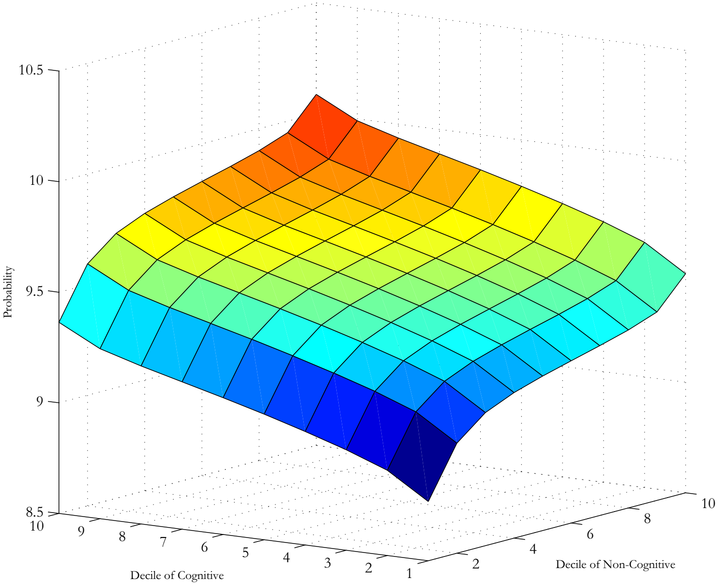
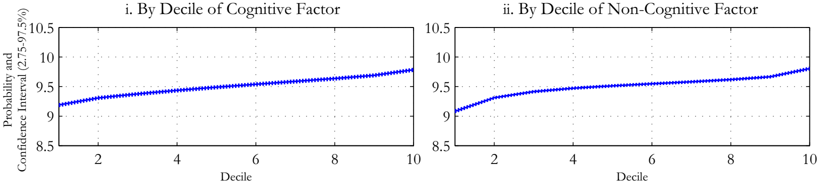


Figure 18B. Mean Work Experience of 4-yr College Graduates by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 19A. Mean Work Experience of 4-yr College Graduates by Age 30 - Females
i. By Decile of Cognitive and Non-Cognitive Factors

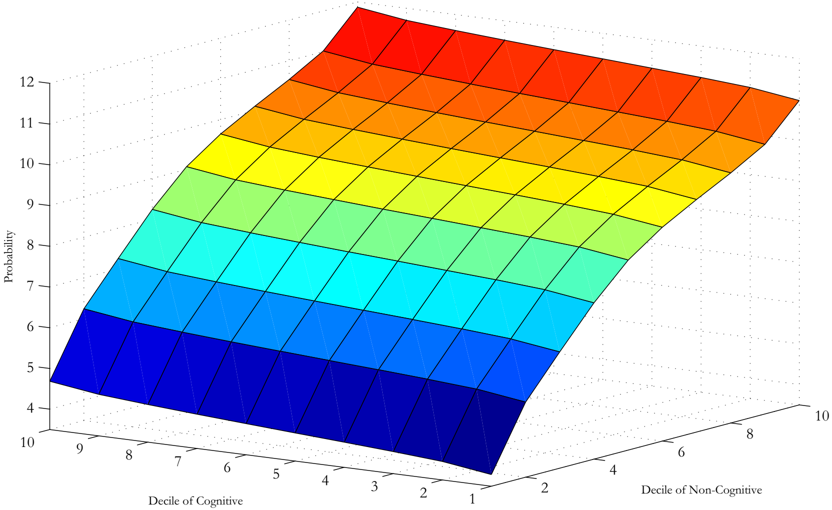
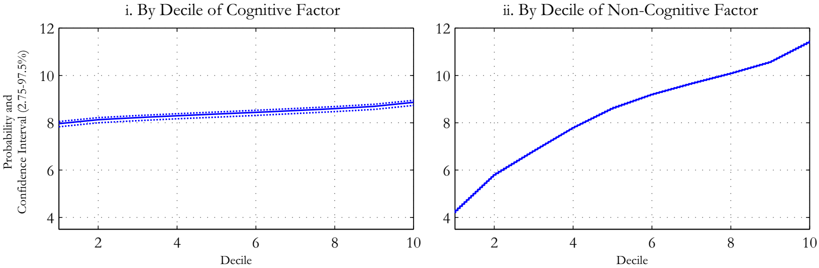


Figure 19B. Mean Work Experience of 4-yr College Graduates by Age 30 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 20A. Probability Of Being a White Collar Worker by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factor

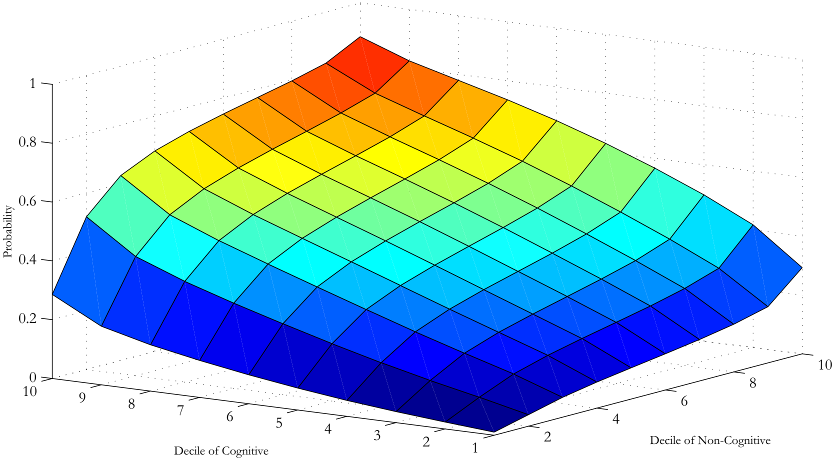
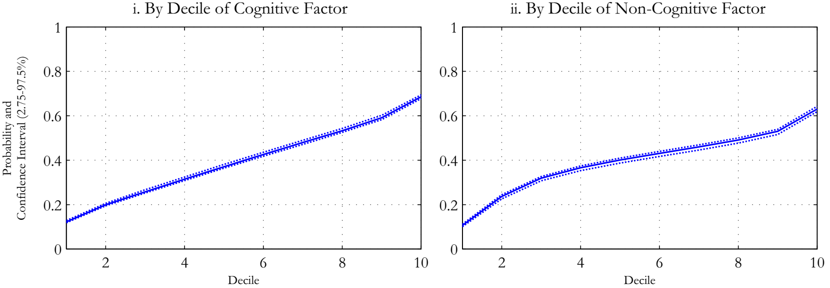


Figure 20B. Probability Of Being a White Collar Worker by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 21A. Probability of Being a High School Dropout by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

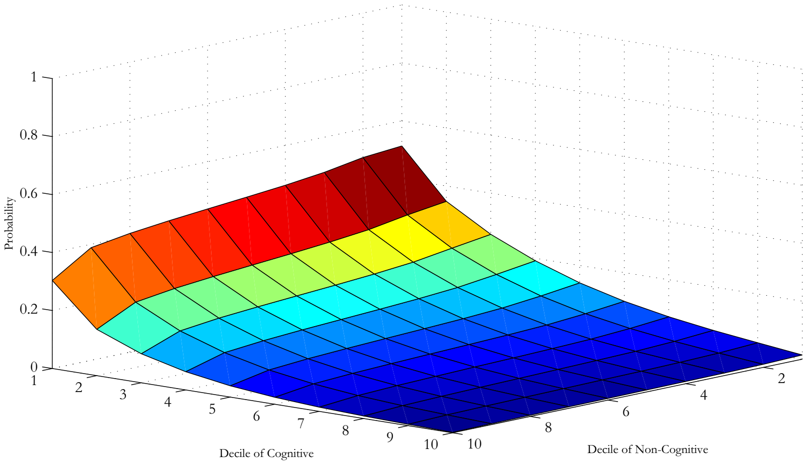
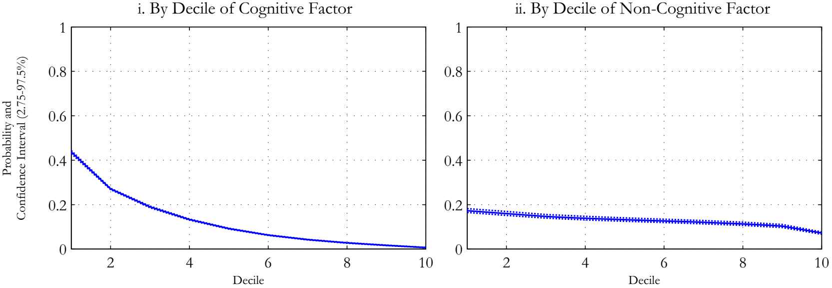


Figure 21B. Probability of Being a High School Dropout by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 22 A. Probability of Being a GED by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

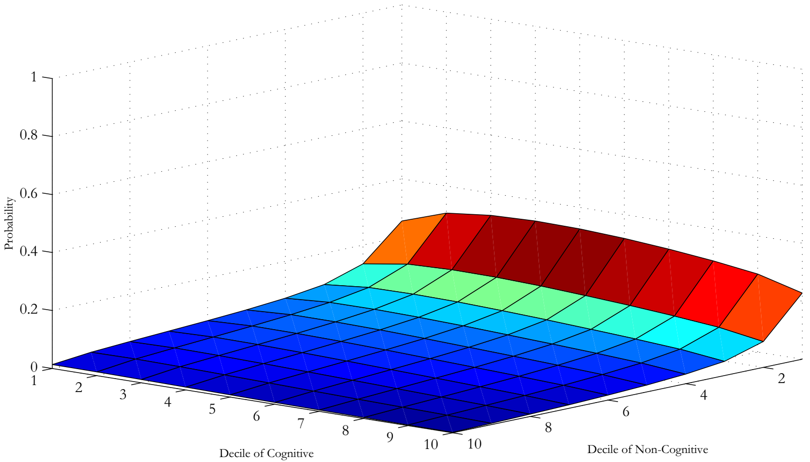
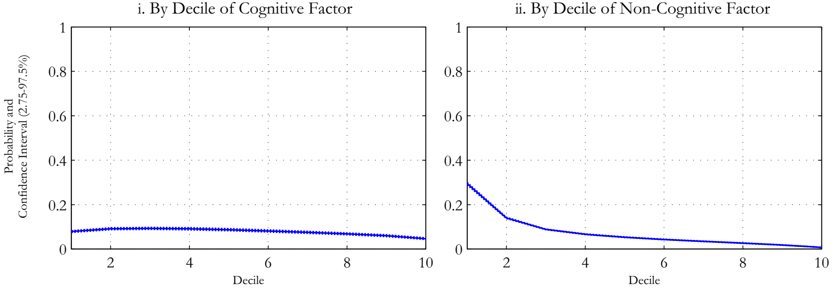


Figure 22 B. Probability of Being a GED by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 23 A. Probability of Being a 2-yr College Graduate by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

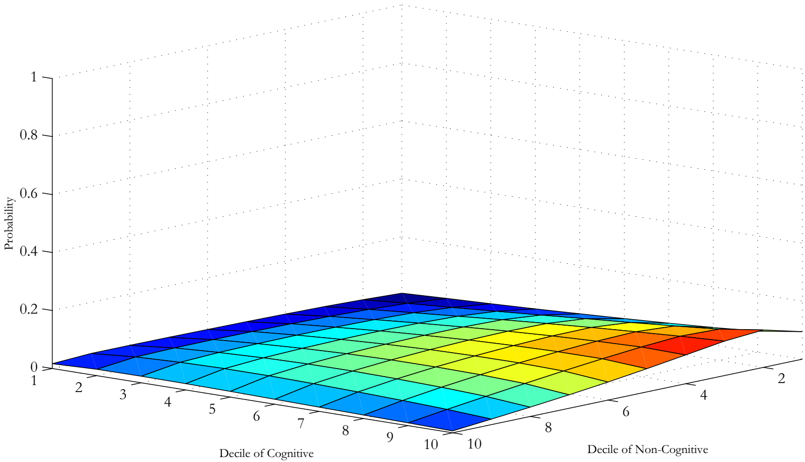
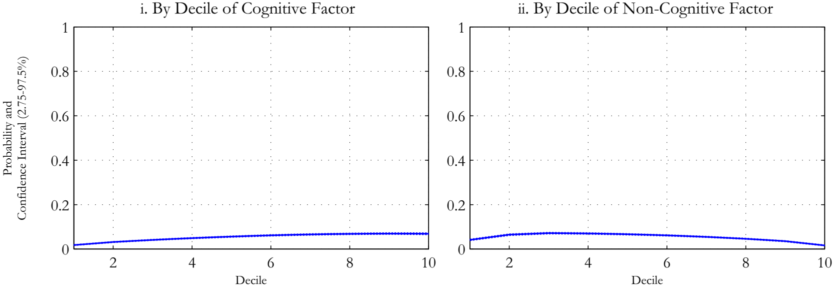


Figure 23 B. Probability of Being a 2-yr College Graduate by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 24 A. Probability of Being a 2-yr College Graduate by Age 30 - Females
i. By Decile of Cognitive and Non-Cognitive Factors

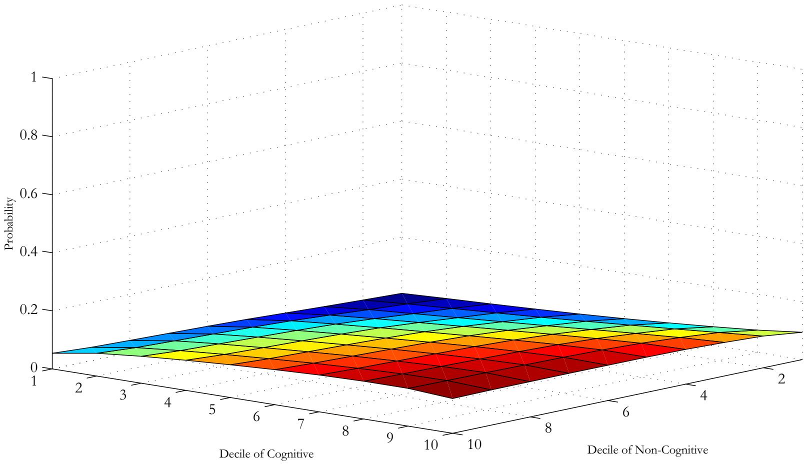
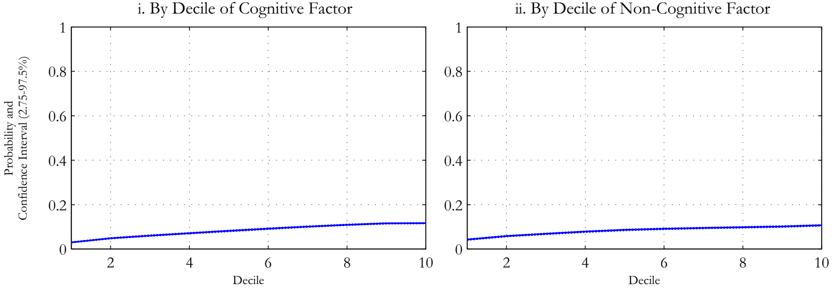


Figure 24B . Probability of Being a 2-yr College Graduate by Age 30 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 25A. Probability of Being a 4-yr College Graduate by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factors

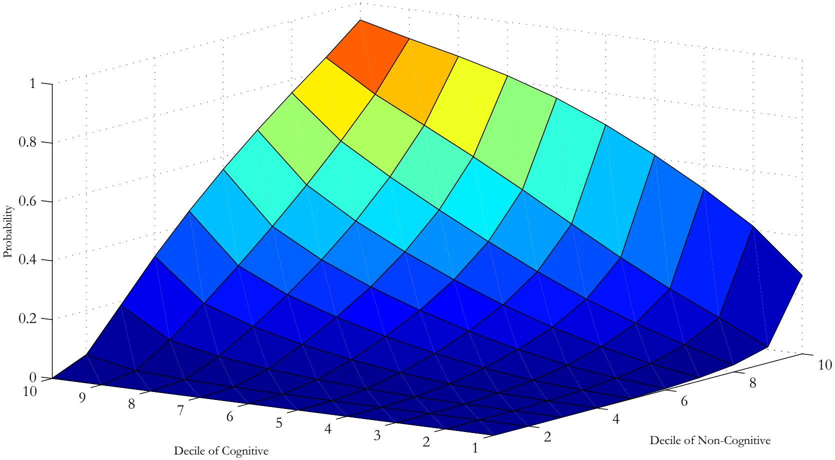
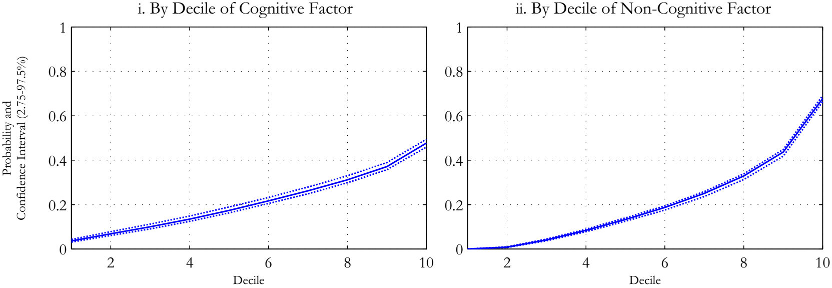


Figure 25 B. Probability of Being a 4-yr College Graduate by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 26 A. Probability of Being a 4-yr College Graduate by Age 30 - Females
i. By Decile of Cognitive and Non-Cognitive Factors

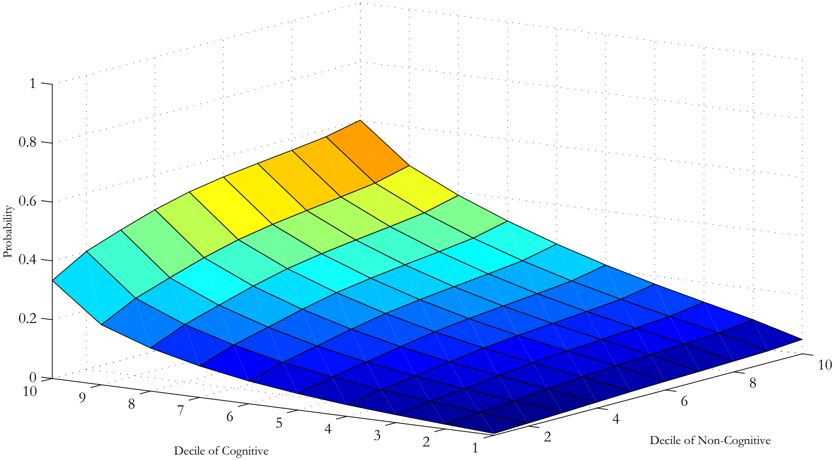
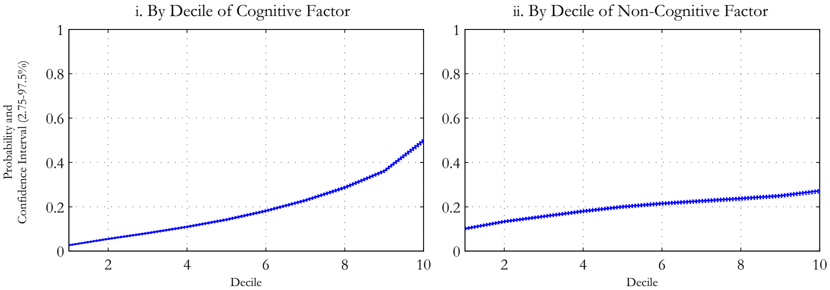


Figure 26 B. Probability of Being a 4-yr College Graduate by Age 30 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 27 A. Probability Of Daily Smoking By Age 18 - Males
i. By Decile of Cognitive and Non-Cognitive Factor

