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

James J. Heckman, *University of Chicago, University College Dublin, and American Bar Foundation*

Jora Stixrud, *University of Chicago*

Sergio Urzua, *University of Chicago*

This article establishes that a low-dimensional vector of cognitive and noncognitive skills explains a variety of labor market and behavioral outcomes. Our analysis addresses the problems of measurement error, imperfect proxies, and reverse causality that plague conventional studies. Noncognitive skills strongly influence schooling decisions and also affect wages, given schooling decisions. Schooling, employment, work experience, and choice of occupation are affected by latent noncognitive and cognitive skills. We show that the same low-dimensional vector of abilities that explains schooling choices, wages, employment, work experience, and choice of occupation explains a wide variety of risky behaviors.

This research was supported by NIH grant R01-HD043411 and a Pew Foundation grant to Heckman. This article was presented at the Mark Berger memorial conference, University of Kentucky, October 2004; as one of the Ely lectures, Johns Hopkins University, April 2005; at University College Dublin, April 2005; at the Institute for Research on Poverty Workshop, Madison, WI, June 2005; at the Second IZA Prize Conference, Berlin, Germany, October 2005; and at the Workshop in Public Policy and Economics at the Harris School of the University of Chicago in September 2005. We also presented it at the American Economics Association annual meeting, Boston, January 2006. We thank participants for their comments. We thank the editors William Johnson and James P. Ziliak and two

[*Journal of Labor Economics,* 2006, vol. 24, no. 3]

0734-306X/2006/2403-0003$10.00

I. Introduction

Numerous studies establish that measured cognitive ability is a strong predictor of schooling attainment and wages.[1] It also predicts a range of social behaviors (see Herrnstein and Murray 1994). Less well investigated is the role of personal preference and personality traits on economic and social behavior.

Common sense suggests that personality traits, persistence, motivation, and charm matter for success in life. Marxist economists (Bowles and Gintis 1976; Edwards 1976) have produced a large body of evidence that employers in low-skill labor markets value docility, dependability, and persistence more than cognitive ability or independent thought (see the survey by Bowles, Gintis, and Osborne [2001]). Sociologists have written extensively about the role of noncognitive skills in predicting occupational attainment and wages (see the essay by Peter Mueser in Jencks [1979]), and several studies in the psychology literature have shown the important role of noncognitive skills on the schooling performance of children and adolescents (Wolfe and Johnson 1995; Duckworth and Seligman 2005).

This article presents an analysis of the effects of both cognitive and noncognitive skills on wages, schooling, work experience, occupational choice, and participation in a range of adolescent risky behaviors. We show that a model with one latent cognitive skill and one latent noncognitive skill explains a large array of diverse behaviors.

Our approach differs from previous methods used to address these issues by accounting for the effects of schooling and family influence on the measurements of the latent skills used in our empirical analysis. We allow the latent skills to determine measured skills and schooling choices, and for schooling to determine measured skills.

We find that both types of latent skills are important in explaining a diverse array of outcomes. The skills are priced differently in different schooling markets. There are important gender differences in the effects of these skills but for most behaviors, both factors play an important role for both men and women.

For a variety of dimensions of behavior and for many labor market outcomes, a change in noncognitive skills from the lowest to the highest level has an effect on behavior comparable to or greater than a corresponding change in cognitive skills. This evidence contradicts the *g* theory of human behavior espoused by Herrnstein and Murray (1994), Jensen

anonymous referees for helpful comments. We also thank Jeff Grogger, Bruce Meyer, and Derek Neal for very helpful comments that led to revisions and clarifications. Supplementary materials are on our Web site at http://jenni.uchicago.edu/noncog/web_supplement.pdf. We thank Federico Temerlin and Tae Ho Whang for very competent research assistance. Contact the corresponding author, James J. Heckman, at jjh@uchicago.edu.

[1] See, e.g., the evidence summarized in Cawley, Heckman, and Vytlacil (2001).

(1998), and others, that focuses on the primacy of cognitive skills in explaining socioeconomic outcomes.

Our evidence has important implications for the literature on labor market signaling as developed by Arrow (1973) and Spence (1973). That literature is based on the notion that schooling only conveys information about a student's cognitive ability and that smarter persons find it less costly to complete schooling. Our findings show that schooling signals multiple abilities. This observation fundamentally alters the predictions of signaling theory (see Araujo, Gottlieb, and Moreira 2004).

Our approach recognizes that test scores measuring both cognitive and noncognitive abilities may be fallible. It also recognizes that a person's schooling and family background at the time tests are taken affect test scores. Observed ability-wage and ability-schooling relationships may be consequences of schooling causing measured ability, rather than the other way around. Building on the analysis of Hansen, Heckman, and Mullen (2004), we correct measured test scores for these problems.

Our analysis supports the commonsense view that noncognitive skills matter. As conjectured by Marxist economists (Bowles and Gintis 1976), we find that schooling determines the measures of noncognitive skills that we study. We find that latent noncognitive skills raise wages through their direct effects on productivity, as well as through their indirect effects on schooling and work experience. Our evidence is consistent with an emerging body of literature that finds that "psychic costs" (which may be determined by noncognitive traits) explain why many adolescents who would appear to financially benefit from schooling do not pursue it (Carneiro, Hansen, and Heckman 2003; Carneiro and Heckman 2003; Cunha, Heckman, and Navarro 2005; Heckman, Lochner, and Todd 2006).

Our evidence bolsters and interprets the findings of Heckman and Rubinstein (2001), who show that General Educational Development (GED) recipients (high school dropouts who exam certify as high school equivalents) have the same achievement test scores as high school graduates who do not go on to college yet earn, on average, the wages of dropouts. The poor market performance of GED recipients is due to their low levels of noncognitive skills, which are lower than those of high school dropouts who do not get the GED. Both cognitive and noncognitive skills are valued in the market. The GED surplus of cognitive skills is not outweighed by the GED deficit in noncognitive skills.

Carneiro and Heckman (2003), Heckman and Masterov (2004), and Cunha, Heckman, Lochner, and Masterov (2006), and the numerous papers they cite, establish that parents play an important role in producing both the cognitive and noncognitive skills of their children. More able and engaged parents have greater success in producing both types of skills. Because cognitive and noncognitive abilities are shaped early in the life cycle, differences in these abilities are persistent, and both are crucial to

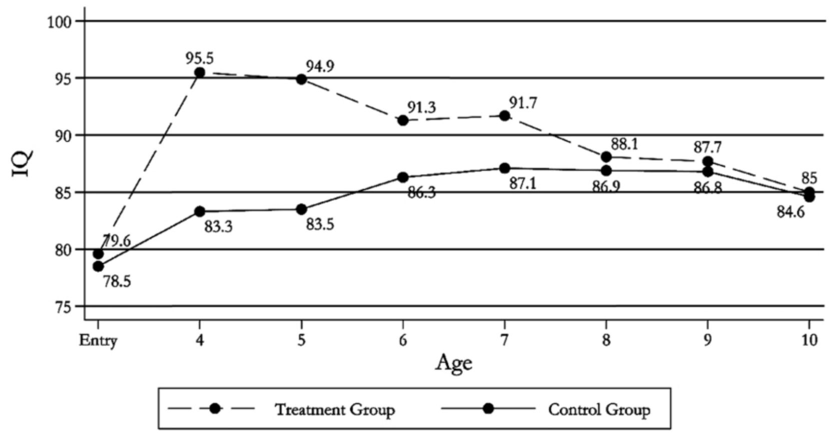


Fig. 1.—Perry Preschool Program: IQ by age and treatment group. Source: Authors' calculations using information from the Perry Preschool Program. IQ is measured on the Stanford-Binet Intelligence Scale (Terman and Merrill 1960). The test was administered at program entry and each of the ages indicated.

social and economic success; gaps among income and racial groups begin early and persist.

Evidence from early interventions also motivates our work. Early interventions, such as enriched child-care centers coupled with home visitations, have been successful in alleviating some of the initial disadvantages of children born into adverse family environments. The success of these interventions is not attributable to IQ improvements of children, but rather to their success in boosting noncognitive skills (Heckman 2005).

As an example, the Perry Preschool Program intervened early in the life cycle of disadvantaged children.[2] Children were randomly assigned to treatment and control groups, and both were followed to age 40. The program did not boost IQ (see fig. 1) but raised achievement test scores, schooling, and social skills. For example, 66% of the individuals in the treatment group graduated from high school by age 18 versus only 45% of the control group, 49% of the individuals in the treatment group performed at or above the lowest 10th percentile in the California Achievement Test (age 14) versus 15% of the control group, and individuals in the treatment groups were significantly less likely to get involved in illegal activities before age 40.[3] This evidence is consistent with the interpretation that noncognitive traits matter for successful social per-

[2] The program was implemented between 1962 and 1967 in Ypsilanti, Michigan.

[3] See figs. S1A and S1B in our Web appendix for more evidence on the Perry Program.

formance and that noncognitive traits were boosted by the program, but not cognitive traits, at least as measured by IQ.

Our analysis explains the phenomenon of correlated risky behaviors using the same low-dimensional model of latent skills that explains wages, employment, and schooling attainment. Biglan (2004) documents that risky behaviors such as antisocial behavior (aggressiveness, violence, and criminality), cigarette smoking, alcohol use, and the like, are pursued by the same cluster of adolescents. We find that latent cognitive and noncognitive skills explain all of these behaviors and the observed clustering pattern.

The plan of this article is as follows. Section II introduces the data used in our analysis and presents empirical analyses using conventional methods. We reproduce key findings reported in the previous literature. We then discuss interpretive problems that plague the conventional approach. Section III presents a model of schooling, employment, occupational choice, work experience, and wages generated by latent skills as well as observables. Section IV extends the model to account for correlated risky behaviors. Section V shows how our econometric model can be interpreted as an approximation to a life-cycle model. Section VI discusses how we empirically implement our model. Section VII presents our evidence. Section VIII relates our analysis to previous work in the literature. Section IX concludes.

II. Some Evidence Using Conventional Approaches

We use data from the National Longitudinal Survey of Youth, 1979 (NLSY79). The NLSY79 data are standard and widely used. It is the data source for the *g* analysis of Herrnstein and Murray (1994). It contains panel data on wages, schooling, and employment for a cohort of young persons, age 14–22 at their first interview in 1979. This cohort has been followed ever since. Important for our purposes, the NLSY79 contains information on cognitive test scores as well as noncognitive measures. Web appendix A (http://jenni.uchicago.edu/noncog/web_supplement.pdf) describes the sampling frame of the data in detail.

Our analysis of test scores uses five measures of cognitive skills (arithmetic reasoning, word knowledge, paragraph comprehension, mathematical knowledge, and coding speed) derived from the Armed Services Vocational Aptitude Battery (ASVAB), which was administered to the sample participants in 1980. A composite score derived from these sections of the test battery can be used to construct an approximate Armed Forces Qualifications Test (AFQT) score for each individual. The AFQT is a general measure of trainability and a primary criterion of eligibility for service in the armed forces. It has been used extensively as a measure of cognitive skills in the literature (see, e.g., Heckman 1995; Neal and Johnson 1996; Cameron and Heckman 1998; Ellwood and Kane 2000; Cameron and Heckman 2001; Osborne-Groves 2006). The noncognitive measures we use in this article are the Rotter Locus of Control Scale (Rotter 1966), which was administered in 1979, and the Rosenberg Self-Esteem Scale (Rosenberg 1965), which was administered in 1980.

The Rotter scale measures the degree of control individuals feel they possess over their life and has been used in previous studies analyzing the role of noncognitive skills on labor outcomes (Osborne-Groves 2006). The Rosenberg scale measures perceptions of self-worth. These tests are discussed in detail in Web appendix A.

This section of the article presents a standard least-squares analysis of the effects of cognitive and noncognitive skills on wages. We obtain the same qualitative results that have been reported by previous analysts (see, e.g., Jencks 1979; Bowles et al. 2001; Osborne-Groves 2006). We use the standardized average of an individual's five ASVAB components as a measure of cognitive skills and the standardized average of the person's scores on the Rotter and Rosenberg scales as a measure of noncognitive skills. Figure 2 presents the distributions of the cognitive and noncognitive measures by gender and final schooling level. The distributions of both measures of skill are ordered by schooling level, with college graduates having the most favorable distribution of skills and high school dropouts the worst.

Conditioning on schooling, both cognitive and noncognitive tests predict wages (see table 1, cols. 1 and 3). However, schooling is a choice variable, and any convincing analysis must account for the endogeneity of schooling. Deleting schooling from the wage equation (see table 1, cols. 2 and 4) produces larger estimated effects of both abilities on wages. Removing the conditioning on schooling solves the problem of endogeneity of schooling in wage equations and produces an estimate of the net effect of the abilities on wages (their direct effects plus their effects through schooling).

Not controlling for schooling, the cognitive ability measure explains 9.0% of the variance of log wages. For men, the noncognitive measure explains only 0.9% of the variance. For women, the corresponding figures are 12.4% and 0.4%. We will show that even though cognitive ability explains a larger share of wage variance than noncognitive ability, both are important in the sense that moving persons from the top to the bottom of the ability distribution has similar effects for both types of abilities.

This evidence suggests that both noncognitive and cognitive abilities significantly affect wages, as an entire literature has found (see Jencks 1979). However, this evidence is not without its problems. First, we note that there is an important distinction between intelligence tests (i.e., IQ tests) and achievement tests. Although IQ is fairly well set by age 8, achievement tests have been demonstrated to be quite malleable. Neal and

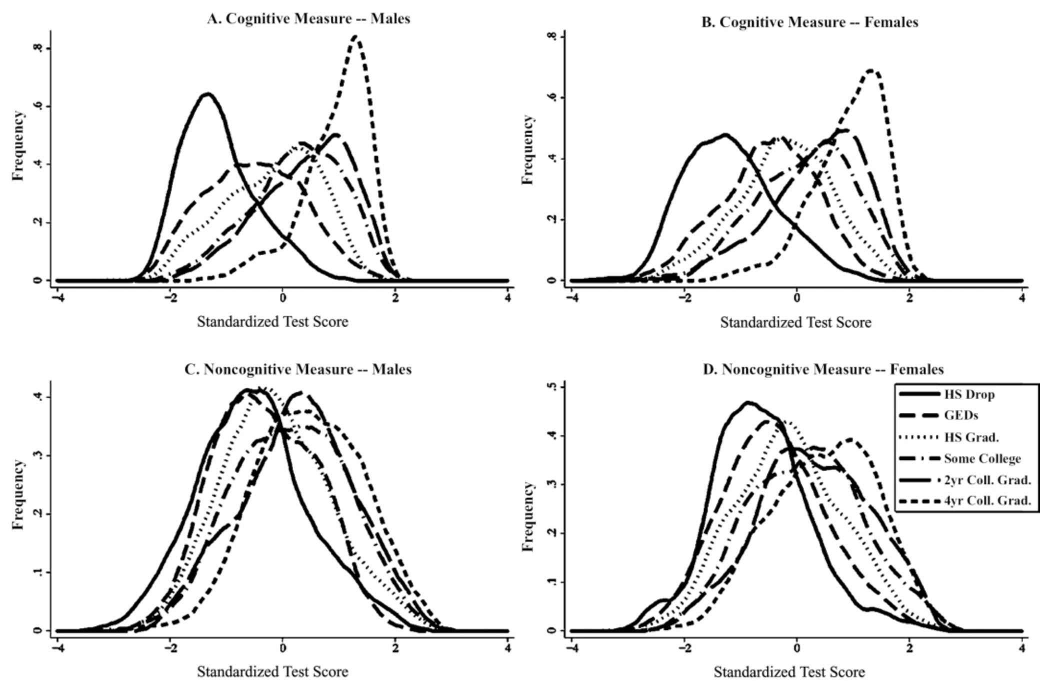


FIG. 2.—Distribution of test scores by gender and schooling level. The cognitive measure represents the standardized average over the ASVAB scores (arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations, and coding speed). The noncognitive measure is computed as a (standardized) 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 app. A for details).

