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

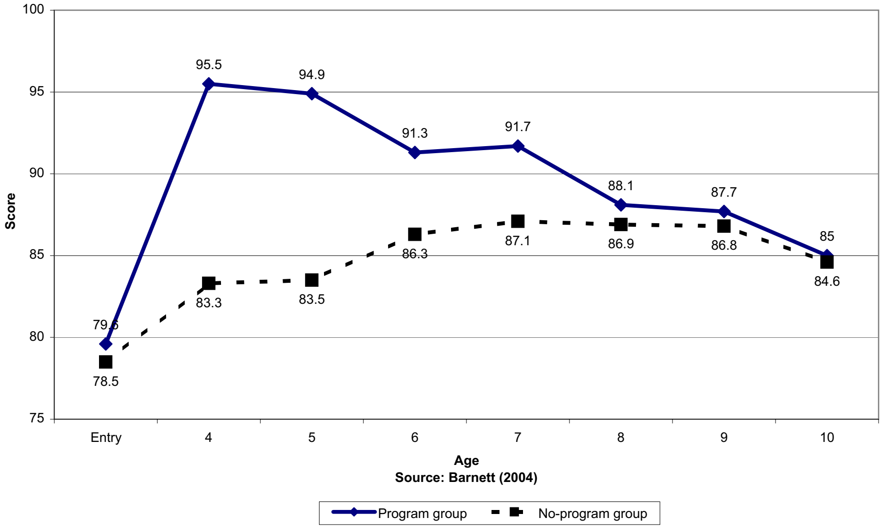
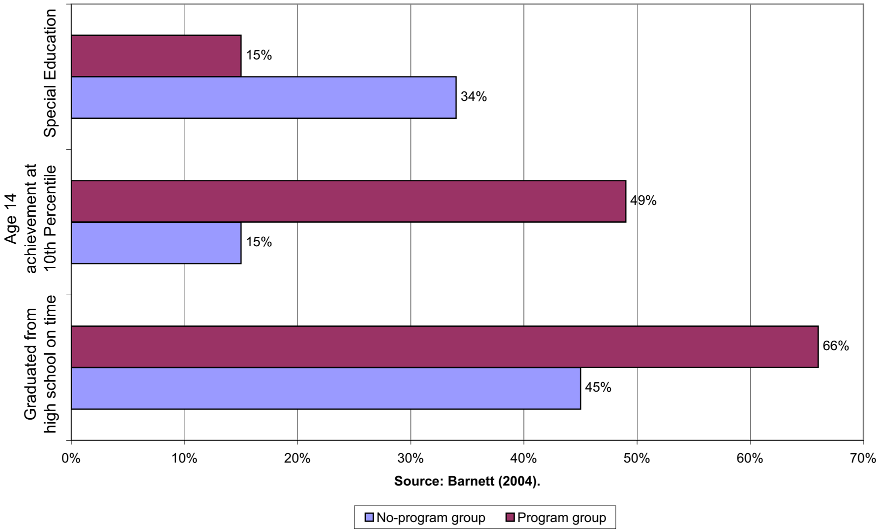
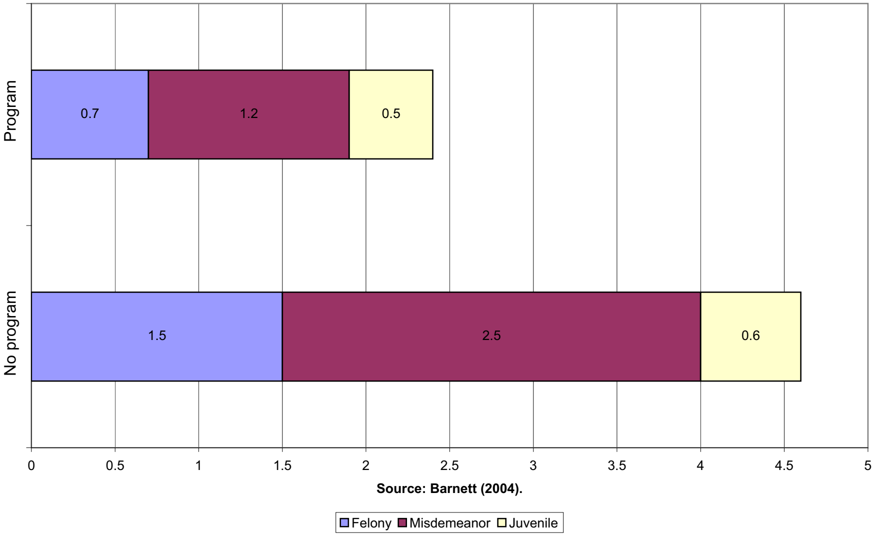


Figure 1B
Perry Preschool: Educational Effects



Source: Barnett (2004).

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



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

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

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.

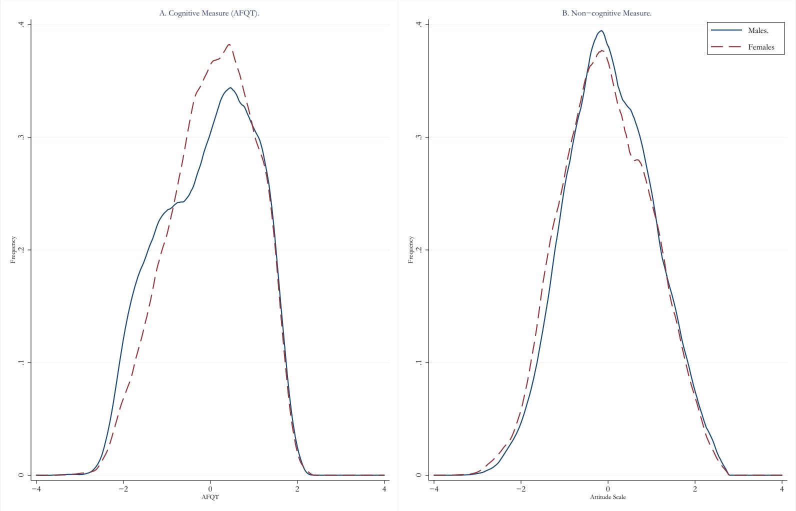
- There are no natural units for latent ability. Therefore, for some C_i ($i = 1, \ldots, n_C$) and N_j ($j = 1, \ldots, n_N$) we set $\alpha_{C_i} = \alpha_{N_i} = 1$.

- 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

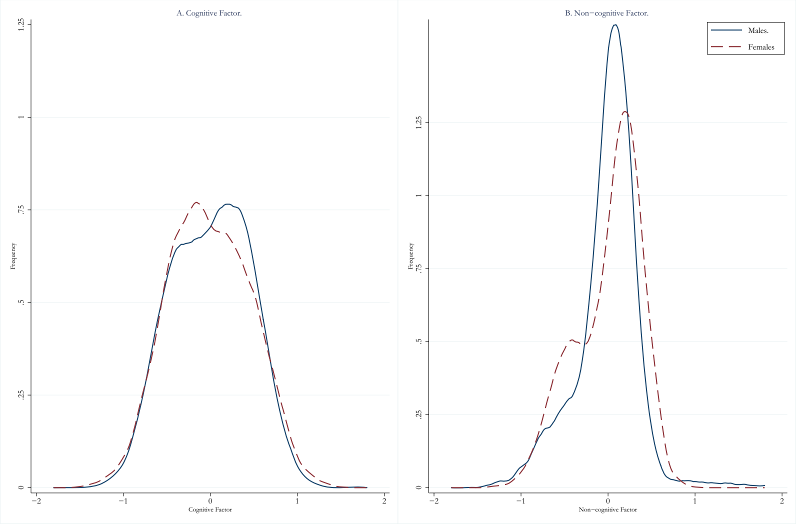
- We use a robust semiparametric approach to estimation.
- We make no distributional assumptions.
- Our evidence argues strongly against normality.
- Male distributions more variable; higher right tail in male cognitive distributions.
- Thicker lower tail for male noncognitive distributions.

Figure 2A. Distribution of Test Scores by Gender



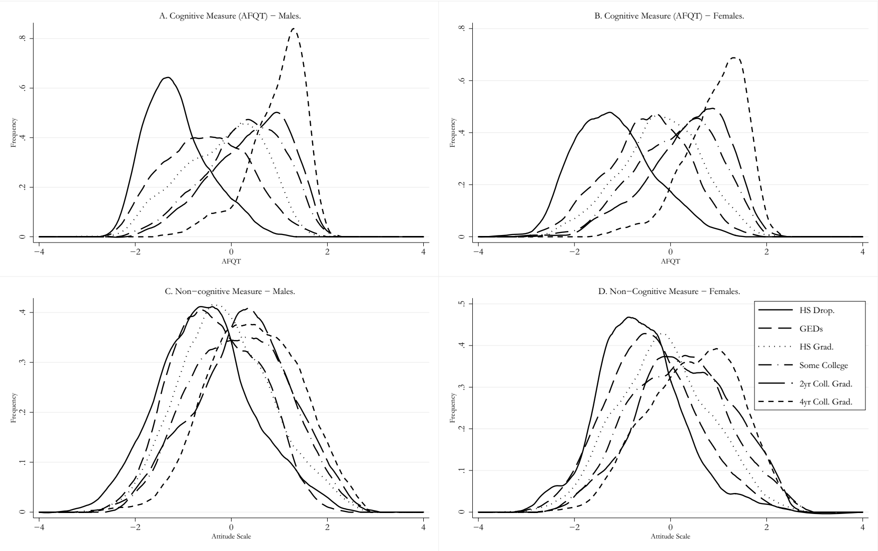
Notes: The AFQT is the mean raw score computed using ASVAB tests.The Attitude Scale is the average raw score between the Rosenberg scale of Self−Steem and the Rotter scale of internal−external locus of control.

Figure 2B. Distribution of Factors by Gender



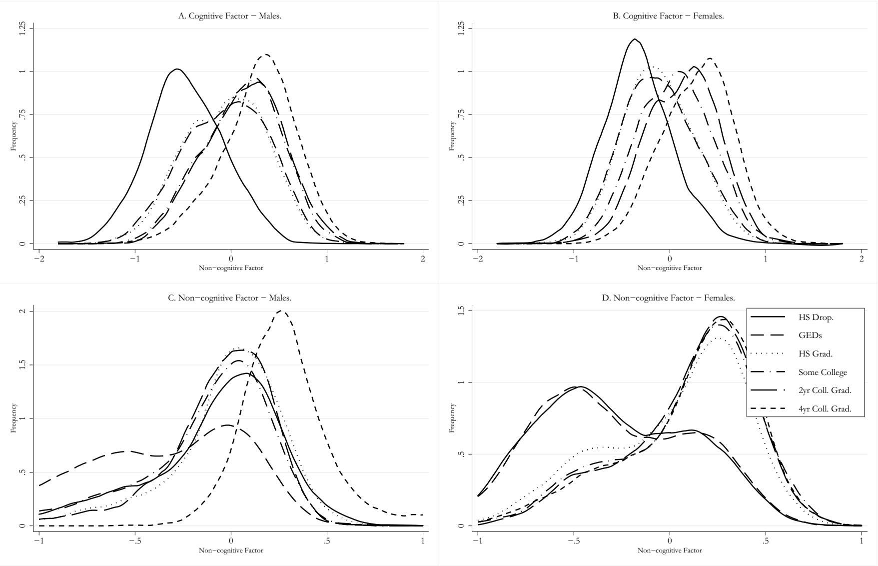
Notes: The factors are simulated from the estimates of the model. The simulated data contain 19,600 observations.

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

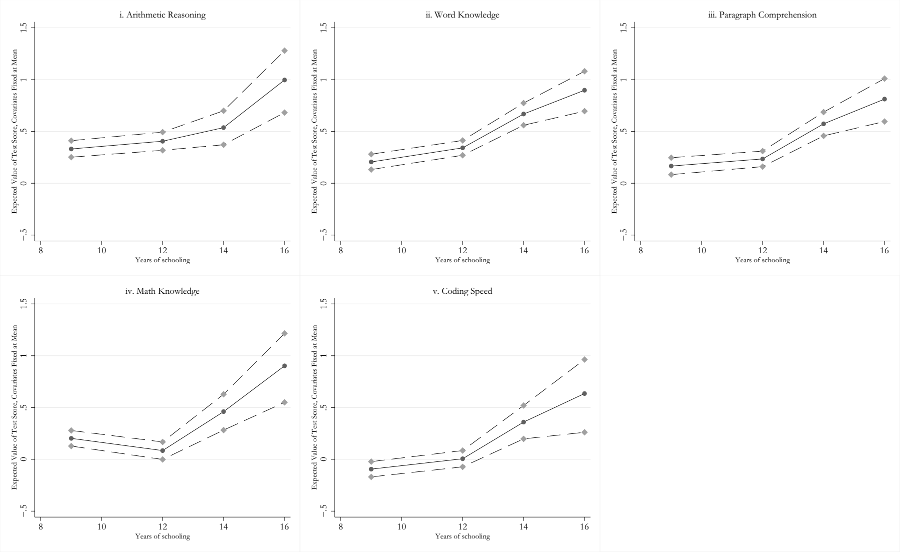


Notes: The factors are simulated from the estimates of the model. The schooling levels represent the predicted schooling level by age 30. These schooling levels are obtained. from the structure and estimates of the model and our sample of the NLSY79 (See Appendix A for details). The simulated data contain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

The Effect of Schooling on Test Scores

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

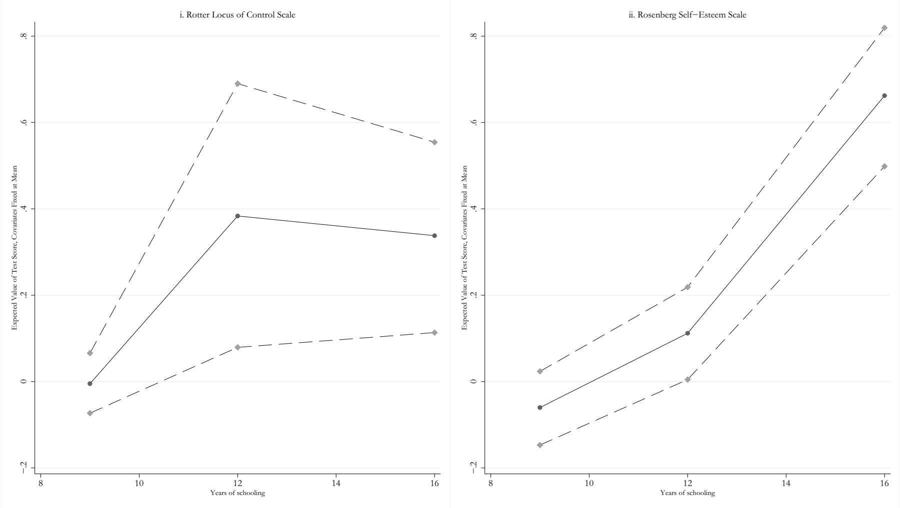
with 95% confidence bands−−Males



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

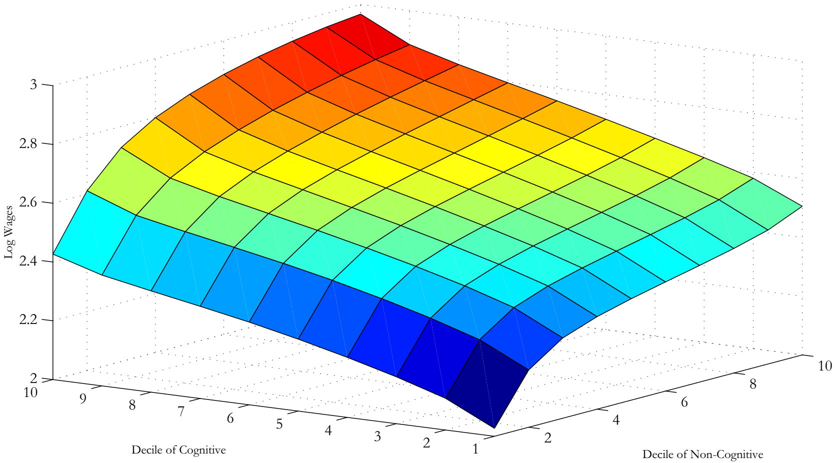
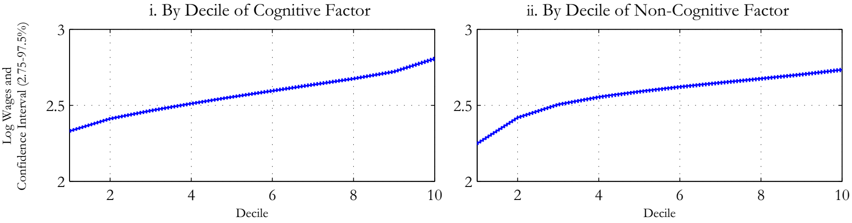