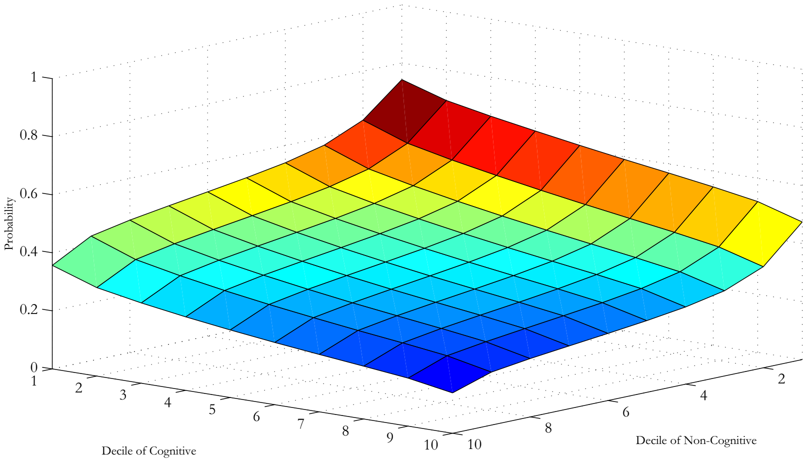
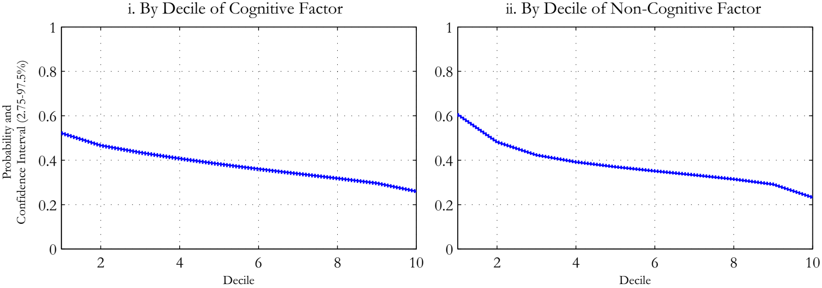


Figure 27 B. Probability Of Daily Smoking By Age 18 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 28 A. Probability Of Daily Smoking By Age 18 - Females
i. By Decile of Cognitive and Non-Cognitive Factor

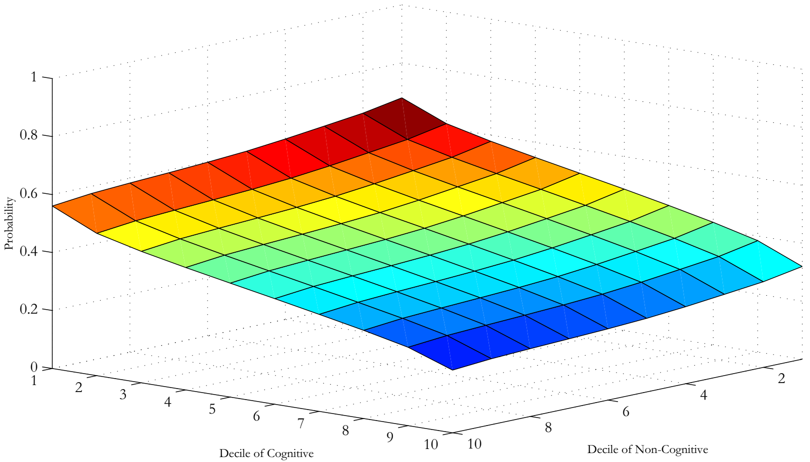
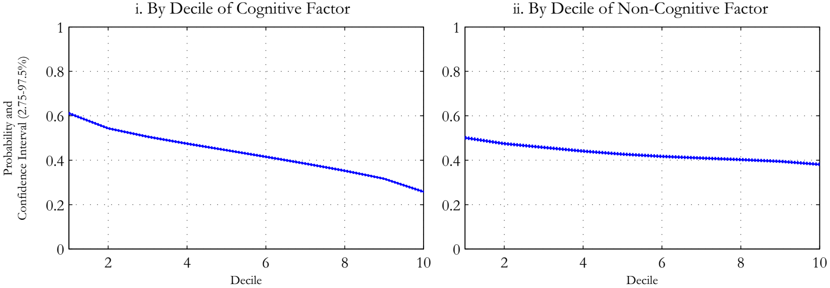


Figure 28 B. Probability Of Daily Smoking By Age 18 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 29 A. Probability of Smoking Marijuana during the Year 1979 - Males
i. By Decile of Cognitive and Non-Cognitive Factor

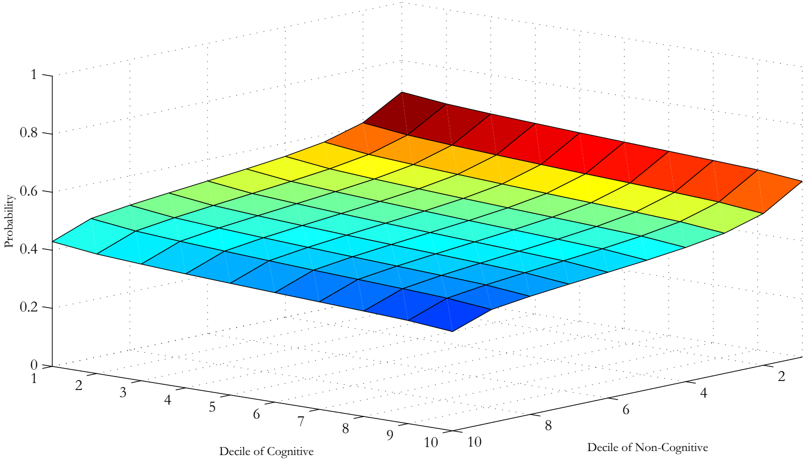
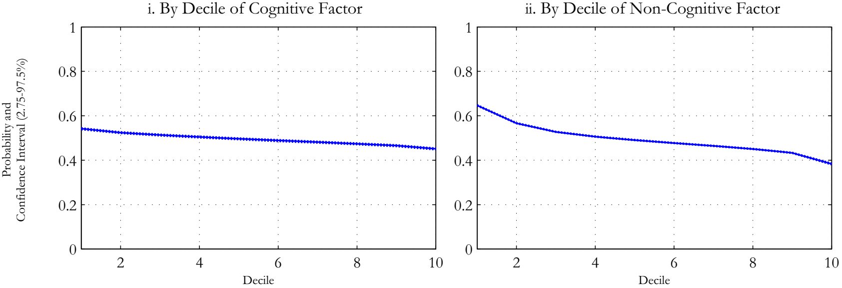


Figure 29 B Probability of Smoking Marijuana during the Year 1979 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 30 A. Probability of Smoking Marijuana during the Year 1979 - Females
i. By Decile of Cognitive and Non-Cognitive Factor

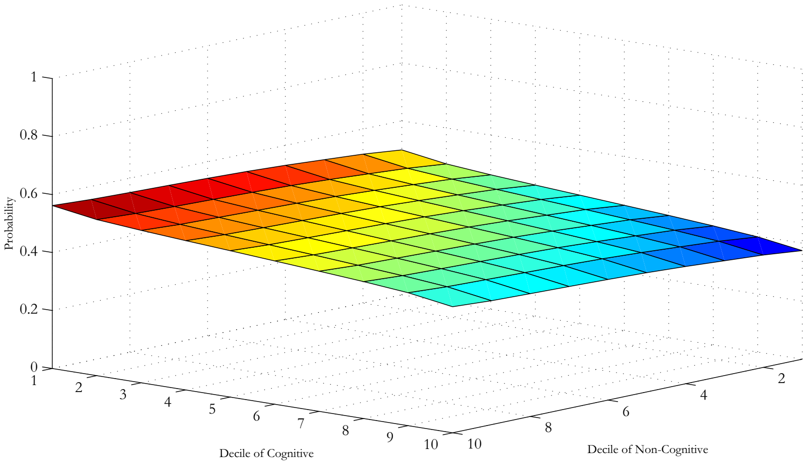
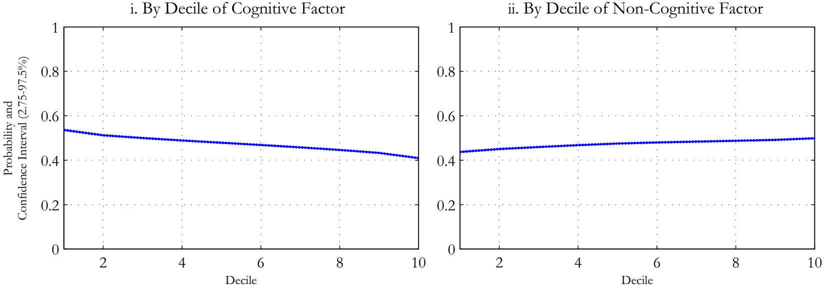


Figure 30 B. Probability of Smoking Marijuana during the Year 1979 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 31 A. Probability of Participating in Illegal Activities during the Year 1979- Males
i. By Decile of Cognitive and Non-Cognitive Factor

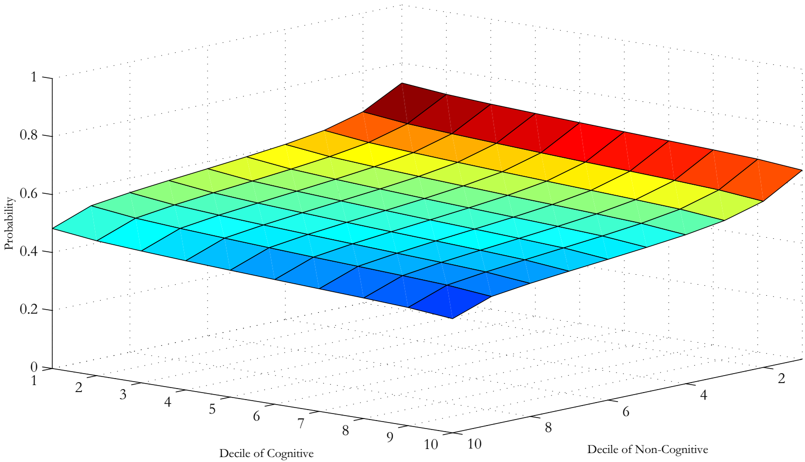
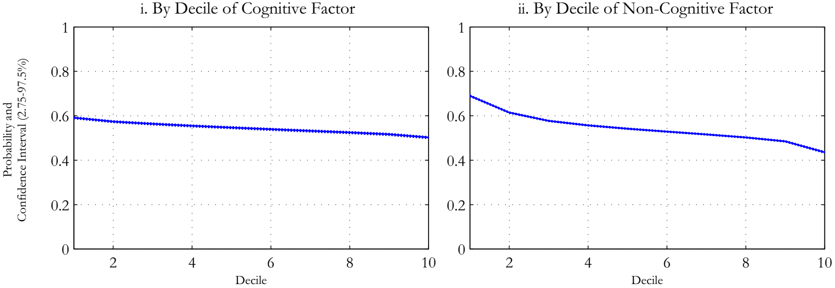


Figure 31 B. Probability of Participating in Illegal Activities during the Year 1979- Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 32 A. Probability of Incarceration by Age 30 - Males
i. By Decile of Cognitive and Non-Cognitive Factor

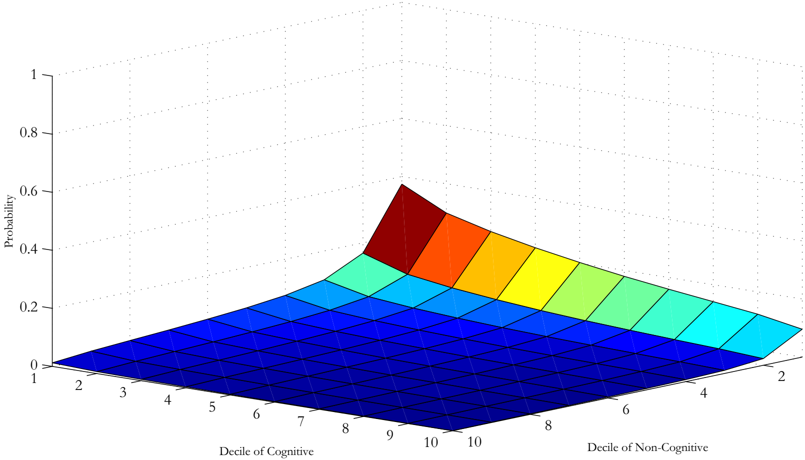
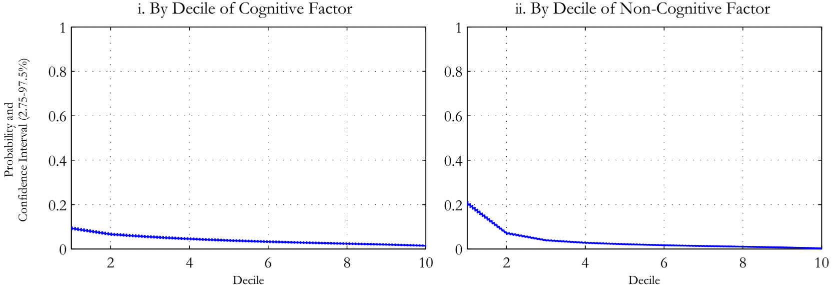


Figure 32B . Probability of Incarceration by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

Figure 33 A. Probability Of Being Single With Child - Females
i. By Decile of Cognitive and Non-Cognitive Factors

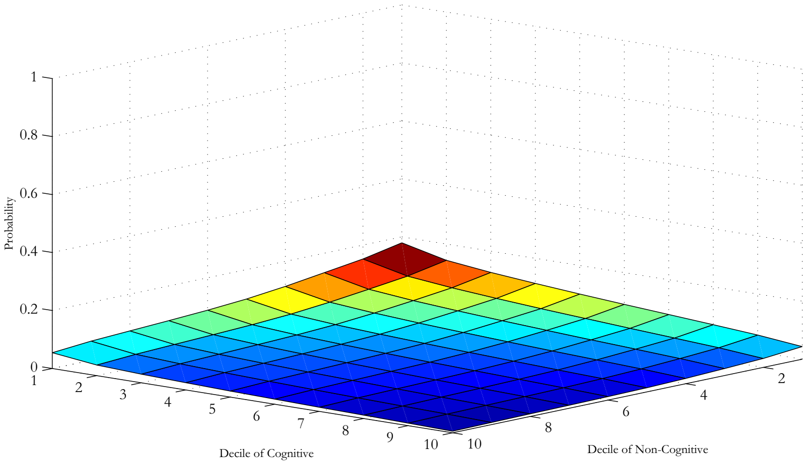
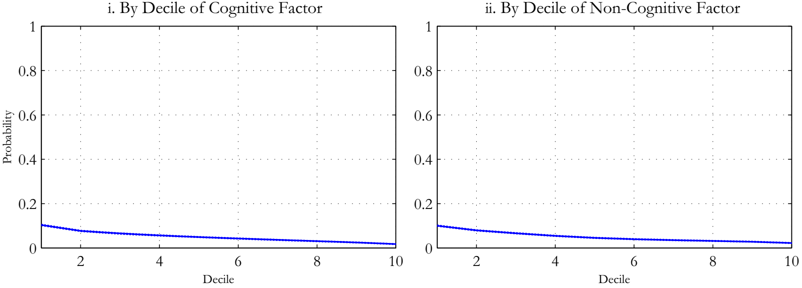


Figure 33 B. Probability Of Being Single With Child - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (50 draws).

9 How Important is the assumption that $f^C \perp\!\!\!\perp f^N$?

- First, observed cognitive and noncognitive test scores can be highly correlated even if factors are not

  (through $\beta_N(s_T, X)$, $\beta_C(s_T, X)$, $\alpha_N(s_T, X)$, $\alpha_C(s_T, X)$).

- C and N are not highly correlated.

- Adjusting for background, the correlation weakens greatly.

10 Where Do Skills Come From?: The Technology of Skill Formation

(Cunha and Heckman, 2005: Technology of Skill Formation, first draft, 2003).

- Using CNLSY data, we estimate determinants of cognitive and noncognitive skills over the life cycle.
- f_t^C denotes the cognitive factor at period t.
- f_t^N denotes the noncognitive factor at period t.
- f_t^{IC} denotes the investment in the cognitive skills at period t.
- f_t^{IN} denotes the investments in noncognitive skills at period t.

- The dynamic factor model is described by:

$$f^C_{t+1} = \gamma^C_1 f^C_t + \gamma^C_2 f^N_t + \left(1 - \gamma^C_1 - \gamma^C_2\right) f^{IC}_t + \eta^C_{t+1}$$

$$f^N_{t+1} = \gamma^N_1 f^C_t + \gamma^N_2 f^N_t + \left(1 - \gamma^N_1 - \gamma^N_2\right) f^{IN}_t + \eta^N_{t+1}$$

$$f^{HOME}_{t+1} = \gamma^{IC}_1 f^C_t + \gamma^{IC}_2 f^N_t + \gamma^I f^{HOME}_t + \eta^{IC}_{t+1}$$

- The estimated equation coefficients are (using CNLSY data):

$$f^C_{t+1} = 0.516 f^C_t + 0.483 f^N_t + 0.001 f^{IC}_t + \eta^C_{t+1},\ var\left(\eta^C_{t+1}\right) = 0.036$$

$$f^N_{t+1} = 0.98 f^N_t + 0.02 f^{IN}_t + \eta^N_{t+1},\ var\left(\eta^N_{t+1}\right) = 0.00184$$

$$\begin{aligned} f^{HOME}_{t+1} &= -0.01 f^C_t + 0.036 f^N_t + 0.8074 f^{HOME}_t + \eta^{HOME}_{t+1}, \\ var\left(\eta^{HOME}_{t+1}\right) &= 0.0042 \end{aligned}$$

- So, more noncognitive skill today increases the stock of cognitive skills tomorrow, but the reverse effect of cognitive skills on noncognitive skills is practically nonexistent.

Table 2A

Correlation Matrix
Dynamic Factor Model - White Children / CNLSY-1979
Initial Covariance - Assumed

	Cognitive	Noncognitive	Home
Cognitive	1.0000	0.0000	0.0000
Noncognitive	0.0000	1.0000	0.0000
Home	0.0000	0.0000	1.0000

Correlation Matrix
Dynamic Factor Model - White Children / CNLSY-1979
Period 2 = Children aged between 7 and 8

	Cognitive	Noncognitive	Home
Cognitive	1.0000	0.1370	0.0023
Noncognitive	0.1370	1.0000	0.0341
Home	0.0023	0.0341	1.0000

Table 2B

Correlation Matrix Dynamic Factor Model - White Children / CNLSY-1979 Period 3 = Children aged between 9 and 10			
	Cognitive	Noncognitive	Home
Cognitive	1.0000	0.0992	0.0016
Noncognitive	0.0992	1.0000	0.0313
Home	0.0016	0.0313	1.0000

Correlation Matrix Dynamic Factor Model - White Children / CNLSY-1979 Period 4 = Children aged between 11 and 12			
	Cognitive	Noncognitive	Home
Cognitive	1.0000	0.0879	0.0012
Noncognitive	0.0879	1.0000	0.0295
Home	0.0012	0.0295	1.0000

Table 2C

Covariance Matrix
Dynamic Factor Model - White Children / CNLSY-1979
Period 5 = Children aged between 13 and 14

	Cognitive	Noncognitive	Home
Cognitive	1.0000	0.0848	0.0010
Noncognitive	0.0848	1.0000	0.0288
Home	0.0010	0.0288	1.0000

11 Conclusion

- Low dimensional model for two latent abilities explains a diverse array of behaviors controlling for reverse causality and selection

- We move beyond looking only at effects of cognitive and noncognitive skills on wages.

- For many dimensions of behavior, noncognitive ability is more important than, or as important as (in the sense of effects of movements from the top to the bottom of the distribution) cognitive ability.

- Noncognitive ability affects acquisition of skills and a variety of behaviors as well as market productivity as measured by wages.

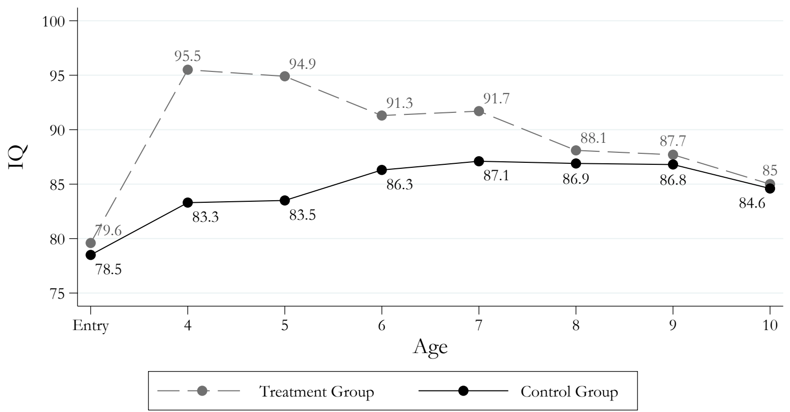
- Cognitive ability affects market productivity, skill acquisition and a variety of behaviors.