Table 1
Estimated Coefficients from Log Hourly Wage Regressions, NLSY79—Males and Females at Age 30

	Males		Females*	
Variable	(1)	(2)	(3)	(4)
GED	.017		−.002	
	(.048)		(.056)	
High school graduate	.087		.059	
	(.035)		(.044)	
Some college	.146		.117	
	(.044)		(.052)	
2-year-college graduate	.215		.233	
	(.058)		(.058)	
4-year-college graduate	.292		.354	
	(.046)		(.054)	
Cognitive measure†	.121	.190	.169	.251
	(.016)	(.013)	(.017)	(.014)
Noncognitive measure‡	.042	.052	.028	.041
	(.011)	(.012)	(.013)	(.013)
Constant	2.558	2.690	2.178	2.288
	(.057)	(.050)	(.063)	(.052)

Note.—Standard errors are in parentheses. We exclude the oversample of blacks, Hispanics, and poor whites, the military sample, and those currently enrolled in college. The model includes a set of cohort dummies, local labor market conditions (unemployment rate), the region of residence, and race. Columns 1 and 3 present the estimates obtained from ordinary least squares (OLS). Columns 2 and 4 present the results from an OLS model in which the schooling dummies are excluded.

* For females we also estimate the equations correcting for selection into the labor force. The results presented in this table are robust to this correction.

† Represents the standardized average over the ASVAB scores (arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge, and coding speed).

‡ Computed as a (standardized) average of the Rosenberg Self-Esteem Scale and Rotter Internal-External Locus of Control Scale.

Johnson (1996) and Hansen et al. (2004) demonstrate that each additional year of schooling increases an individual's measured AFQT score by 2–4 percentage points, on average. This creates a reverse causality problem. The least-squares estimates reported in table 1 cannot distinguish whether higher "ability" (as proxied by our cognitive measure) causes higher wages or whether additional years of schooling cause both higher measured cognitive scores and higher wages. Least squares estimates likely overstate the contribution of ability to wages and understate the contribution of schooling to wages (see Carneiro, Heckman, and Masterov 2005).

The analysis of Bowles and Gintis (1976) suggests that a similar phenomenon may be at work for noncognitive skills. They claim that schooling builds traits that are useful in the workplace. In their language, schooling produces a docile proletariat. In addition, scores on the attitude scales used to proxy noncognitive ability, as well as the cognitive scores, are likely to be affected by family background characteristics and are at best imperfect measures of an individual's true noncognitive and cognitive abilities. The least-squares estimates reported in table 1 will be biased and inconsistent unless the measures used are perfect proxies for cognitive and noncognitive skills.

Standard IV methods for addressing measurement error and simultaneity in test scores are also subject to important limitations. First, the instruments selected are often controversial. Second, in a model with heterogeneous responses, it is far from clear how instrumental variables can solve these problems (Heckman and Vytlacil 2005; Heckman, Urzua, and Vytlacil 2006). The empirical strategy presented in this article, unlike the IV strategy, is able to account for the problems of reverse causality, measurement error, and heterogeneous responses.[4]

We develop an alternative to IV that postulates a low-dimensional vector of latent cognitive and noncognitive abilities that generates measured cognitive and noncognitive test scores and that is the source of dependence among not only test scores, schooling choices, and wages but also employment, occupational choice, and behavioral outcomes. Controlling for the latent skills solves the problems of endogeneity and measurement error. Our method extends the LISREL model of Jöreskog (1977) and the MIMIC model of Jöreskog and Goldberger (1975) to account for the effects of choice variables (schooling) and background variables on the measurements of cognitive and noncognitive skills where the schooling, in turn, depends on the latent factors. We estimate a factor model with endogenous factor loadings. Our methodology is a form of matching where the match variables that create the conditional independence are unobserved and their distributions are estimated nonparametrically. Carneiro et al. (2003) and Hansen et al. (2004) develop this method. We now present our model.

III. A Model of Schooling, Employment, Work Experience, Occupational Choice, and Wages Based on Latent Skills

Cognitive and noncognitive skills can affect the endowments of persons, their preferences, their technology of skill formation (see Cunha, Heck-

[4] Table S1 in our Web appendix extends the analysis presented in table 1 to consider other labor market and behavioral outcomes. It presents estimates of the effects of the measured abilities on schooling, occupational choice, smoking, drug use, incarceration, participation in illegality (whether an individual participated in any of the following illegal activities in 1979 or 1980: attempting to con someone, taking a vehicle without the owner's permission, shoplifting, intentionally damaging another's property, or using force to obtain things), work experience, and premarital pregnancy. These models are estimated using probit analysis and multinomial choice models. At a purely descriptive level, both measured cognitive and noncognitive traits are associated with a variety of behavioral outcomes for males and females. At issue is whether the relationships in table S1 have any causal status. The same issue applies to the results presented in table 1. Simple IV strategies that might be useful for linear outcome models do not apply in analyzing the nonlinear (discrete choice/discrete outcome) models analyzed in table S1.

man, Lochner, and Masterov 2006), or all three. They might affect risk preference, time preference, and the efficiency of human-capital production without necessarily being direct determinants of market wages. Cognitive and noncognitive skills might also raise the productivity of workers and directly affect wages. Our empirical analysis suggests that both cognitive and noncognitive skills play multiple roles.

We postulate the existence of two underlying factors representing latent cognitive and noncognitive ability. Conditioning on the observables, these factors account for all of the dependence across choices and outcomes. The levels of an individual's factors may result from some combination of inherited ability, the quality of the environment provided by his parents, and the effects of any early interventions. We assume that the factors are known by each individual but not by the researcher and that they are fixed by the time the individual makes his schooling and behavioral choices.

Let f^C and f^N denote the levels of latent cognitive and noncognitive abilities, respectively. We assume that latent abilities are mutually independent ($f^C \perp\!\!\!\perp f^N$), and both determine the individual's wage, schooling, employment, work experience, and occupational decisions.[5]

The assumption that one latent factor captures cognitive ability is traditional in the literature (see, e.g., Jensen 1998). The g theory used by Herrnstein and Murray (1994) and many others is based on it. Heckman (1995) shows that it applies to the NLSY79 data we use. The assumption that one latent factor captures noncognitive ability is less traditional. Since there are many aspects of noncognitive skills—self-control, time preference, sociability, and so forth—it is less likely that one trait captures all aspects of these behaviors.[6] Nonetheless, a model with one factor for cognitive skills and one for noncognitive skills is a useful starting point, and we use it throughout this article.[7]

The assumption of independence between f^C and f^N is motivated by the evidence presented in our Web appendix A.[8] Table S3 shows that correlations of test scores within the batteries of cognitive tests and noncognitive tests are much stronger than they are across cognitive and non-

[5] We can identify a model with correlated factors. See Carneiro et al. (2003) or Cunha and Heckman (2006c). The independence assumption is a normalization and a convenient point of departure. Cunha and Heckman (2006c) estimate models with correlated factors and establish identification.

[6] The evidence in our Web app. A, table S2, argues against the existence of only one latent factor that summarizes all aspects of noncognitive ability. For cognitive scores, one factor explains 77% of the trace of the cognitive test-score correlation matrix for males. The second factor explains only 9% of the trace. For noncognitive skills, one factor explains only 31% of the trace of the correlation matrix. The second factor explains 9% of the trace.

[7] We relax this assumption in work under way.

[8] See Cunha and Heckman (2006b, 2006c), who relax this assumption in both theoretical and empirical work.

cognitive tests. The cross-correlations weaken further when we condition on family background variables. We can account for the dependence across cognitive and noncognitive test scores, even invoking independence between f^C and f^N, by allowing observables to affect means and factor loadings. In addition, both factors affect schooling. In our model, the factor loadings in the test-score equations depend on schooling at the time of the test. Therefore, for those who complete their schooling by the time of the test, both latent factors affect both cognitive and noncognitive tests, albeit in an indirect way. We now present our model for wages and work experience.

A. A Hedonic Model for Wages and Work Experience

We allow for the possibility that different schooling groups operate in different labor markets. Both latent abilities and observable variables determine wages in the different schooling markets and may be priced differently in different markets. Denote by s the schooling level attained by the individual. Wages are given by a linear-in-the-parameters specification:

$$Y_s = \beta_{Y,s} X_Y + \alpha_{Y,s}^C f^C + \alpha_{Y,s}^N f^N + e_{Y,s} \quad \text{for} \quad s = 1, \dots, \bar{S},$$

where X_Y is a vector of observed controls, $\beta_{Y,s}$ is the vector of returns associated with X_Y, $\alpha_{Y,s}^C$ and $\alpha_{Y,s}^N$ are the parameters associated with the cognitive and noncognitive factors (i.e., factor loadings), respectively, and $e_{Y,s}$ represents an idiosyncratic error term such that $e_{Y,s} \perp\!\!\!\perp (f^N, f^C, X_Y)$ for $s = 1, \dots, \bar{S}$. This equation allows for separate prices for workers of different schooling categories, who operate in different labor markets.

We estimate a parallel equation for work experience:

$$R_s = \beta_{R,s} X_R + \alpha_{R,s}^C f^C + \alpha_{R,s}^N f^N + e_{R,s} \quad \text{for} \quad s = 1, \dots, \bar{S},$$

where X_R is a vector of observed controls, $\beta_{R,s}$ is the vector of returns associated with X_R, $\alpha_{R,s}^C$ and $\alpha_{R,s}^N$ are the cognitive and noncognitive loadings, respectively, and $e_{R,s}$ represents an idiosyncratic error term such that $e_{R,s} \perp\!\!\!\perp (f^N, f^C, X_R)$ for $s = 1, \dots, \bar{S}$.

B. The Model for Schooling

Each agent chooses the level of schooling, among $\bar{S}$ possibilities, that maximizes his benefit. Let I_s represent the net benefit associated with each schooling level s ($s = \{1, \dots, \bar{S}\}$), and assume the following linear-in-the-parameters model for the benefit of schooling level s:

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

where X_s is a vector of observed variables affecting schooling, β_s is its associated vector of parameters, α_s^C and α_s^N are the factor loadings associated with the cognitive and noncognitive latent abilities, respectively, and e_s represents an idiosyncratic component assumed to be independent of f^N, f^C, and X_s. The individual components $\{e_s\}_{s=1}^{\bar{S}}$ are mutually independent. All of the dependence across these choices comes through the observable, X_s, and the common factors f^N and f^C. The I_s's solve out the effects of wages and other benefits on the utility associated with schooling.

The agent chooses the level of schooling with the highest benefit. Formally,

$$D_S = \arg\max_{s\in\{1,\ldots,\bar{S}\}} \{I_S\}, \tag{2}$$

where D_S denotes the individual's chosen schooling level. Notice that conditional on X_s (with $s = 1, \ldots, \bar{S}$), equations (1) and (2) produce a standard discrete choice model with a factor structure.[9]

The assumption of linearity in the parameters and separability of the factors simplifies the analysis. In more tightly specified economic models, the factors would be nonlinear and inseparable as, for example, time preference parameters, risk aversion parameters, human-capital production function parameters, and endowment parameters in dynamic models of skill accumulation (see, e.g., Cunha and Heckman 2006b; Cunha, Heckman, Lochner, and Masterov 2006). We interpret f^N and f^C as approximations to the basic parameters of preferences, technology, and endowments that generate the outcomes we study. We discuss a more tightly specified model in Section V. We next develop the equation for employment.

C. The Model for Employment

Let I_E denote the net benefit associated with working, and assume a linear-in-the-parameters specification:

$$I_E = \beta_E X_E + \alpha_E^C f^C + \alpha_E^N f^N + e_E, \tag{3}$$

where β_E, X_E, α_E^C, α_E^N, and e_E are defined as in the schooling model. Then, $D_E = 1(I_E > 0)$ is a binary variable that equals one if the individual is employed and zero otherwise (where 1 is an indicator function, $1(A) = 1$ if A is true, and $1(A) = 0$ otherwise). The error term e_E is such that $e_E \perp\!\!\!\perp (f^N, f^C, X_E)$.

[9] See Heckman (1981), where this model was first introduced.

D. The Model for Occupational Choice

Let I_O denote the latent utility associated with choosing a white-collar occupation (where the alternative is a blue-collar occupation). We postulate the following linear model for I_O:

$$I_O = \beta_O X_O + \alpha_O^C f^C + \alpha_O^N f^N + e_O, \tag{4}$$

where β_O, X_O, α_O^C, α_O^N, and e_O are defined analogously to the model of equation (3). And $D_O = 1(I_O > 0)$ is an indicator of choice of white-collar occupational status. The error term in equation (4) is such that $e_O \perp\!\!\!\perp (f^N, f^C, X_O)$.

Further, assume that $e_{Y,s} \perp\!\!\!\perp e_{R,s'} \perp\!\!\!\perp e_{s''} \perp\!\!\!\perp e_E \perp\!\!\!\perp e_O$ for any schooling levels s, s', and s'' and that all of the error terms are independent of both factors (f^C and f^N) and all the observables (X variables with subscripts) in our model.

E. A Measurement System That Accounts for Simultaneity in Cognitive and Noncognitive Test Scores

Identification of the model of Sections III.A–III.D is established using the strategy developed in Carneiro et al. (2003) and elaborated in Hansen et al. (2004). For the sake of brevity, in this article we summarize their results without repeating their proofs.[10]

Our identification strategy assumes the existence of two sets of measurements (each with at least two elements), with one set measuring cognitive skills and the other set measuring noncognitive skills.[11] In our case, latent cognitive ability is only allowed to affect scores on cognitive measures, and latent noncognitive ability is only allowed to affect scores on noncognitive measures.[12]

Building on the analysis of Hansen et al. (2004), we address the possibility of reverse causality between schooling and cognitive and noncognitive test scores. In the context of this article, the problem is likely to arise since our measures of cognitive and noncognitive abilities were administered to all sample members in 1979 and 1980, when they were between 14 and 23 years of age. Many had finished their schooling. Consequently, the observed measures may not be fully informative about the

[10] A more technical discussion of aspects of identification is presented in our Web app. B.