Figure 5B. Mean Log Wages by Age 30 - Males



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

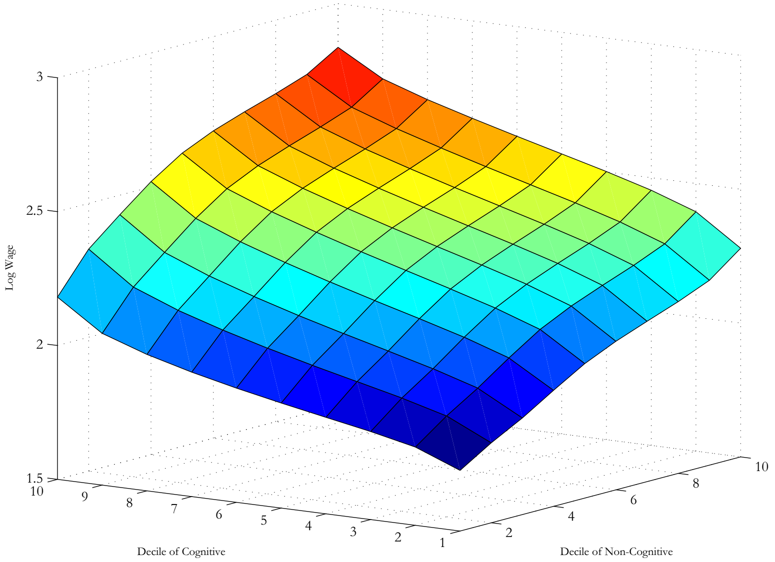
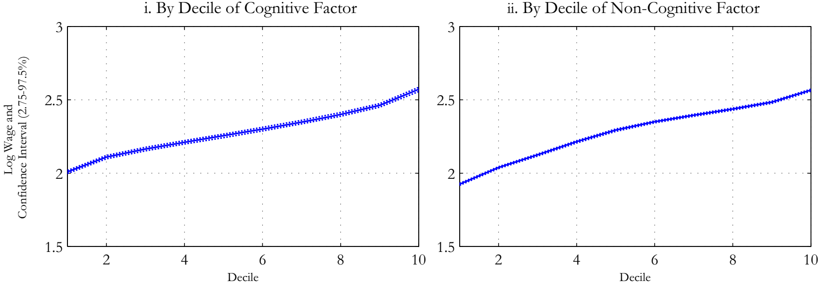


Figure 6B. Mean Log Wages by Age 30 - Females



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

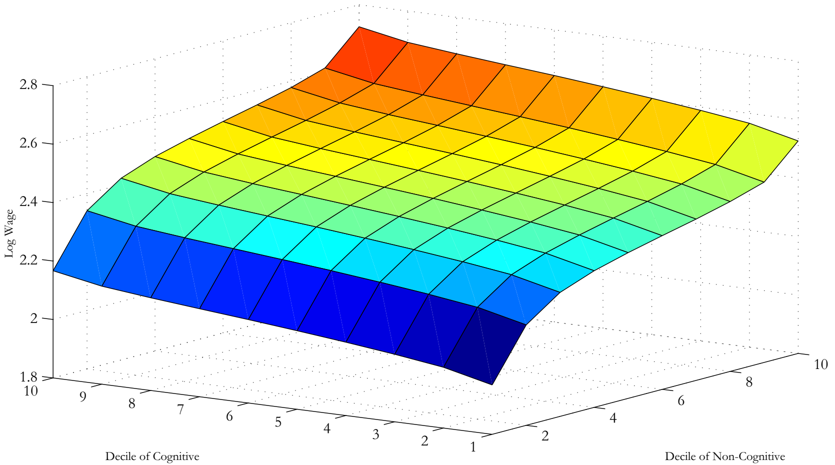
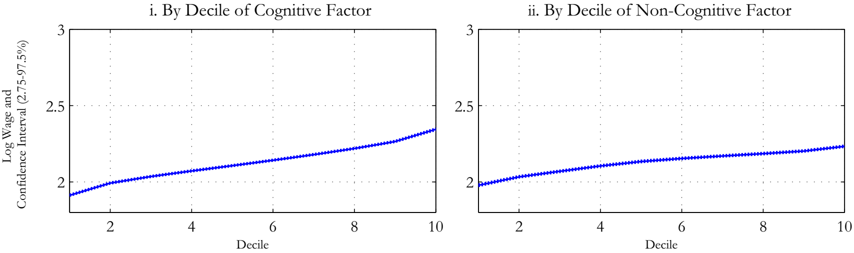


Figure 7B. Mean Log Wages of High School Dropouts by Age 30 - Males



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

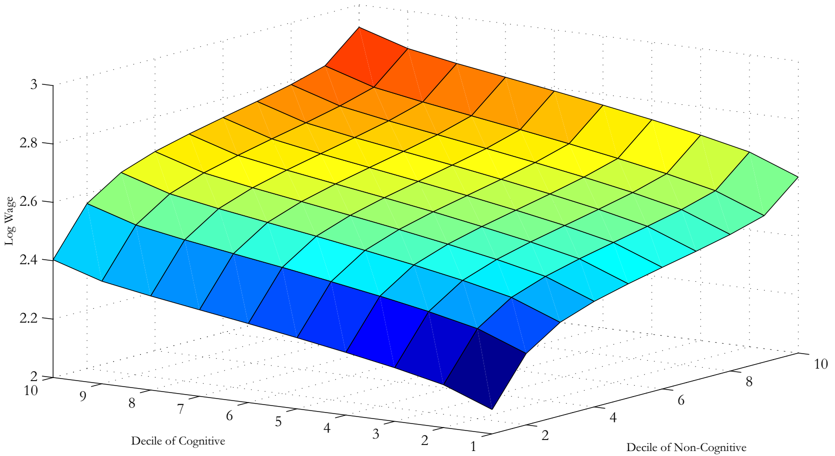
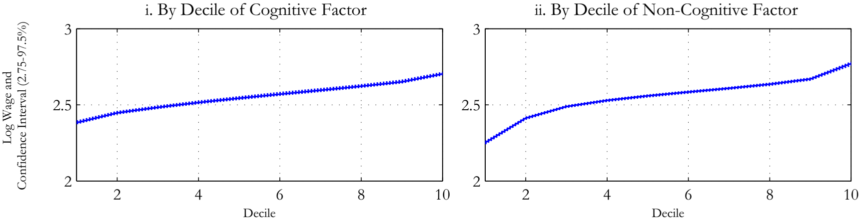


Figure 8B. Mean Log Wages of GEDs by Age 30 - Males



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

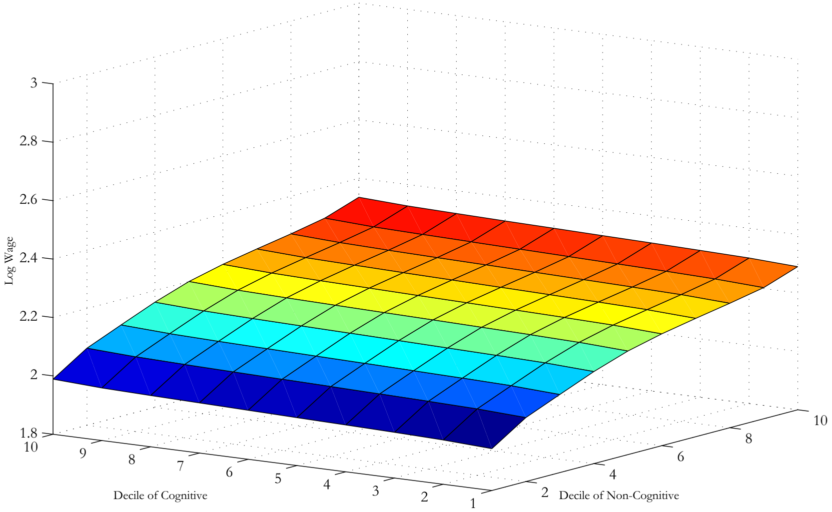
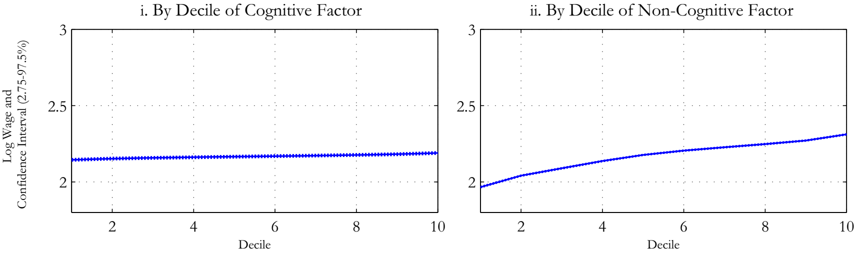


Figure 9B. Mean Log Wages of GEDs by Age 30 - Females



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

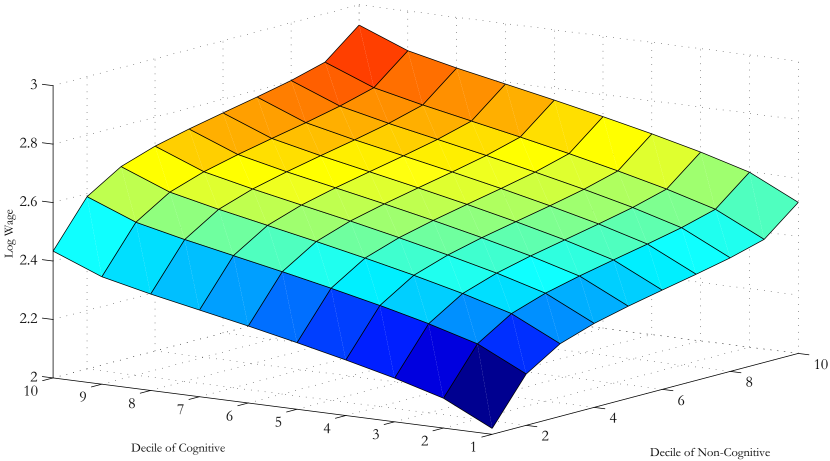
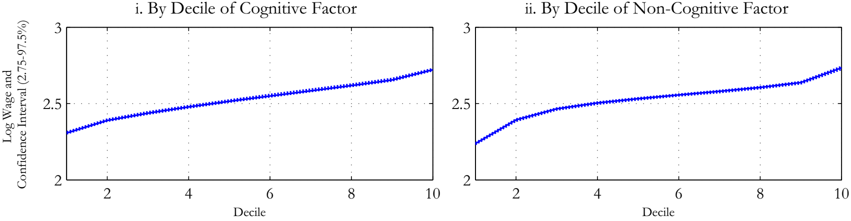


Figure 10B. Mean Log Wages of High School Graduates by Age 30 - Males



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

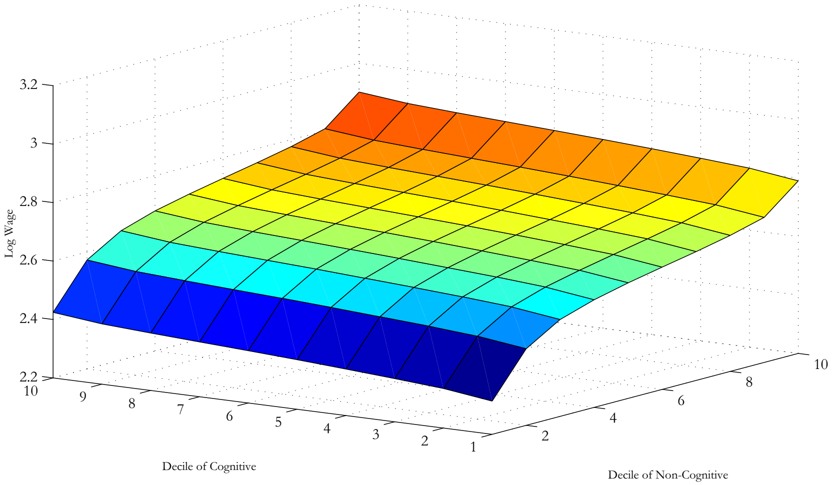
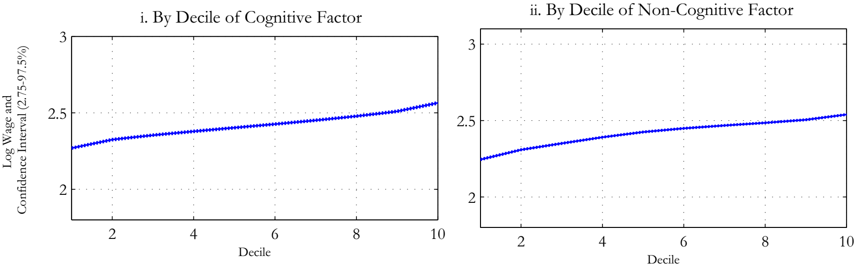


Figure 11B. Mean Log Wages of 2-yr College Graduates by Age 30 - Males



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

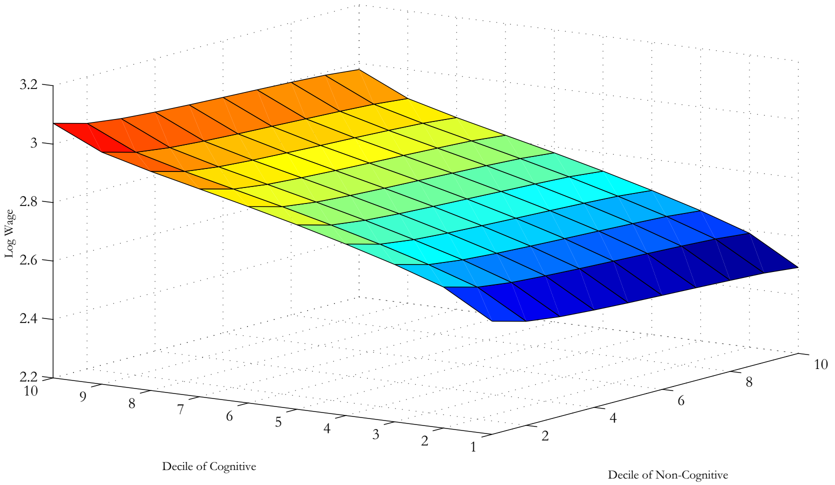
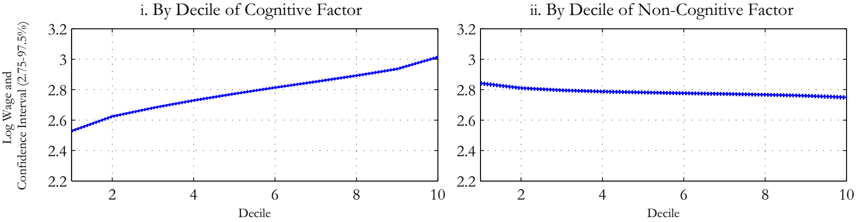


Figure 12 B. Mean Log Wages of 4-yr College Graduates by Age 30 - Males



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

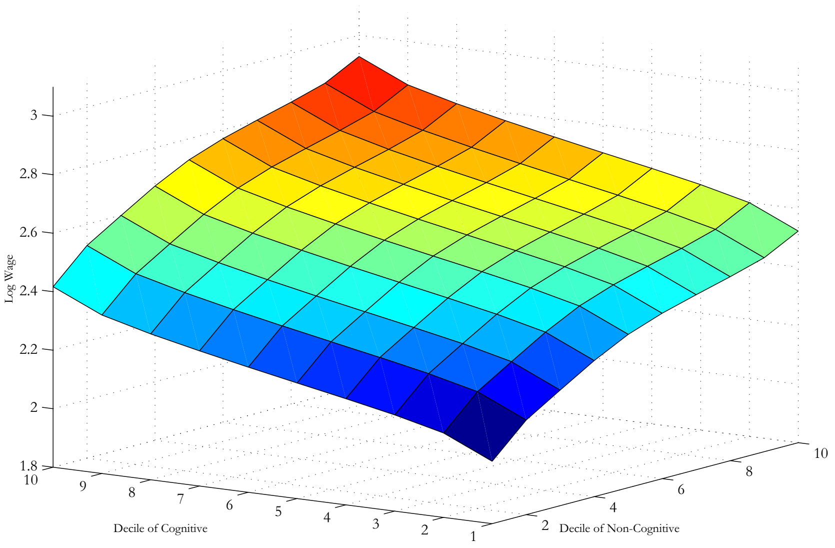
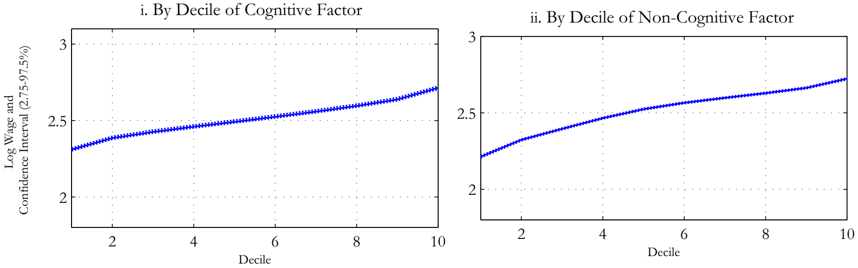


Figure 13 B. Mean Log Wages of 4-yr College Graduates by Age 30 - Females



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

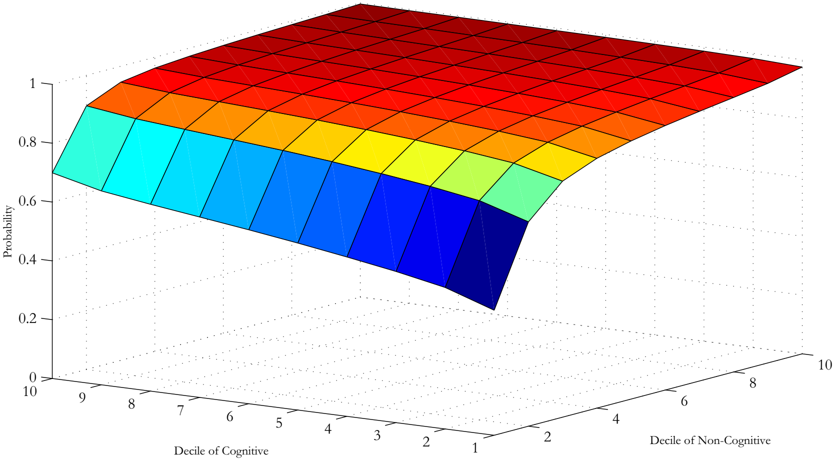
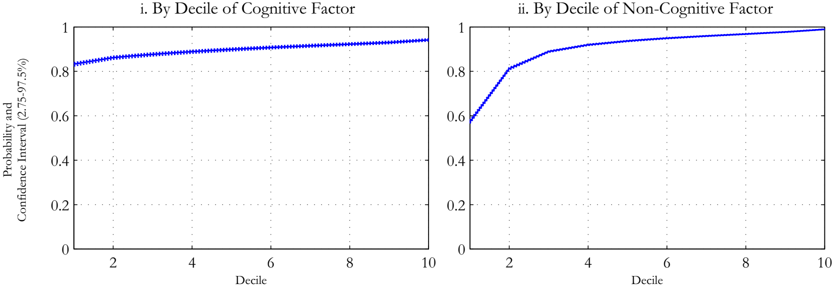