- Schooling affects both cognitive and noncognitive skills.
- Existence of multiple skills alters signalling theory which is based on assuming a single ability.
- Single crossing property is violated.
- Araujo, Gottlieb and Moreira (2004) developed this theory in response to our evidence on the GED.
- They explore implications of the GED as a mixed signal. GEDs have higher cognitive skills than dropouts, but lower noncognitive skills than graduates.
- One interpretation of high "psychic costs" found in the recent literature, is that it represents noncognitive ability.
- High psychic costs explain sluggish response of schooling to increases in wages.
- Race differences. Evidence that noncognitive components are very important in determining the wages of blacks.

- Some evidence that multiple noncognitive factors required to fit the data.
- Cunha and Heckman (2006) relax independence of factors in a dynamic model of skill formation.
- They show that noncognitive skills promote cognitive skill formation but not vice versa.

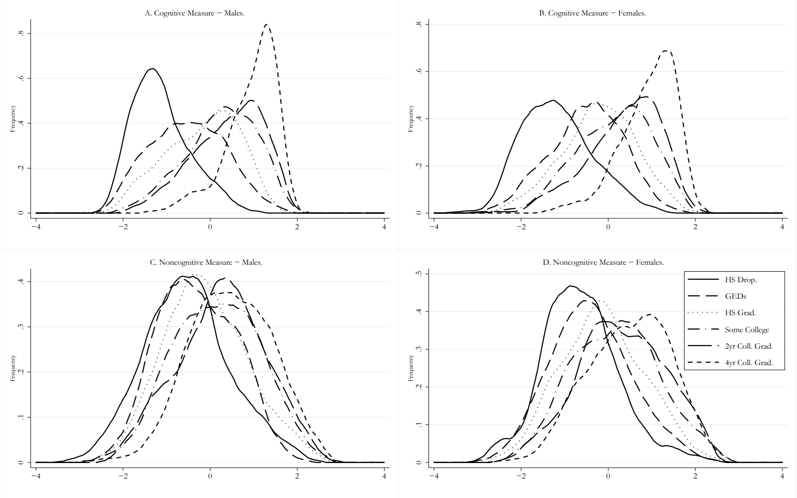
Figure 1

Perry Preschool Program: IQ, by Age and Treatment Group



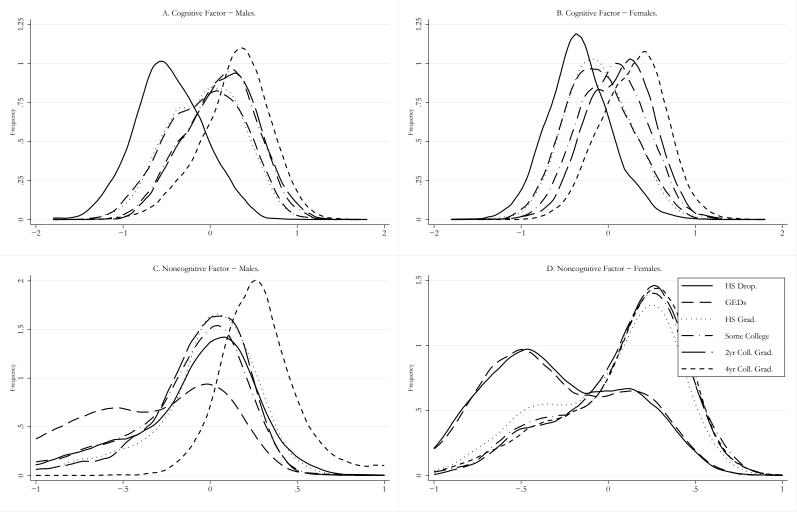
Source: Perry Preschool Program. IQ measured on the Stanford−Binet Intelligence Scale (Terman & Merrill, 1960).
Test was administered at program entry and each of the ages indicated.

Figure 2. Distribution of Test Scores by Gender and Schooling Level



Notes: The cognitive measure represents the standardized average over the ASVAB scores (arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations and coding speed). The noncognitive measure is computed as a (standardized) average of the Rosenberg Self−Esteem Scale and Rotter Internal−External Locus of Control Scale. The schooling levels represent the observed schooling level by age 30 in the NLSY79 sample (See Web Appendix A for details).

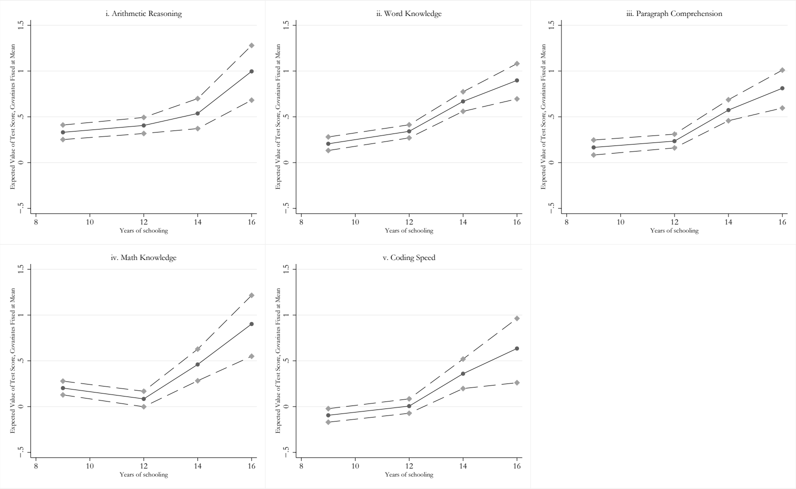
Figure 3. Distribution of Factors by Gender and Schooling Level



Notes: The factors are simulated from the estimates of the model. The schooling levels represent the predicted schooling level by age 30. These schooling levels are obtained from the structure and estimates of the model and our sample of the NLSY79 (See Web Appendix A for details). The simulated data contain 19,600 observations.

Figure 4A. Effect of schooling on ASVAB Components for person with average ability

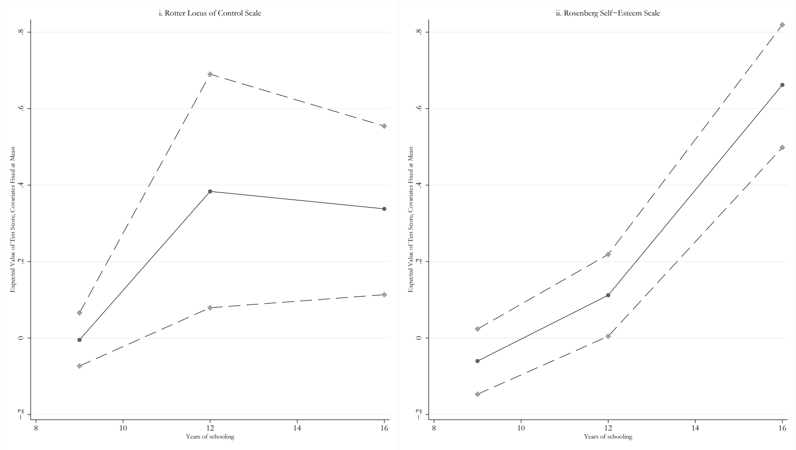
with 95% confidence bands−−Males



Notes: We standardize the test scores to have within−sample mean 0, variance 1. The model is estimated using the Age 30 NLSY79 Sample (See Web Appendix A for details).

Figure 4B. Effect of schooling on Noncognitive scales for person with average ability

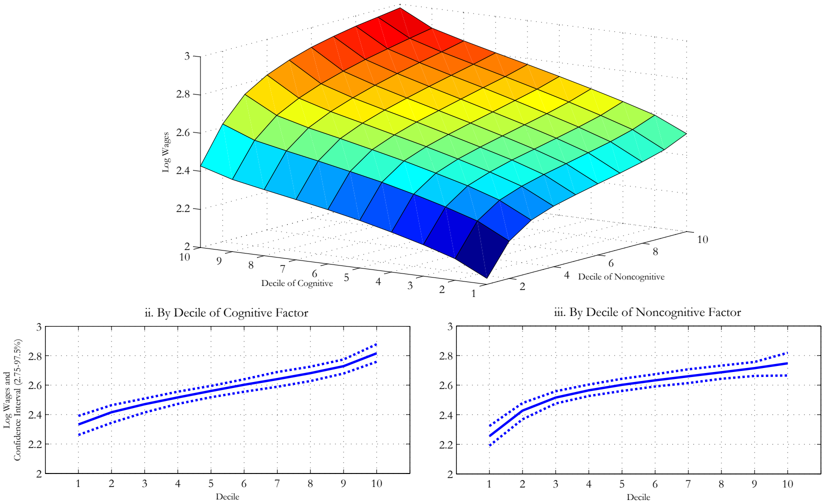
with 95% confidence bands−−Males



Notes: The locus of control scale is based on the four−item abbreviated version of the Rotter Internal−External Locus of Control Scale. This scale is designed to measure the extent to which individuals believe they have control over their lives through self−motivation or self−determination (internal control) as opposed to the extent that the enviroment controls their lives (external control). The self−esteem scale is based on the 10−item Rosenberg Self−Esteem Scale. This scale describes a degree of approval or disapproval toward oneself. In both cases, we standardize the test scores to have within−sample mean 0 and variance 1, after taking averages over the respective sets of scales. The model is estimated using the Age 30 NLSY79 Sample (See Web Appendix A for details).

Figure 5A. Mean Log Wages by Age 30 - Males

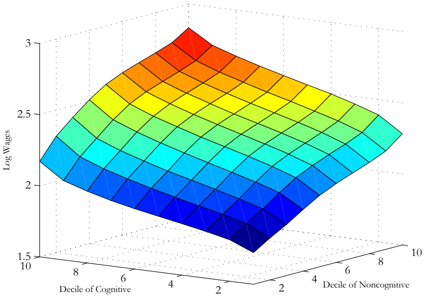
i. By Decile of Cognitive and Noncognitive Factors



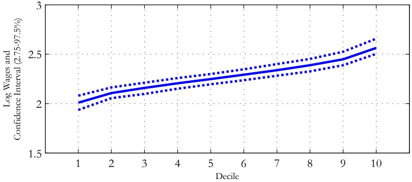
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

Figure 5B. Mean Log Wages by Age 30 - Females

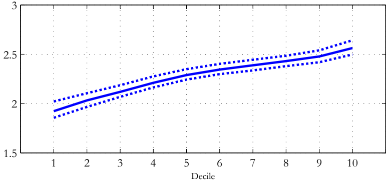
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor



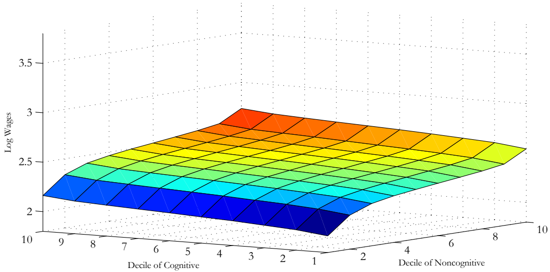
iii. By Decile of Noncognitive Factor



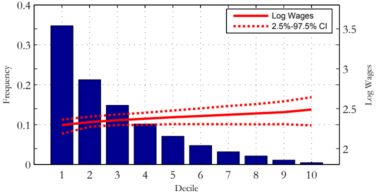
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 6A. Mean Log Wages of High School Dropouts by Age 30 - Males

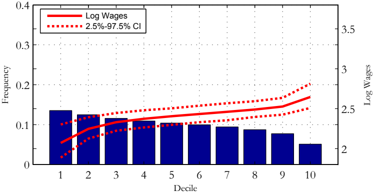
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor



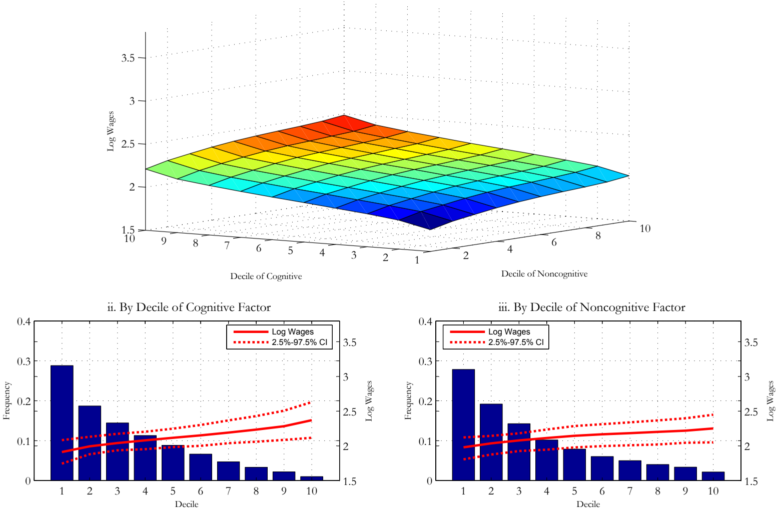
iii. By Decile of Noncognitive Factor



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

Figure 6B. Mean Log Wages of High School Dropouts by Age 30 - Females

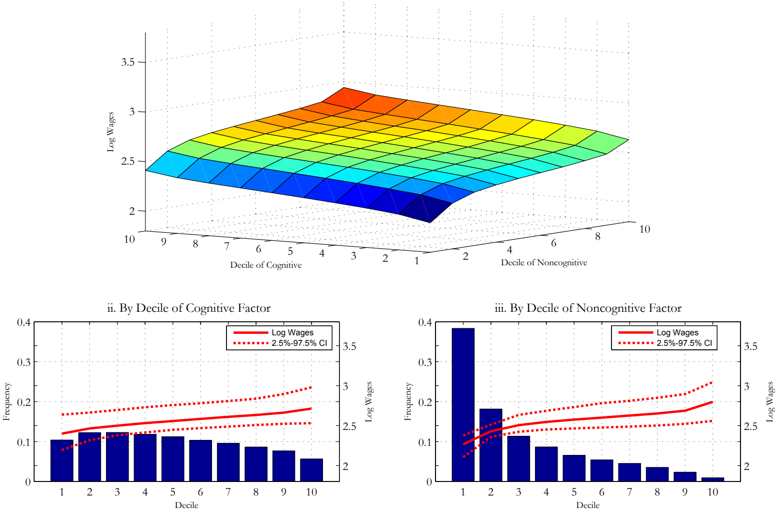
i. By Decile of Cognitive and Noncognitive Factors



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

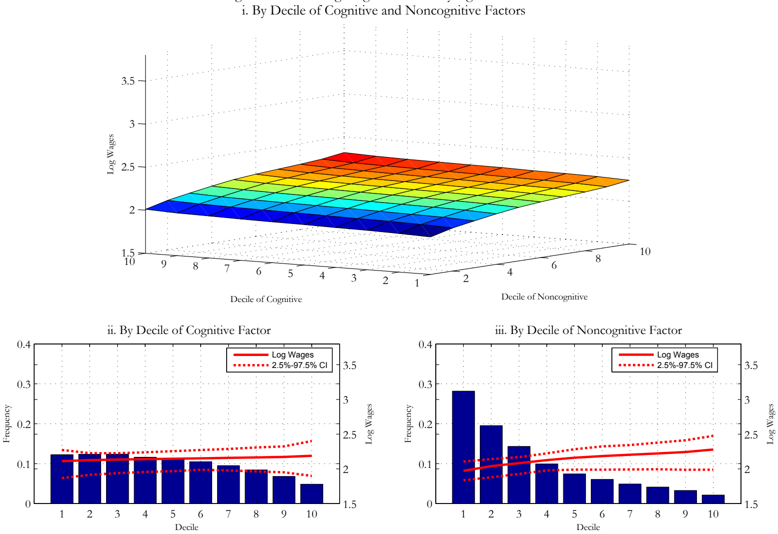
Figure 7A. Mean Log Wages of GEDs by Age 30 - Males

i. By Decile of Cognitive and Noncognitive Factors



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

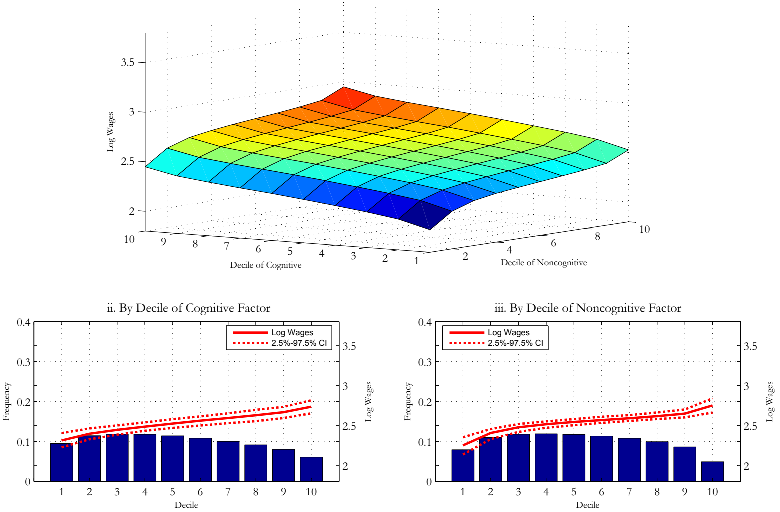
Figure 7B. Mean Log Wages of GEDs by Age 30 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

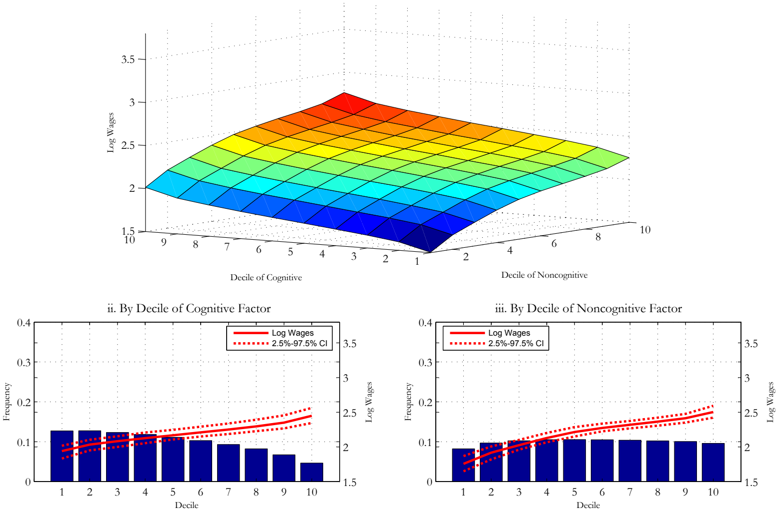
Figure 8A. Mean Log Wages of High School Graduates by Age 30 - Males

i. By Decile of Cognitive and Noncognitive Factors



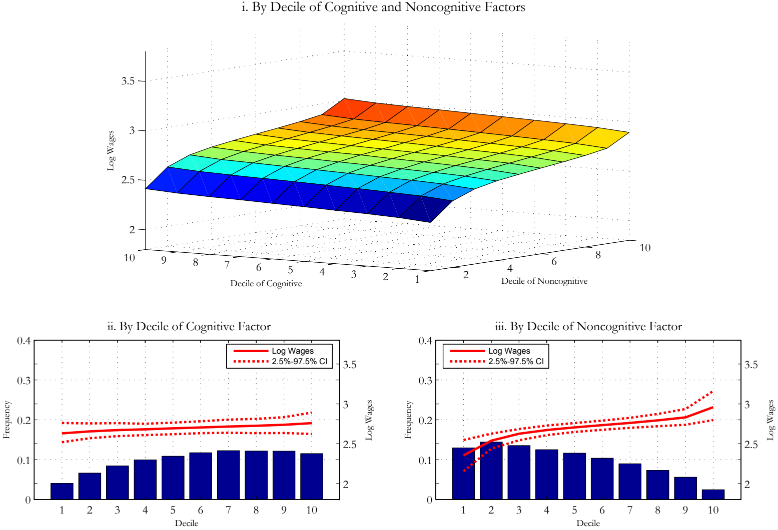
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