[11] We can weaken the number of required measurements if we assume nonnormality of the factors following the analysis in Bonhomme and Robin (2004) and the discussion in Heckman and Navarro (2006).

[12] These conditions are sufficient but not necessary. See Carneiro et al. (2003, n. 18) for identification of a factor system where all but one test can depend on both factors.

latent cognitive and noncognitive skills of the individuals, since they may be influenced by the schooling level at the date of the test.

Our procedure allows each individual's schooling level at the time of the test to affect the coefficients of the measurement system. Thus, if we denote by s_T the schooling level at the time of the test ($s_T = 1, \ldots, \bar{S}_T$), the model for the cognitive measure C_i ($i = 1, \ldots, n_C$) is

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

$$\text{for} \quad i = 1, \ldots, n_C \quad \text{and} \quad s_T = 1, \ldots, \bar{S}_T,$$

where $e_{C_i}(s_T) \perp\!\!\!\perp (f^C, X_C)$ and $e_{C_i}(s_T) \perp\!\!\!\perp e_{C_j}(s'_T)$ for any $i,j \in \{1, \ldots, n_C\}$ and schooling levels s_T and s'_T, such that $i \neq j$ for any (s_T, s'_T) or $s_T \neq s'_T$ for any (i,j).[13]

Likewise, the model for the noncognitive measure N_i ($i = 1, \ldots, n_N$) is

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

$$\text{for} \quad i = 1, \ldots, n_N \quad \text{and} \quad s_T = 1, \ldots, \bar{S}_T,$$

where $e_{N_i}(s_T) \perp\!\!\!\perp (f^N, X_N)$ and $e_{N_i}(s_T) \perp\!\!\!\perp e_{N_j}(s'_T)$ for any $i,j \in \{1, \ldots, n_N\}$ and schooling levels s_T and s'_T, such that $i \neq j$ for any (s_T, s'_T) or $s_T \neq s'_T$ for any (i,j). Again, all error terms (e variables with subscripts) are mutually independent, independent of (f^N, f^C) and all the observable X's.

Since there are no intrinsic units for the latent ability measures, one α coefficient devoted to each ability must be normalized to unity to set the scale of each ability. Therefore, for some C_i ($i = 1, \ldots, n_C$) in C and N_j ($j = 1, \ldots, n_N$) in N, we set $\alpha_{C_i}(s_T) = 1$ and $\alpha_{N_j}(s'_T) = 1$. Carneiro et al. (2003) establish that these assumptions provide enough structure to semiparametrically identify the model, including the distributions of the factors and the equation-specific shocks, provided that the regressors have sufficient support.

Our assumptions imply that conditional on the X variables, the dependence across all measurements, choices, and outcomes comes through f^N and f^C. If we control for this dependence, we control for the endogeneity in the model.[14] If the (f^N, f^C) were observed, we could use matching to control for this dependence. Instead, we assume that the match variables are unobserved and estimate their distributions, along with the other parameters of the model.

[13] Our procedure includes the case where s_T is final schooling. See Hansen et al. (2004).

[14] Recall that the factor loadings in the measurement equations can depend on schooling at the time of the test and hence that the dependence is more complicated than in the standard factor analysis model.

IV. Incorporating Behavioral Outcomes into the Model

Much of the literature estimating the impact of cognitive and noncognitive abilities has focused on the effects of these abilities on educational and labor market outcomes (e.g., Bowles et al. 2001; Cameron and Heckman 2001; Segal 2005; Osborne-Groves 2006). Herrnstein and Murray (1994) present evidence on the correlation between levels of cognitive ability and different dimensions of social behavior (e.g., marriage, out-of-wedlock birth, and crime). They consider only the predictive power of cognitive ability measures. We use our model to consider the predictive power of both cognitive and noncognitive measures. We establish that noncognitive factors are important in explaining numerous labor market outcomes and social behaviors.

We investigate the effects of both types of latent abilities on individuals' decisions regarding teenage pregnancy and marital status and whether or not to smoke daily by age 18, use marijuana in 1979 or 1980, participate in activities that lead to incarceration by age 30, and participate in other illegal activities in 1979 or 1980. Our model assumes that each of these decisions is jointly determined by latent cognitive and noncognitive abilities, as well as by observable variables and outcome-specific shocks.

The models that we fit are all in the form of linear-in-the-parameters index models that generate discrete outcomes of the sort analyzed in Section III. Let I_j be the linear-in-the-parameters index for behavior j, with associated vector X_j and coefficient vector β_j. Let α_j^C be the loading on the cognitive factor and α_j^N the loading on the noncognitive factor. The latent index generating choices is

$$I_j = \beta_j X_j + \alpha_j^C f^C + \alpha_j^N f^N + e_j; \tag{5}$$

$$D_j = 1(I_j \geq 0); \tag{6}$$

where e_j is independent of f^N, f^C and X_j, and f^N and f^C are independent of X_j.

We analyze daily smoking, marijuana use, imprisonment, and illegal activities using this framework. We study teenage pregnancy and marriage for women using a multinomial choice model. Let I_p denote the latent utility associated with the decision p ($p = 1$ [single with no child], $p = 2$ [married with a child], $p = 3$ [married with no child], and $p = 4$ [single with a child]). We postulate the following linear-in-the-parameters model for I_p:

$$I_p = \beta_p X_p + \alpha_p^C f^C + \alpha_p^N f^N + e_p \quad \text{for} \quad p = 1, \ldots, 4, \tag{7}$$

where β_p, X_p, α_p^C, α_p^N, and e_p are defined analogously to the previous cases. From equation (7) we define the outcome selected by

$$D_P = \arg \max_{p \in \{1,\dots,4\}} \{I_p\}$$

so that D_P denotes the individual's chosen marital and pregnancy status. We assume that the X's are independent of f^N, f^C and the e_p's. The f^N, f^C are independent of the e_p's, and the components of the e_p's are mutually independent. Again, all of the dependence across equations comes from the X's and the factors f^N, f^C. All distinctly subscripted e variables (across all labor market and behavioral outcomes) are mutually independent and independent of f^N, f^C, and all subscripted X variables. Again, this is a form of matching where the unobserved components of the match variables are independent of the observed components, and we estimate their distribution.

V. Interpreting Our Model as an Approximation to an Explicit Economic Model

Our statistical model is an approximation to a simple life-cycle model of youth and adult decision making over horizon T. We now sketch that model. Let consumption and labor supply at period t be $c(t)$ and $l(t)$, respectively. Consumption is a vector and includes a variety of behaviors, such as smoking, drug use, and so on. Let the vector $P(t)$ denote the market prices of the consumption goods. Utility is $U(c(t), l(t); \eta)$ where the η are preference parameters. The agent discounts utility at time preference rate ρ. Human capital in period t is $h(t)$, which can be a vector. It is produced by the human-capital production function

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

where the τ are productivity parameters, $I(t)$ is investment at t, and $\dot{h}(t)$ denotes the rate of change of the human-capital stock. The initial condition is given by $h(0)$.

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

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

Assuming perfect credit markets at interest rate r, the law of motion for assets at period t ($A(t)$), given initial condition $A(0)$ and ignoring taxes, is

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

The agent maximizes

$$\int_0^T \exp(-\rho t) U(c(t), l(t); \eta) dt$$

subject to the laws of motion of assets and human capital.

In this specification, cognitive and noncognitive skills can affect preferences ($\eta = \eta(f^C, f^N)$, $\rho = \rho(f^C, f^N)$), human-capital productivity ($\tau = \tau(f^C, f^N)$), and direct market productivity ($\theta = \theta(f^C, f^N)$). They might also affect initial conditions $h(0) = h_0(f^C, f^N)$ and $A(0) = A_0(f^C, f^N)$.

Our econometric model is a linear-in-the-parameters approximation to this general model. In this article, we do not estimate relationships for each of the channels through which cognitive and noncognitive abilities might operate. Noncognitive abilities affect some combination of η, ρ, τ, and θ (market productivity). Cognitive abilities operate through θ as well as some combination of η, ρ, and τ.[15]

An open question, which we plan to address in other work, is the relationship between the psychologist's measure of noncognitive skills as elicited from test scores and the fundamental parameters of risk aversion, time preference, and human-capital productivity, which can be estimated from behaviors (see, e.g., Browning, Hansen, and Heckman 1999). In principle, one can determine which factors are common across tests and preference parameters. Test scores and behaviors can be used interchangeably to proxy factors. This task is left for future work.

VI. Implementing the Model

We use Bayesian Markov chain Monte Carlo methods to compute the sample likelihood. Our use of Bayesian methods is only a computational convenience. Our identification analysis is strictly classical.[16] Under our

[15] Cunha and Heckman (2006b) estimate a more general model in which the (f^C, f^N) evolve over time and are consequences of investment behavior.

[16] The analysis in Carneiro et al. (2003), Hansen et al. (2004), and Heckman and Navarro (2006) establishes conditions on the support of the regressors that allow for semiparametric identification of the model. Figure S2 presents evidence on the support conditions for both males and females. It graphs the sample distributions of probabilities of different schooling attainment levels. For the support conditions for semiparametric identification to hold, the support of the distribution of each probability should be the unit interval [0,1]. It is evident from fig. S2 that this condition is not met, although for 4-year-college graduation the condition is nearly satisfied. This evidence suggests that the empirical results that we generate are identified from the parametric structure of the model. However, we use a robust mixture of normal approximation to the underlying distributions. Varying the components of the mixtures (adding more components beyond the ones we report) does not change our empirical estimates. Our estimates are not artifacts of normality assumptions, and relaxing normality is essential in obtaining a good fit to the data.

assumptions, the priors we use are asymptotically irrelevant. Explanatory variables and exclusion restrictions are reported in tables 2 and 3.

Our empirical model has six schooling levels ($\bar{S} = 6$): high school dropout, GED, high school graduate, some college and no degree, 2-year-college degree, and 4-year-college degree. To facilitate identification of the educational choice model, we assume that tuition at 2- and 4-year colleges only affects the benefits of obtaining those degrees and that the cost of obtaining the GED only affects the benefit of obtaining that degree.[17] We also assume that local-area wages and unemployment rates at age 17 for individuals with each final schooling level (i.e., high school dropouts, high school graduates, some college, and college graduates) partly determine the opportunity cost and expectations of returns associated with the final schooling level. Family background characteristics, race and cohort dummies, as well as both factors, are also allowed to affect educational choices.

Wage equations at age 30 are estimated for individuals at each final schooling level. Race and ethnicity dummies, cohort dummies, local labor market conditions, and region of residence dummies are included in these equations, as well as the cognitive and noncognitive factors.[18] We assume that, fixing these variables, family background characteristics and childhood residence do not affect adult wages. The local labor market variables are based on the Bureau of Economic Affairs database discussed in Cameron and Heckman (2001), updated for our sample year.

The employment and occupational choice latent indices are assumed to depend on the same list of variables that determine adult wages.[19] Family background characteristics, race and cohort dummies, and both factors enter into the index functions determining daily smoking, marijuana use, incarceration, participation in illegal activities, and teenage pregnancy. Family background characteristics, race and cohort dummies, and both factors also enter into the equations determining work experience by age 30.

Our theoretical model is static and does not consider the timing of decisions. We analyze smoking and marital-pregnancy decisions (for

[17] Exclusions are required for semiparametric identification of the choice equations unless curvature restrictions are introduced (see Cameron and Heckman 1998; Heckman and Navarro 2006). Alternatively, we can invoke a parametric distributional assumption.

[18] Urzua (2006) presents race ethnicity specific estimates.

[19] The blue- and white-collar distinction is made according to the following definition. The following occupations are classified as white-collar: professional, technical, and kindred; managers, officials, and proprietors; sales workers; farmers and farm managers; and clerical and kindred. The following occupations are classified as blue-collar: craftsmen, foremen, and kindred; operatives and kindred; laborers, except farm; farm laborers and foremen; and service workers.

women only) as of age 18, marijuana use and participation in illegal activities in 1979 or 1980,[20] and incarceration by age 30 (for men only). Labor market outcomes and schooling decisions are studied as of age 30.

Following the analysis in Section III.E, our cognitive and noncognitive measures are allowed to depend on the cognitive (f^C) and noncognitive (f^N) factors, respectively. Each equation is estimated allowing the highest grade attained at the time of the test to affect means and factor loadings and includes as controls family background characteristics and cohort dummies.[21] Our cognitive measures are five ASVAB test scores. We use two attitudinal scales, the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale, as our noncognitive measures. We choose these measures because of their availability in the NLSY79. Ideally, it would be better to use a wider array of psychological measurements and, as previously noted, to connect them with more conventional measures of preference parameters in economics.

As explained in Section III.E, two normalizations are required to assure identification of the model. These set the scale of the factors. We normalize the loadings ($\alpha_{C_i}(s_T)$, $\alpha_{N_j}(s'_T)$) of the cognitive (f^C) and noncognitive (f^N) factors to be equal to one in the equations associated with coding speed and the Rosenberg Self-Esteem Scale for individuals in grades 9–11 at the time of the test, respectively.

The distributions of the unobservables are identified nonparametrically. The factors are estimated as three-component mixtures of normals. The uniquenesses (the e) of the wage equations are distributed as three-component mixtures of normals.[22] The other uniquenesses are normally distributed. When we permit them to be nonnormal, the fit of the model does not improve.

VII. Evidence from the Model

Estimates of the parameters of the equations of the model are presented in Web appendix tables S4–S20. The model fits the data on wages and other outcomes.[23] Overall goodness-of-fit tests are passed for most

[20] The definition of illegal activities is given in the note to table 2.

[21] The schooling levels at test date considered in the estimation of the cognitive measurement system are grades 9–11, grade 12, 13–15 years of schooling, and 16 or more years of schooling. For the noncognitive measurement system the schooling levels are grades 9–11, grade 12, and 13 or more years of schooling. This difference is due to the years in which the different tests were administered. See Web app. A for details.

[22] Models for wages with fewer mixture components do not fit the data as well.

[23] See figs. S3A and S3B in our Web appendix at http://jenni.uchicago.edu/noncog/web_supplement.pdf.