Figure 14B. Probability of Employment by Age 30 - Males



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

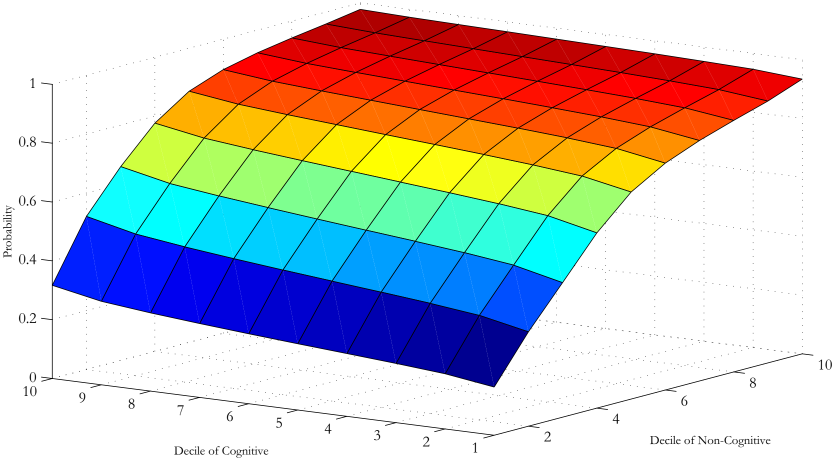
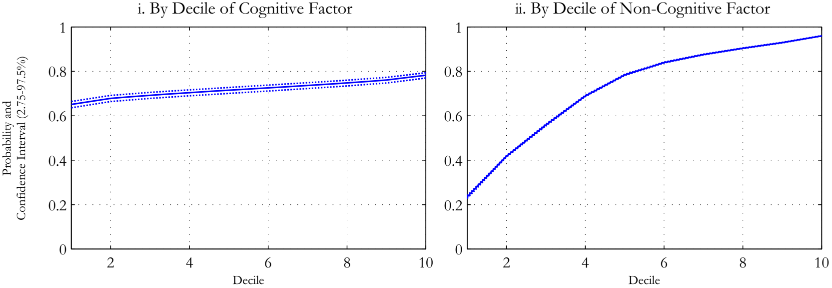


Figure 15B. Probability of Employment by Age 30 - Females



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

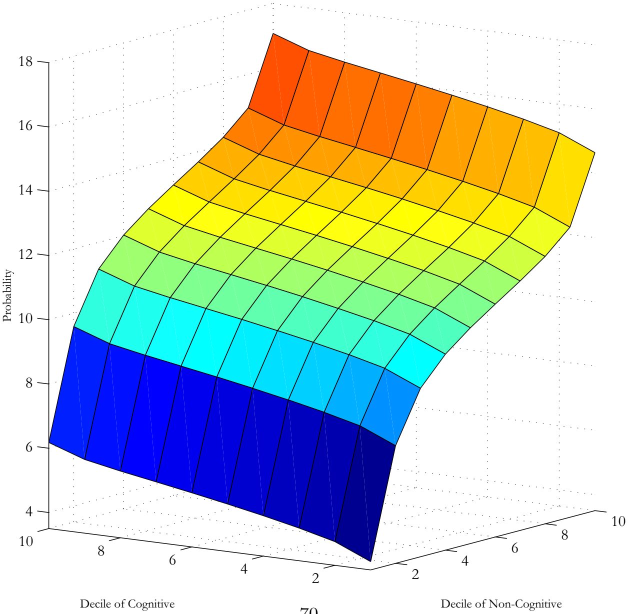
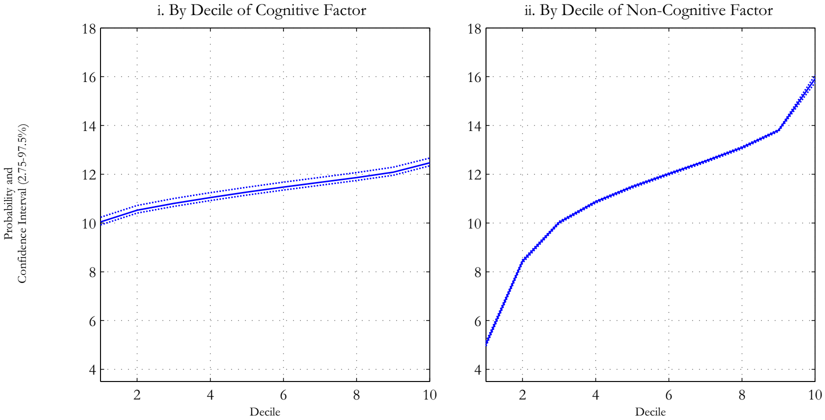


Figure 16B. Mean Work Experience of High School Dropouts by Age 30 - Males



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

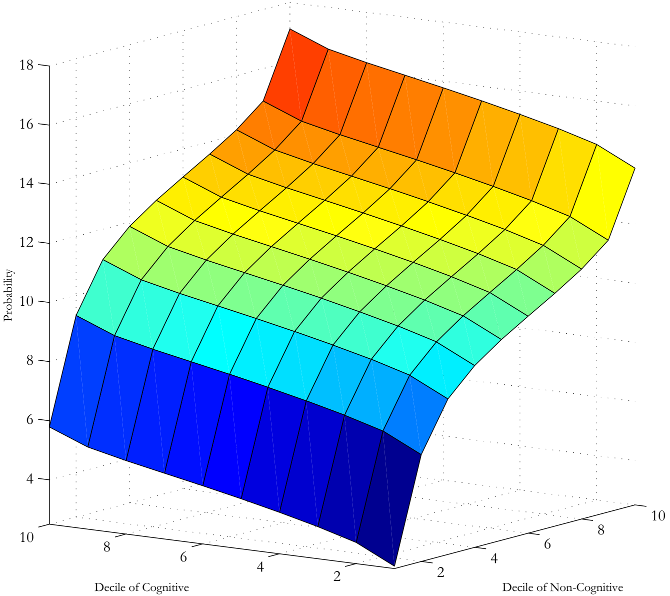
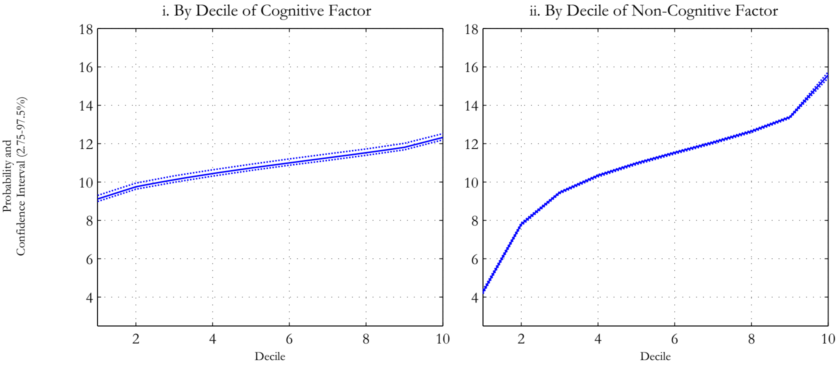


Figure 17B Mean Work Experience of High School Graduates by Age 30 - Males



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

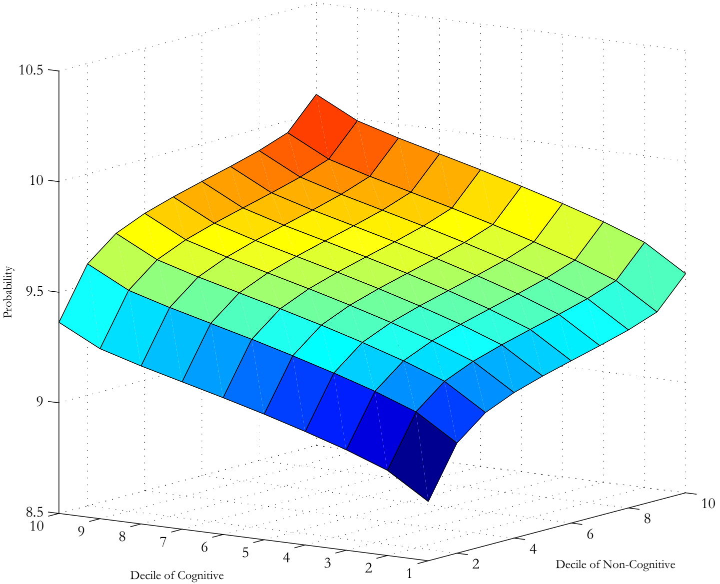
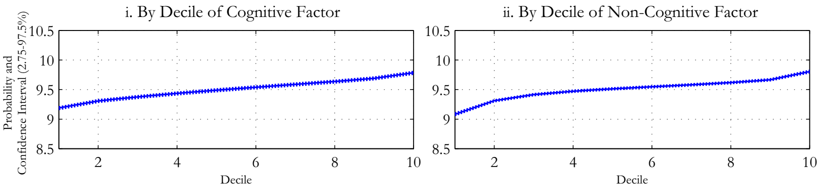


Figure 18B. Mean Work Experience of 4-yr College Graduates by Age 30 - Males



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

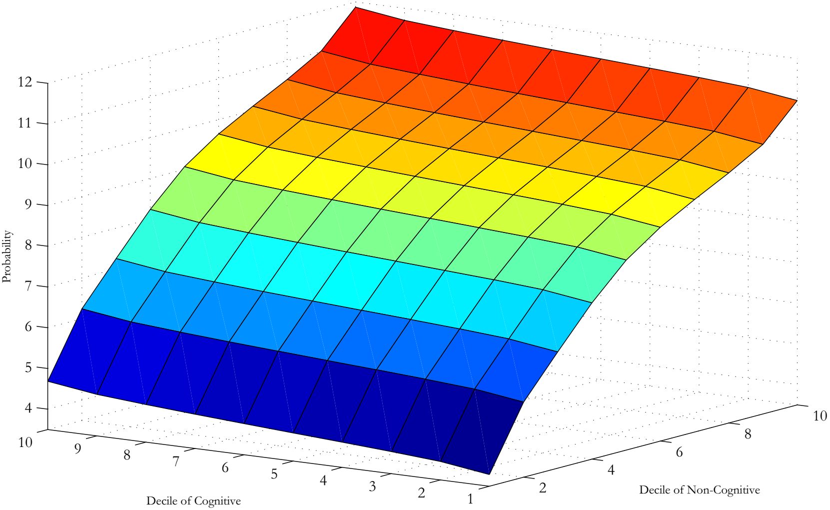
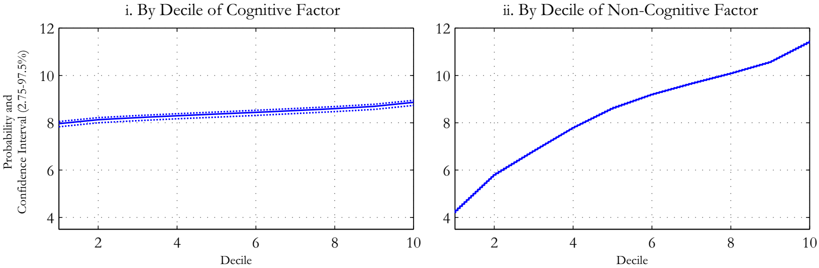


Figure 19B. Mean Work Experience of 4-yr College Graduates by Age 30 - Females



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

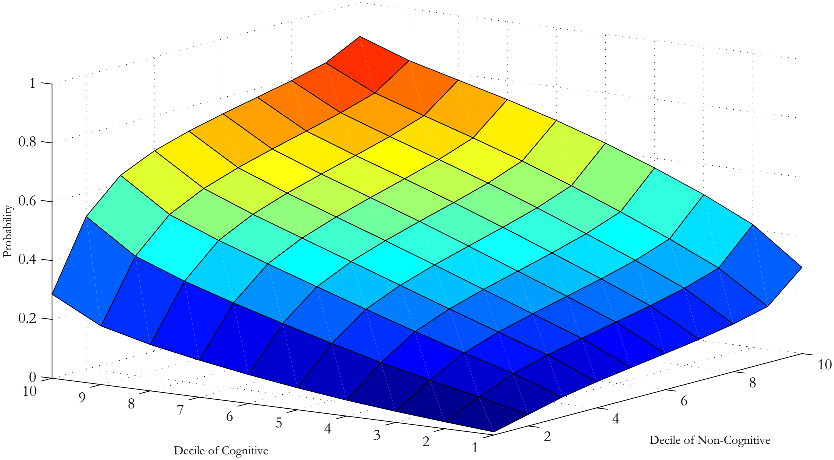
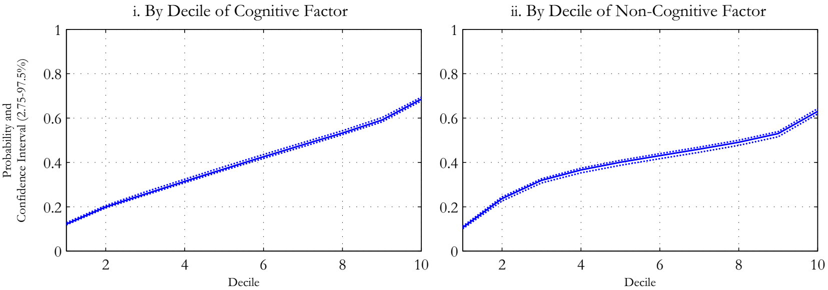


Figure 20B. Probability Of Being a White Collar Worker by Age 30 - Males



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

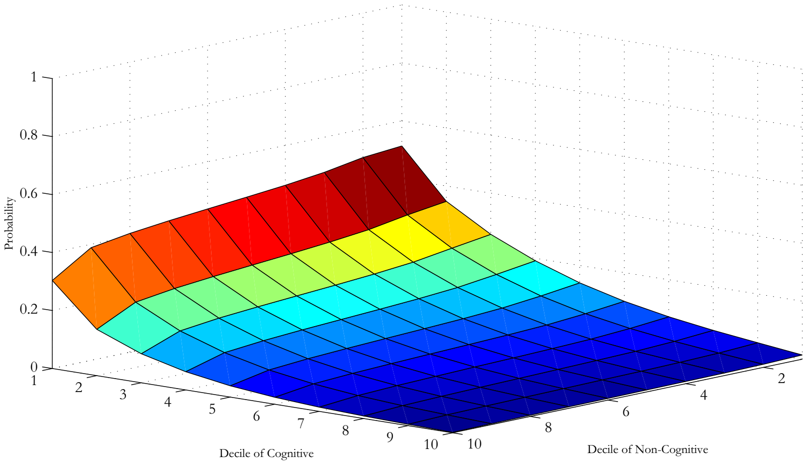
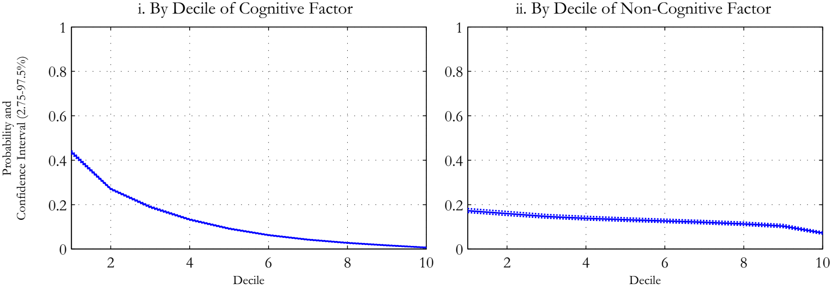


Figure 21B. Probability of Being a High School Dropout by Age 30 - Males



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

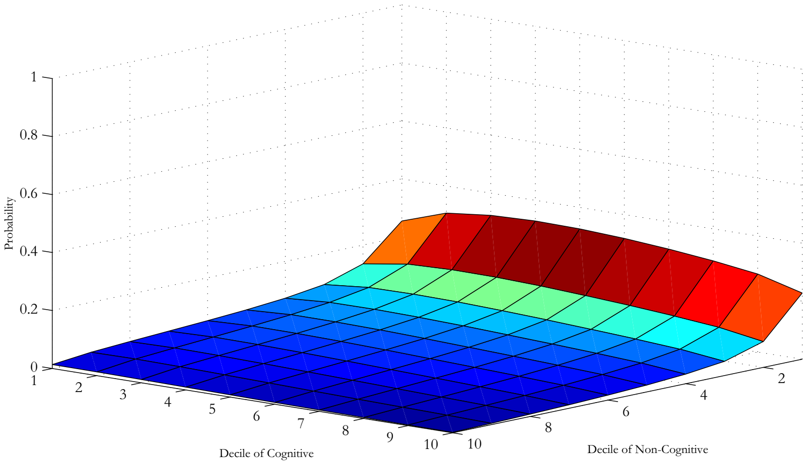
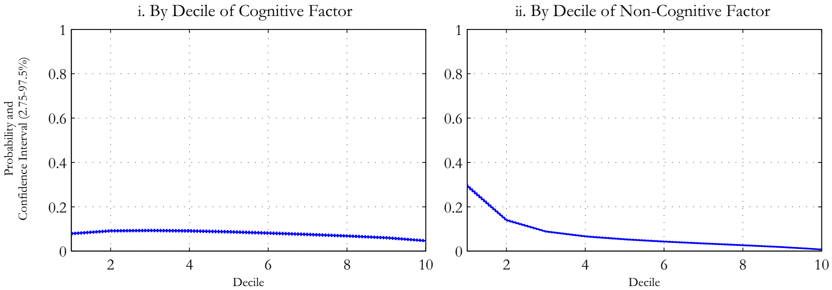