Figure 8B. Mean Log Wages of High School Graduates by Age 30 - Females
i. By Decile of Cognitive and Noncognitive Factors



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

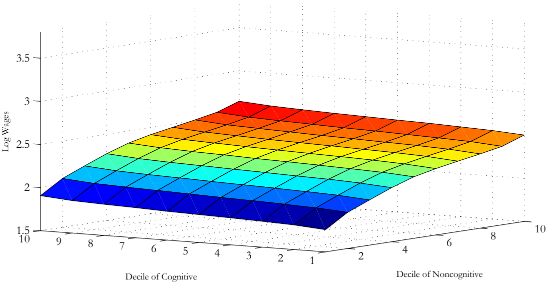
Figure 9A. Mean Log Wages of Some College Attenders by Age 30 - Males



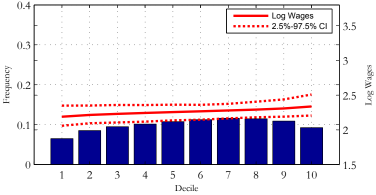
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

Figure 9B. Mean Log Wages of Some College Attenders by Age 30 - Females

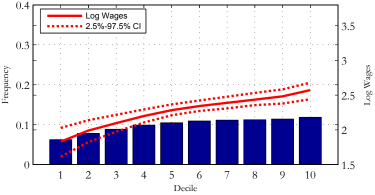
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor

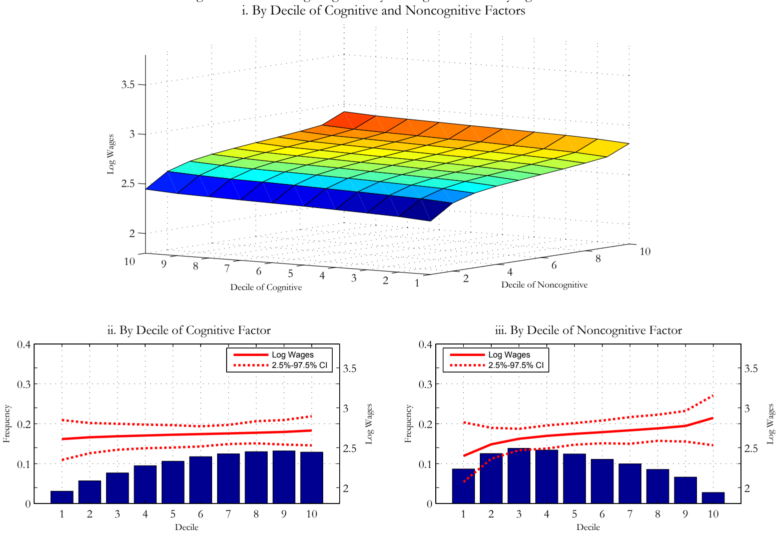


iii. By Decile of Noncognitive Factor



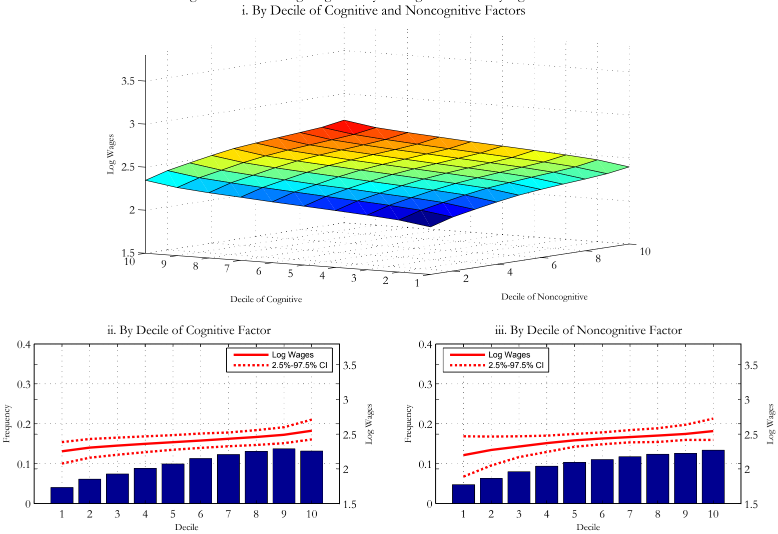
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

Figure 10A. Mean Log Wages of 2-yr College Graduates by Age 30 - Males



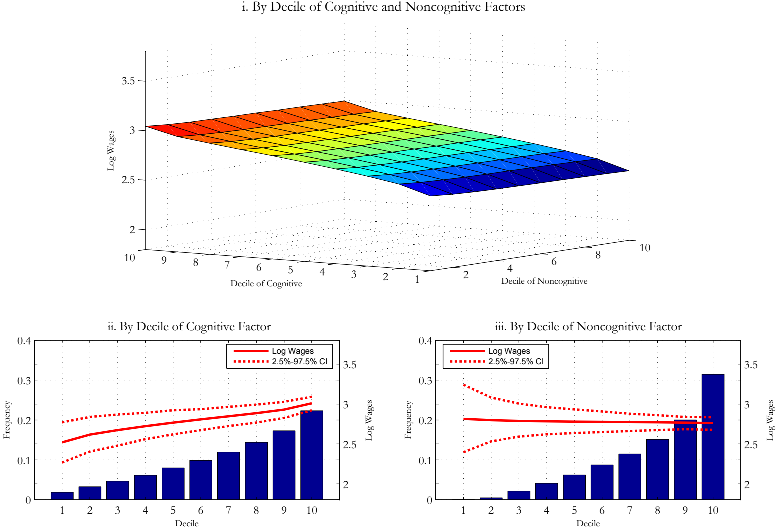
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

Figure 10B. Mean Log Wages of 2-yr College Graduates by Age 30 - Females



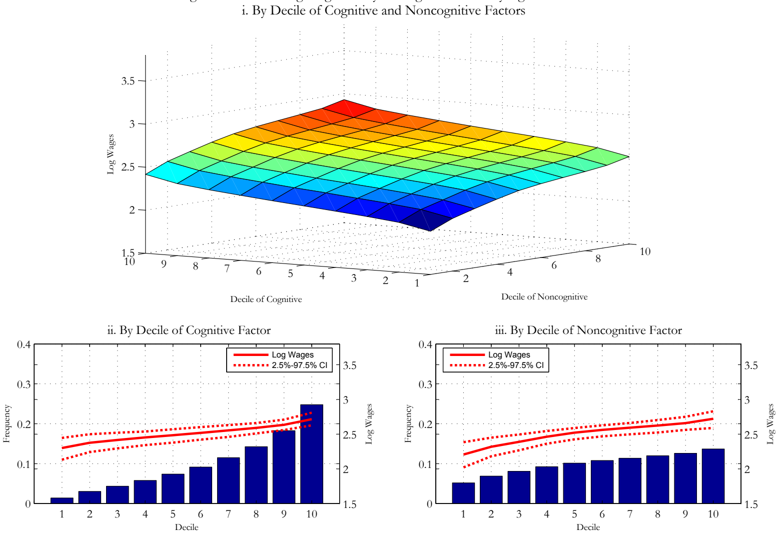
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

Figure 11A. Mean Log Wages of 4-yr College Graduates by Age 30 - Males



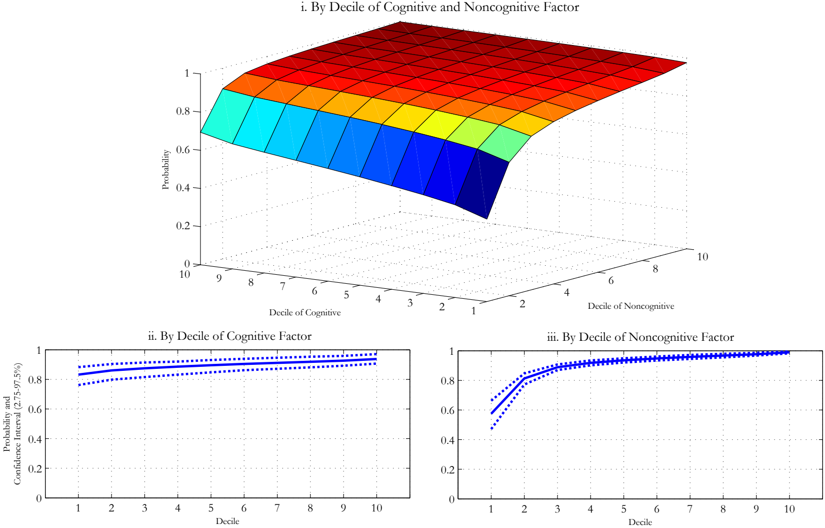
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

Figure 11B. Mean Log Wages of 4-yr College Graduates by Age 30 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

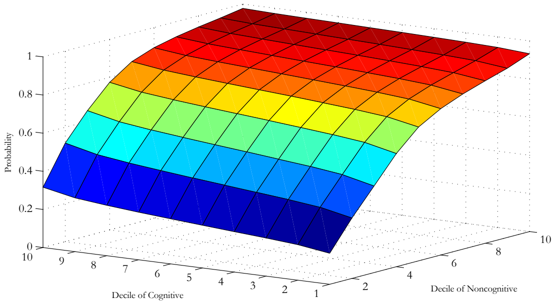
Figure 12A. Probability of Employment at Age 30 - Males



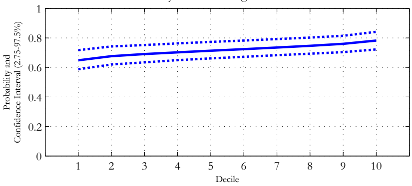
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

Figure 12B. Probability of Employment at Age 30 - Females

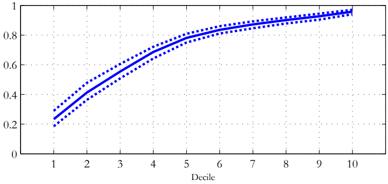
i. By Decile of Cognitive and Noncognitive Factor



ii. By Decile of Cognitive Factor

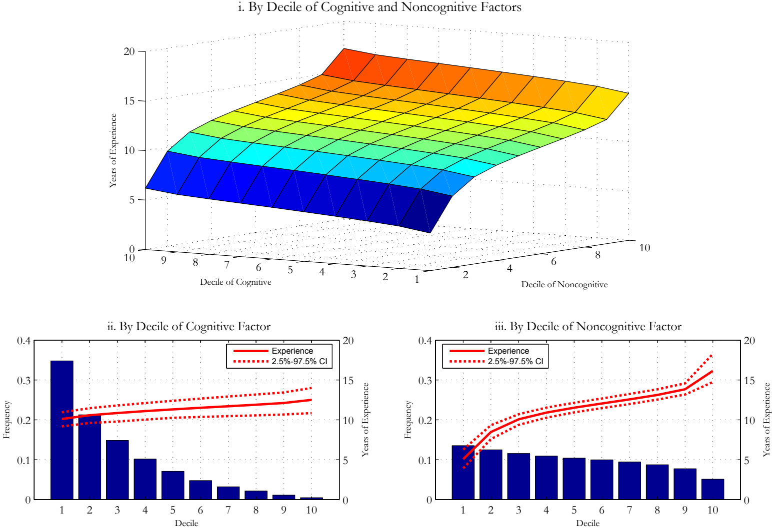


iii. By Decile of Noncognitive Factor



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

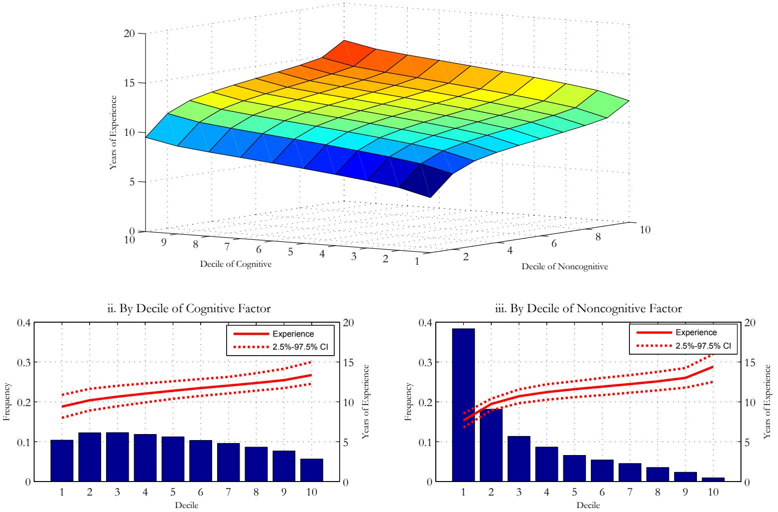
Figure 13A. Mean Work Experience of High School Dropouts by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

Figure 13B. Mean Work Experience of GEDs by Age 30 - Males

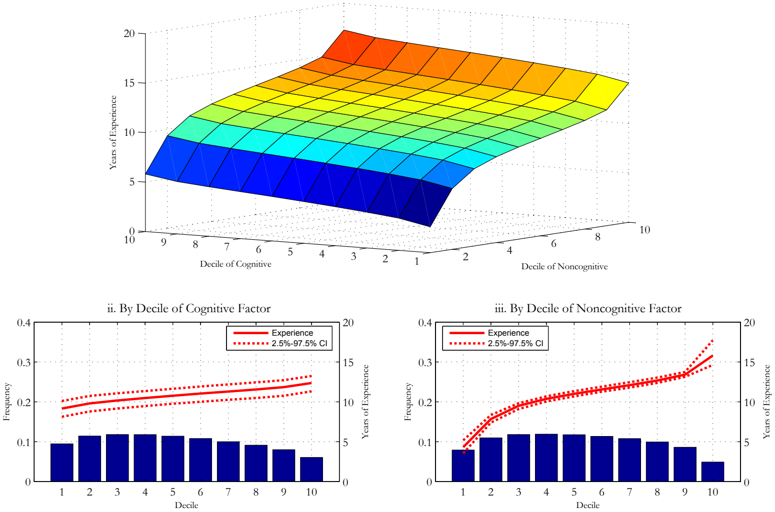
i. By Decile of Cognitive and Noncognitive Factors



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

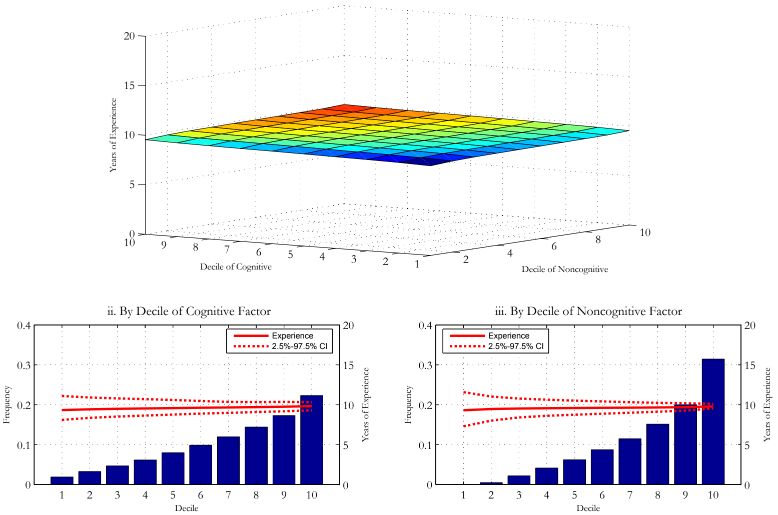
Figure 13C. Mean Work Experience of High School Graduates by Age 30 - Males

i. By Decile of Cognitive and Noncognitive Factors



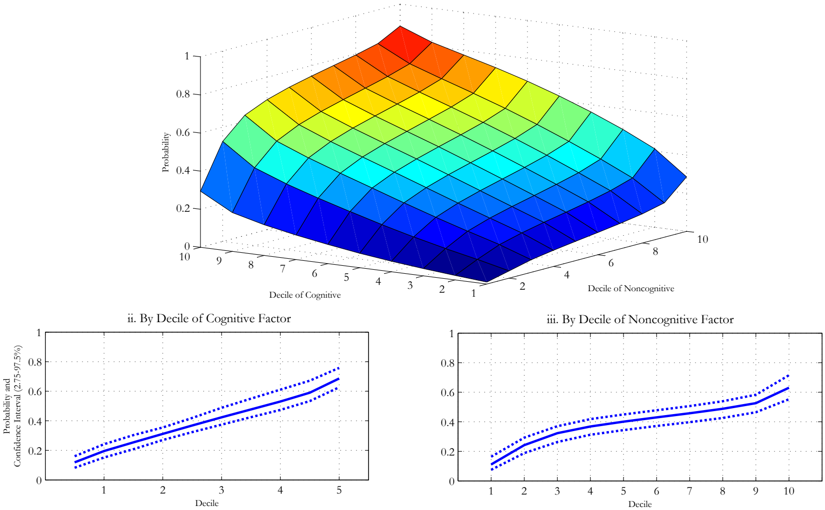
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

Figure 13D. Mean Work Experience of 4-yr College Graduates by Age 30 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals with the indicated level of education whose abilities lie in the indicated decile of the distribution.