Table 2
Variables in the Empirical Implementation of the Model, Outcome Equations

	Log of Hourly Wage* Employment† and Occupational Choice‡ Models	Educational Choice Model§ (Multinomial Probit): High School Dropouts	GED Recipients	High School Graduates	Some College, No Degree	2-Year Degree	4-Year Degree	Behavioral Outcomes,‖ Work Experience,# and Fertility Choice Model**
Variable:								
Black (dummy)	Yes	Yes	Yes	Yes	Yes	Yes	...	Yes
Hispanic (dummy)	Yes	Yes	Yes	Yes	Yes	Yes	...	Yes
Region of residence (dummy variables)	Yes	...	...	...	...	...	...	...
Urban residence (dummy)	Yes	...	...	...	...	...	...	...
Local unemployment rate at age 30	Yes	...	...	...	...	...	...	...
Living in an urban area at age 14 (dummy)	...	Yes	Yes	Yes	Yes	Yes	...	Yes
Living in the South at age 14 (dummy)	...	Yes	Yes	Yes	Yes	Yes	...	Yes
Family income in 1979	...	Yes	Yes	Yes	Yes	Yes	...	Yes
Broken home at age 14 (dummy)	...	Yes	Yes	Yes	Yes	Yes	...	Yes
No. of siblings at age 17 (dummy)	...	Yes	Yes	Yes	Yes	Yes	...	Yes
Mother's highest grade completed at age 17	...	Yes	Yes	Yes	Yes	Yes	...	Yes
Father's highest grade completed at age 17	...	Yes	Yes	Yes	Yes	Yes	...	Yes
Local wage of high school dropouts at age 17	...	Yes	...	...	...	...	...	...
Local unemployment rate of high school dropouts at age 17	...	Yes	...	...	...	...	...	...
Local wage of high school graduates at age 17	...	...	...	Yes	...	...	...	...

Local unemployment rate of high school graduates at age 17	. . .	. . .	. . .	Yes	. . .	. . .	. . .	. . .
Local wage of attendees of some college at age 17	. . .	. . .	. . .	. . .	Yes	. . .	. . .	. . .
Local unemployment rate of attendees of some college at age 17	. . .	. . .	. . .	. . .	Yes	. . .	. . .	. . .
Local wage for college graduates at age 17	. . .	. . .	. . .	. . .	. . .	. . .	Yes	. . .
Local unemployment rate for college graduates at age 17	. . .	. . .	. . .	. . .	. . .	. . .	Yes	. . .
Tuition at 2-year college at age 17	. . .	. . .	. . .	. . .	. . .	Yes	. . .	. . .
Tuition at 4-year college at age 17	. . .	. . .	. . .	. . .	. . .	. . .	Yes	. . .
GED costs	. . .	. . .	Yes	. . .	. . .	. . .	. . .	. . .
Cohort dummies	Yes	Yes	Yes	Yes	Yes	Yes	. . .	Yes
Factor:								
Cognitive	Yes	Yes	Yes	Yes	Yes	Yes	. . .	Yes
Noncognitive	Yes	Yes	Yes	Yes	Yes	Yes	. . .	Yes

* The log hourly wage model is estimated for six different categories: high school dropouts, GED recipients, high school graduates, some college but no degree, 2-year-college graduates, and 4-year-college graduates. Hourly wages are measured at age 30.

† Employment is at age 30.

‡ Occupational choice is white-collar or blue-collar, conditional on being employed at age 30.

§ The educational choice model is estimated considering six different categories: high school dropouts, GED recipients, high school graduates, some college but no degree, 2-year-college graduates, and 4-year-college graduates.

‖ Four behavioral choices are estimated: whether an individual smokes daily by age 18; whether an individual smoked marijuana in 1979 or 1980; whether an individual has been incarcerated by age 30 (estimated only for men); and whether an individual participated in any of the following illegal activities in 1979 or 1980: attempting to con someone, taking a vehicle without the owner's permission, shoplifting, intentionally damaging another person's property, or using force to obtain things.

Experience is measured as total years of work experience by age 30.

** The fertility choice model is a multinomial probit. It is estimated only for women and considers four choices for marital/fertility status by age 18: single with child, single with no child, married with child, and married with no child.

Table 3
Variables in the Empirical Implementation of the Model, Auxiliary Measures

	Test Scores (Cognitive Measures)	Attitude Scales (Noncognitive Measures)
Variable:		
Black (dummy)	Yes	Yes
Hispanic (dummy)	Yes	Yes
Living in an urban area at age 14 (dummy)	Yes	Yes
Living in the South at age 14 (dummy)	Yes	Yes
Mother's highest grade completed at age 17	Yes	Yes
Father's highest grade completed at age 17	Yes	Yes
Number of siblings at age 17 (dummy)	Yes	Yes
Family income in 1979	Yes	Yes
Broken home (dummy)	Yes	Yes
Cohort dummies	Yes	Yes
Factor:		
Cognitive	Yes	. . .
Noncognitive	. . .	Yes

Note.—The included cognitive measures are arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge, and coding speed. The included noncognitive measures are Rotter Locus of Control Scale and Rosenberg Self-Esteem Scale. 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 that they have control over their lives through self-motivation or self-determination (internal control), as opposed to the extent to which individuals believe that the environment 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 zero and variance one, after taking averages over the respective sets of scales.

outcome and choice equations.[24] The loadings on both cognitive and noncognitive factors are statistically significant in most equations. Both factors are required to produce a model that passes goodness-of-fit tests.[25] The estimated distributions of the factors are highly nonnormal. Standard normality assumptions would produce seriously biased estimates of the true factors and force symmetry onto highly asymmetric data.[26] We find strong evidence that schooling affects both measured cognitive ability and measured noncognitive ability.[27] The first finding

[24] See Web appendix tables S21A and S21B for men and women. In the case of experience, however, the model does not pass the overall goodness-of-fit test. By schooling level, the performance is much better, especially for males.

[25] Table S22 in the Web appendix shows that we reject the null hypotheses that either cognitive or noncognitive factors do not belong in the outcome and choice equations.

[26] See Web appendix table S23 and figs. S4A and S4B.

[27] For males, the χ^2 test for the null that schooling does not affect measured cognitive tests (means and factor loadings) is 431.65 with 150 degrees of freedom. Hence, we reject the null (the critical values are 179.58 [95%], 172.58 [90%]). The χ^2 test for the null that schooling does not affect the means and factor loadings of the latent noncognitive tests is 116.53 with 40 degrees of freedom. Hence, we

corroborates the earlier analyses of Neal and Johnson (1996), Hansen et al. (2004), and Heckman, Larenas, and Urzua (2004). The second result is new and corroborates the claims of Marxist economists (see, e.g., Bowles and Gintis 1976).

Because our model is nonlinear and multidimensional, the best way to understand it is to simulate it. Figure 3 plots the densities of the estimated cognitive and noncognitive factors by schooling level for men and women. These are to be compared with the densities of the raw test scores presented in figure 2. The distributions of f^N and f^C are clearly nonnormal. For the cognitive factor, the sorting patterns are about the same in figures 2 and 3, although the shapes are different. More cognitively able people attain higher levels of education. The GED recipients are smarter than dropouts, and their distribution of the cognitive trait is very close to that of high school graduates who do not go on to college.

Our estimated distribution of noncognitive ability reverses the pattern for dropouts and GED recipients that is found in the raw data reported in figure 2. Male GED recipients have a worse noncognitive ability distribution than dropouts. For females, dropouts and GED recipients have similar distributions of noncognitive skills. Thus, male GED recipients are the same or worse than high school dropouts in terms of noncognitive factors, but are better in cognitive terms. This confirms a hypothesis of Heckman and Rubinstein (2001) that GED recipients are as smart as high school graduates who go on to college but they have much lower noncognitive skills.

Figure 4 summarizes the estimated effects of schooling at the date of the test (s_T) on components of the ASVAB for males of average cognitive and noncognitive ability. Since the means of f^N and f^C are zero, these figures isolate the effect of schooling on the intercepts of the test-score equations. Schooling raises measured test scores. Figure 5 summarizes, for men, the effect of schooling at the test date on the noncognitive measures. Schooling raises scores on the Rotter Scale at lower levels of schooling. For the Rosenberg Scale, scores are raised across all grades of schooling.[28]

Figures 6–27 graphically summarize the main implications of our model for a variety of outcome measures. We report results for both men and women when there are differences by gender. Otherwise, we only report the results for men, posting the results for women at our Web appendix. The structure of these figures is the same across all outcomes. Each figure has three panels. Panel i displays the joint dis-

reject that hypothesis as well (the critical values are 55.75 [95%], 51.80 [90%]). For females, we obtain similar results. Table S24 in the Web appendix presents these results.

[28] The results for women are comparable and can be found at http://jenni.uchicago.edu/noncog/web_supplement.pdf. See figs. S5A and S5B.

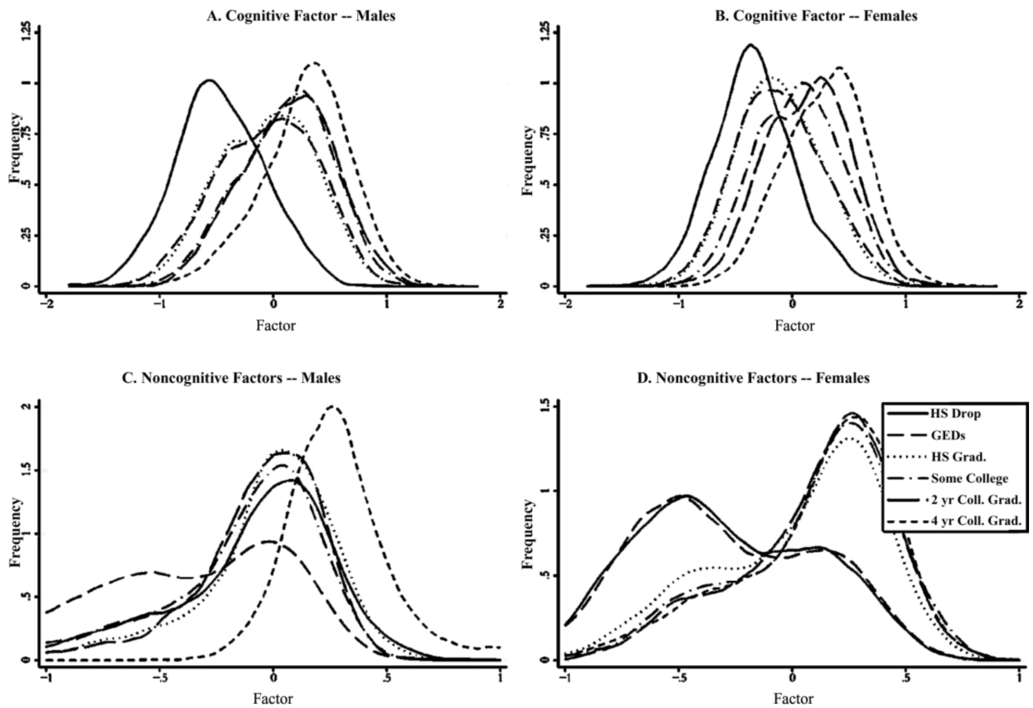


FIG. 3.—Distribution of factors by gender and schooling level. The factors are simulated from the estimates of the model. The schooling levels represent the simulated schooling level at age 30. The simulated data contain 19,600 observations.

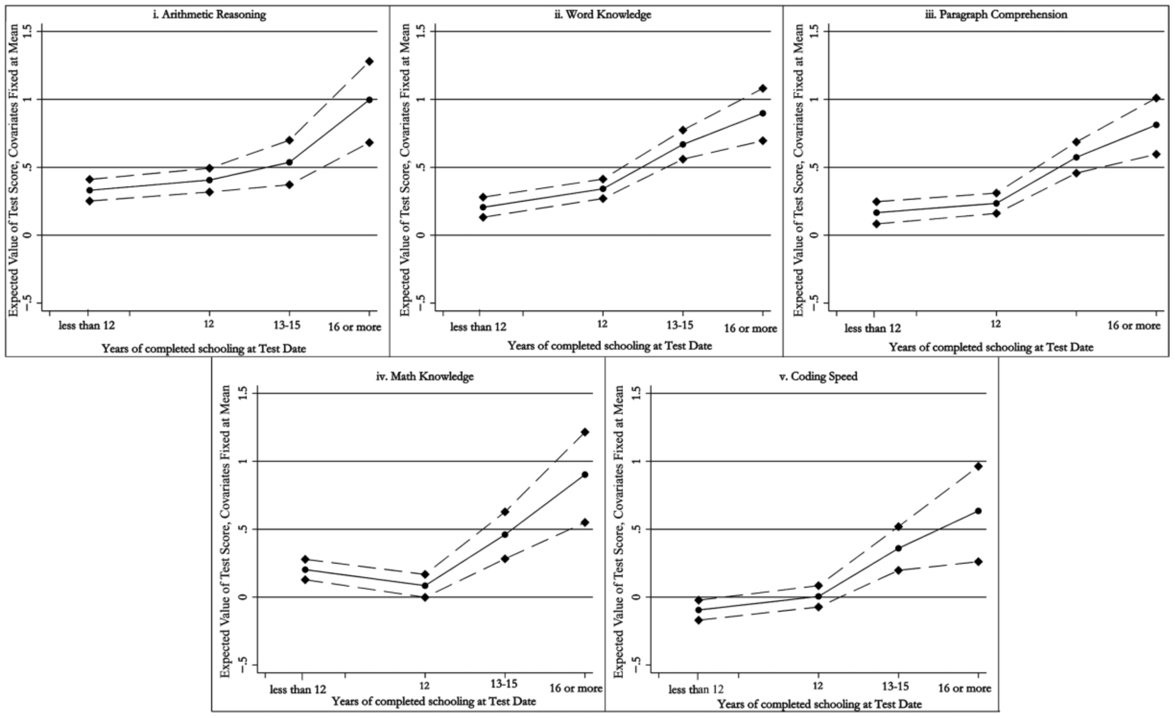


Fig. 4.—Effect of schooling on ASVAB components for males with average ability, with 95% confidence bands. We standardize the test scores to have within-sample mean zero, variance one. The model is estimated using the NLSY79 sample (see Web app. A for details). Solid lines depict average test scores, and dashed lines, 2.5%–97.5% confidence intervals.