Figure 22 B. Probability of Being a GED by Age 30 - Males



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

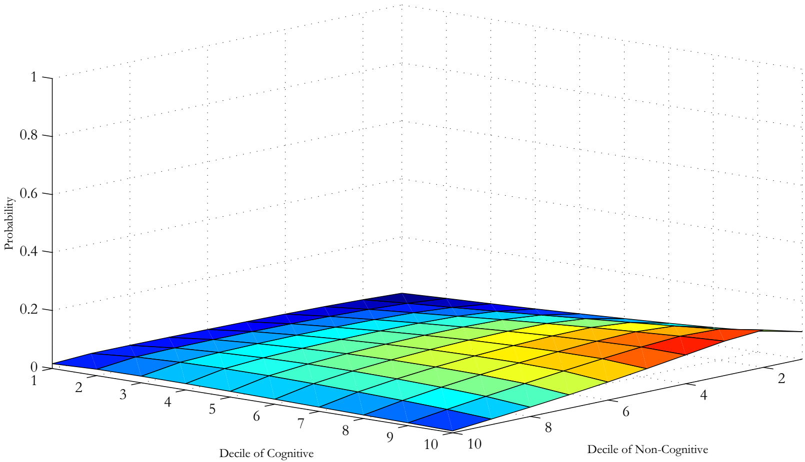
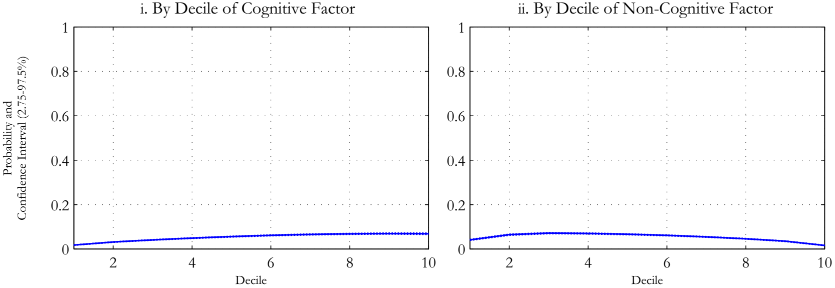


Figure 23 B. Probability of Being a 2-yr College Graduate by Age 30 - Males



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

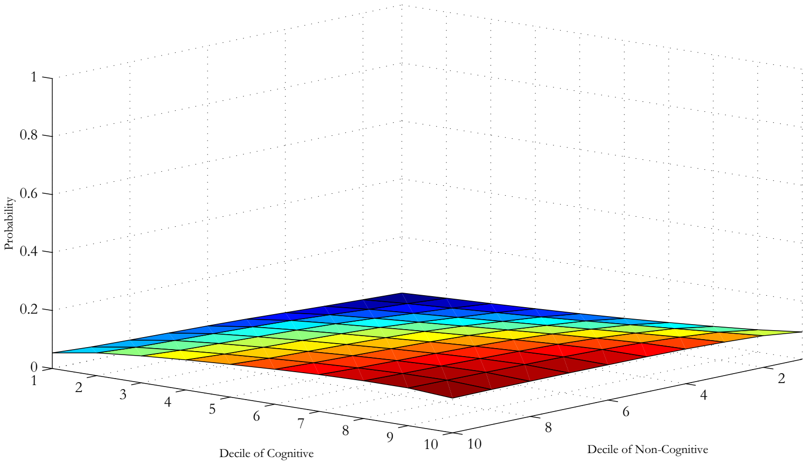
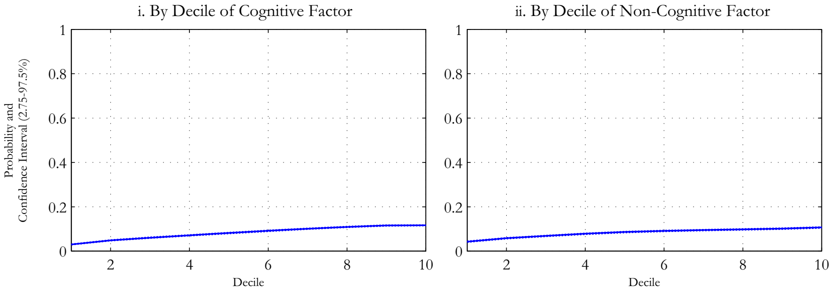


Figure 24B . Probability of Being a 2-yr College Graduate by Age 30 - Females



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

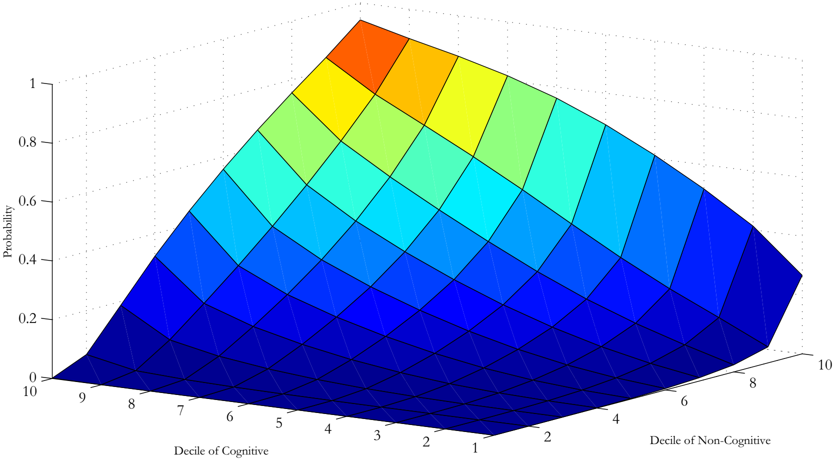
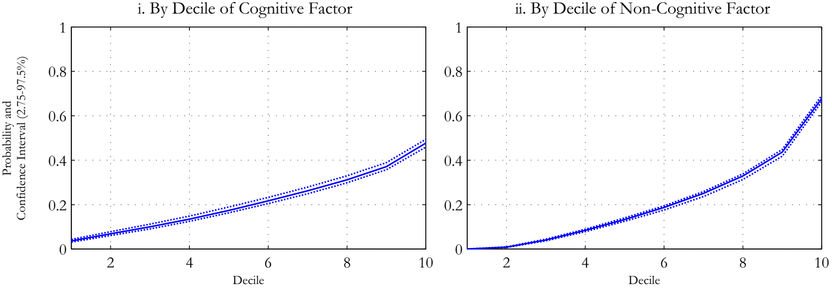


Figure 25 B. Probability of Being a 4-yr College Graduate by Age 30 - Males



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

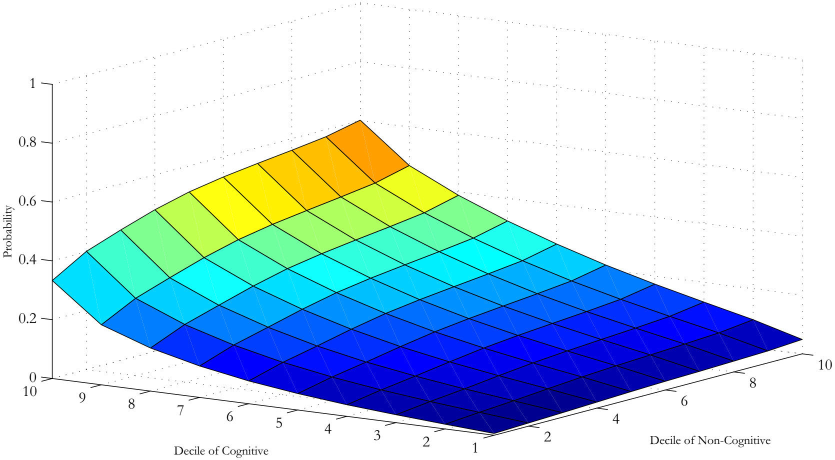
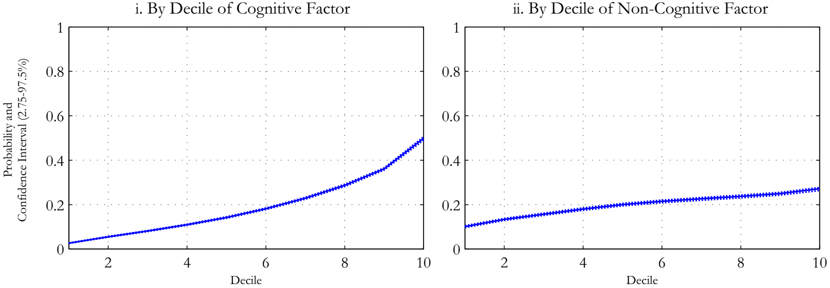


Figure 26 B. Probability of Being a 4-yr College Graduate by Age 30 - Females



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

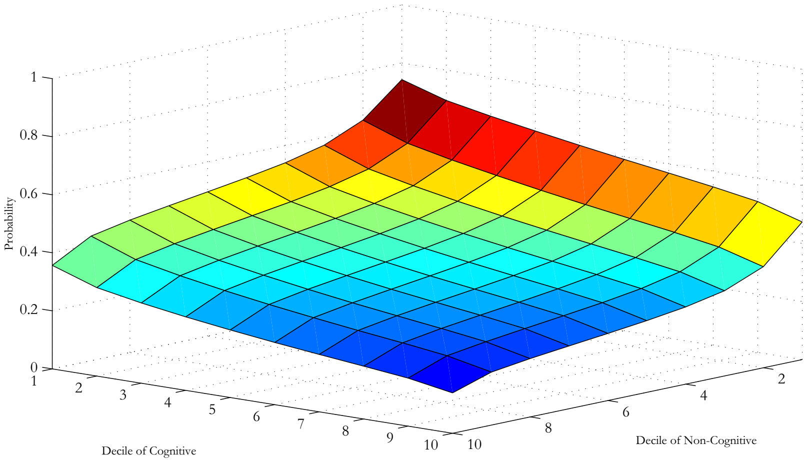
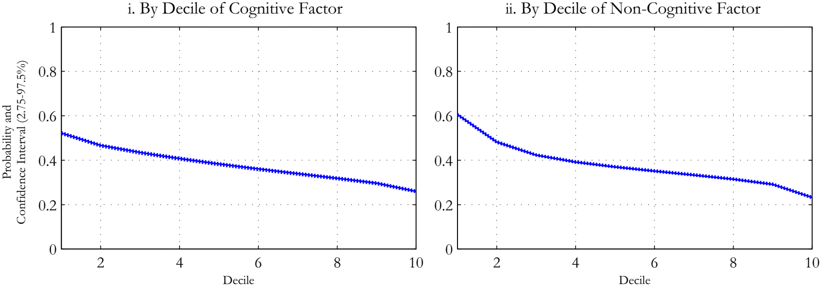


Figure 27 B. Probability Of Daily Smoking By Age 18 - Males



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

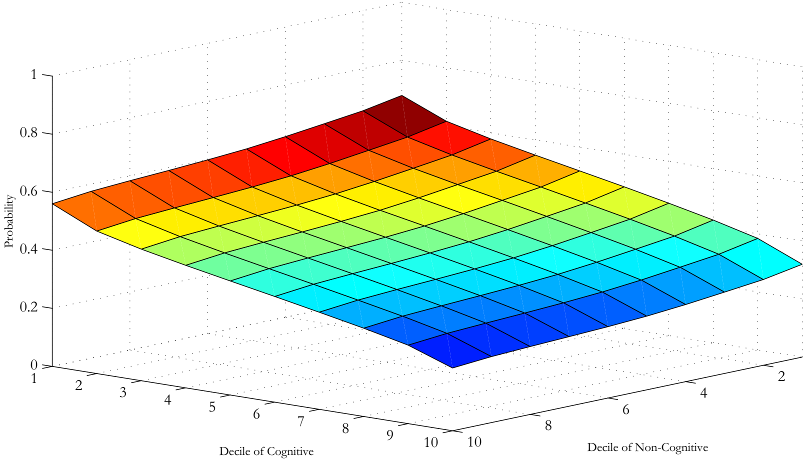
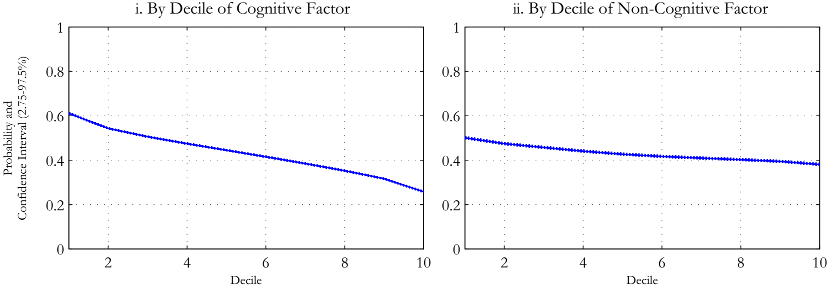


Figure 28 B. Probability Of Daily Smoking By Age 18 - Females



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

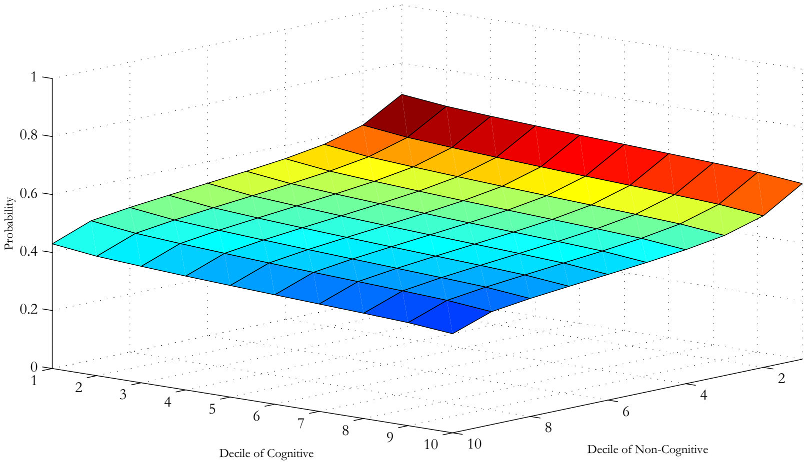
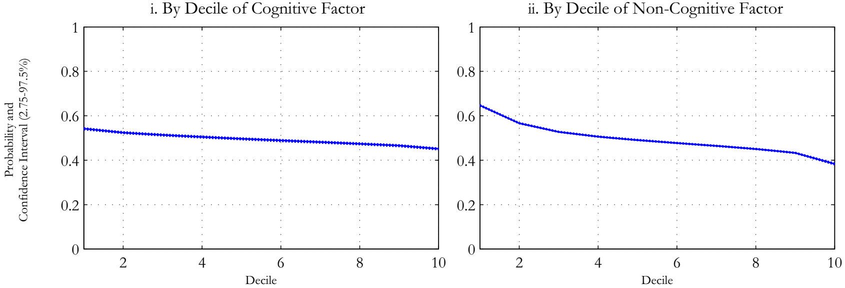


Figure 29 B Probability of Smoking Marijuana during the Year 1979 - Males



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

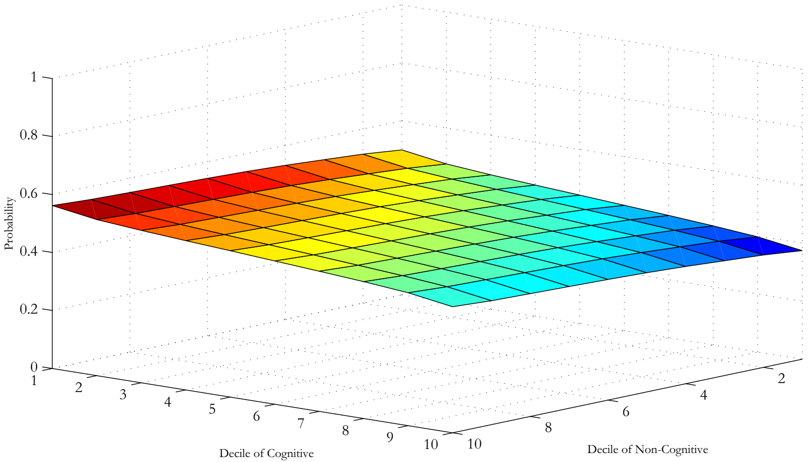
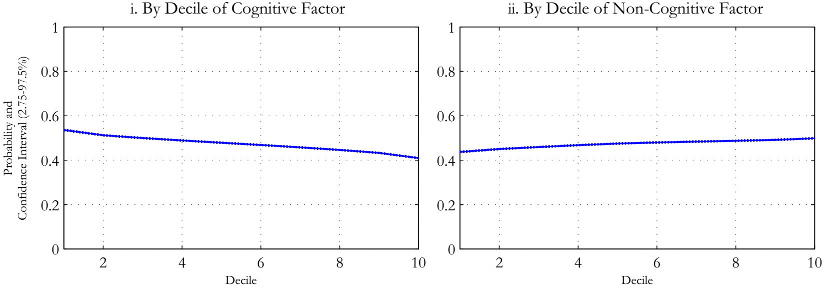