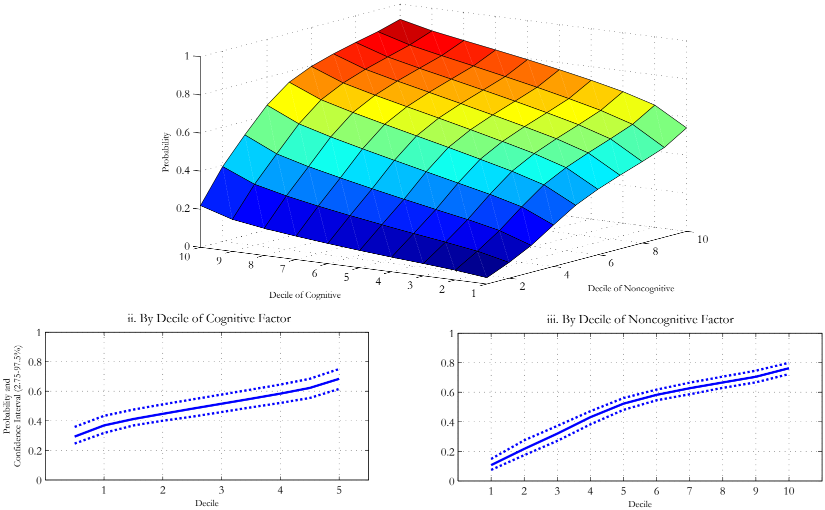
Figure 14A. Probability Of Being a White Collar Worker by Age 30 - Males
i. By Decile of Cognitive and Noncognitive Factor



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

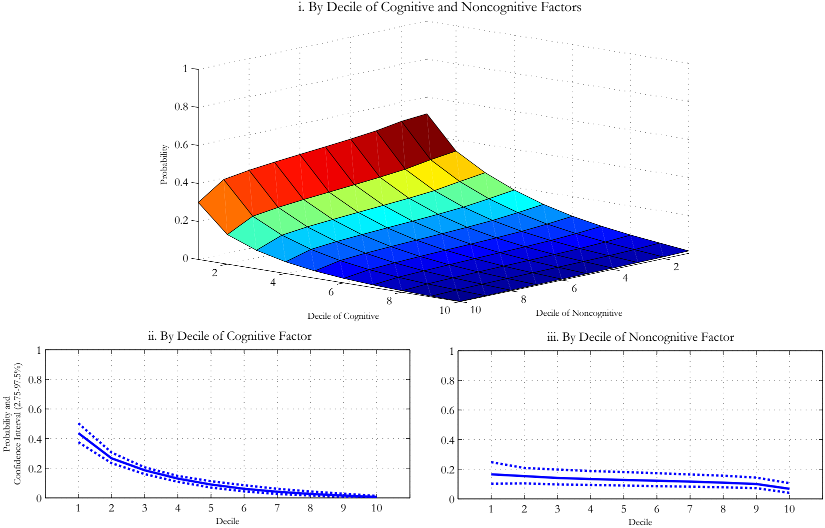
Figure 14B. Probability Of Being a White Collar Worker by Age 30 - Females

i. By Decile of Cognitive and Noncognitive Factor



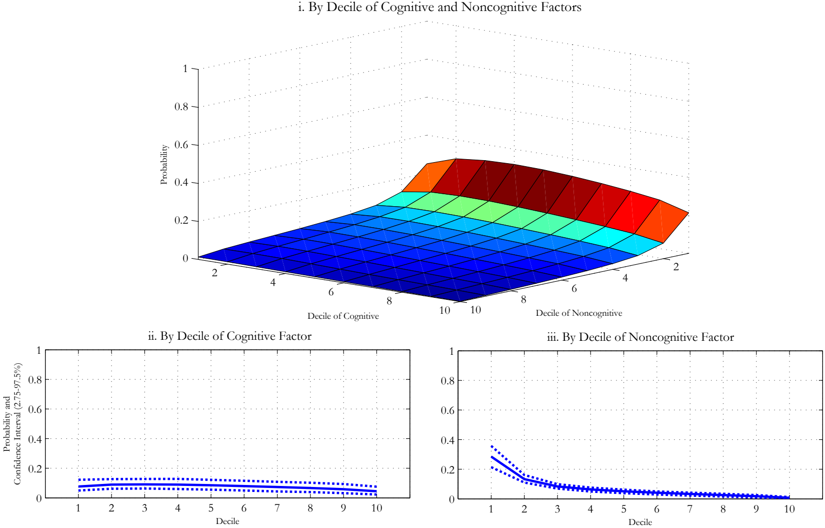
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

Figure 15. Probability of Being a High School Dropout by Age 30 - Males



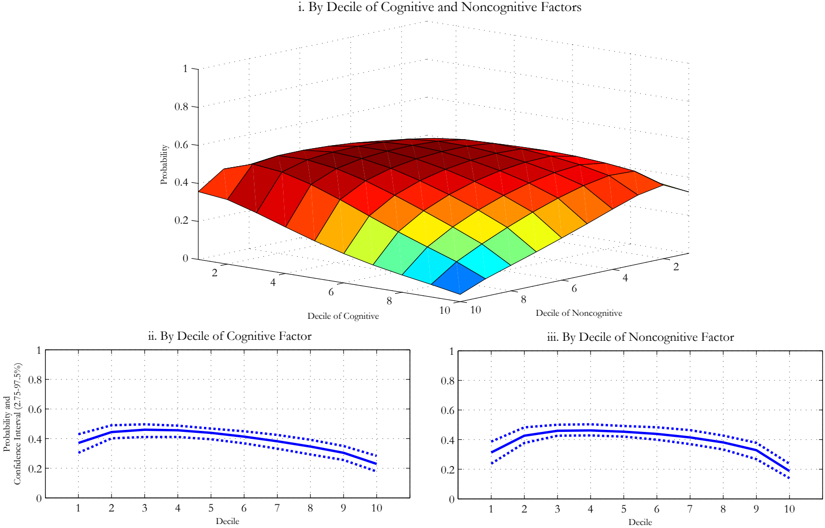
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 16. Probability of Being a GED by Age 30 - Males



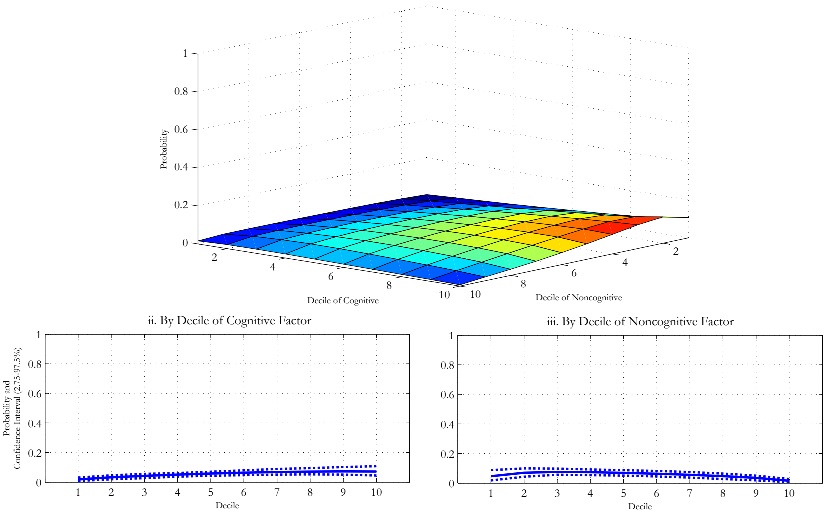
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 17. Probability of Being a High School Graduate by Age 30 - Males



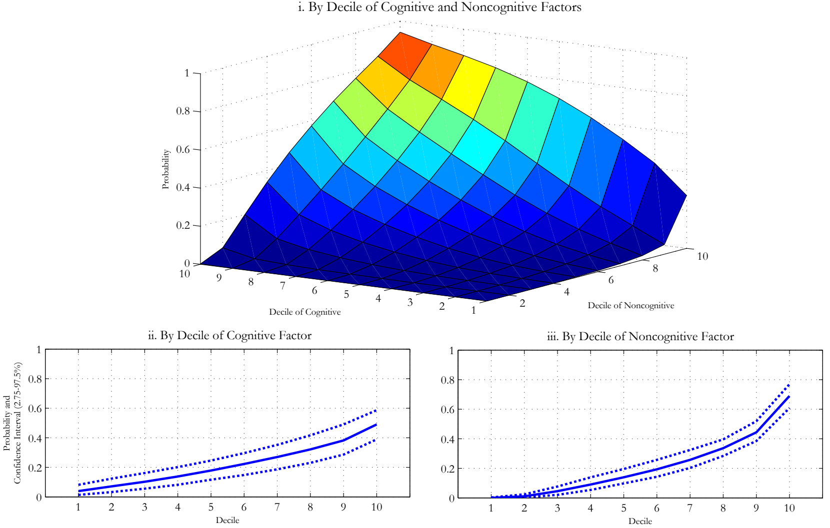
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 18. Probability of Being a 2-yr College Graduate by Age 30 - Males
i. By Decile of Cognitive and Noncognitive Factors



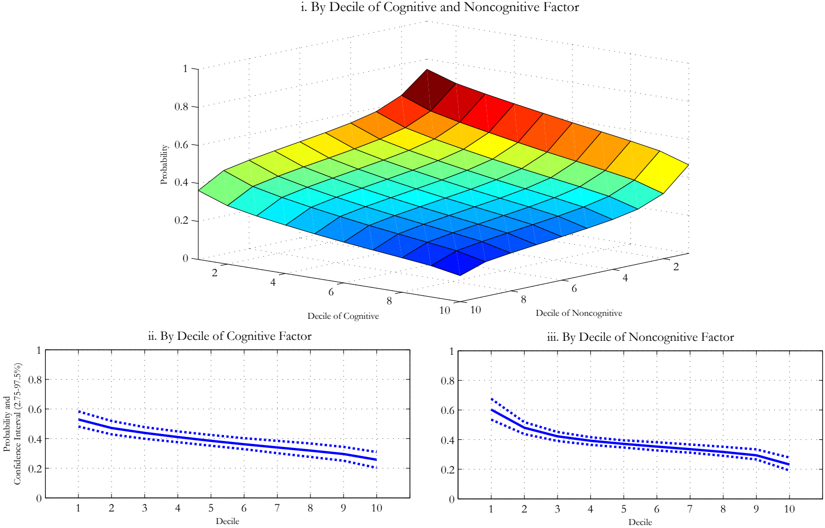
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

Figure 19. Probability of Being a 4-yr College Graduate by Age 30 - Males



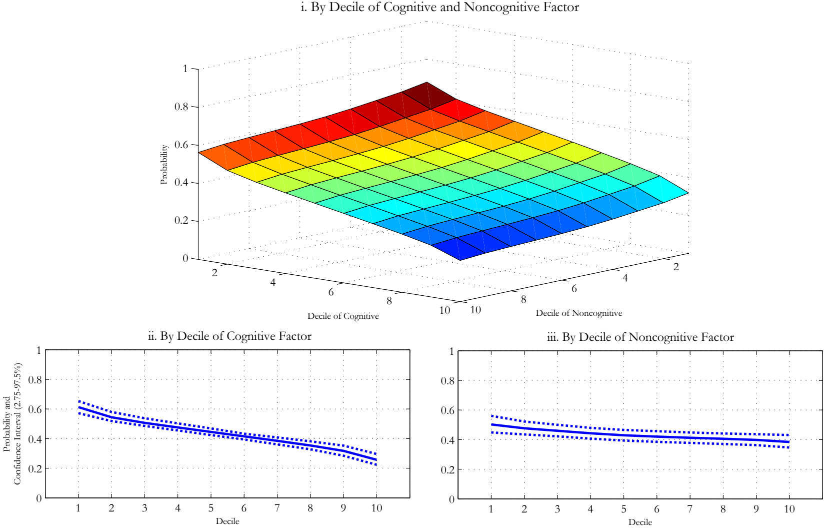
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 20A. Probability Of Daily Smoking By Age 18 - Males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

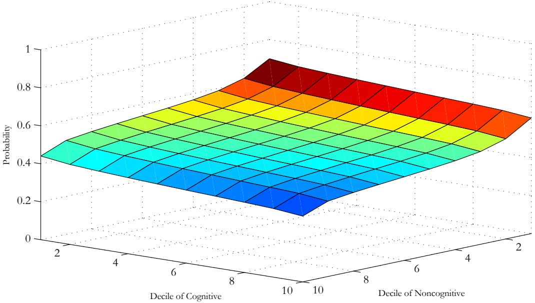
Figure 20B. Probability Of Daily Smoking By Age 18 - Females



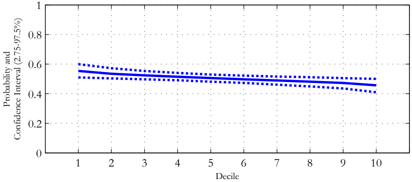
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 21. Probability of Smoking Marijuana during the Year 1979 - Males

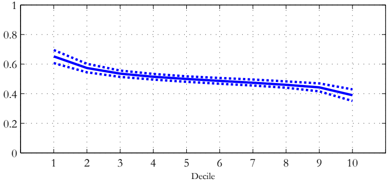
i. By Decile of Cognitive and Noncognitive Factor



ii. By Decile of Cognitive Factor

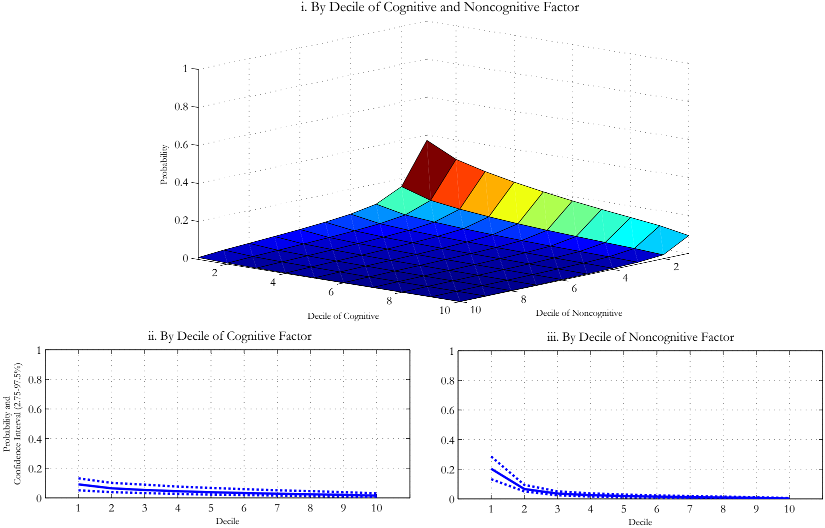


iii. By Decile of Noncognitive Factor



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

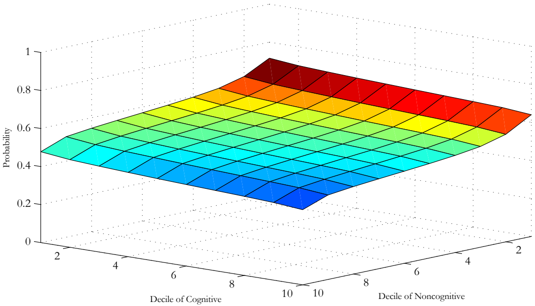
Figure 22. Probability of Incarceration by Age 30 - Males



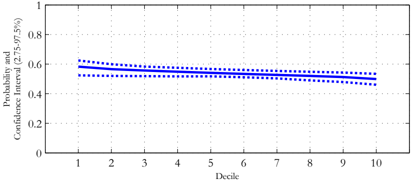
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 23. Probability of Participating in Illegal Activities during the Year 1979- Males

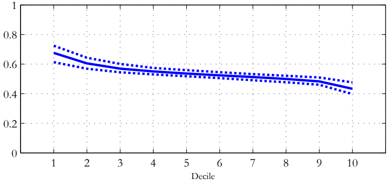
i. By Decile of Cognitive and Noncognitive Factor



ii. By Decile of Cognitive Factor

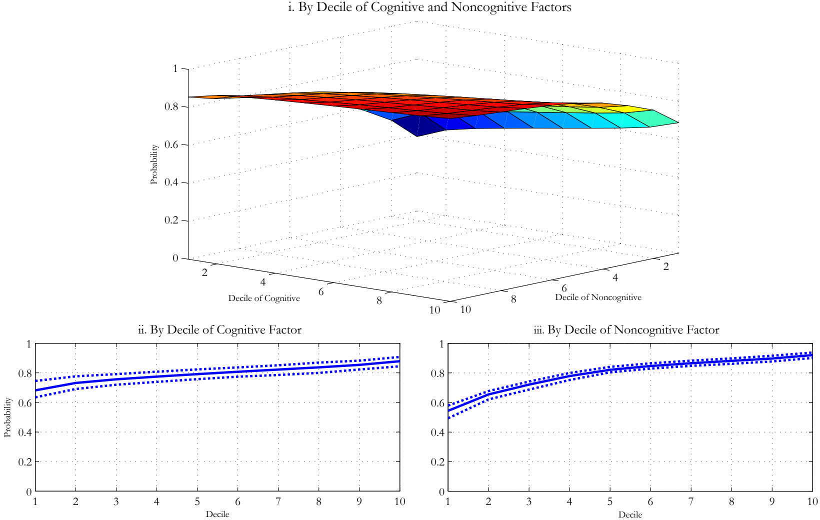


iii. By Decile of Noncognitive Factor



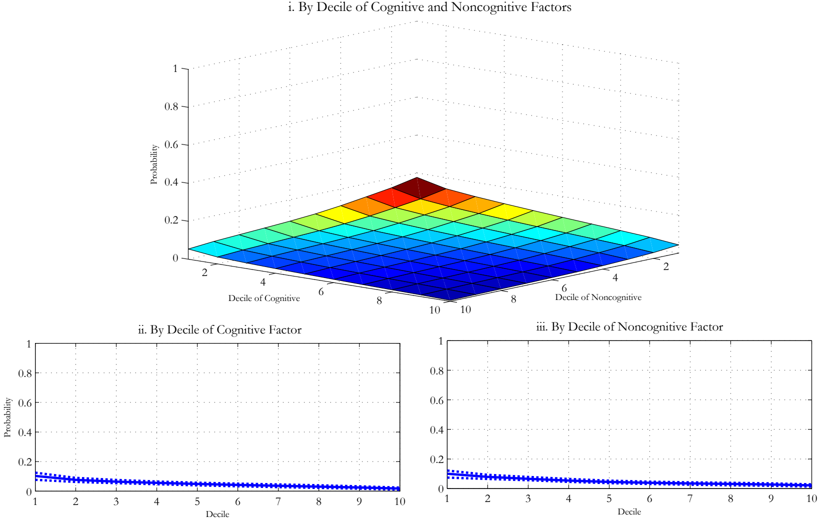
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 24. Probability Of Being Single With No Child at Age 18 - Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 25. Probability Of Being Single With Child at Age 18- Females



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Estimation of Dynamic Discrete Choice Models by Maximum Likelihood and Simulated Method of Moments

Fit structural model to *observed* data. $\hat{\psi}^{obs}$

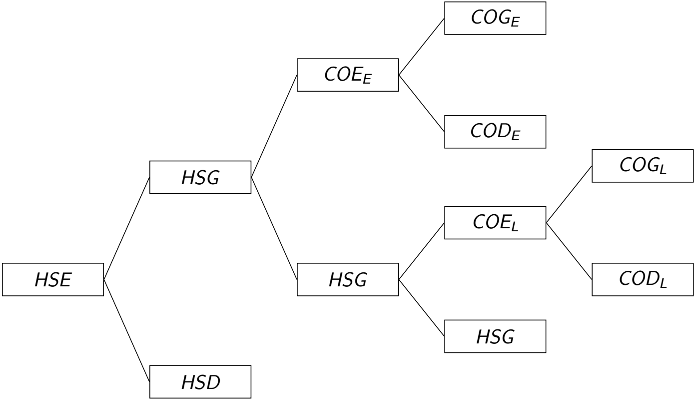
Simulate synthetic dataset using estimates.