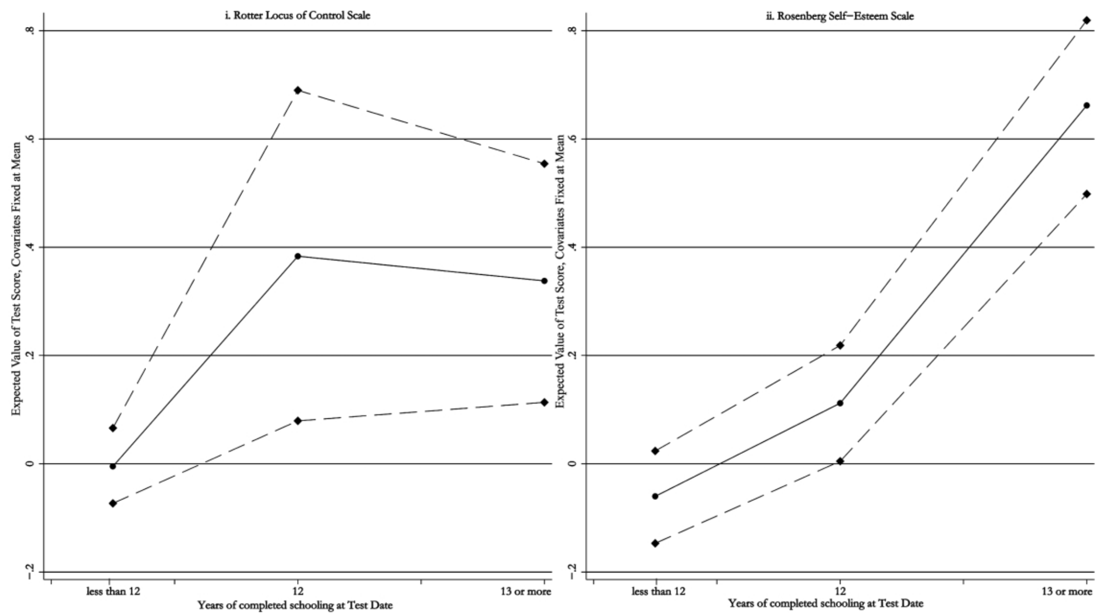


Fig. 5.—Effect of schooling on noncognitive scales for males with average ability, with 95% confidence bands. The locus of control scale is based on the four-item abbreviated version of the Rotter Internal-External Locus of Control Scale. This scale is designed to measure the extent to which individuals believe that they have control over their lives through self-motivation or self-determination (internal control) as opposed to the extent to which individuals believe that the environment 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 zero and variance one, after taking averages over the respective sets of scales. The model is estimated using the NLSY79 sample (see Web app. A for details). Solid lines depict average test scores, and dashed lines, 2.5%–97.5% confidence intervals.

tribution of the outcome reported by deciles of the cognitive and noncognitive factors, while panels ii and iii display the marginal effects of one factor integrating out the effect of the other factor.

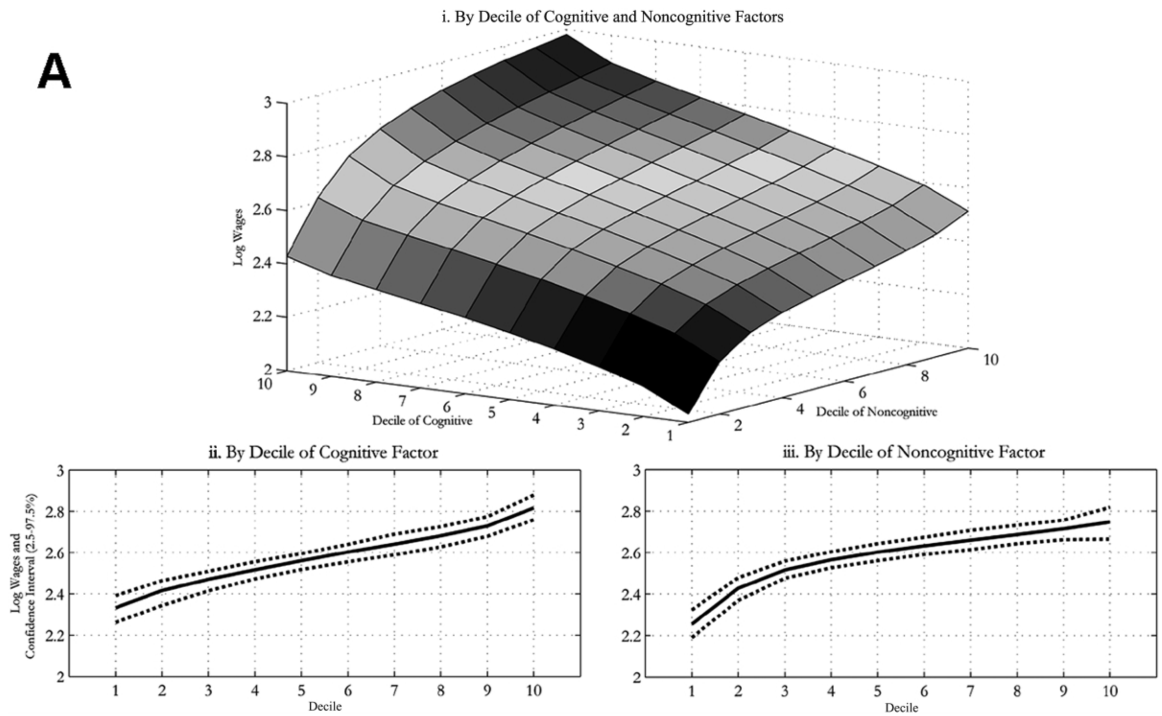
Mean log hourly wages by decile of cognitive and noncognitive ability for men and women are displayed in figure 6*A* and *B*, respectively. In this figure we display log wages as a function of the deciles of the factors. Standard error bands are presented along with the main graphs. For both men and women, cognitive skills have about the same effect on wages as noncognitive skills. The effect of noncognitive skills for men is slightly less strong, as measured by the slope of the log wage–ability decile curve, than it is for women.

Figure 6 displays the net effect of increases in the abilities on log wages inclusive of the direct effect of ability on log wages holding schooling fixed, the effect of ability on schooling, and the generated effect of schooling on log wages. Table 4 shows that the factor loadings (hedonic prices) on latent skills vary substantially across schooling levels. Noncognitive traits are not valued in the labor market for male 4-year-college graduates, although they are for female college graduates. In most of the educational labor markets, noncognitive factors are valued for both genders. For men, noncognitive traits are valued more highly in low-skill markets. For women, noncognitive traits are more uniformly valued.

Figures 7–12 show the valuation of each type of skill in different schooling labor markets jointly (panel i) and integrating out the factor not being studied (panels ii and iii). Panels ii and iii also display the proportion of individuals with the indicated level of schooling whose cognitive (panel ii) and noncognitive (panel iii) abilities lie in each decile of the overall distribution. We integrate out each of the regressors in performing this simulation. When the proportions are small, the standard error bands are larger. Across schooling markets different factors are priced differently. Thus, in the male-dropout market, the log wage gradient for noncognitive ability is greater than it is for cognitive ability. The opposite pattern is found for females. In the GED market, the gradient for noncognitive ability is greater than that of cognitive ability. For the high school market, the gradients are similar across skills for men and women, but the gradients are much steeper for women.

For those attending some college, the noncognitive gradients are much steeper than the cognitive gradients. In the market for 2-year-college graduates, the gradients are about equally strong across skills and across sex groups. For males in the 4-year-college market, noncognitive skills have little marginal value, while cognitive skills have a strong gradient. For females in the 4-year-college market, both skills command high marginal prices.

Figure 13*A* and *B* displays the effects of cognitive and noncognitive skills on employment for men and women, respectively. For both genders,


A
i. By Decile of Cognitive and Noncognitive Factors
Log Wages
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
Log Wages and Confidence Interval (2.5-97.5%)
Decile
iii. By Decile of Noncognitive Factor
Decile

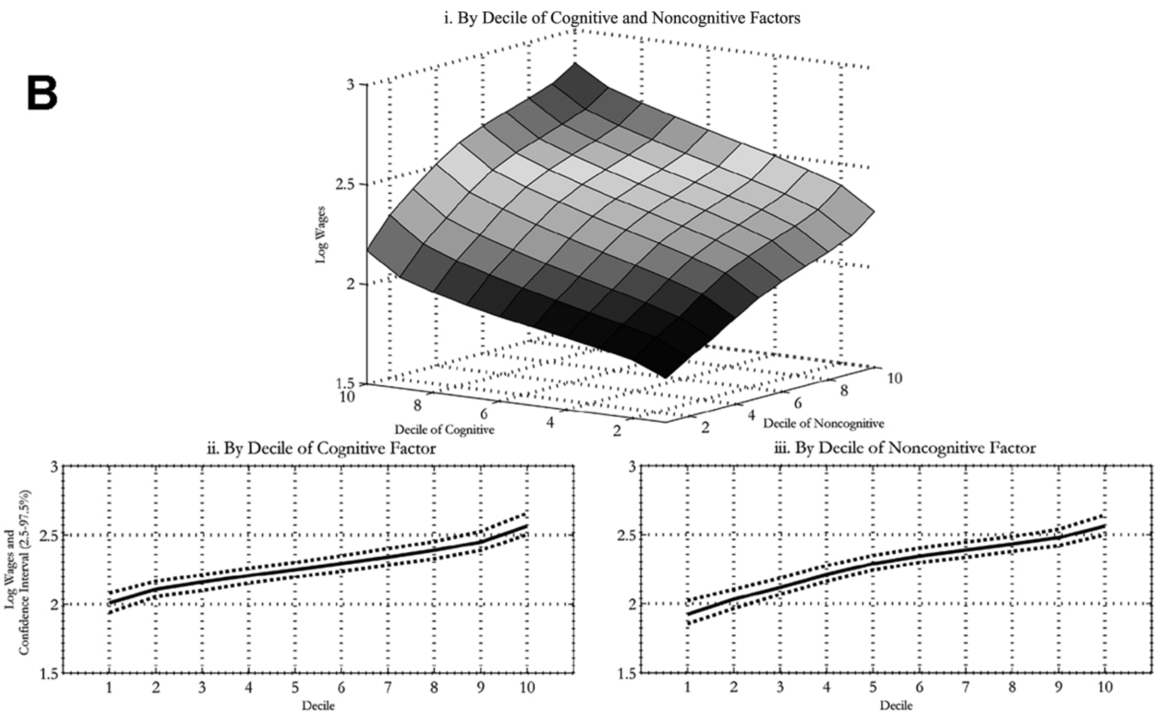


FIG. 6.—Mean log wages by age 30 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict overall (log) wages, and dashed lines, 2.5%–97.5% confidence intervals.

Table 4
Estimated Coefficients of the Cognitive and Noncognitive Factors for the Log Hourly Wage Model

	Males		Females	
Schooling Level	Cognitive	Noncognitive	Cognitive	Noncognitive
High school dropout	.113	.424	.322	.208
	(.076)	(.092)	(.125)	(.103)
GED	.175	.357	.020	.242
	(.107)	(.117)	(.137)	(.153)
High school graduate	.259	.360	.341	.564
	(.041)	(.059)	(.049)	(.056)
Some college, no degree	.069	.401	.093	.569
	(.086)	(.110)	(.084)	(.116)
2-year-college degree	.039	.368	.206	.279
	(.138)	(.209)	(.096)	(.145)
4-year-college degree	.296	−.060	.290	.379
	(.075)	(.175)	(.066)	(.103)

Note.—Standard errors are in parentheses. Sample from NLSY79 males and females at age 30. We exclude the oversample of blacks, Hispanics, and poor whites, the military sample, and those currently enrolled in college. The cognitive measure represents the standardized average over the raw ASVAB scores (arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge, 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 model also includes a set of cohort dummies, local labor market conditions (unemployment rate), and the region of residence.

the gradient on noncognitive skills is greater than it is for cognitive skills. The pattern is especially pronounced for women.

The effects of both cognitive and noncognitive ability on employment cumulate over the life cycle into effects on work experience. Figures 14 and 15 show the effects of both cognitive and noncognitive ability on work experience for male workers in different educational labor markets. Except for the market for 4-year-college graduates—the highest skill market we study—the gradient for noncognitive skills is much steeper than it is for cognitive skills. If anything, the results are more dramatic for women (see Web appendix). For both genders, cognitive and noncognitive abilities are important determinants of the choice of white- versus blue-collar occupations (see fig. 16).

We next consider the effects of cognitive and noncognitive abilities on schooling decisions. For the sake of brevity, we report results for selected schooling levels. We report results for women when they are different from those of men.

Figure 17 shows the effects of latent abilities on the high school dropout decision. Those at the top of the cognitive ability distribution are very unlikely to drop out. Both types of ability have strong effects on the dropout decision, but cognitive ability is more important in the sense of having a steeper gradient than noncognitive ability.[29] For the decision to

[29] The results for women show a steeper gradient for noncognitive skills (see fig. S6 at our Web appendix).

drop out from high school and attain a GED and not continue on to college, the opposite is the case (see fig. 18). For a man with cognitive ability in the lowest decile, increasing his noncognitive ability from the lowest to the highest decile decreases the probability that he will obtain a GED. The cognitive ability–GED curve is flat. Noncognitive factors play a strong role, with those who have high noncognitive skills unlikely to attain a GED.

The effects of both cognitive and noncognitive ability on attaining a high school degree and stopping there are not monotonic (see fig. 19 for men). At the lowest deciles of both abilities, increasing either ability raises the probability of graduating from high school and obtaining no further schooling. At higher levels, it decreases the probability as more able people (in both senses of ability) do not stop their education at high school but go on to attain higher levels of schooling. Similar phenomena appear for persons who attend (but do not graduate from) college. See figures S9 and S10 posted in our Web appendix.

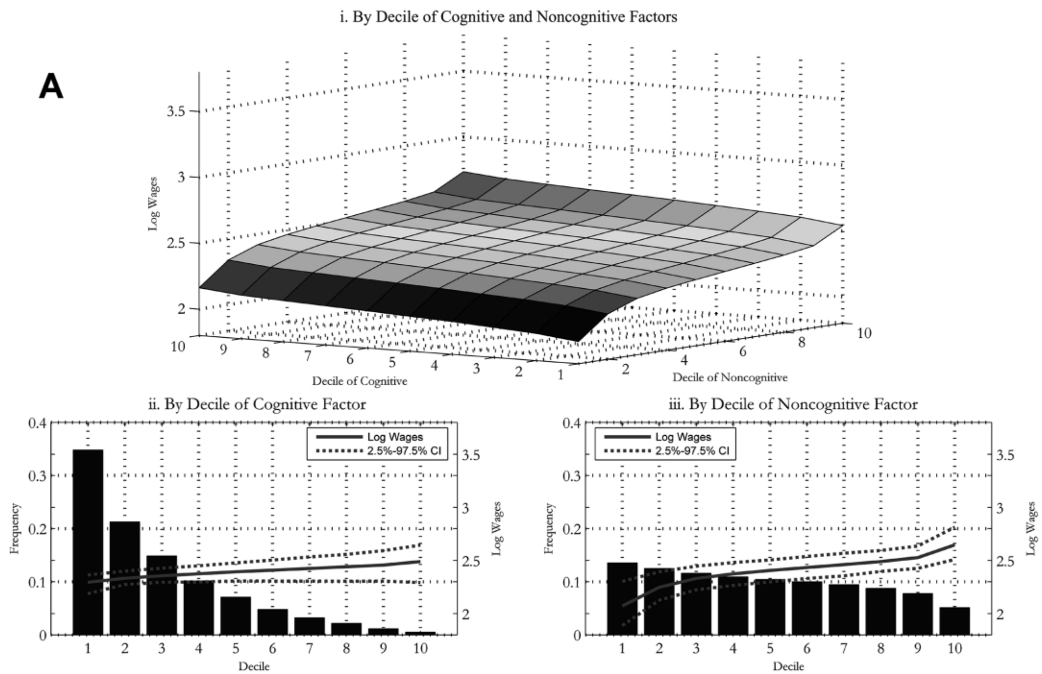
The effects of cognitive and noncognitive ability on the probability of graduating from a community college are weak (see fig. 20). The effects of noncognitive abilities are nonmonotonic. Figure 21 shows that both cognitive and noncognitive abilities have strong effects on graduating from a 4-year college. The gradient of noncognitive ability on the probability of graduating from a 4-year college is smaller for women (see fig. S12 in our Web appendix).