Figure 30 B. Probability of Smoking Marijuana during the Year 1979 - Females



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

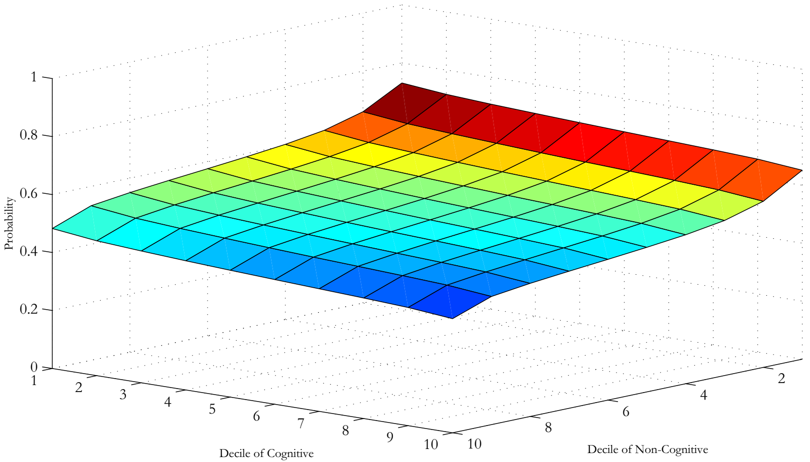
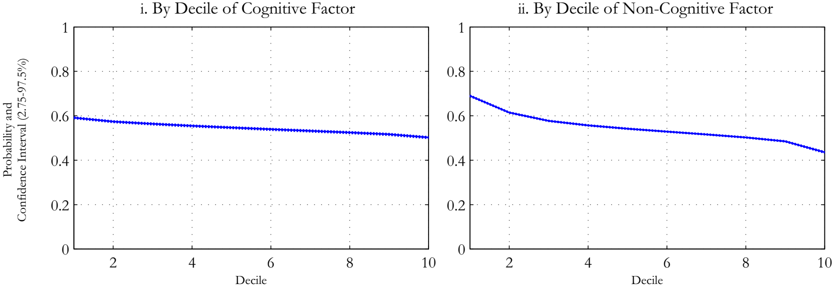


Figure 31 B. Probability of Participating in Illegal Activities during the Year 1979- Males



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

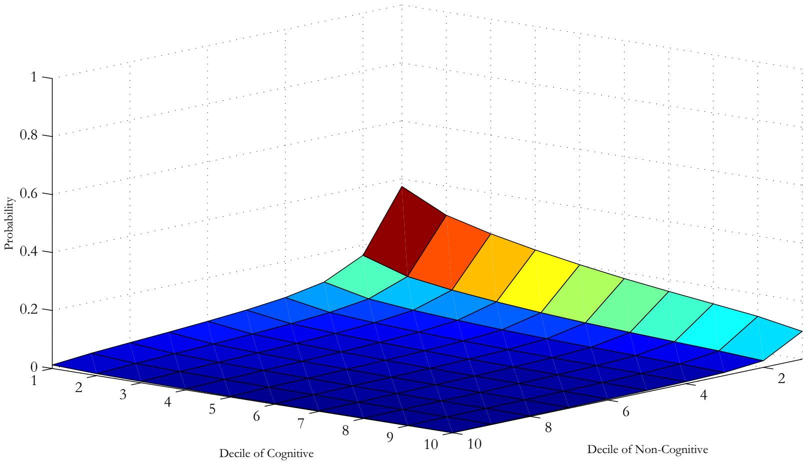
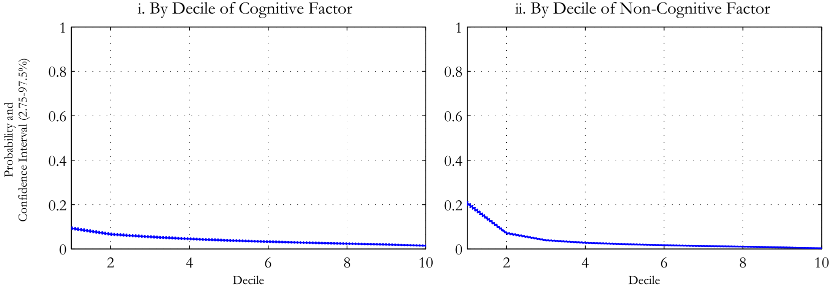


Figure 32B . Probability of Incarceration by Age 30 - Males



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

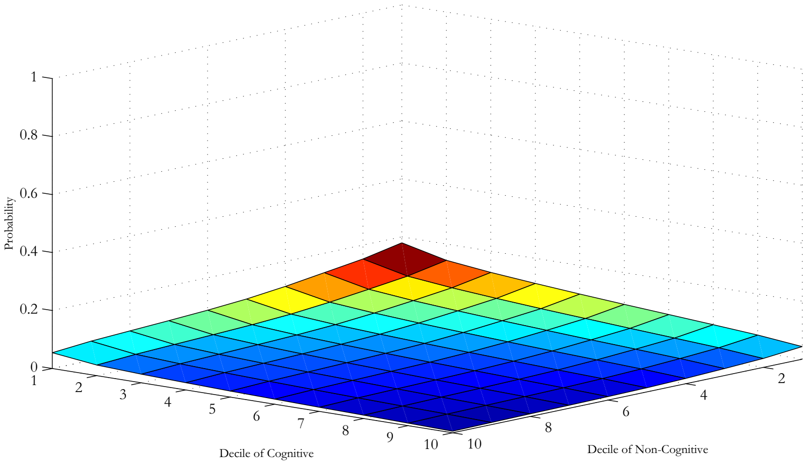
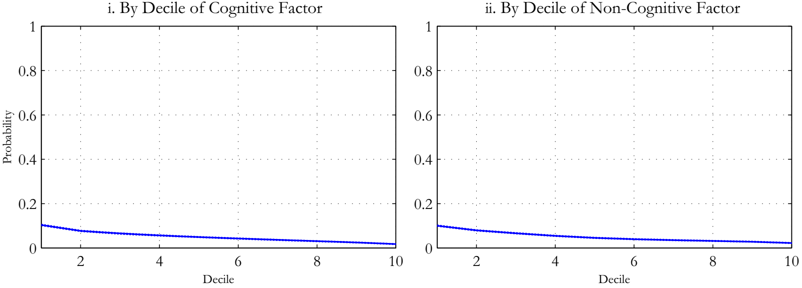


Figure 33 B. Probability Of Being Single With Child - Females



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:

$$f_{t+1}^{C} = \gamma_1^C f_t^C + \gamma_2^C f_t^N + \left(1 - \gamma_1^C - \gamma_2^C\right) f_t^{IC} + \eta_{t+1}^C$$

$$f_{t+1}^{N} = \gamma_1^N f_t^C + \gamma_2^N f_t^N + \left(1 - \gamma_1^N - \gamma_2^N\right) f_t^{IN} + \eta_{t+1}^N$$

$$f_{t+1}^{HOME} = \gamma_1^{IC} f_t^C + \gamma_2^{IC} f_t^N + \gamma^I f_t^{HOME} + \eta_{t+1}^{IC}$$

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

$$f_{t+1}^{C} = 0.516 f_t^C + 0.483 f_t^N + 0.001 f_t^{IC} + \eta_{t+1}^C,\ var\left(\eta_{t+1}^C\right) = 0.036$$

$$f_{t+1}^{N} = 0.98 f_t^N + 0.02 f_t^{IN} + \eta_{t+1}^N,\ var\left(\eta_{t+1}^N\right) = 0.00184$$

$$\begin{aligned} f_{t+1}^{HOME} &= -0.01 f_t^C + 0.036 f_t^N + 0.8074 f_t^{HOME} + \eta_{t+1}^{HOME}, \\ var\left(\eta_{t+1}^{HOME}\right) &= 0.0042 \end{aligned}$$

- 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

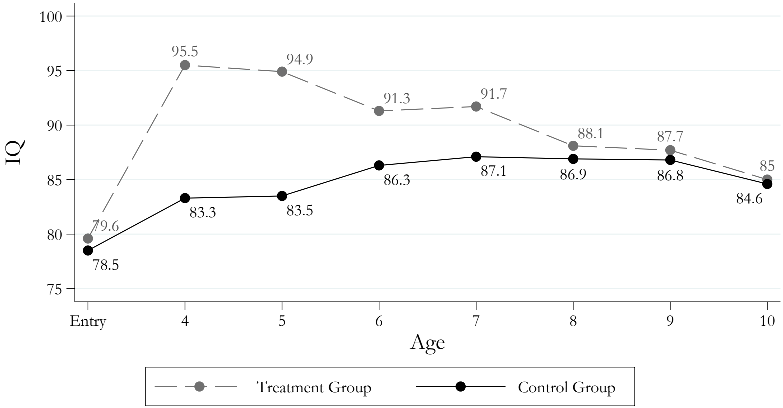
- Low dimensional model for two latent abilities explains a diverse array of behaviors controlling for reverse causality and selection
- We move beyond looking only at effects of cognitive and noncognitive skills on wages.
- For many dimensions of behavior, noncognitive ability is more important than, 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.

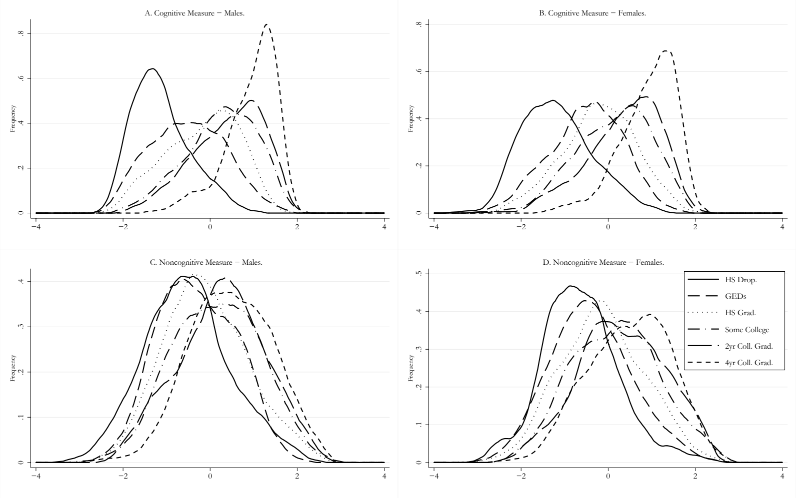
Figure 1

Perry Preschool Program: IQ, by Age and Treatment Group



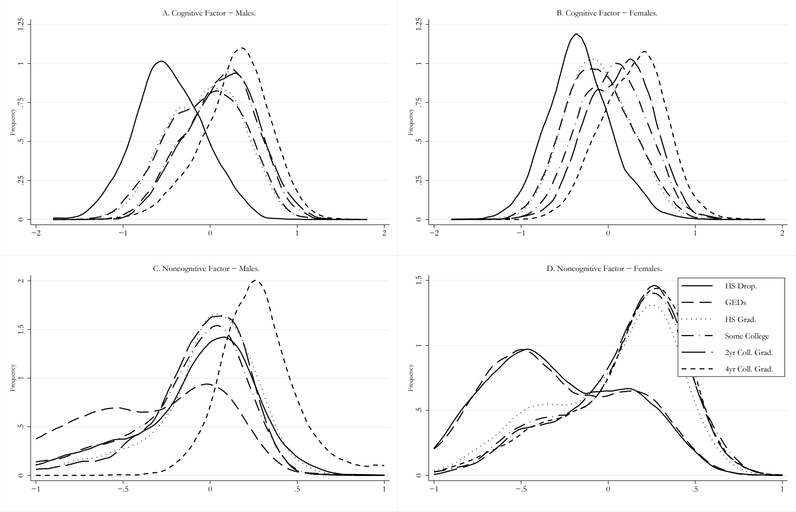
Source: Perry Preschool Program. IQ measured on the Stanford−Binet Intelligence Scale (Terman & Merrill, 1960). Test was administered at program entry and each of the ages indicated.

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).

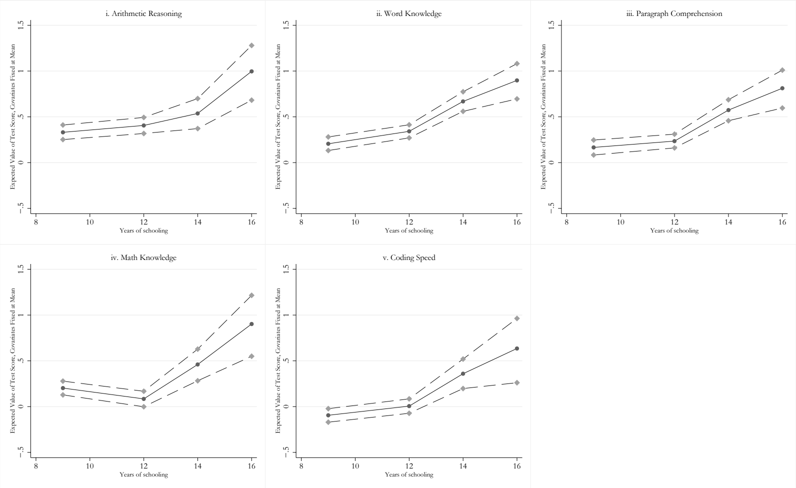
Figure 3. Distribution of Factors by Gender and Schooling Level



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

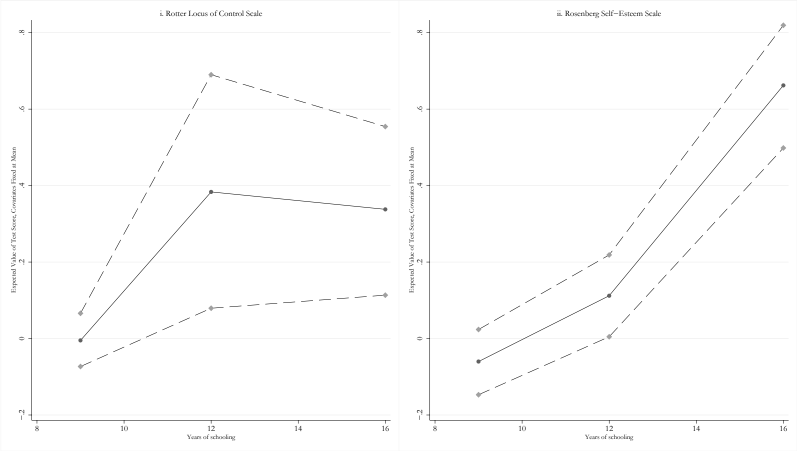
with 95% confidence bands−−Males



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

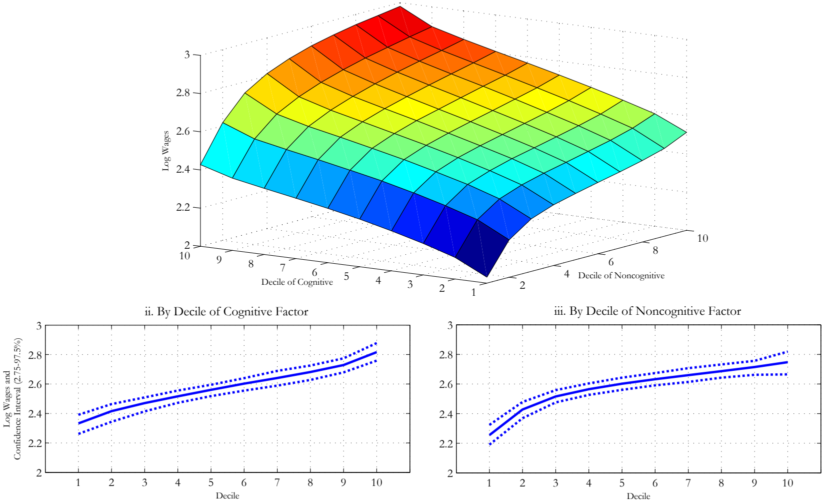
with 95% confidence bands−−Males



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

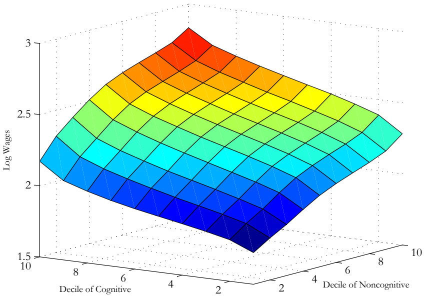
i. By Decile of Cognitive and Noncognitive Factors



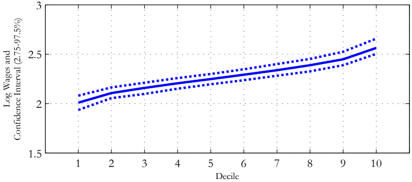
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

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

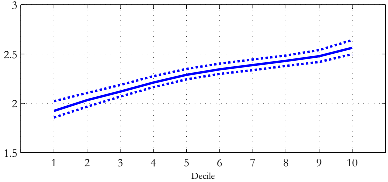
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor

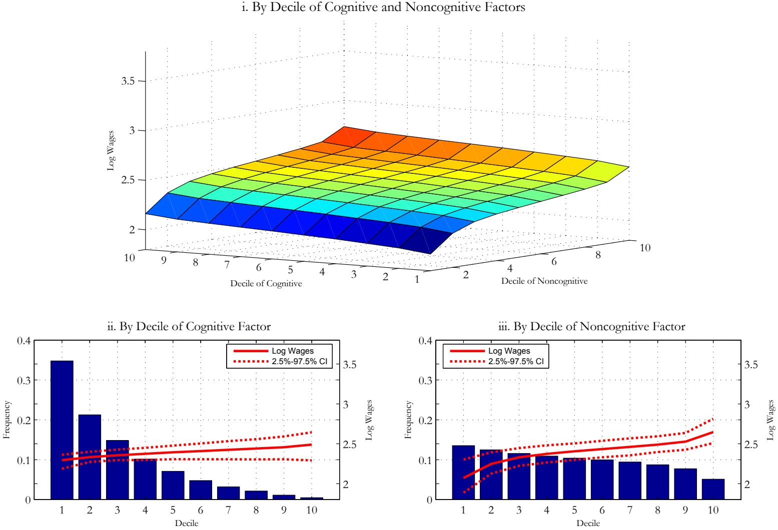


iii. By Decile of Noncognitive Factor



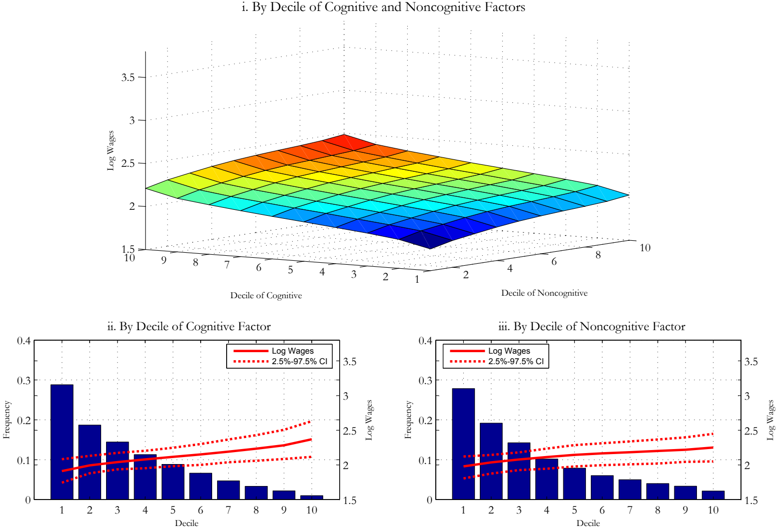
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

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



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.