Compare

Fit structural model to *synthetic* data. $\hat{\psi}^{syn}$

Model

Figure: Decision Tree



Agent Behavior

- Objective
- Constraints
  - Institution
  - Information

Value Function

$$V(s \mid \mathcal{I}(s)) = Y(s) + \max_{s' \in \Omega(s)} \left\{ \frac{1}{1+r} \Big(- C(s', s) + \mathbb{E}[V(s' \mid \mathcal{I}(s')) \,\big|\, \mathcal{I}(s) \,] \Big) \right\}$$

Policy Function

$$s' = \begin{cases} \hat{s}' & \text{if} \quad \mathbb{E}\left[V(\hat{s}') \,\middle|\, \mathcal{I}(s)\right] - C(\hat{s}', s) > \mathbb{E}\left[V(\tilde{s}') \,\middle|\, \mathcal{I}(s)\right] \\ \tilde{s}' & \text{otherwise} \end{cases}$$

Objects of Interest

- Net Return
- Gross Return

Net Return

$$NR(\hat{s}', \tilde{s}', s) = \frac{\mathbb{E}\left[V(\hat{s}') - V(\tilde{s}') \,\middle|\, \mathcal{I}(s)\right] - C(\hat{s}', s)}{\mathbb{E}\left[V(\tilde{s}') \,\middle|\, \mathcal{I}(s)\right]}$$

Gross Return

$$GR(\hat{s}', \tilde{s}', s) = \frac{\mathbb{E}\left[\tilde{V}(\hat{s}') - \tilde{V}(\tilde{s}') \,\middle|\, \mathcal{I}(s)\right]}{\mathbb{E}\left[\tilde{V}(\tilde{s}') \,\middle|\, \mathcal{I}(s)\right]}$$

Parametrization

- Benefits
- Costs
- Measurements

Functional Forms

$$
\begin{array}{rcll}
Y(s) & = & X(s)'\beta_s + \theta'\alpha_s + \epsilon(s) & \forall \quad s \in \mathcal{S} \\
C(\hat{s}', s) & = & Q(\hat{s}', s)'\delta_{\hat{s}', s} + \theta'\varphi_{\hat{s}', s} + \eta(\hat{s}', s) & \forall \quad s \in \mathcal{S}^c
\end{array}
$$

$$
\begin{array}{rcll}
M(j) & = & X(j)'\kappa_j + \theta'\gamma_j + \nu(j) & \forall \quad j \in M
\end{array}
$$

Distributions of Unobservables

$$
\begin{aligned}
\theta &\sim \mathcal{N}(0, \sigma_\theta) && \forall \quad \theta \in \Theta \\
\epsilon(s) &\sim \mathcal{N}(0, \sigma_{\epsilon(s)}) && \forall \quad s \in \mathcal{S} \\
\eta(\hat{s}', s) &\sim \mathcal{N}(0, \sigma_{\eta(\hat{s}', s)}) && \forall \quad s \in \mathcal{S}^c \\
\nu(j) &\sim \mathcal{N}(0, \sigma_{\nu(j)}) && \forall \quad j \in \mathcal{M}
\end{aligned}
$$

Individual Likelihood

$$\int_{\underline{\Theta}} \left[\prod_{j \in M} \underbrace{f\Big(M(j) \,\Big|\, D, \theta; \psi\Big)}_{\text{Measurement}} \times \right.$$

$$\left. \prod_{s \in \mathcal{S}} \left\{ \underbrace{f\Big(Y(s) \,\Big|\, D, \theta; \psi\Big)}_{\text{Outcome}} \underbrace{\Pr\Big(G(s) = 1 \,\Big|\, D, \theta; \psi\Big)}_{\text{Transition}} \right\}^{\mathbb{1}\{s \in \Gamma\}} \right] dF(\theta)$$

Results

Table: Cross Section Model Fit

	Average Earnings		State Frequencies	
State	Observed	ML	Observed	ML
High School Graduates	4.29	3.84	0.30	0.32
High School Dropouts	2.29	2.59	0.17	0.14
Early College Graduates	6.73	7.46	0.29	0.29
Early College Dropouts	4.55	3.87	0.12	0.12
Late College Graduates	4.84	6.22	0.06	0.07
Late College Dropouts	4.89	4.88	0.06	0.06

Objects of Interest

- Choice Probabilities
- Gross Return
- Net Return
- Schooling Attainment

Figure: Choice Probability, Early College Enrollment

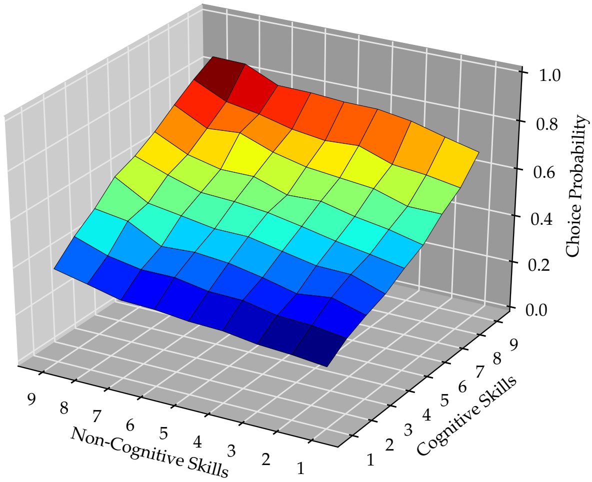


Figure: Gross Return, Early College Enrollment

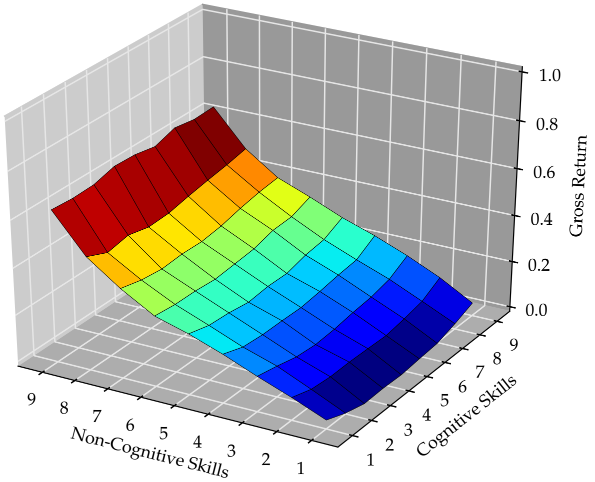


Figure: Net Return, Early College Enrollment

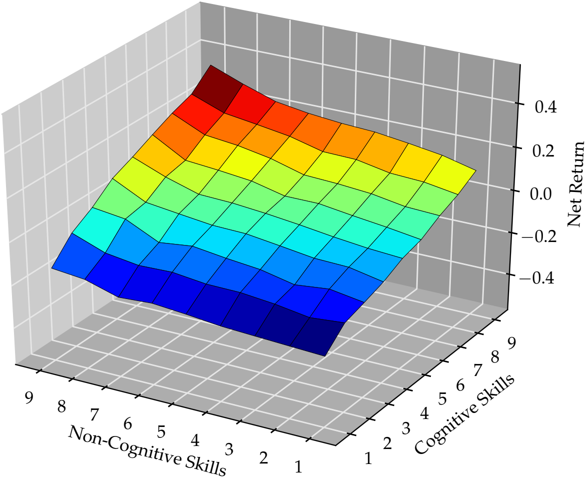


Figure: Schooling Attainment by Cognitive Skills

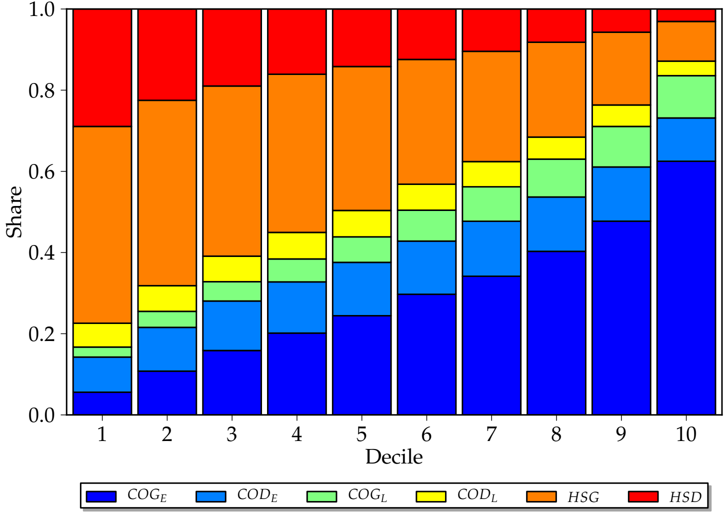
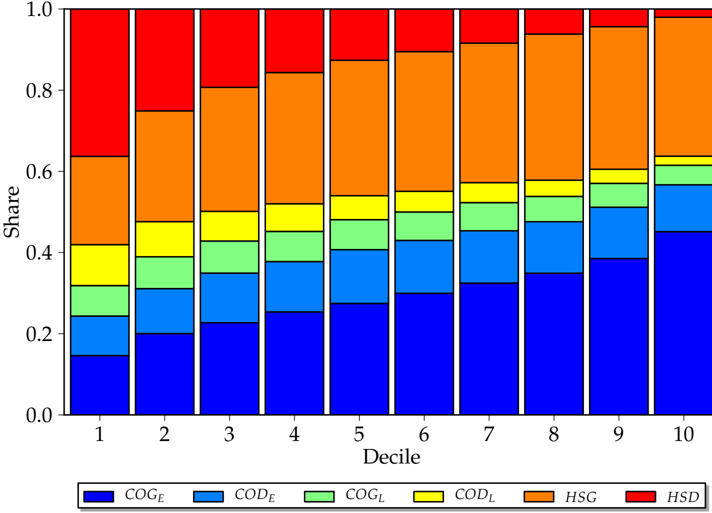


Figure: Schooling Attainment by Non-Cognitive Skills



Monte Carlo Study

Synthetic Agents	5,000
Structural Parameters	192
Free Parameters	138

Simulated Method of Moments

Criterion Function

$$\arg\min_{\psi} \quad \Lambda(\psi) = \left[\check{f} - \hat{f}(\psi)\right]' W^{-1} \left[\check{f} - \hat{f}(\psi)\right]$$

$$, \text{ where } \quad \hat{f}(\psi) = \frac{1}{R} \sum_{r=1}^{R} \hat{f}_r(u_r; \psi)$$

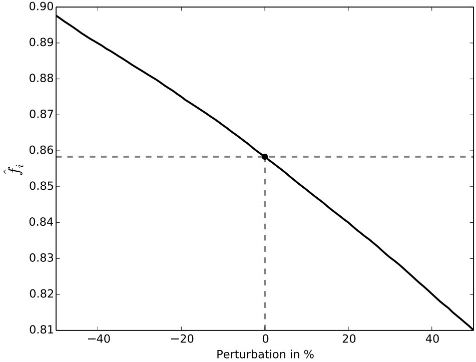
Number of Moments	250
Number of Replications	50
Weighting Matrix	Diagonal Matrix with Variances
Optimization Algorithm	TAO POUNDerS

Moment Conditions

- Benefit Equations
    - Means
    - Standard Deviations
    - Ordinary Least Squares Models
- Cost Equations
    - State Frequencies
    - Linear Probability Models

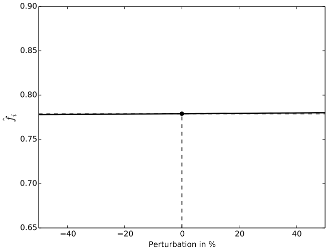
Parameter Perturbations

Figure: Parameter Perturbation, Cost

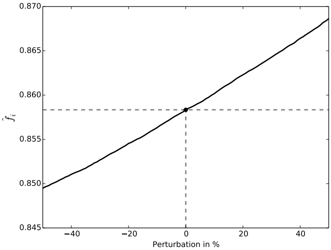


High School Graduation, State Frequency

Figure: Parameter Perturbation, Benefit



HS Graduates, Average Wages



HS Graduation, State Frequency

Model Fit

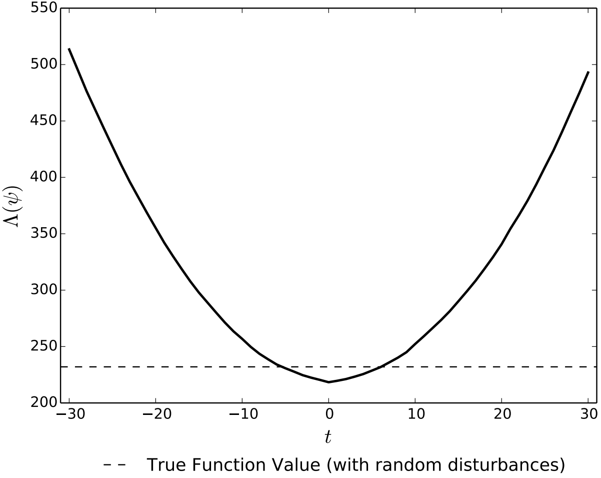
Table: Cross Section Model Fit

	Average Earnings		
State	True	ML	SMM
High School Graduates	3.87	3.87	3.86
High School Dropouts	2.51	2.52	2.53
Early College Graduates	6.80	6.78	6.86
Early College Dropouts	3.90	3.95	3.92
Late College Graduates	6.03	6.14	6.24
Late College Dropouts	5.10	5.07	5.08
RMSE		0.05	0.07

Table: Cross Section Model Fit

	State Frequencies		
State	True	ML	SMM
High School Graduates	0.32	0.32	0.33
High School Dropouts	0.14	0.14	0.13
Early College Graduates	0.29	0.29	0.29
Early College Dropouts	0.12	0.12	0.13
Late College Graduates	0.07	0.07	0.05
Late College Dropouts	0.06	0.06	0.07
RMSE		0.00	0.00

Figure: Function Value



Economic Implications

Table: Economic Implications

	Gross Return		
State	True	ML	SMM
High School Graduation	30%	41%	42%
Early College Graduation	88%	96%	96%
Late College Graduation	29%	28%	36%
RMSE		0.09	0.13

Table: Economic Implications

	Net Return		
State	True	ML	SMM
High School Graduation	63%	61%	212%
Early College Graduation	57%	51%	125%
Late College Graduation	14%	12%	38%
RMSE		0.05	0.75

Table: Standard Deviations

		$\hat{\sigma}_{\eta_{(\hat{s}',s)}}$	
State	True	ML	SMM
High School Graduation	0.27	0.24	0.85
Early College Graduation	0.61	0.59	1.49
Late College Graduation	0.61	0.59	1.49
RMSE		0.01	0.68

Tuning Parameters

- Moment Conditions
- Replications
- Optimization Algorithm

Moment Conditions

Table: Set of Moments

Sets	*Dynamic (Panel) Moments* Base	Alt. A	Alt. B	*Cross Section Moments* Base
Benefit Models				
Means	✓	✓	✓	✓
Standard Deviations	✓	✓	✓	✓
Ordinary Least Squares	✓	✓	✓	✓
Correlations			✓	
Choice Models				
State Frequencies	✓	✓	✓	✓
Linear Probability				
- cross section				✓
- dynamic	✓	✓	✓	
Probit				
- dynamic		✓	✓	
Correlations			✓	

Table: Set of Moments

	Dynamic (Panel) Moments			*Cross Section Moments*
Sets	Base	Alt. A	Alt. B	Base
Overall Statistics				
Number of Moments	250	339	545	222
Number of Replications	50	50	50	50
Weighting Matrix	Diagonal Matrix with Variances			
Optimization Algorithm	TAO POUNDerS			
Quality of Fit Measures				
$\Lambda(\hat{\psi})$	218.32	277.80	447.60	632.07
$\Lambda(\psi^*)$	232.66	291.56	471.95	215.12

Table: Robustness of Economic Implications

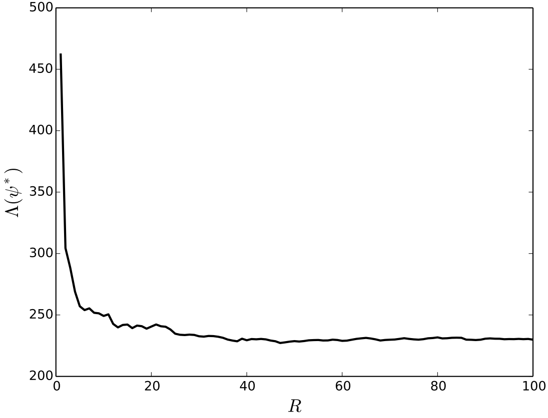
		Dynamic (Panel) Moments		
State	True	Base	Alt. A	Alt. B
Gross Return				
High School Graduation	30%	42%	37%	37%
Early College Graduation	88%	96%	72%	73%
Late College Graduation	29%	36%	18%	18%
Net Return				
High School Graduation	63%	212%	203%	194%
Early College Graduation	57%	125%	107%	112%
Late College Graduation	14%	38%	30%	35%

Table: Robustness of Economic Implications

		Cross Section Moments
State	True	Base
	Gross Return	
High School Graduation	30%	16%
Early College Graduation	88%	57%
Late College Graduation	29%	16%
	Net Return	
High School Graduation	63%	215%
Early College Graduation	57%	79%
Late College Graduation	14%	26%

Replications

Figure: Role of Replications



Optimization Algorithm

TAO POUNDerS

Practical Optimization Using No Derivatives for Sums of Squares

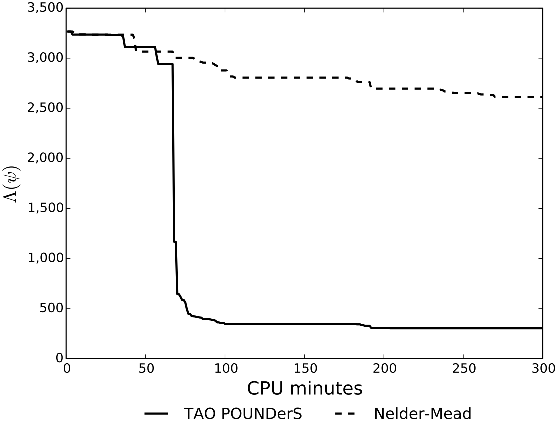
Nonlinear Least-Squares

$$\Lambda(\psi) = \sum_{i=1}^{I} \bar{f}_i(\psi)^2 = \sum_{i=1}^{I} \left(\frac{\check{f}_i - \hat{f}_i(\psi)}{\hat{\sigma}_i} \right)^2$$

Derivative-Free Trust-Region Algorithm

$$\min\{m_k(\psi) : ||\psi - \psi_k||_\infty \leq \Delta_k\}$$

Figure: Optimization Algorithms



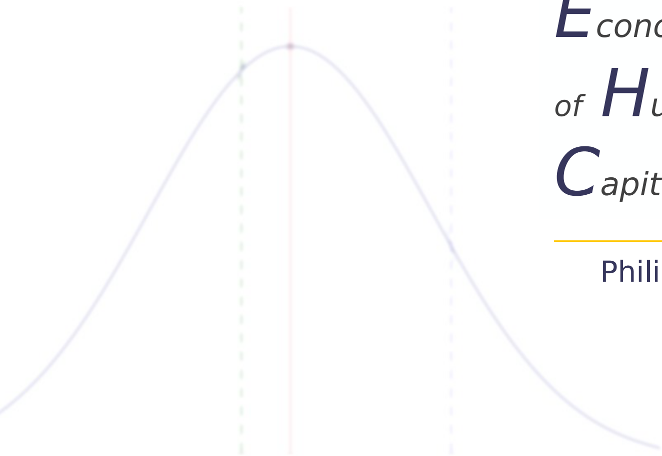
Appendix

References

Becker, G. S. (1964). *Human capital* (1st ed.). New York City, NY: Columbia University Press.

Eisenhauer, P., Heckman, J. J., & Mosso, S. (2015). Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. *International Economic Review*, *56*(2), 331–357.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*(3), 411–482.



Economics of Human Capital

Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

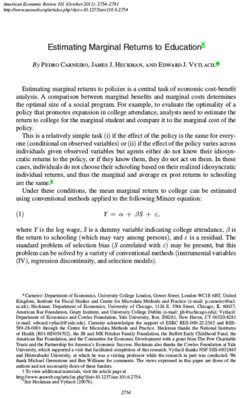
Economics of Human Capital

Static model of educational choice

Philipp Eisenhauer

Introduction

Figure: Motivation



American Economic Review 101 (October 2011): 2754–2781
http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.6.2754

Estimating Marginal Returns to Education†

By PEDRO CARNEIRO, JAMES J. HECKMAN, AND EDWARD J. VYTLACIL*

Estimating marginal returns to policies is a central task of economic cost-benefit analysis. A comparison between marginal benefits and marginal costs determines the optimal size of a social program. For example, to evaluate the optimality of a policy that promotes expansion in college attendance, analysts need to estimate the return to college for the marginal student and compare it to the marginal cost of the policy.

This is a relatively simple task (i) if the effect of the policy is the same for everyone (conditional on observed variables) or (ii) if the effect of the policy varies across individuals given observed variables but agents either do not know their idiosyncratic returns to the policy, or if they know them, they do not act on them. In these cases, individuals do not choose their schooling based on their realized idiosyncratic individual returns, and thus the marginal and average ex post returns to schooling are the same.[1]

Under these conditions, the mean marginal return to college can be estimated using conventional methods applied to the following Mincer equation:

$$Y = \alpha + \beta S + \varepsilon, \tag{1}$$

where Y is the log wage, S is a dummy variable indicating college attendance, β is the return to schooling (which may vary among persons), and ε is a residual. The standard problem of selection bias (S correlated with ε) may be present, but this problem can be solved by a variety of conventional methods (instrumental variables (IV), regression discontinuity, and selection models).

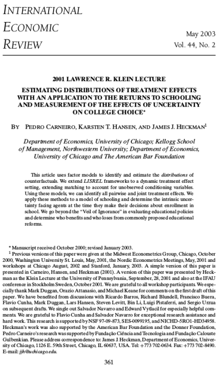
* Carneiro: Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom, Institute for Fiscal Studies and Centre for Microdata Methods and Practice (e-mail: p.carneiro@ucl.ac.uk); Heckman: Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637, American Bar Foundation, Geary Institute, and University College Dublin (e-mail: jjh@uchicago.edu); Vytlacil: Department of Economics and Cowles Foundation, Yale University, Box 208281, New Haven, CT 06520-8281 (e-mail: edward.vytlacil@yale.edu). Carneiro acknowledges the support of ESRC RES-000-22-2542 and RES-589-28-0001 through the Centre for Microdata Methods and Practice. Heckman thanks the National Institutes of Health (R01-HD054702), the JB and MK Pritzker Family Foundation, the Buffett Early Childhood Fund, the American Bar Foundation, and the Committee for Economic Development with a grant from The Pew Charitable Trusts and the Partnership for America's Economic Success. Heckman also thanks the Cowles Foundation at Yale University, which supported a visit that facilitated completion of this research. Vytlacil thanks NSF SES-0832845 and Hitotsubashi University, at which he was a visiting professor while the research in part was conducted. We thank Michael Greenstone and Ben Williams for comments. The views expressed in this paper are those of the authors and not necessarily those of these funders.