For daily smoking by age 18, an equivalent decile movement in the noncognitive factor induces a larger change in behavior for males than does a change in the cognitive factor. For women, the opposite is true (see fig. 22). For men, increasing noncognitive ability from the lowest to the highest decile decreases their probability of using marijuana (see fig. 23). Cognitive skills are not strong predictors of marijuana use.

Figure 24 displays the probability of incarceration by age 30 for males.[30] Although both factors are important, we find that the noncognitive factor induces a much larger change in behavior than a comparable decile change in the cognitive factor. For males in the lowest decile of the cognitive distribution, moving from the lowest to the highest decile of the noncognitive distribution substantially decreases the probability of incarceration. In comparison, taking the same males who are in the lowest deciles of both distributions and moving them to the highest decile of the cognitive distribution only slightly decreases their probability of incarceration. Contrary to claims made by Herrnstein and Wilson (1985) and Herrnstein and Murray (1994), it is noncognitive ability, not cognitive ability, that is the dominant factor in explaining different rates of participation in crime.

[30] For females, incarceration is not an empirically important phenomenon.

A


i. By Decile of Cognitive and Noncognitive Factors
Log Wages
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
iii. By Decile of Noncognitive Factor
Log Wages
2.5%-97.5% CI
Frequency
Decile

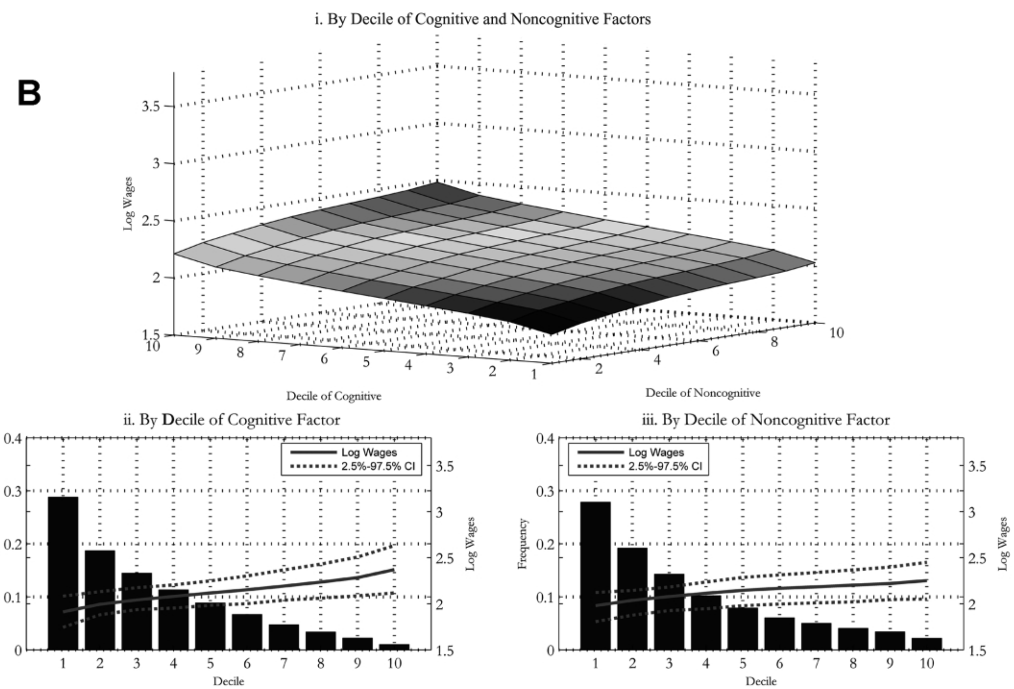
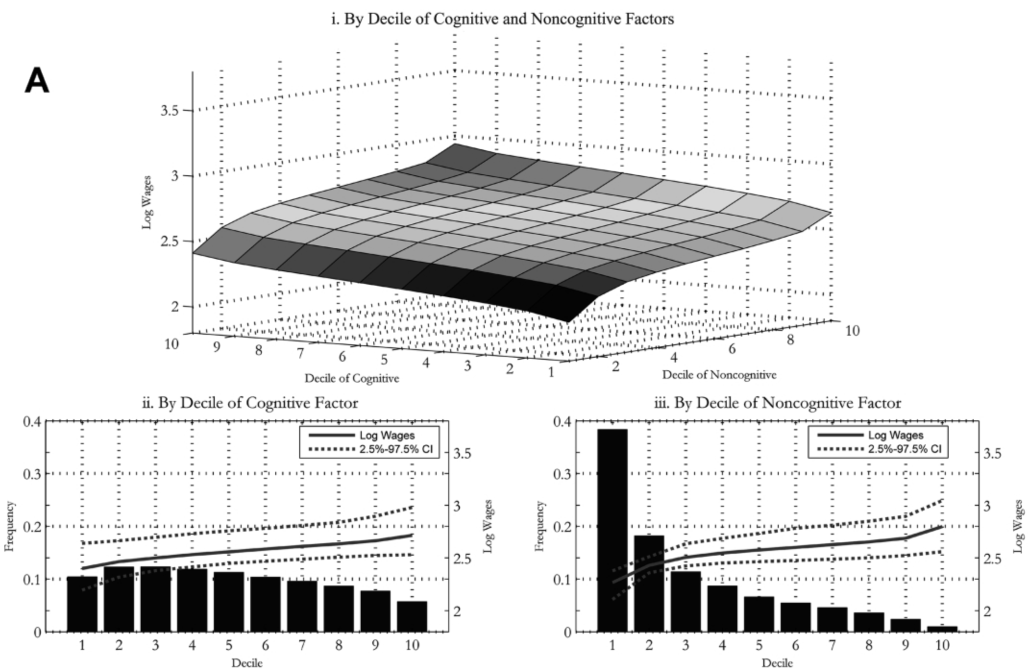


FIG. 7.—Mean log wages of high school dropouts at age 30 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are 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.


A
i. By Decile of Cognitive and Noncognitive Factors
Log Wages
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
iii. By Decile of Noncognitive Factor
Frequency
Log Wages
Decile
Log Wages
2.5%-97.5% CI

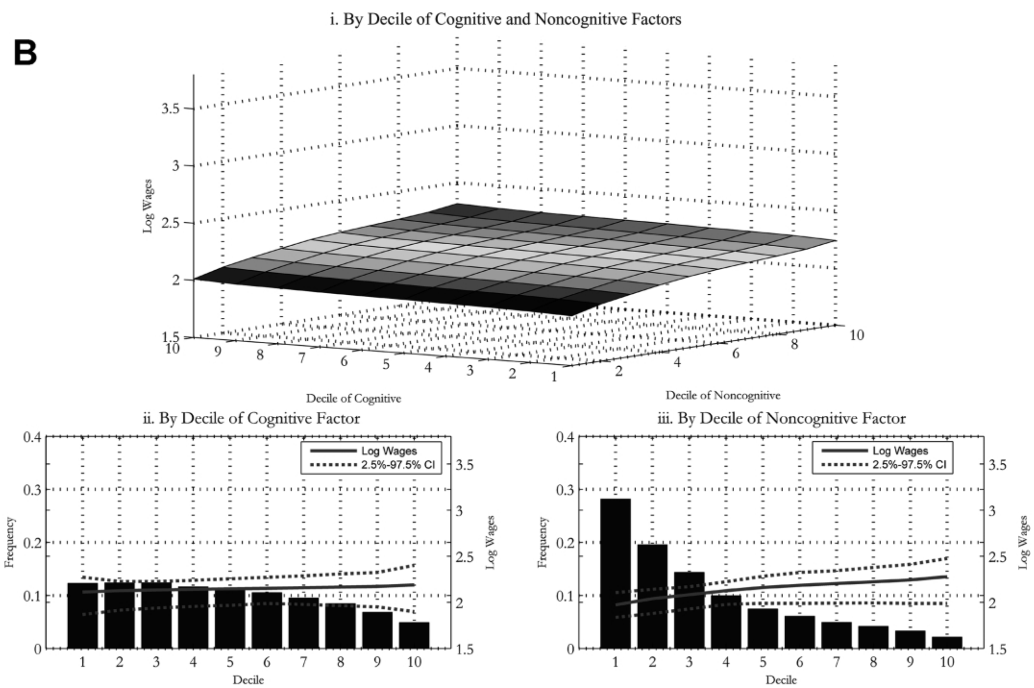
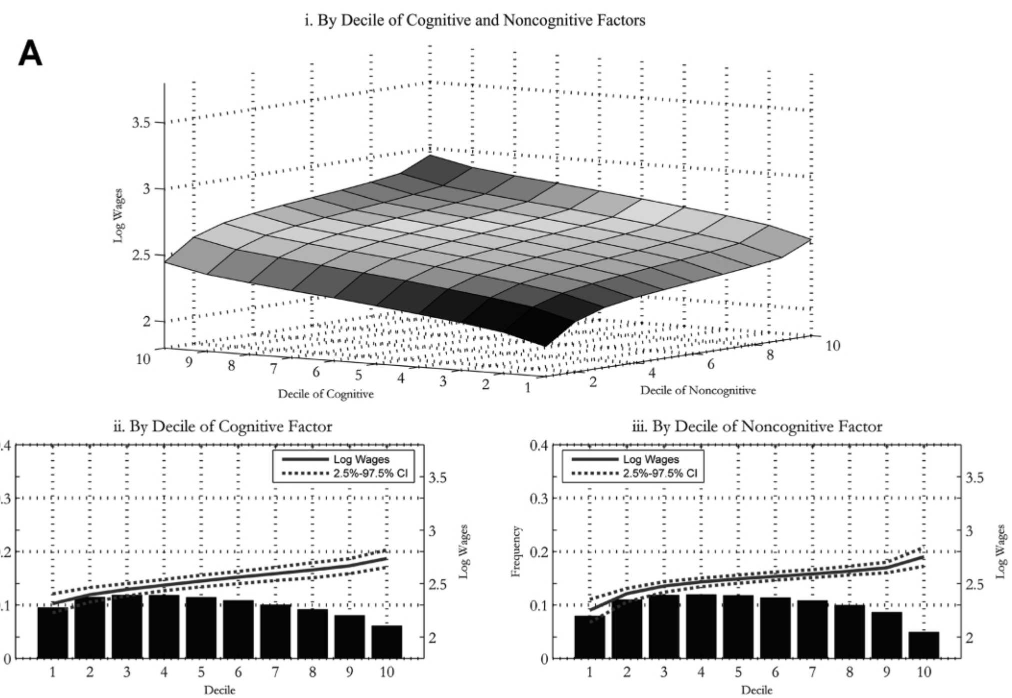


Fig. 8.—Mean log wages of GED recipients at age 30 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed 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.


A
i. By Decile of Cognitive and Noncognitive Factors
Log Wages
3.5
3
2.5
2
10
9
8
7
6
5
4
3
2
1
Decile of Cognitive
2
4
6
8
10
Decile of Noncognitive
ii. By Decile of Cognitive Factor
iii. By Decile of Noncognitive Factor
Log Wages
2.5%-97.5% CI
Frequency
0.4
0.3
0.2
0.1
0
Decile
1 2 3 4 5 6 7 8 9 10

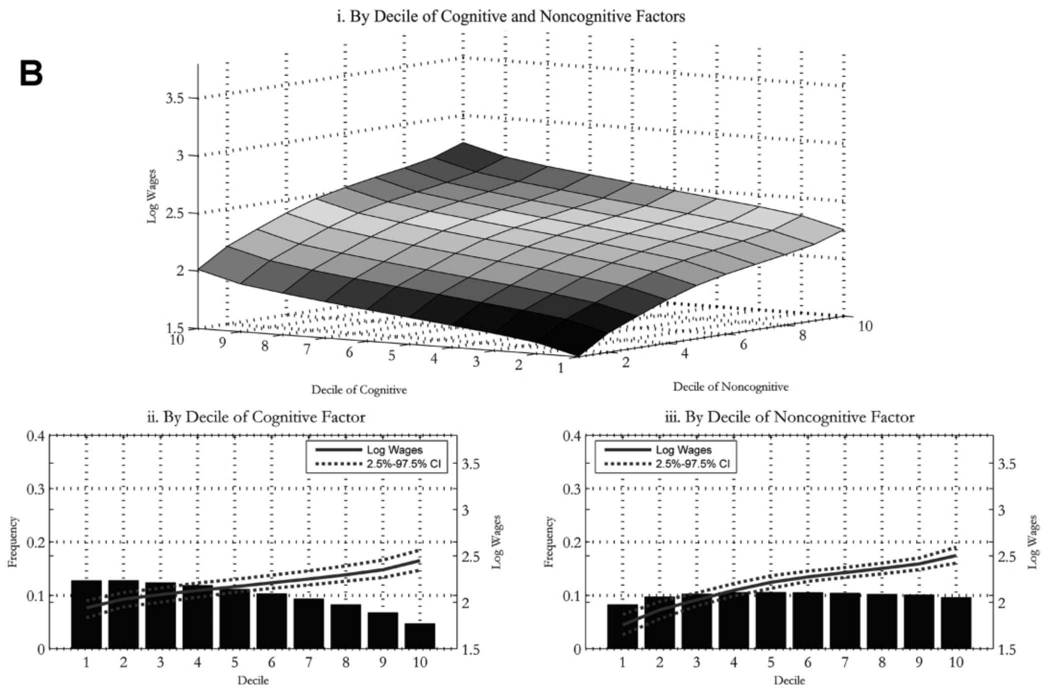
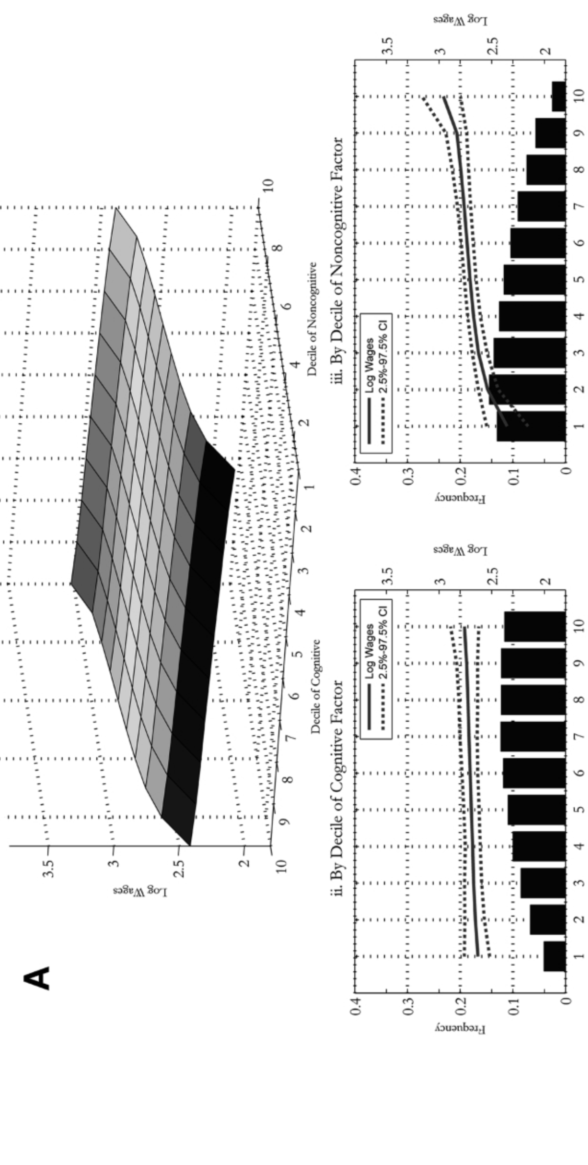


FIG. 9.—Mean log wages of high school graduates at age 30 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are 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.