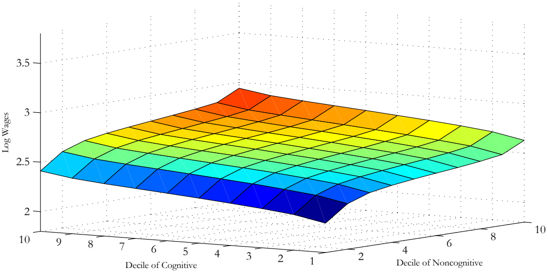
Figure 6B. Mean Log Wages of High School Dropouts by Age 30 - Females



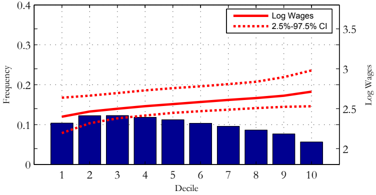
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals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

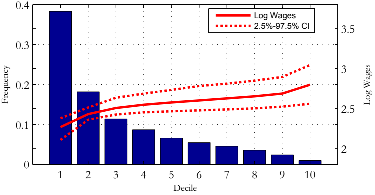
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor



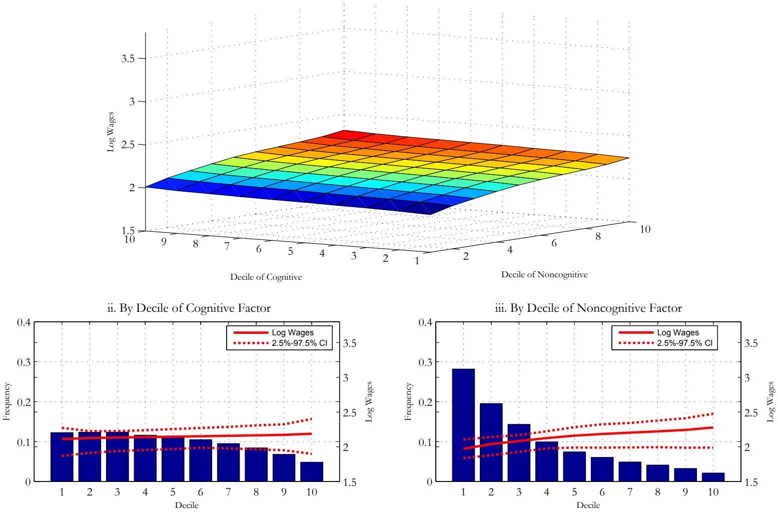
iii. By Decile of Noncognitive Factor



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.

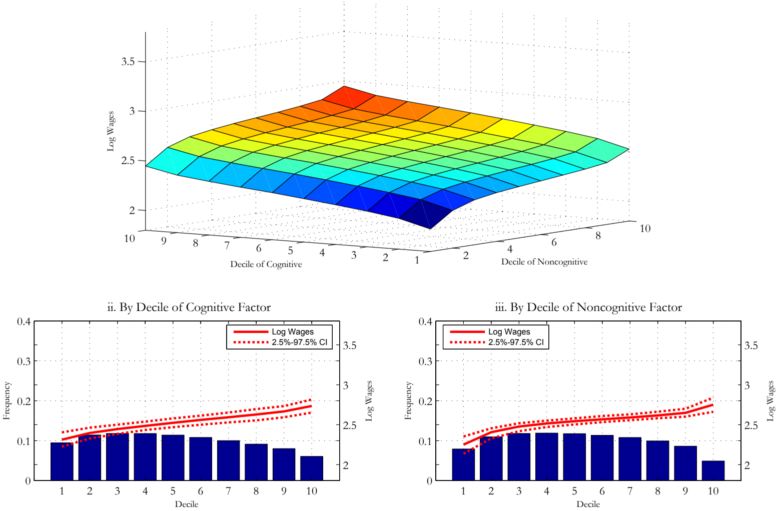
Figure 7B. Mean Log Wages of GEDs by Age 30 - Females

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.

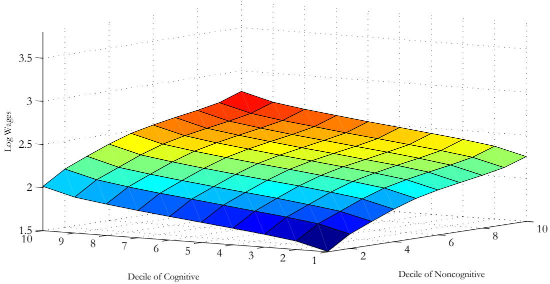
Figure 8A. Mean Log Wages of High School Graduates by Age 30 - Males
i. By Decile of Cognitive and Noncognitive Factors



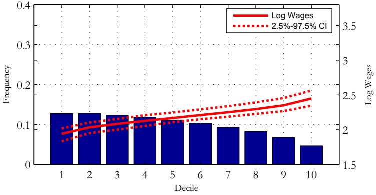
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals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

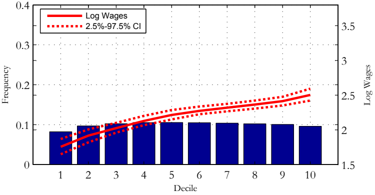
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor



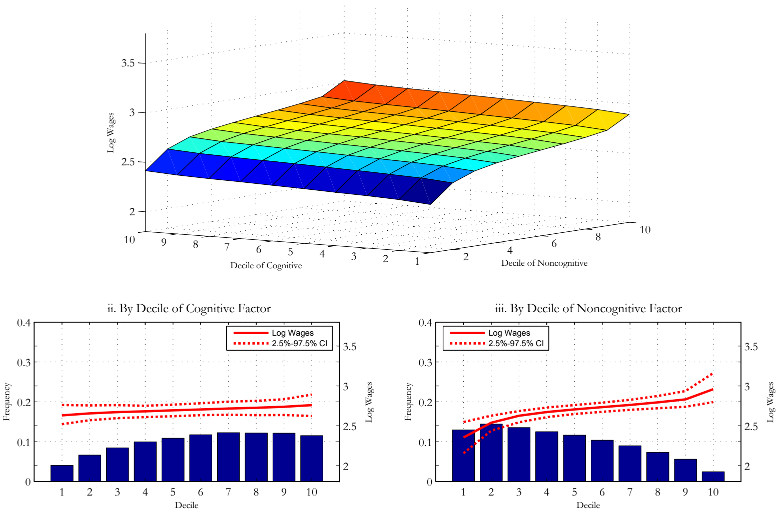
iii. By Decile of Noncognitive Factor



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

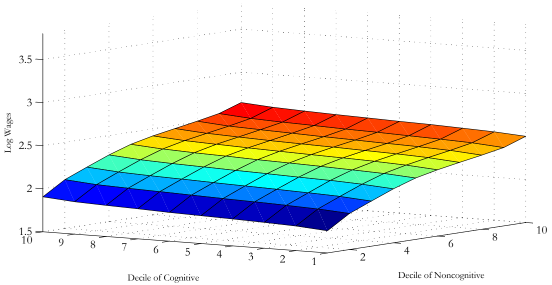
i. By Decile of Cognitive and Noncognitive Factors

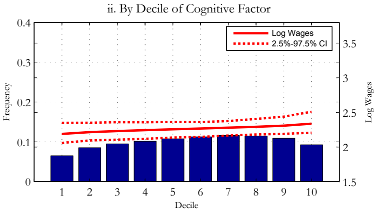


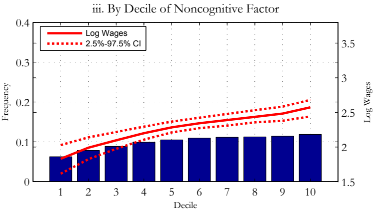
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals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

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

i. By Decile of Cognitive and Noncognitive Factors



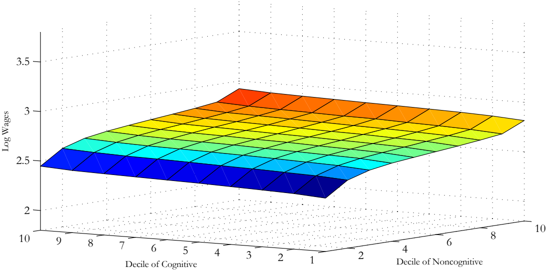




Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals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

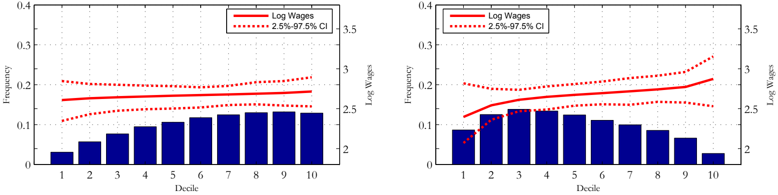
Figure 10A. Mean Log Wages of 2-yr College Graduates by Age 30 - Males

i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor

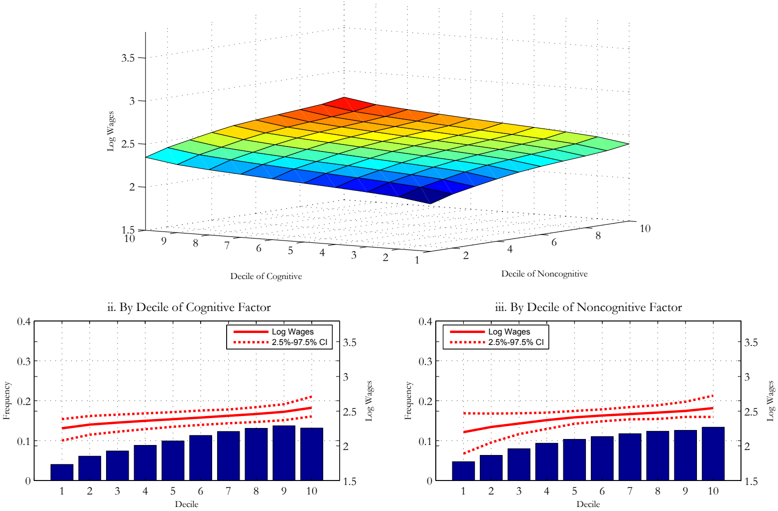
iii. By Decile of Noncognitive Factor



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.

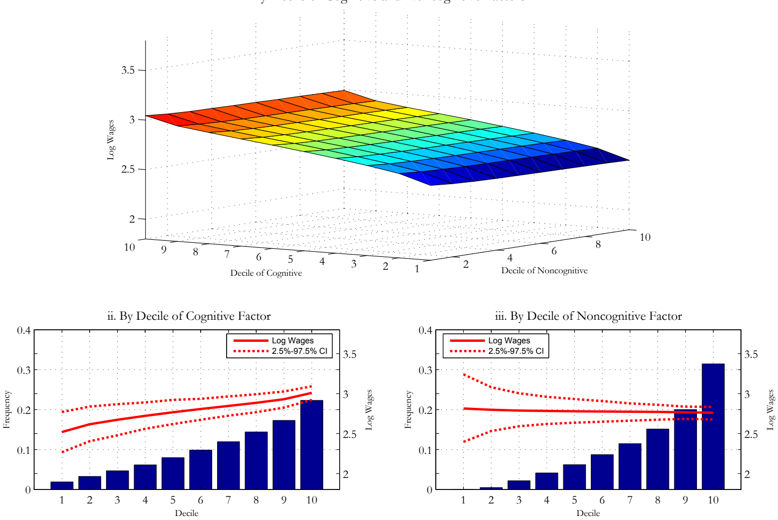
Figure 10B. Mean Log Wages of 2-yr College Graduates by Age 30 - Females

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.

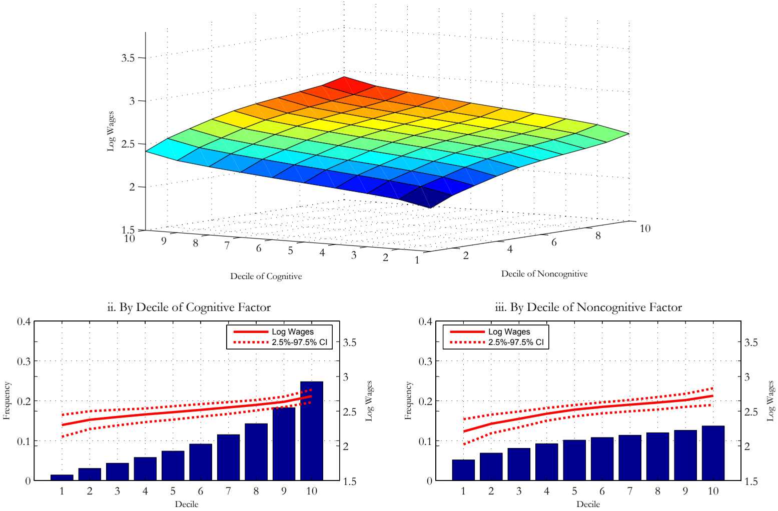
Figure 11A. Mean Log Wages of 4-yr College Graduates by Age 30 - Males



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.

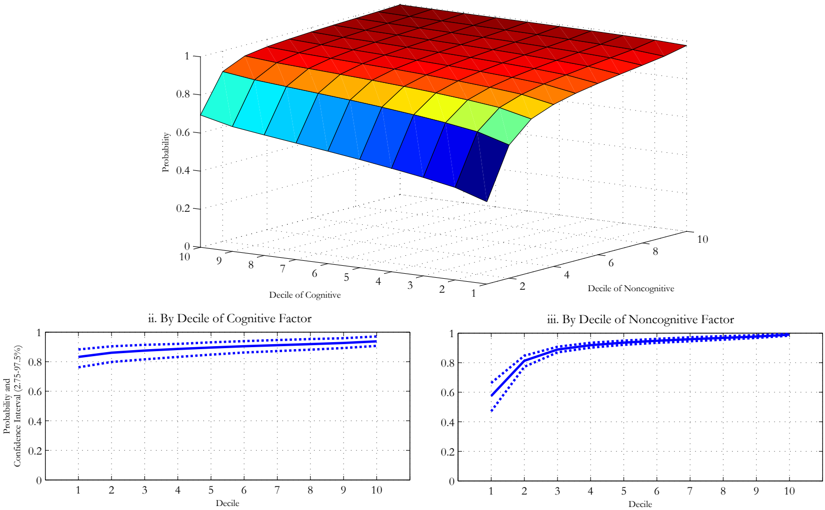
Figure 11B. Mean Log Wages of 4-yr College Graduates by Age 30 - Females

i. By Decile of Cognitive and Noncognitive Factors



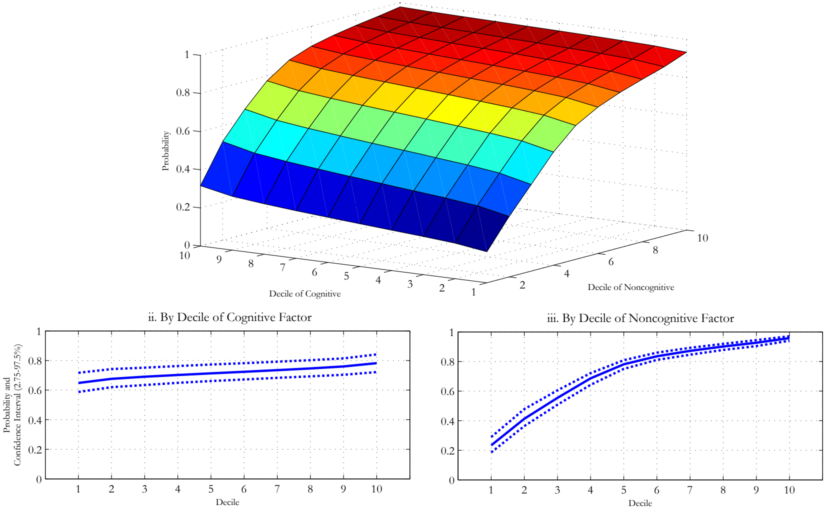
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals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

Figure 12A. Probability of Employment at 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).