† To view additional materials, visit the article page at http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.6.2754.

[1] See Heckman and Vytlacil (2007b).

2754

Carneiro & al. (2011)



INTERNATIONAL ECONOMIC REVIEW

May 2003
Vol. 44, No. 2

2001 LAWRENCE R. KLEIN LECTURE

ESTIMATING DISTRIBUTIONS OF TREATMENT EFFECTS WITH AN APPLICATION TO THE RETURNS TO SCHOOLING AND MEASUREMENT OF THE EFFECTS OF UNCERTAINTY ON COLLEGE CHOICE*

BY PEDRO CARNEIRO, KARSTEN T. HANSEN, AND JAMES J. HECKMAN[1]

Department of Economics, University of Chicago; Kellogg School of Management, Northwestern University; Department of Economics, University of Chicago and The American Bar Foundation

This article uses factor models to identify and estimate the *distributions* of counterfactuals. We extend *LISREL* frameworks to a dynamic treatment effect setting, extending matching to account for unobserved conditioning variables. Using these models, we can identify all pairwise and joint treatment effects. We apply these methods to a model of schooling and determine the intrinsic uncertainty facing agents at the time they make their decisions about enrollment in school. We go beyond the "Veil of Ignorance" in evaluating educational policies and determine who benefits and who loses from commonly proposed educational reforms.

* Manuscript received October 2000; revised January 2003.

[1] Previous versions of this paper were given at the Midwest Econometrics Group, Chicago, October 2000, Washington University St. Louis, May, 2001, the Nordic Econometrics Meetings, May, 2001 and workshops at Chicago August, 2002 and Stanford, January, 2003. A simple version of this paper is presented in Carneiro, Hansen, and Heckman (2001). A version of this paper was presented by Heckman as the Klein Lecture at the University of Pennsylvania, September, 28, 2001 and also at the IFAU conference in Stockholm Sweden, October 2001. We are grateful to all workshop participants. We especially thank Mark Duggan, Orazio Attanasio, and Michael Keane for comments on the first draft of this paper. We have benefited from discussions with Ricardo Barros, Richard Blundell, Francisco Buera, Flavio Cunha, Mark Duggan, Lars Hansen, Steven Levitt, Bin Li, Luigi Pistaferri, and Sergio Urzua on subsequent drafts. We single out Salvador Navarro and Edward Vytlacil for especially helpful comments. We are grateful to Flavio Cunha and Salvador Navarro for exceptional research assistance and hard work. This research is supported by NSF 97-09-873, SES-0099195, and NICHD-5RO1-HD34958. Heckman's work was also supported by the American Bar Foundation and the Donner Foundation, Pedro Carneiro's research was supported by Fundação Ciência and Tecnologia and Fundação Calouste Gulbenkian. Please address correspondence to: James J. Heckman, Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637, USA. Tel: +773 702-0634. Fax: +773 702-8490. E-mail: *jjh@uchicago.edu*.

361

Carneiro & al. (2003)

Heckman (2008) defines three policy evaluation tasks:

- Evaluating the impact of historical interventions on outcomes including their impact in terms of well-being of the treated and the society at large.
- Forecasting the impact of historical interventions implemented in one environment in other environments, including their impact in terms of well-being.
- Forecasting the impacts of interventions never historically experienced to various environments, including their impact on well-being.

Econometrics of policy evaluation

- is important
- is complicated
- is multifaceted

Numerous applications

- labor economics
- development economics
- industrial economics
- health economics

Numerous effects

- conventional average effects
- policy-relevant average effects
- marginal effects
- distributional effects
- effects on distributions

Numerous estimation strategies

- instrumental variables
- (quasi-)experimental methods
- matching

Model

Generalized Roy model

Potential Outcomes

$Y_1 = \mu_1(X) + U_1$

$Y_0 = \mu_0(X) + U_0$

Observed Outcome

$Y = DY_1 + (1 - D)Y_0$

Choice

$D = \mathrm{I}[S > 0]$

$S = \mu_D(X, Z) - V$

- S is the overall surplus from treatment participation
- V captures the individual's unobservable dislike of treatment

Individual Heterogeneity

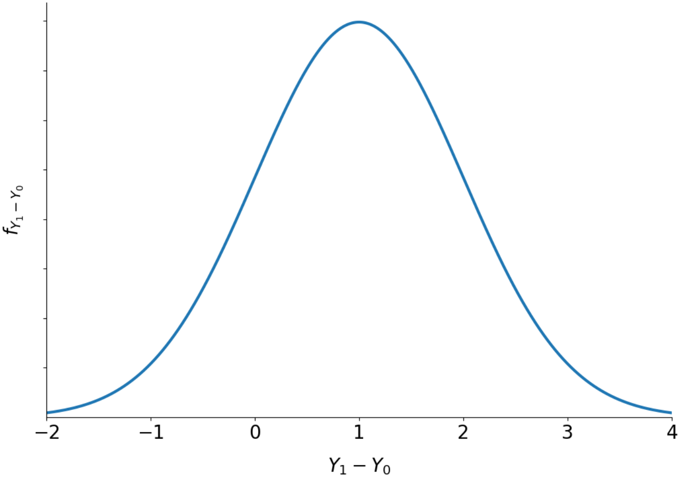
Individual-specific benefit of treatment

$$Y_1 - Y_0 = (\mu_1(X) - \mu_0(X)) + (U_1 - U_0)$$

Sources of Heterogeneity

- Difference in observables
- Difference in unobservables
    - Uncertainty
    - Private information

Figure: Distribution of benefits



Econometric problems

- **Evaluation problem**, we only observe an individual in either the treated or untreated state.
- **Selection problem**, individuals that select into treatment differ from those that do not.

Essential Heterogeneity

Definition: Individuals select their treatment status based on gains unobservable by the econometrician. More formally,

$$Y_1 - Y_0 \not\perp\!\!\!\perp D \quad | X = x.$$

$\Rightarrow$ consequences for the choice of the estimation strategy

Objects of interest

Useful Notation

$$P(X,Z) = \Pr(D = 1 \mid X,Z) = F_V(\mu_D(X,Z))$$
$$U_D = F_V(V)$$

Figure: First-stage unobservable

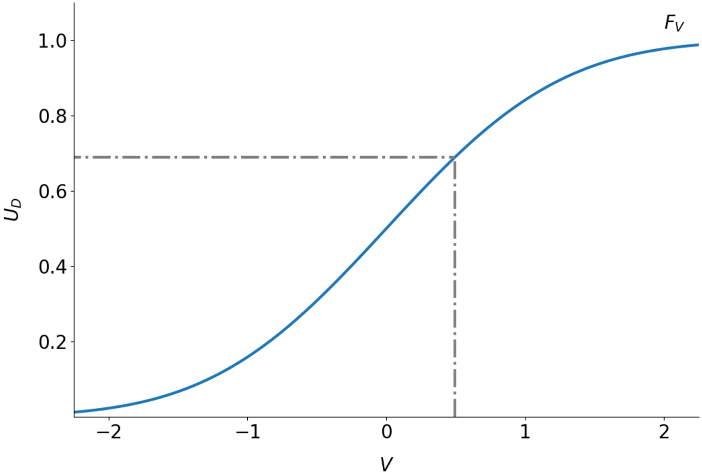


Figure: Support

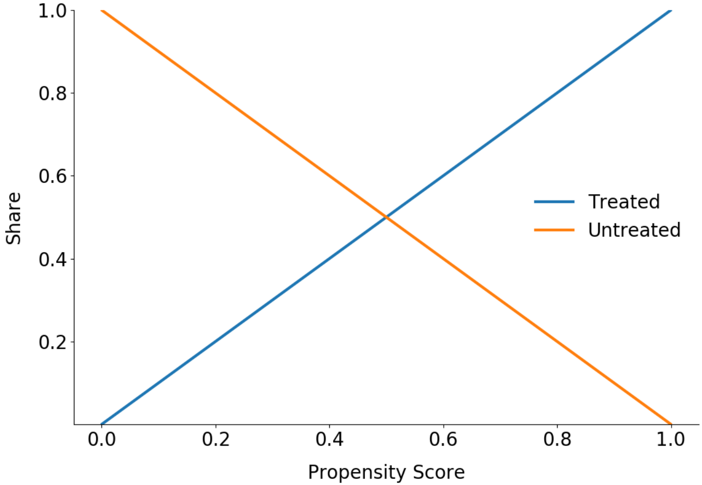


Figure: Distribution of benefits

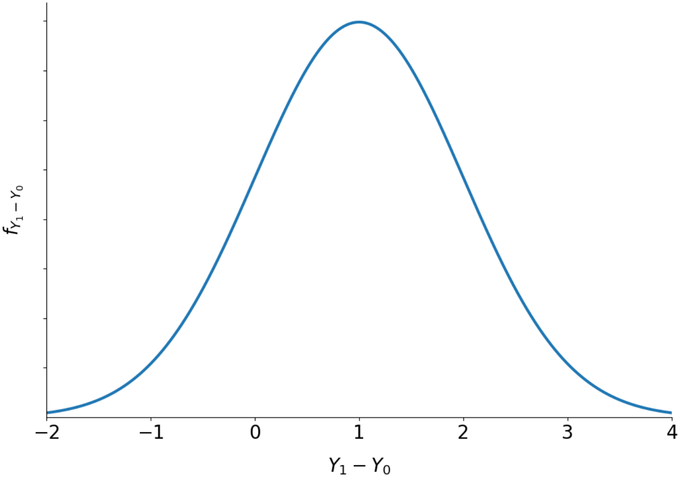
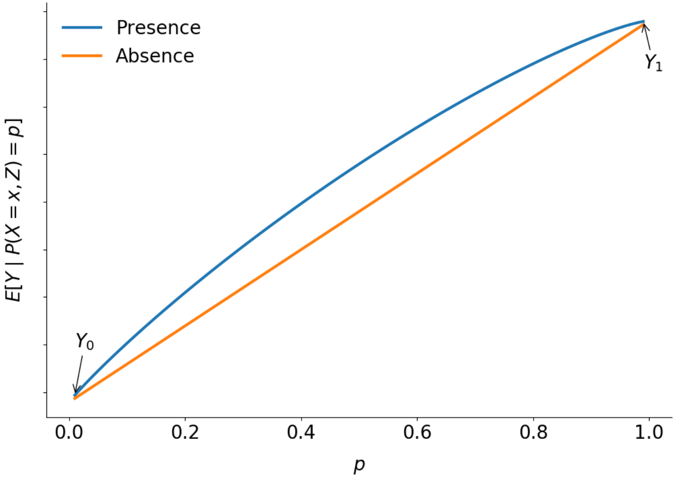


Figure: Conditional expectation and essential heterogeneity



Conventional Average Treatment Effects

Conventional Average Treatment Effects

$$\begin{aligned} B^{ATE} &= E[Y_1 - Y_0] \\ B^{TT} &= E[Y_1 - Y_0 \mid D = 1] \\ B^{TUT} &= E[Y_1 - Y_0 \mid D = 0] \end{aligned}$$

$\Rightarrow$ correspond to *extreme* policy alternatives

Selection Problem

$$E[Y \mid D=1] - E[Y \mid D=0] = \underbrace{E[Y_1 - Y_0 \mid D=1]}_{B^{TT}} + \underbrace{E[Y_0 \mid D=1] - E[Y_0 \mid D=0]}_{\text{Selection bias}}$$

$$
\begin{aligned}
E[Y \mid D=1] - E[Y \mid D=0] = {} & \underbrace{E[Y_1 - Y_0]}_{B^{ATE}} \\
& + \underbrace{E[Y_1 - Y_0 \mid D=1] - E[Y_1 - Y_0]}_{\text{Sorting on gains}} \\
& + \underbrace{E[Y_0 \mid D=1] - E[Y_0 \mid D=0]}_{\text{Sorting on levels}}
\end{aligned}
$$

- bias depends on the parameter of interest
- selection bias as sorting on levels

Figure: Distribution of effects with essential heterogeneity

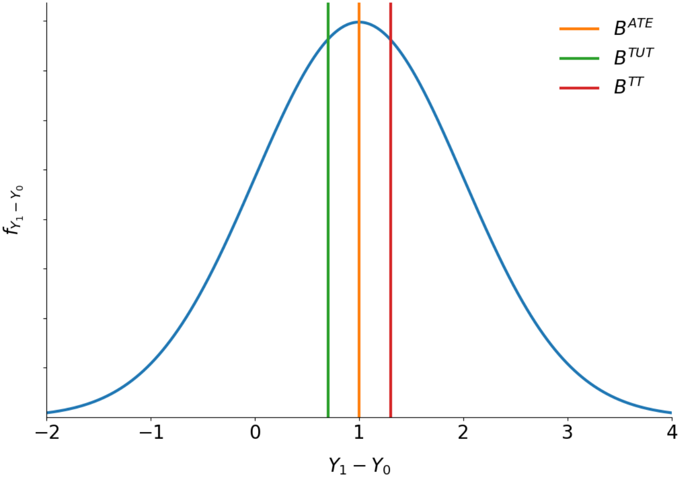
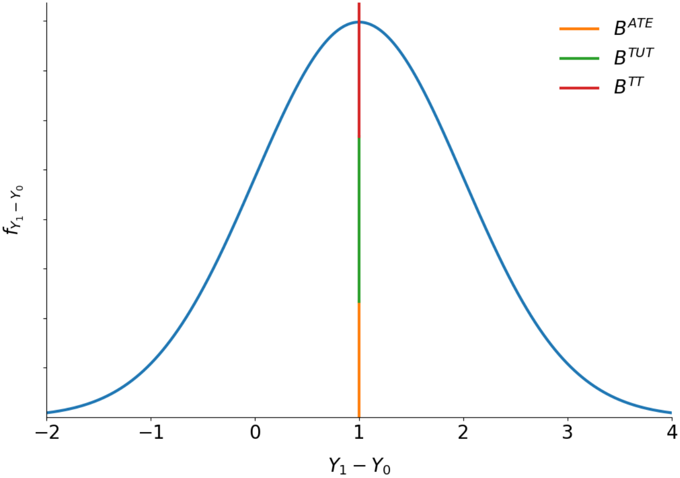


Figure: Distribution of effects without essential heterogeneity



Policy-Relevant Average Treatment Effects

Observed Outcomes

$$Y_B = D_B Y_1 + (1 - D_B) Y_0$$

$$Y_A = D_A Y_1 + (1 - D_A) Y_0$$

Effect of Policy

$$B^{PRTE} = \frac{1}{E[D_A] - E[D_B]} (E[Y_A] - E[Y_B])$$

Marginal Benefit of Treatment

Marginal Benefit of Treatment

$$B^{MTE}(x, u_D) = E[Y_1 - Y_0 \mid X = x, U_D = u_D]$$

Intuition: Mean gross return to treatment for persons at quantile u_D of the first-stage unobservable V or a willingness to pay for individuals at the margin of indifference.

Figure: Margin of indifference

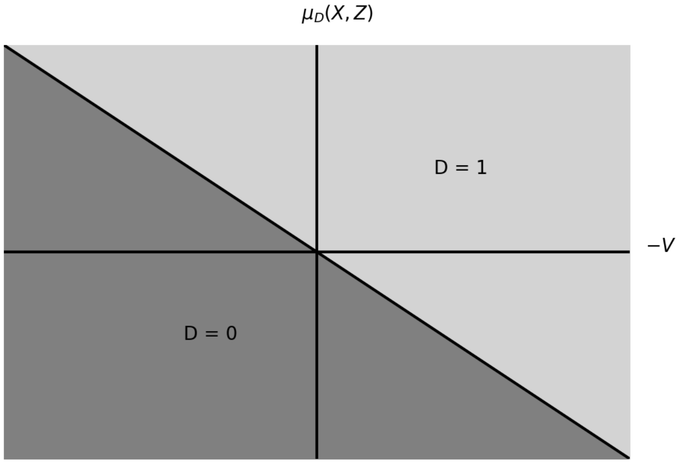
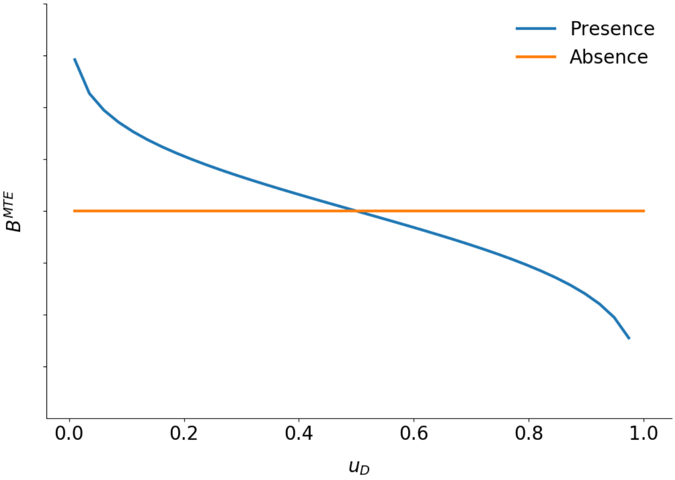


Figure: B^{MTE} and essential heterogeneity



Effects of treatment as weighted averages Parameter Δ_j, can be written as a weighted average of the $B^{MTE}(x, u_D)$.

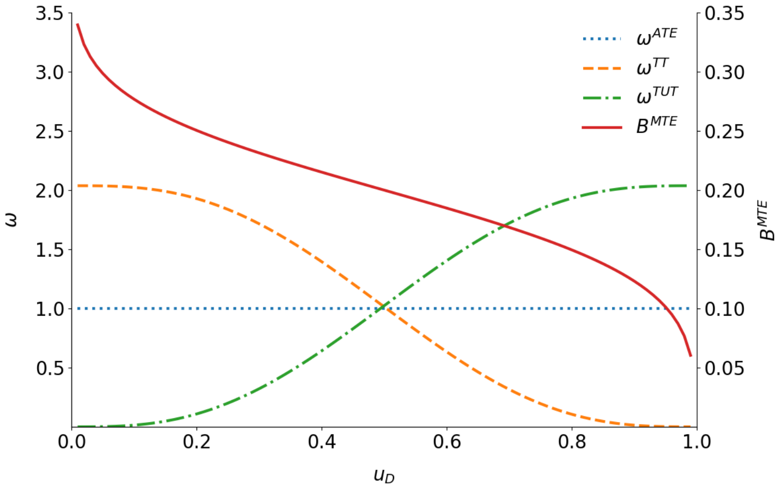
$$\Delta_j(x) = \int_0^1 B^{MTE}(x, u_D)\omega^j(x, u_D)du_D,$$

where the weights $\omega^j(x, u_D)$ are specific to parameter j and integrate to one.

Weights

$$
\begin{aligned}
\omega^{ATE}(x,u_D) &= 1 \\
\omega^{TT}(x,u_D) &= \frac{1-F_{P|X=x}(u_D)}{E[P \mid X=x]} \\
\omega^{TUT}(x,u_D) &= \frac{F_{P|X=x}(u_D)}{E[1-P \mid X=x]}
\end{aligned}
$$

Figure: Effects of treatment as weighted averages



Local Average Treatment Effect

Local Average Treatment Effect

- **Local Average Treatment Effect:** Average effect for those induced to change treatment because of a change in the instrument.

  ⇒ instrument-dependent parameter

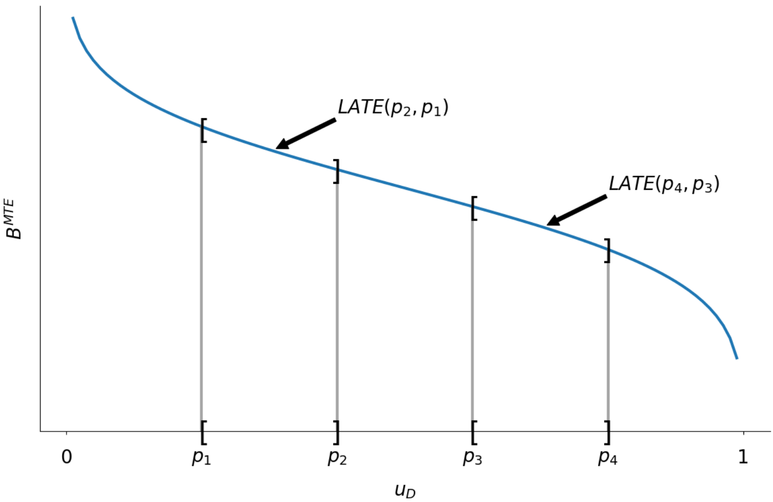
- **Marginal Treatment Effect:** Average effect for those individuals with a given unobserved desire to receive treatment.