A
i. By Decile of Cognitive and Noncognitive Factors
Log Wages
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
iii. By Decile of Noncognitive Factor
Log Wages
2.5%-97.5% CI
Frequency
Decile

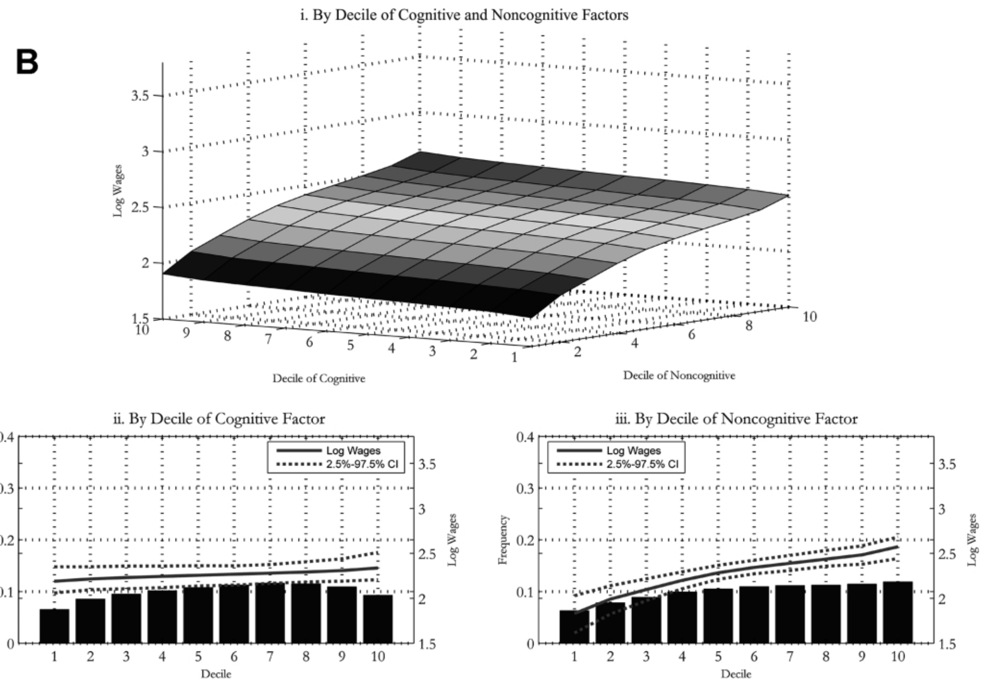
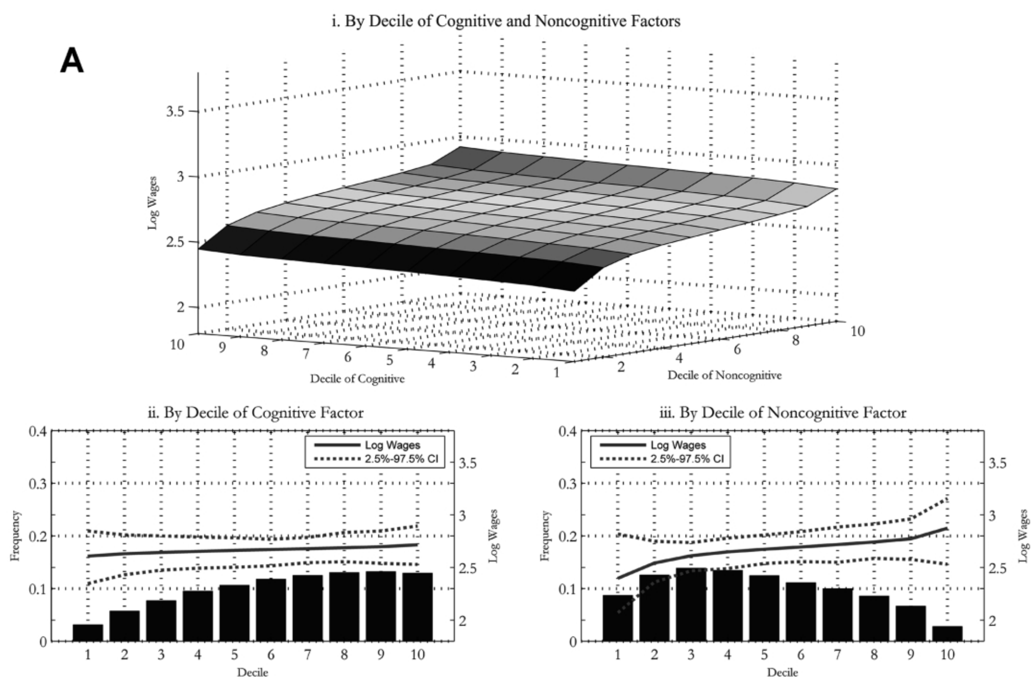


Fig. 10.—Mean log wages of some-college attenders at age 30 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are 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.


A
i. By Decile of Cognitive and Noncognitive Factors
Log Wages
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
iii. By Decile of Noncognitive Factor
Log Wages
2.5%-97.5% CI
Frequency
Decile

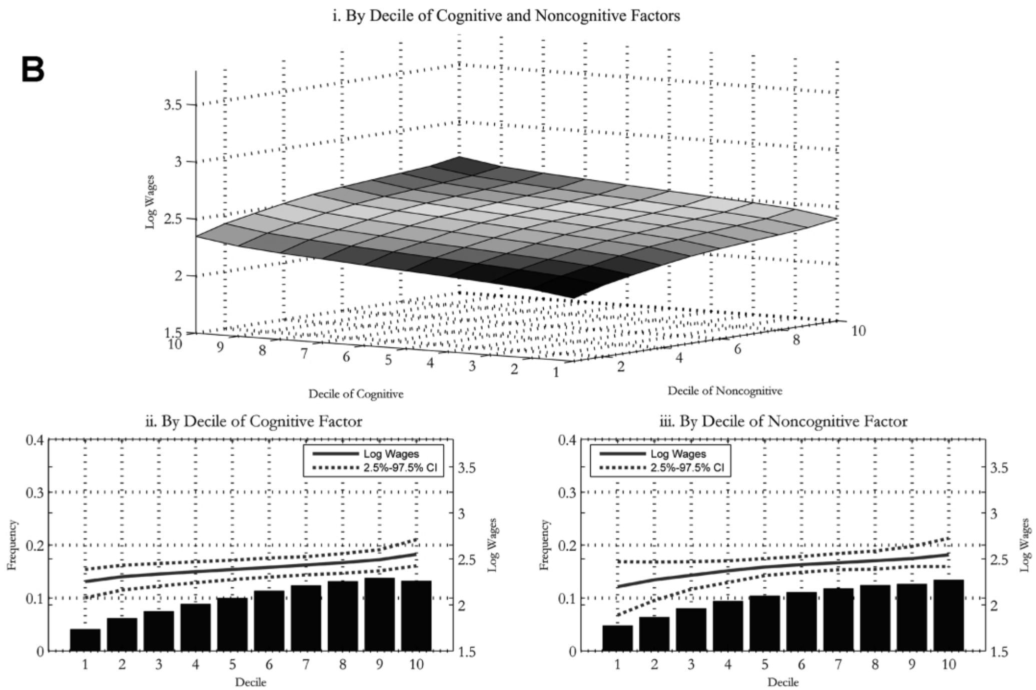
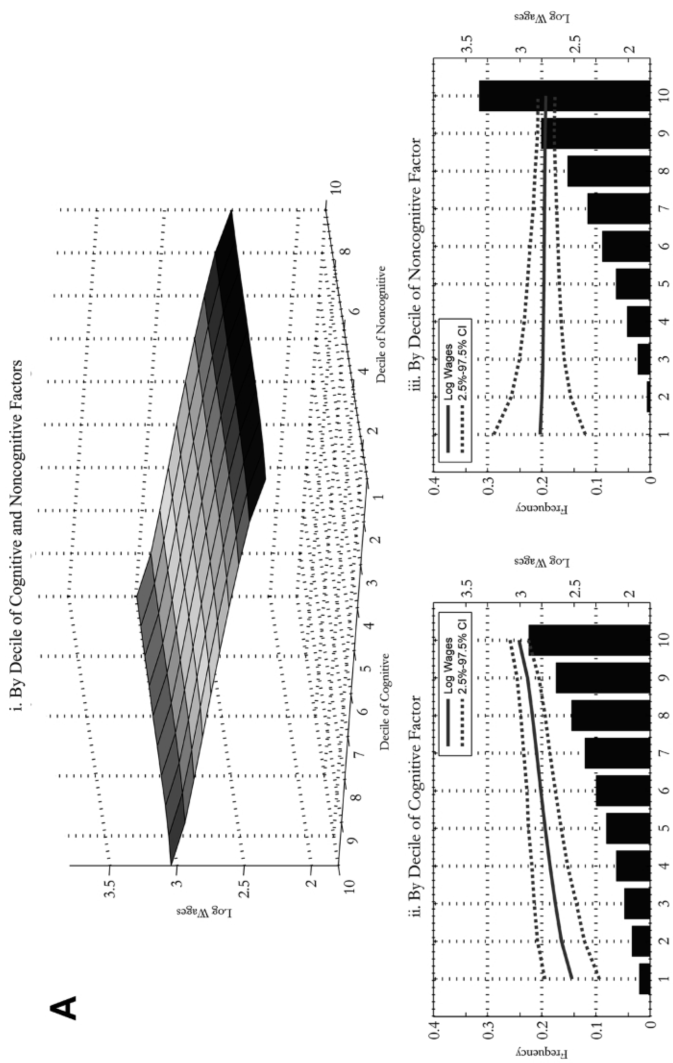


Fig. 11.—Mean log wages of 2-year-college graduates at age 30 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals 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.


A
i. By Decile of Cognitive and Noncognitive Factors
Log Wages
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
Log Wages
2.5%-97.5% CI
Frequency
Decile
iii. By Decile of Noncognitive Factor
Log Wages
2.5%-97.5% CI
Frequency
Decile

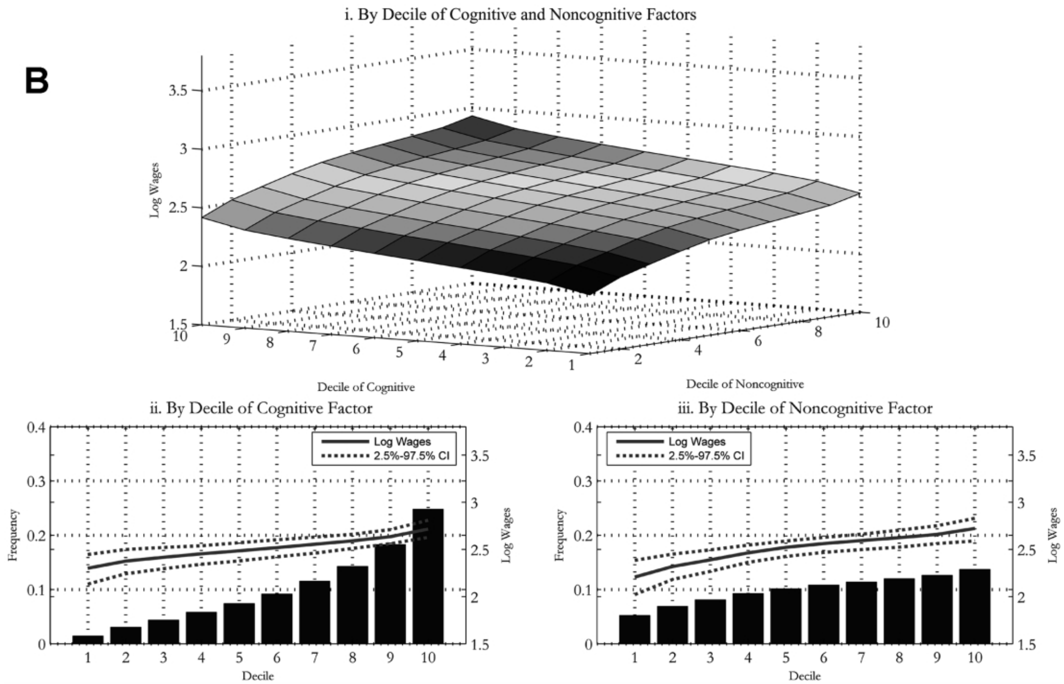
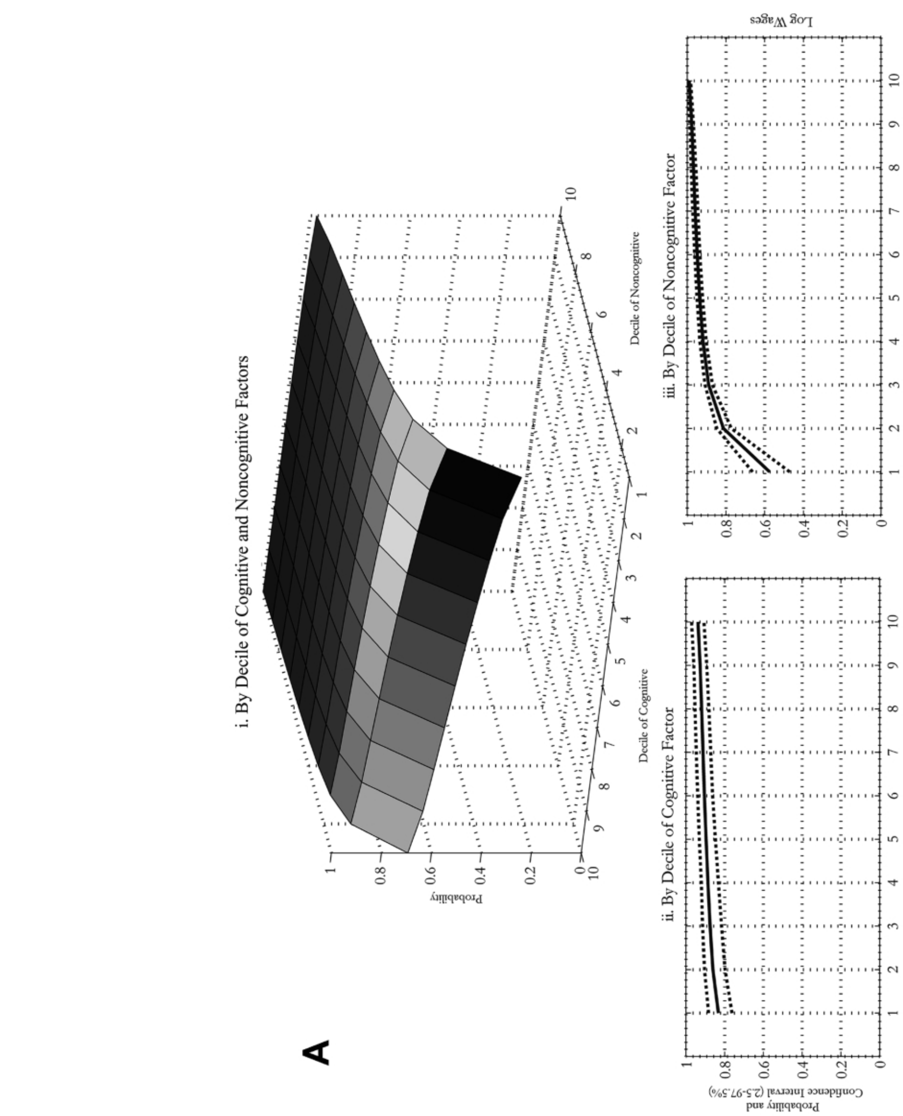


Fig. 12.—Mean log wages of 4-year-college graduates at age 30 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals 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.