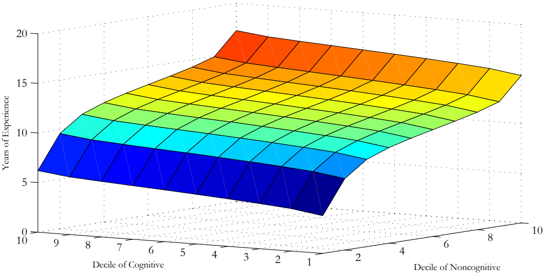
Figure 12B. Probability of Employment at Age 30 - Females
i. By Decile of Cognitive and Noncognitive Factor



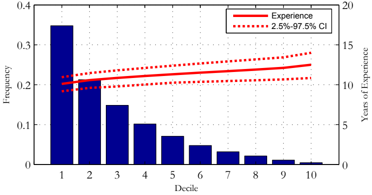
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

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

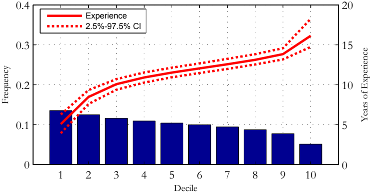
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor



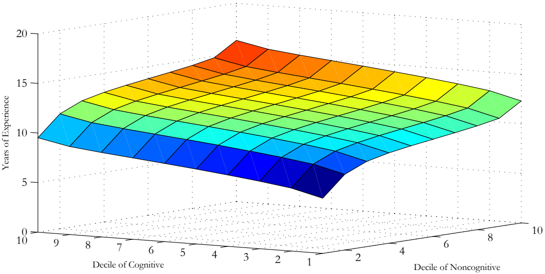
iii. By Decile of Noncognitive Factor



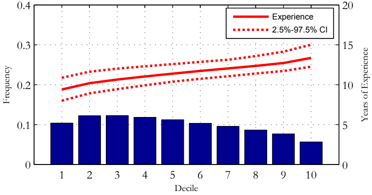
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals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

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

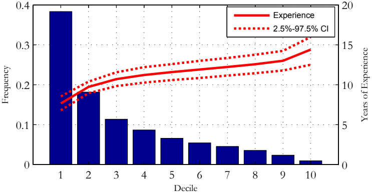
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor



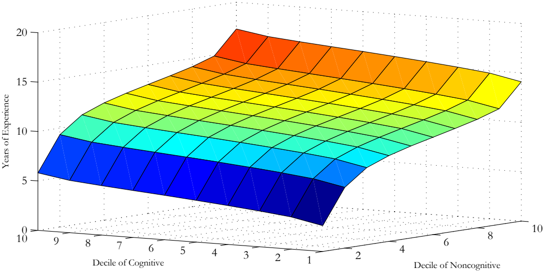
iii. By Decile of Noncognitive Factor



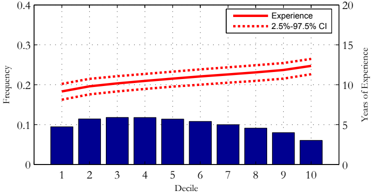
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals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

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

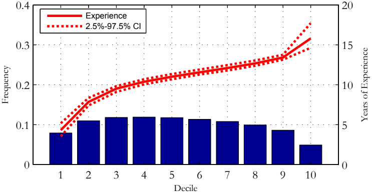
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor

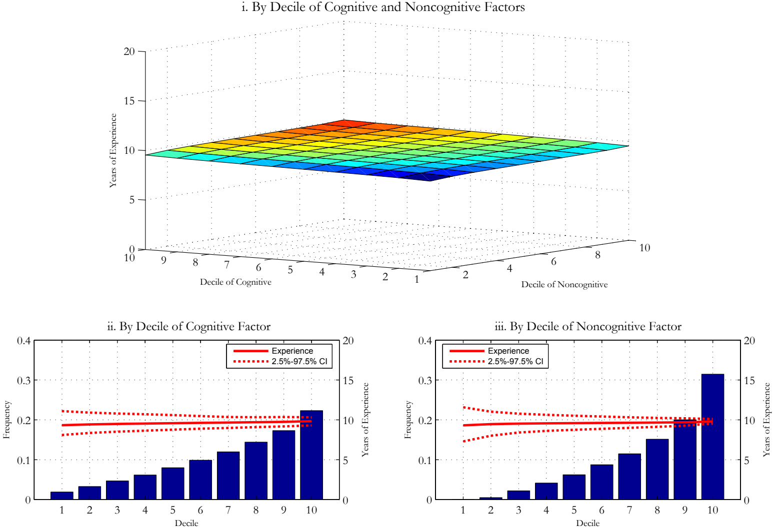


iii. By Decile of Noncognitive Factor



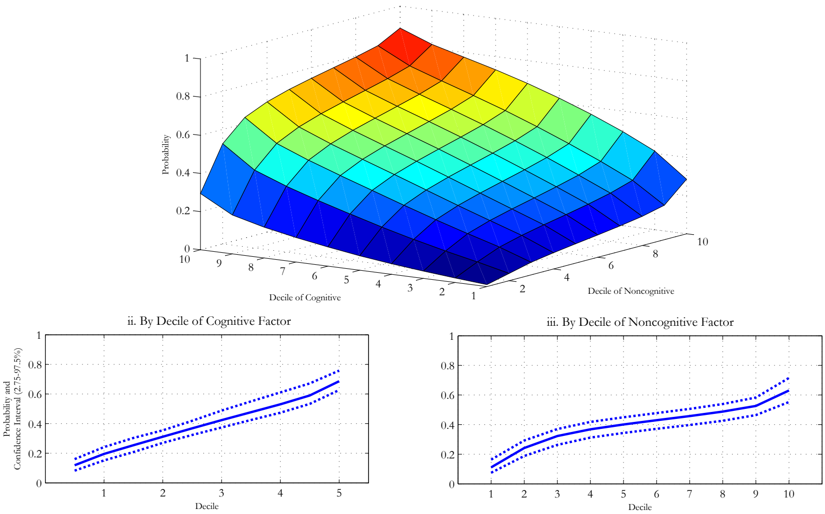
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals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

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



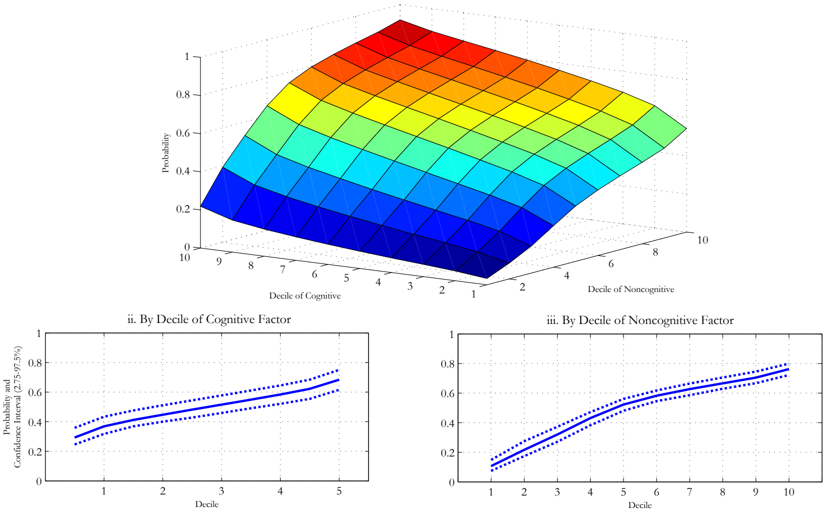
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Frequency indicates proportion of individuals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



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).

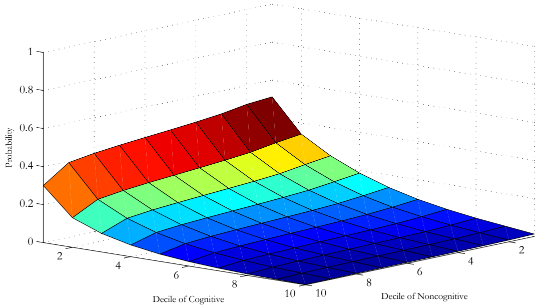
Figure 14B. Probability Of Being a White Collar Worker by Age 30 - Females
i. By Decile of Cognitive and Noncognitive Factor



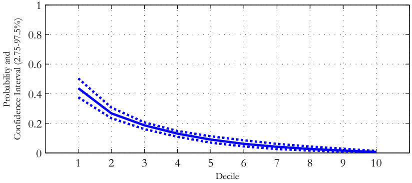
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

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

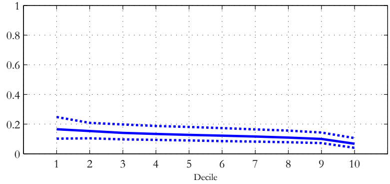
i. By Decile of Cognitive and Noncognitive Factors



ii. By Decile of Cognitive Factor

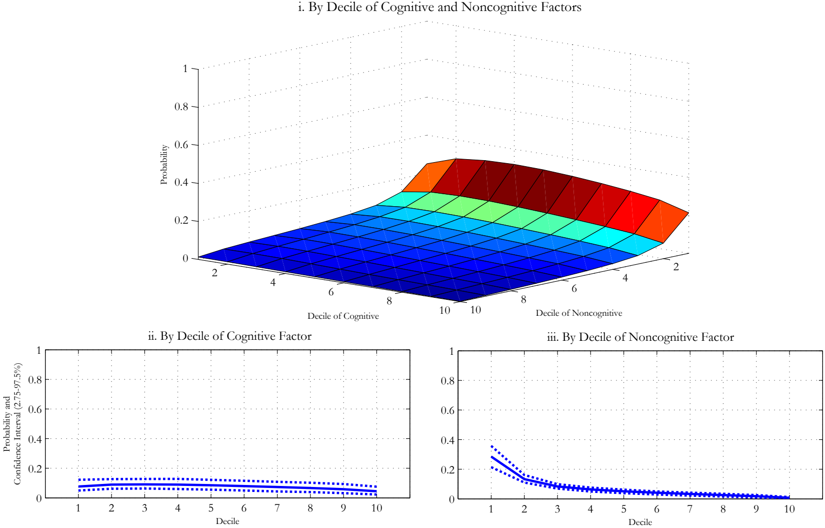


iii. By Decile of Noncognitive Factor



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).

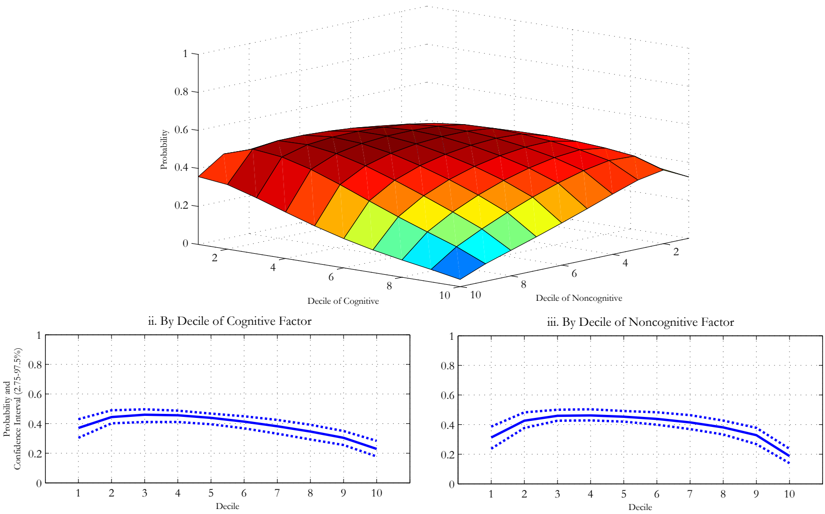
Figure 16. Probability of Being a GED by Age 30 - Males



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).

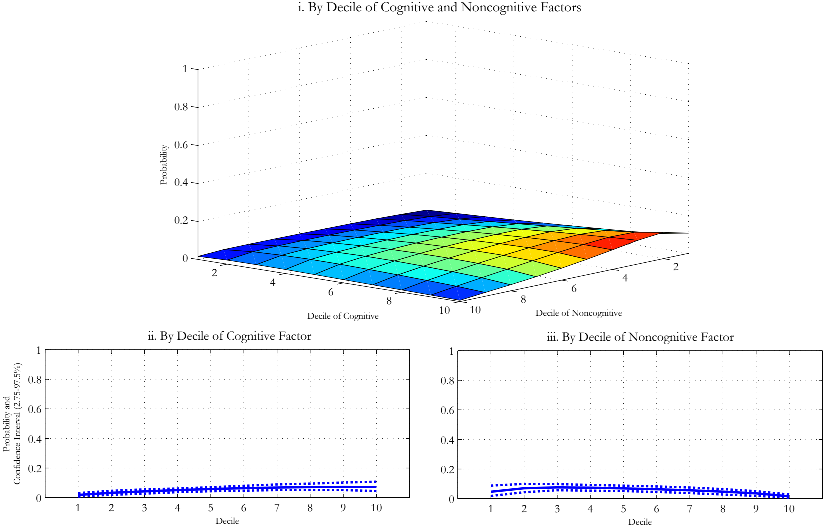
Figure 17. Probability of Being a High School Graduate by Age 30 - Males

i. By Decile of Cognitive and Noncognitive Factors



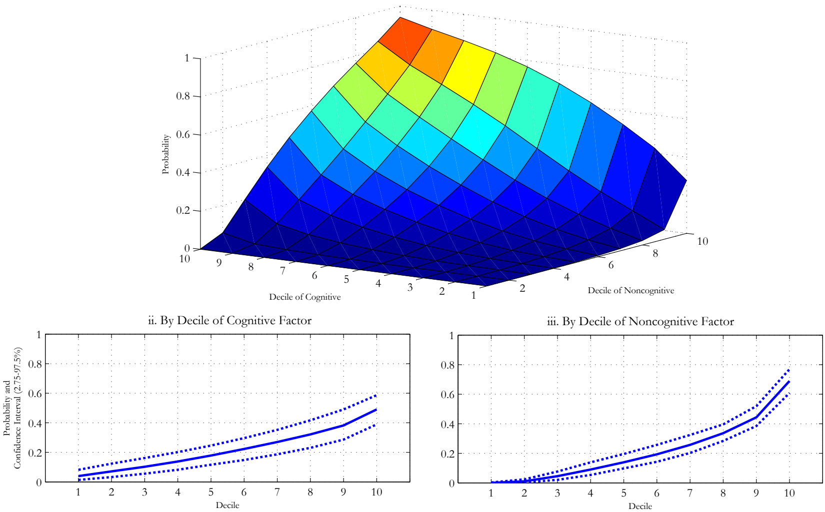
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

Figure 18. Probability of Being a 2-yr College Graduate by Age 30 - Males



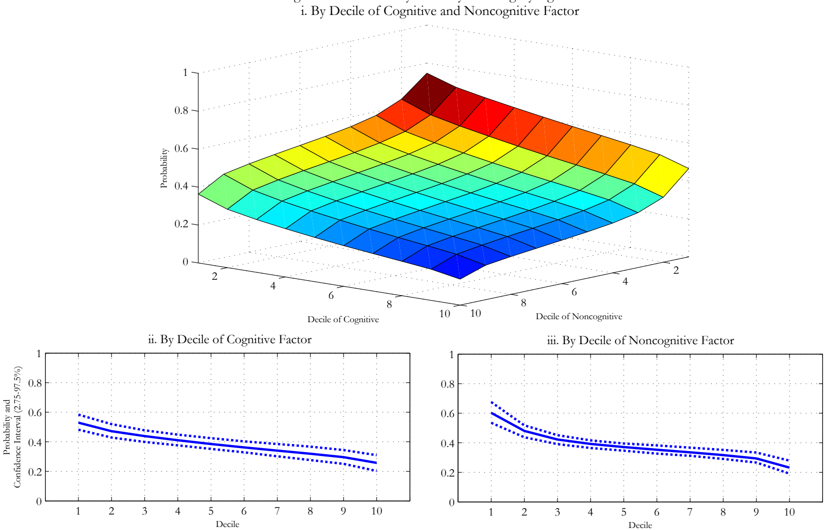
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

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



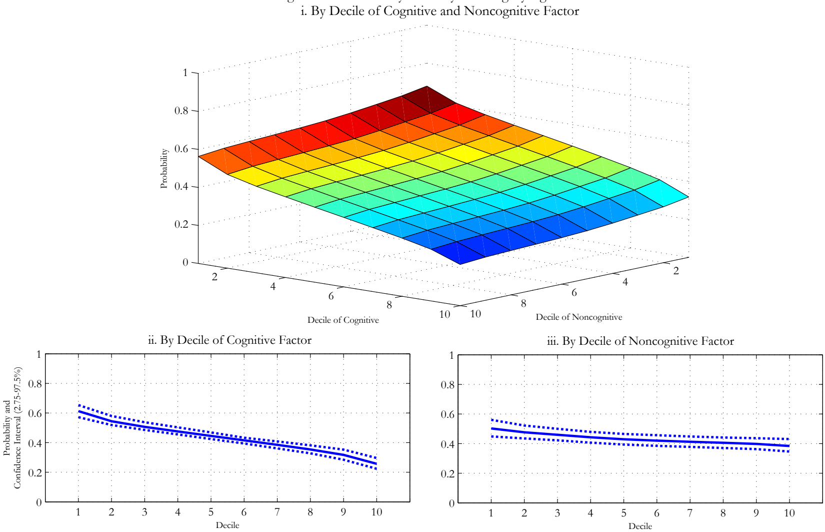
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable.
The confidence intervals are computed using bootstrapping (200 draws).

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).

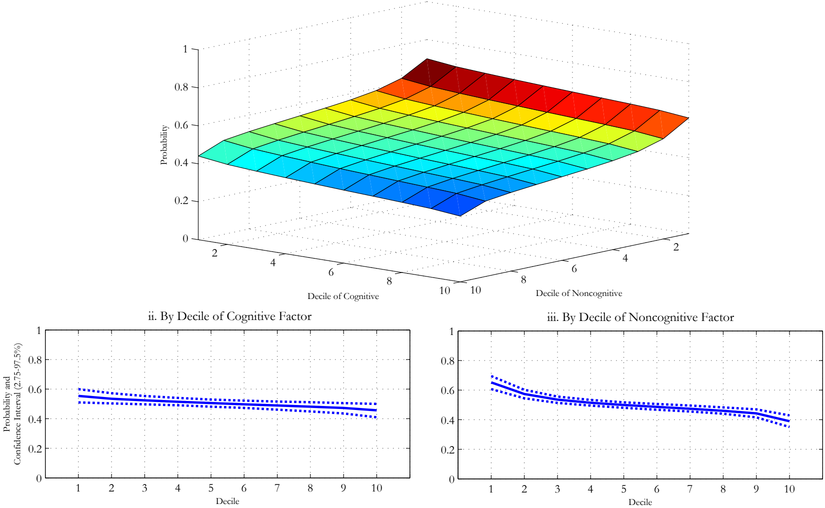
Figure 20B. Probability Of Daily Smoking By Age 18 - Females



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).