  ⇒ deep economic parameter

$$B^{LATE} = \frac{E[Y \mid Z = z] - E[Y \mid Z = z']}{P(z) - P(z')}$$

$$B^{LATE}(x, u_D, u_{D'}) = \frac{1}{u_D - u_{D'}} \int_{u_D}^{u_{D'}} B^{MTE}(x, u) du,$$

Figure: Local average treatment effect



Distributions of Effects

Distrbutions of Effects

- marginal distribution of benefits
- joint distribution of potential outcomes
- joint distribution of benefits and surplus

Figure: Distribution of benefits

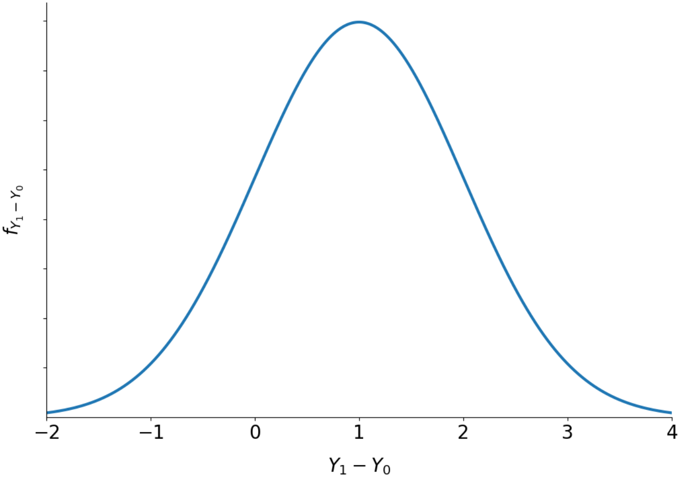


Figure: Distribution of potential outcomes

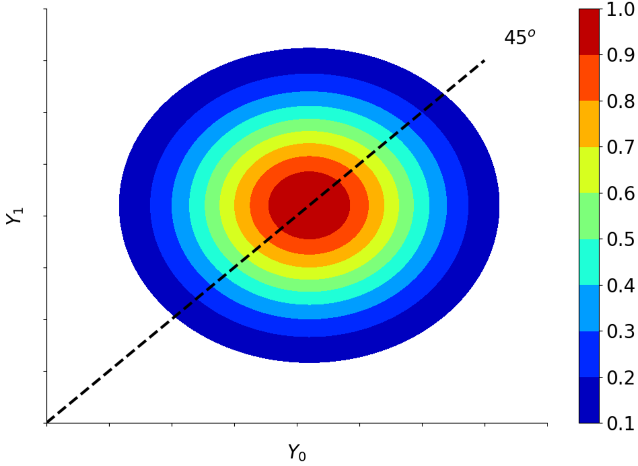
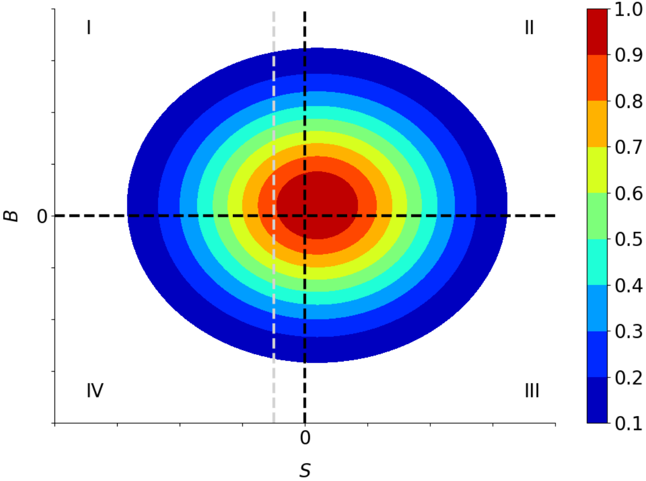


Figure: Distribution of benefits and surplus



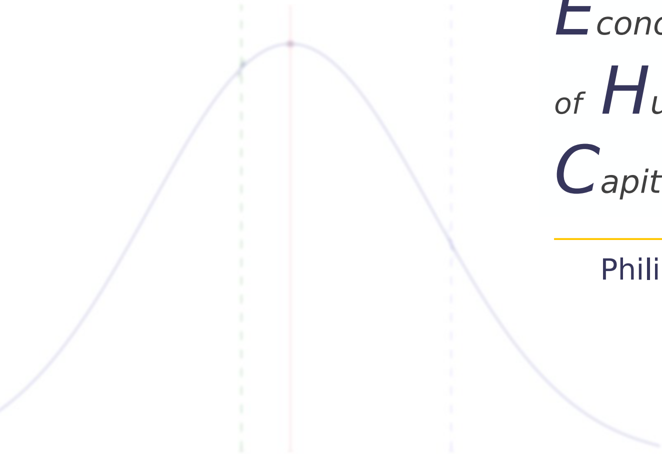
Conclusion

Appendix

References

Becker, G. S. (1964). *Human capital* (1st ed.). New York City, NY: Columbia University Press.

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiry*, *46*, 289–324.



Economics of Human Capital

Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

Economics of Human Capital

Dynamic model of human capital accumulation

Philipp Eisenhauer

Introduction

We build on the following seminal paper:

- Keane, M. P., & Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, *105*(3), 473–522.

Roadmap

- Economic Model
- Mathematical Model
- Data
- Computational Model
- Results

Economic Model

Decision Problem

$t = 1, \ldots, T$	decision period
$s \in S$	state
$a \in A$	action
$d_t(s)$	decision rule
$r_t(s, a)$	immediate reward

Timing of Events

Learn	Choose	Receive
$\{r_t(s,a)\}_{a\in A}$	$d_t(s)$	$r_t(s, d_t(s))$

t

Learn	Choose	Receive
$\{r_{t+1}(s',a)\}_{a\in A}$	$d_{t+1}(s')$	$r_{t+1}(s', d_{t+1}(s'))$

$t+1$

$\pi = (d_1, \ldots, d_T)$	policy
$h_t = (s_1, a_1, \ldots, s_t)$	history
δ	discount factor
$p_t(s, a)$	conditional distribution

Individual's Objective under Risk

$$v_1^{\pi^*}(s) = \max_{\pi \in \Pi} \mathrm{E}_s^{\pi}\left[\sum_{t=1}^{T} \delta^{t-1} r_t(X_t, d_t(X_t))\right]$$

Mathematical Model

Policy Evaluation

$$v_t^{\pi}(s) = \mathsf{E}_s^{\pi}\left[\sum_{\tau=t}^{T} \delta^{\tau-t} r_\tau(X_\tau, d_\tau(X_\tau))\right]$$

Inductive Scheme

$$v_t^{\pi}(s) = r_t(s, d_t(s)) + \delta \mathsf{E}_s^{\pi}\left[v_{t+1}^{\pi}(X_{t+1})\right]$$

Optimality Equations

$$v_t^{\pi^*}(s) = \max_{a \in A} \left\{ r_t(s,a) + \delta \mathsf{E}_s^p \left[v_{t+1}^{\pi^*}(X_{t+1}) \right] \right\}.$$

Backward Induction Algorithm for MDP

for $t = T, \ldots, 1$ **do**

if t == T **then**

$$v_T^{\pi^*}(s) = \max_{a \in A} \Big\{ r_T(s, a) \Big\} \qquad \forall \quad s \in S$$

else

Compute $v_t^{\pi^*}(s)$ for each $s \in S$ by

$$v_t^{\pi^*}(s) = \max_{a \in A} \Big\{ r_t(s, a) + \delta \mathrm{E}_s^p \Big[v_{t+1}^{\pi^*}(X_{t+1}) \Big] \Big\}$$

and set

$$d_t^{\pi^*}(s) = \arg\max_{a \in A} \Big\{ r_t(s, a) + \delta \mathrm{E}_s^p \Big[v_{t+1}^{\pi^*}(X_{t+1}) \Big] \Big\}$$

end if

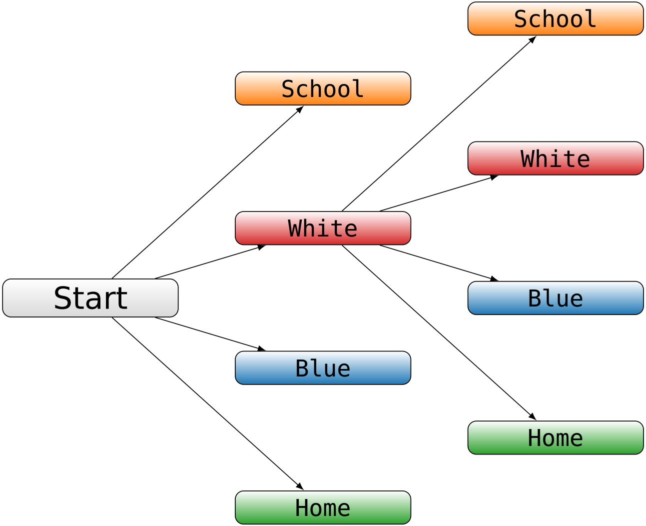
end for

Data

National Longitudinal Survey of Youth (1979)

- 1,373 white males starting at age 16
- life-cycle histories
  - school attendance
  - occupation-specific work status
  - real wages

Figure: Decision Tree



Descriptives

Figure: Sample Size

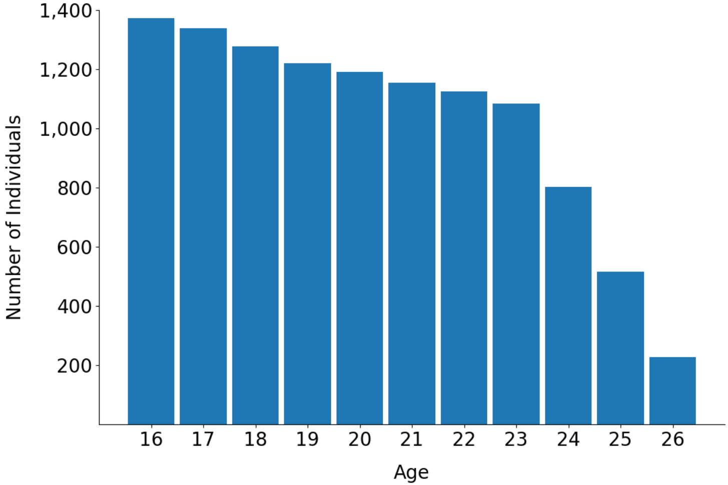


Figure: Observed Choices

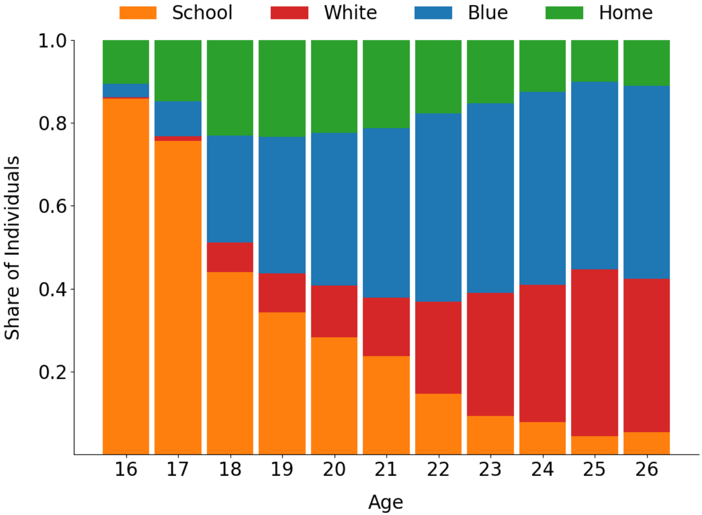


Table: Observed Real Wages

	White		Blue	
Age	Obs.	Mean	Obs.	Mean
16	2	.	26	10,287
20	128	5,499	349	14,432
25	201	16,540	222	21,991

Figure: Observed Transitions

Figure: Initial Schooling

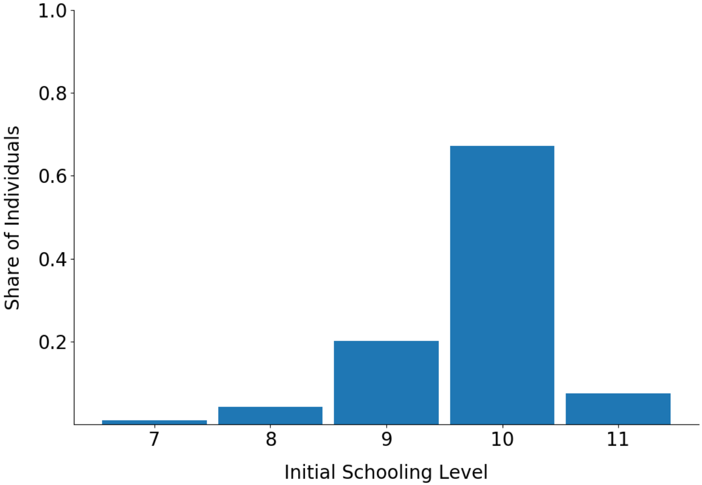


Table: Activities by Initial Schooling

	Initial Schooling				
Alternatives	7	8	9	10	11
School	0.69	0.86	2.48	3.37	2.83
White	0.08	0.38	0.65	1.36	2.04
Blue	3.69	3.62	3.05	2.40	1.98
Home	4.23	4.19	1.91	1.10	1.32
Total	8.69	9.05	8.09	8.24	8.17

Reduced-form Analysis

Table: Mincer Regressions

	Log Real Wages	
Intercept	8.314***	8.329***
Schooling	0.086***	0.077***
	Work Experience	
- linear	0.132***	0.125***
- squared	−0.005***	−0.003***
	Corrected AFQT	
- linear	—	0.002***
Adj-R^2	0.21	0.22
Observations	4,420	4,232

Table: Mincer Regressions

	Log Real Wages	
	White	Blue
Intercept	7.748***	8.790***
Schooling	0.128***	0.044***
	Own Experience	
- linear	0.146***	0.129***
- squared	−0.003	−0.005***
	Other Experience	
- linear	0.096***	0.085***
- squared	0.002	−0.003
Adj-R^2	0.28	0.17
Observations	1,468	2,952

Open Issues

- distinction between ex ante and ex post returns
- role of psychic costs
- nonlinearities in the return
- role of uncertainty

Computational Model

Additional Structure

t	age
k	unobserved type
$x_{j,t}$	experience in occupation j at age t
a_t	action at age j
g_t	level of schooling at age t

Skill Production Function

$$
\begin{aligned}
e_{j,k,t} = \exp\{ & e_{j,k,16} + \underbrace{\alpha_{j,1} g_t + \alpha_{j,2} \mathrm{I}[g_t \geq 12] + \alpha_{j,3} \mathrm{I}[g_t \geq 16]}_{\text{schooling}} \\
& + \underbrace{\alpha_{j,4} x_{j,t} + \alpha_{j,5} x_{j,t}^2 + \alpha_{j,6} \mathrm{I}[x_{j,t} > 0] + \alpha_{j,7} x_{j \neq j',t}}_{\text{work experience}} \\
& + \underbrace{\alpha_{j,8} \mathrm{I}[a_{t-1} \neq j]}_{\text{depreciation}} + \alpha_{j,9}(t-16) + \alpha_{j,10} \mathrm{I}[t < 18] + \epsilon_{j,t} \}
\end{aligned}
$$

with $j, j' = 1, 2$, $k = 1, \ldots, 4$, and $t = 16, \ldots, 65$

Labor Market

$$r_{j,k,t} = w_{j,k,t} + \underbrace{\kappa_1 \mathrm{I}[g_t \geq 12] + \kappa_2 \mathrm{I}[g_t \geq 16]}_{\text{common returns}} + \beta_{j,1}$$

$$+ \underbrace{\beta_{j,2} \mathrm{I}[x_{j,t} > 0, a_{t-1} \neq j] + \beta_{j,3} \mathrm{I}[x_{j,t} = 0, a_{t-1} \neq j]}_{\text{entry cost}}$$

with $w_{j,k,t} = r_j e_{j,k,t}$

School

$$
\begin{aligned}
r_{3,k,t} = e_{3,k,16} &+ \underbrace{\gamma_1 \mathrm{I}[g_t \geq 12] + \gamma_2 \mathrm{I}[g_t \geq 16]}_{\text{monetary and psychic cost}} \\
&+ \underbrace{\gamma_3 \mathrm{I}[a_{t-1} \neq 3, g_t \leq 11] + \gamma_4 \mathrm{I}[a_{t-1} \neq 3, g_t \geq 12]}_{\text{reenrollment cost}} \\
&+ \gamma_5 (t - 16) + \gamma_6 \mathrm{I}[t \leq 18] + \underbrace{\kappa_1 \mathrm{I}[g_t \geq 12] + \kappa_2 \mathrm{I}[g_t \geq 16]}_{\text{common returns}} \\
&+ \epsilon_{3,t}
\end{aligned}
$$

Home

$$r_{4,k,t} = e_{4,k,16} + \zeta_1 \mathrm{I}[18 \leq t \leq 20] + \zeta_2 \mathrm{I}[t \geq 21]$$
$$+ \underbrace{\kappa_1 \mathrm{I}[g_t \geq 12] + \kappa_2 \mathrm{I}[g_t \geq 16]}_{\text{common returns}} + \epsilon_{4,t}$$

State Space

- at time t

$$s_t = \{g_t, \{x_{j,t}\}_{j=1,2}, a_{t-1}, \{\epsilon_{j,t}\}_{j=1,\dots,4}\}$$
$$\bar{s}_t = \{g_t, \{x_{j,t}\}_{j=1,2}, a_{t-1}\}$$

- laws of motion

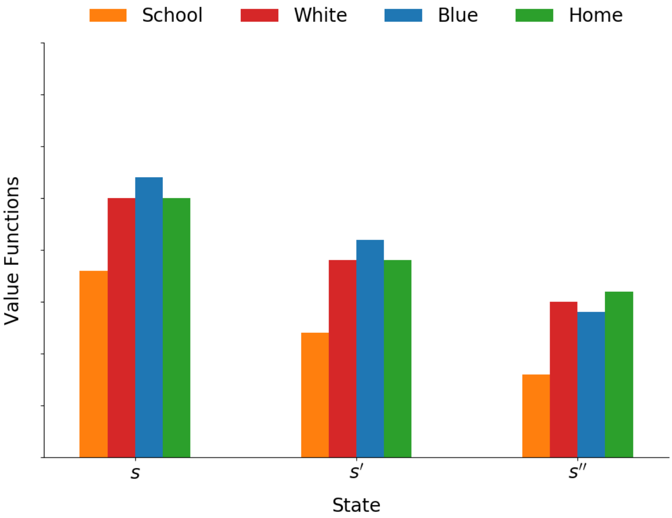
$$x_{j,t+1} = x_{j,t} + \mathrm{I}[a_t = j] \qquad \forall \quad j \in \{1, 2\}$$

$$g_{t+1} = g_t + \mathrm{I}[a_t = 3]$$

Distribution of shocks

$$[\epsilon_{1,t}, \epsilon_{2,t}, \epsilon_{3,t}, \epsilon_{4,t}]^T \sim \mathcal{N}_0(\mathbf{0}, \Sigma)$$

Figure: Value Functions



Computational Tool

https://respy.readthedocs.io

- Technical Documentation
  - Numerical Methods, Source Codes, Test Suite
- User Documentation
  - Tutorial

⇒ Transparency, Recomputability, and Extensibility

Conclusion

Appendix

References

Becker, G. S. (1964). *Human capital* (1st ed.). New York City, NY: Columbia University Press.

Keane, M. P., & Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, *105*(3), 473–522.