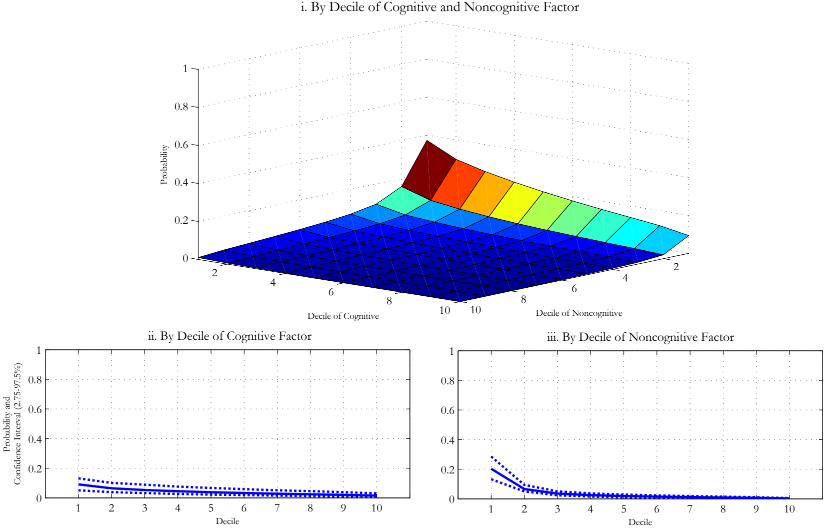
Figure 21. Probability of Smoking Marijuana during the Year 1979 - Males

i. By Decile of Cognitive and Noncognitive Factor



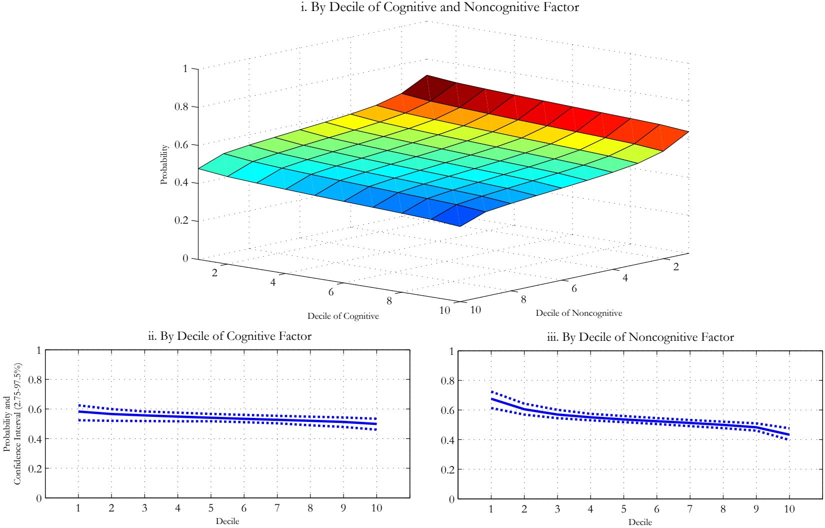
Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws).

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).

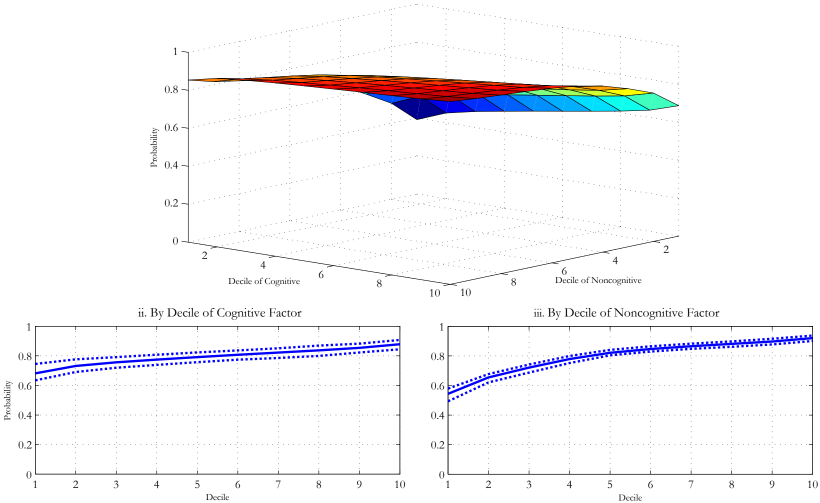
Figure 23. Probability of Participating in Illegal Activities during the Year 1979- Males



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

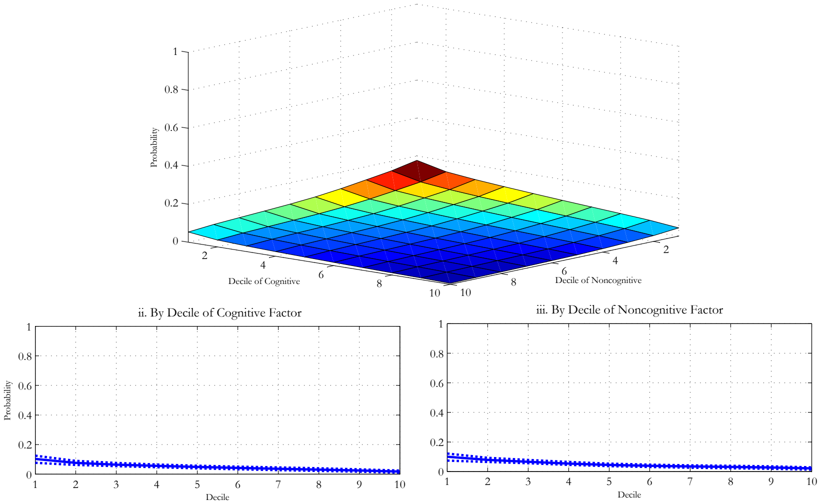
i. By Decile of Cognitive and Noncognitive Factors



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

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

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

$\downarrow$

Simulate synthetic dataset using estimates.

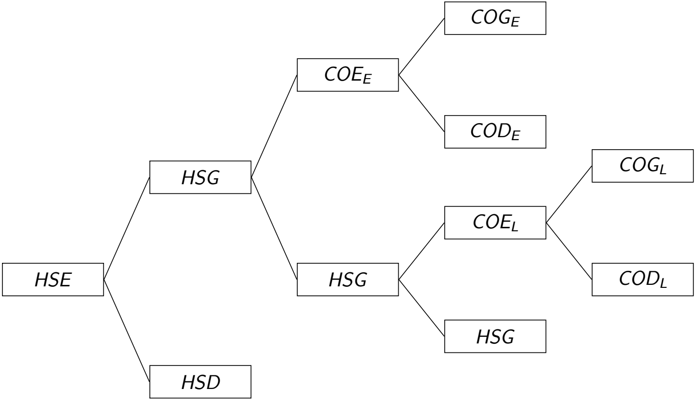
$\downarrow$

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

Compare

Model

Figure: Decision Tree



Agent Behavior

- Objective
- Constraints
  - Institution
  - Information

Value Function

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

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

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

$$
M(j) = X(j)'\kappa_j + \theta'\gamma_j + \nu(j) \qquad \forall \quad j \in M
$$

Distributions of Unobservables

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

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

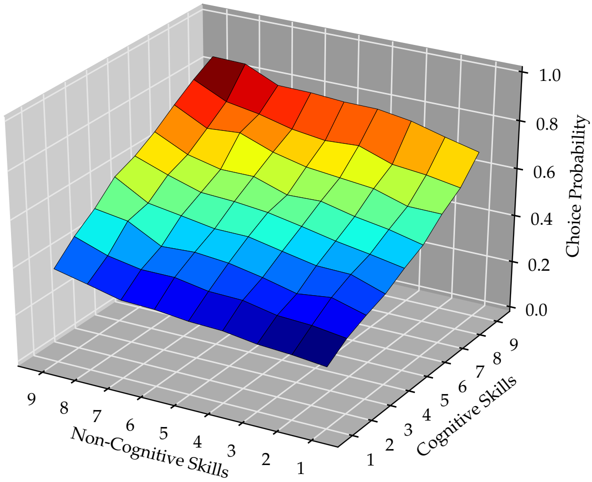


Figure: Gross Return, Early College Enrollment

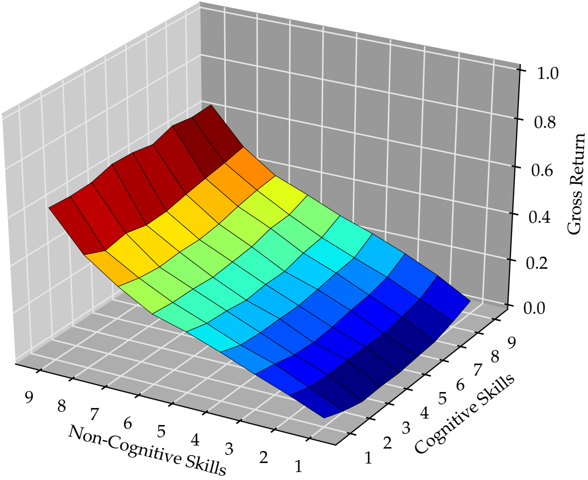


Figure: Net Return, Early College Enrollment

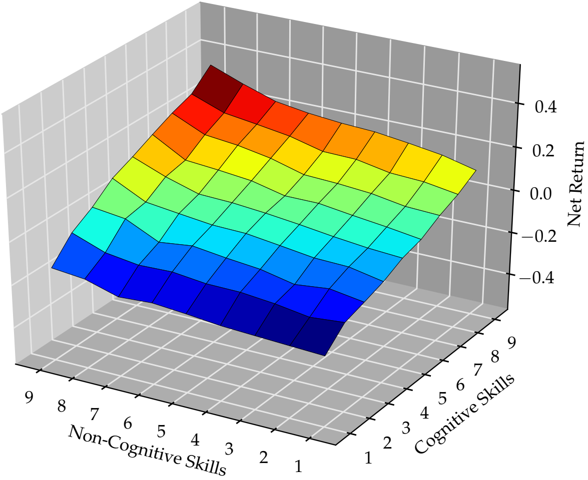


Figure: Schooling Attainment by Cognitive Skills

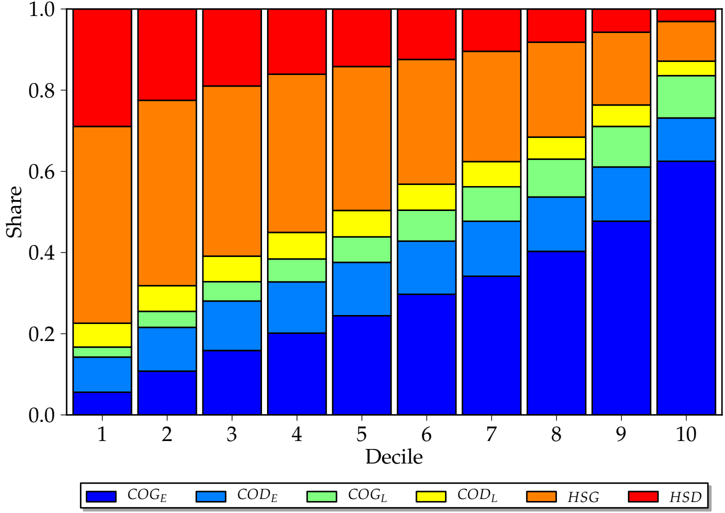
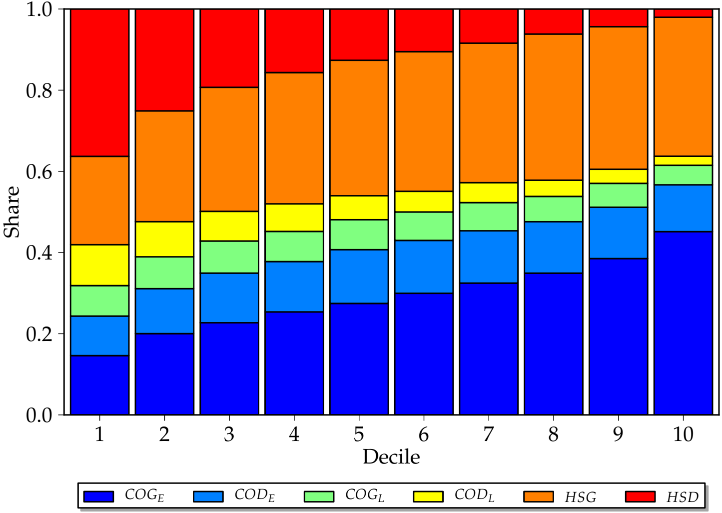


Figure: Schooling Attainment by Non-Cognitive Skills



Monte Carlo Study

Synthetic Agents	5,000
Structural Parameters	192
Free Parameters	138

Simulated Method of Moments

Criterion Function

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

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

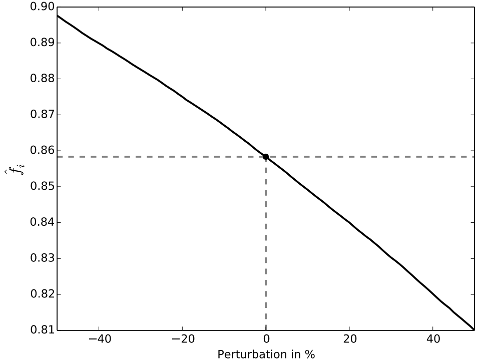
Number of Moments	250
Number of Replications	50
Weighting Matrix	Diagonal Matrix with Variances
Optimization Algorithm	TAO POUNDerS

Moment Conditions

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

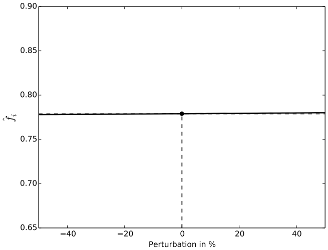
Parameter Perturbations

Figure: Parameter Perturbation, Cost

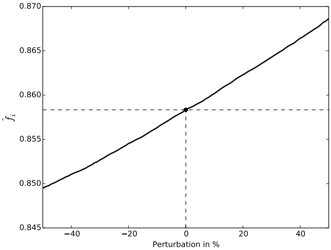


High School Graduation, State Frequency

Figure: Parameter Perturbation, Benefit



HS Graduates, Average Wages



HS Graduation, State Frequency

Model Fit

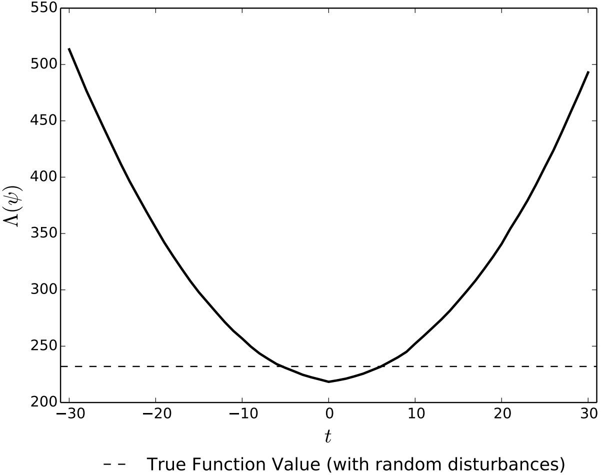
Table: Cross Section Model Fit

	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	True	ML	SMM
	State Frequencies		
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

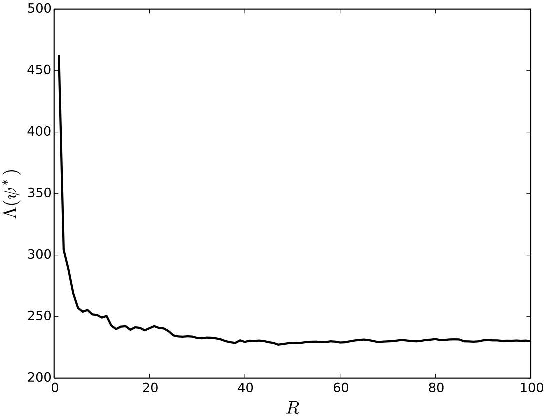
		Dynamic (Panel) Moments		
State	True	Base	Alt. A	Alt. B
		Gross Return		
High School Graduation	30%	42%	37%	37%
Early College Graduation	88%	96%	72%	73%
Late College Graduation	29%	36%	18%	18%
		Net Return		
High School Graduation	63%	212%	203%	194%
Early College Graduation	57%	125%	107%	112%
Late College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

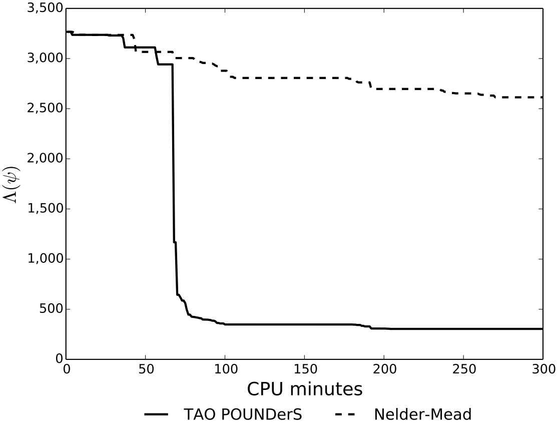
Nonlinear Least-Squares

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

Derivative-Free Trust-Region Algorithm

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

Figure: Optimization Algorithms



Appendix

References

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.

Heckman, J. J., & Vytlacil, E. J. (2007a). Econometric evaluation of social programs, part I: Causal effects, structural models and econometric policy evaluation. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (Vol. 6B, pp. 4779–4874). Amsterdam, Netherlands: Elsevier Science.

Heckman, J. J., & Vytlacil, E. J. (2007b). Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative economic estimators to evaluate social programs and to forecast their effects in new environments. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (Vol. 6B, pp. 4875–5144). Amsterdam, Netherlands: Elsevier Science.