A
i. By Decile of Cognitive and Noncognitive Factors
Probability
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
Probability and Confidence Interval (2.5-97.5%)
Decile
iii. By Decile of Noncognitive Factor
Decile
Log Wages

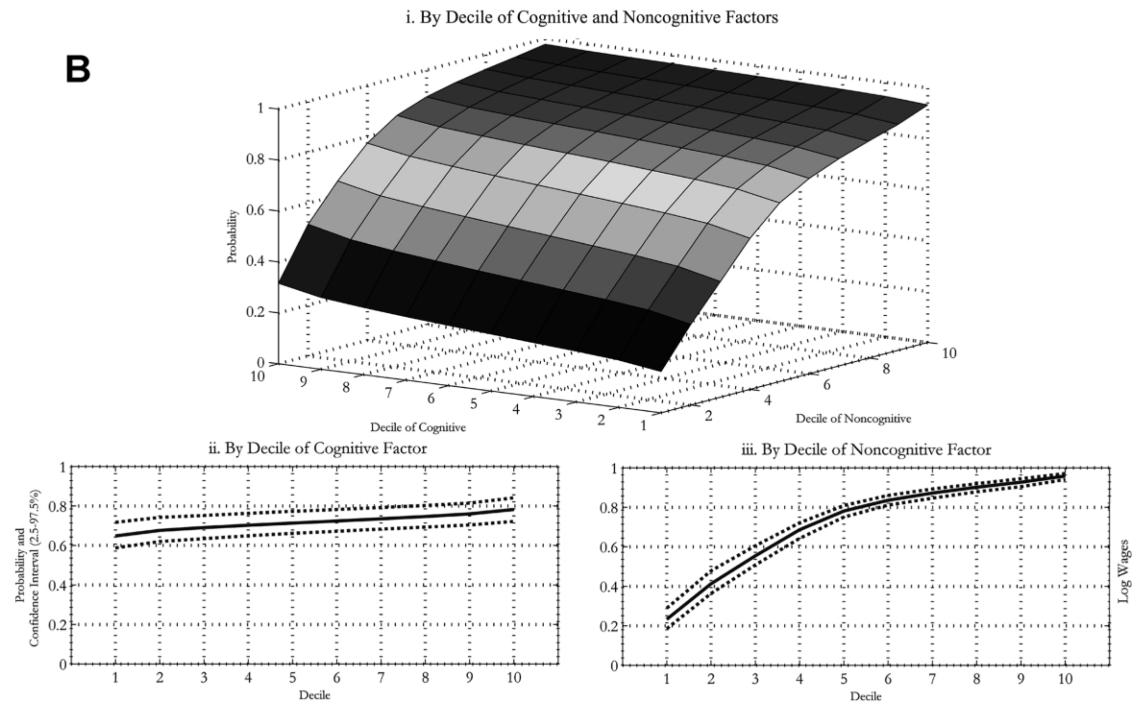
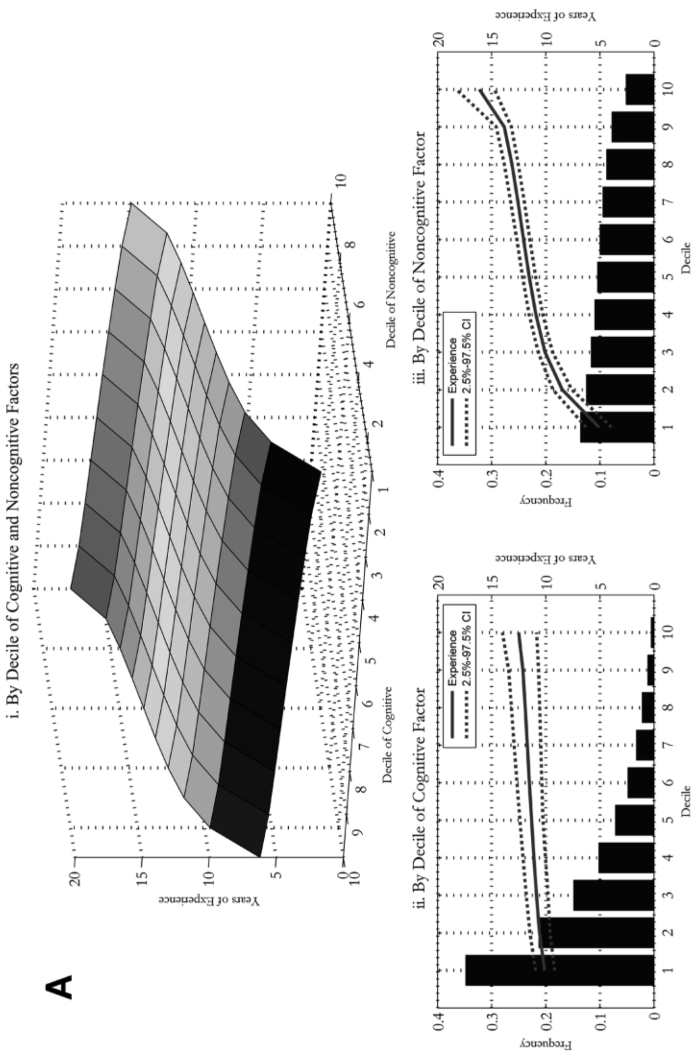


FIG. 13.—Probability of employment at age 30 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.



A
i. By Decile of Cognitive and Noncognitive Factors
Years of Experience
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
Experience
2.5%-97.5% CI
Frequency
Years of Experience
Decile
iii. By Decile of Noncognitive Factor
Experience
2.5%-97.5% CI
Frequency
Years of Experience
Decile

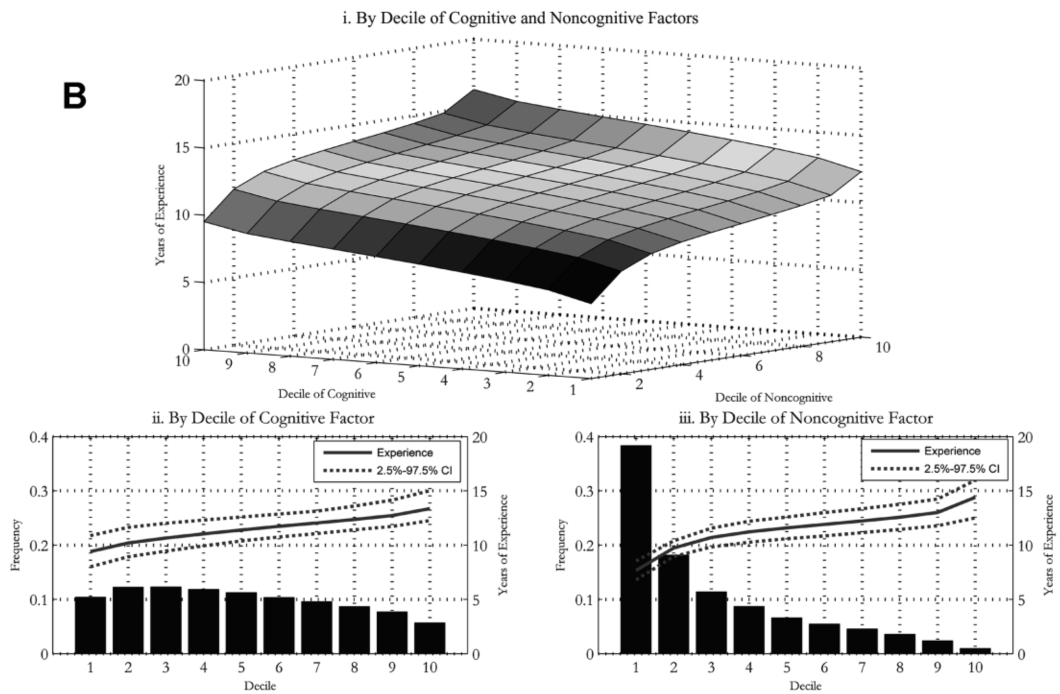
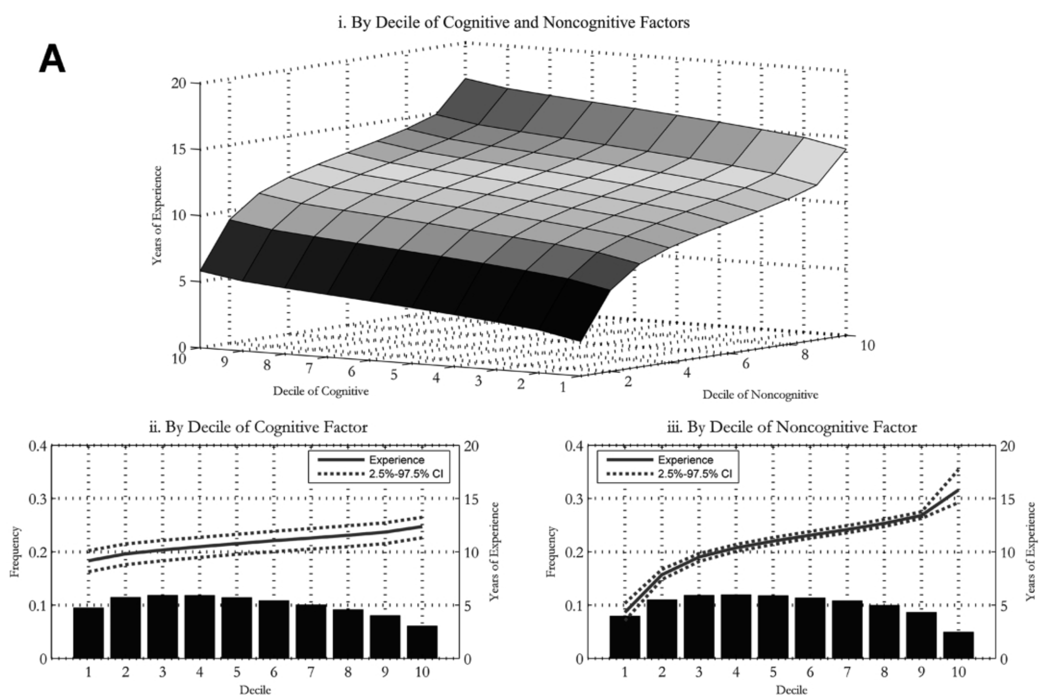


Fig. 14.—Mean work experience of high school dropouts at age 30 (*A*) and of GED recipients by age 30 (*B*) for males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The 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.


A
i. By Decile of Cognitive and Noncognitive Factors
Years of Experience
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
Experience
2.5%-97.5% CI
Frequency
Years of Experience
Decile
iii. By Decile of Noncognitive Factor
Experience
2.5%-97.5% CI
Frequency
Years of Experience
Decile

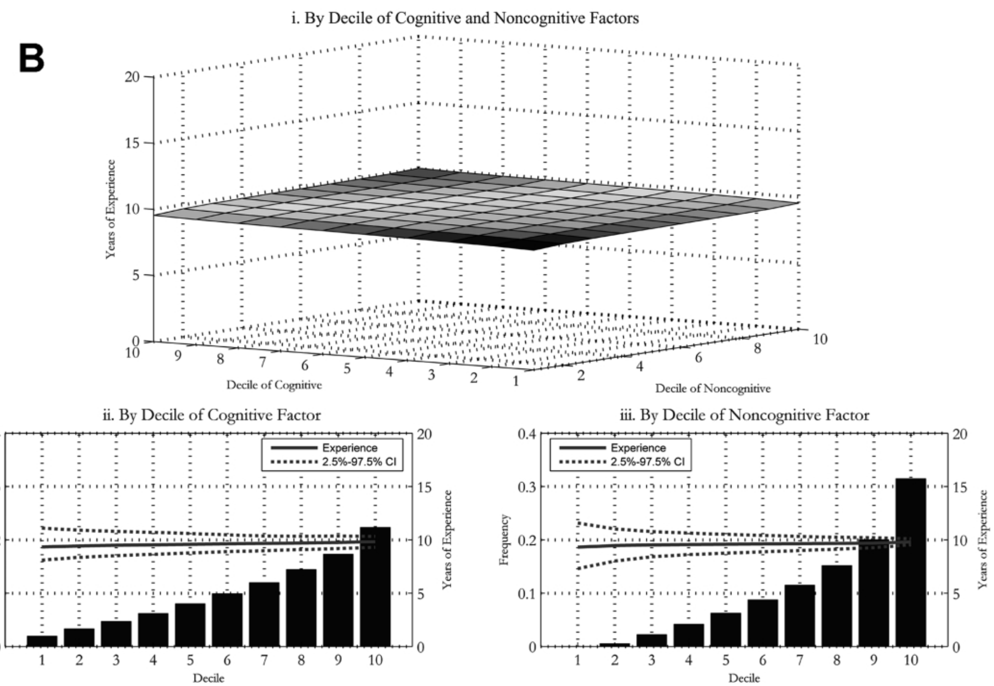
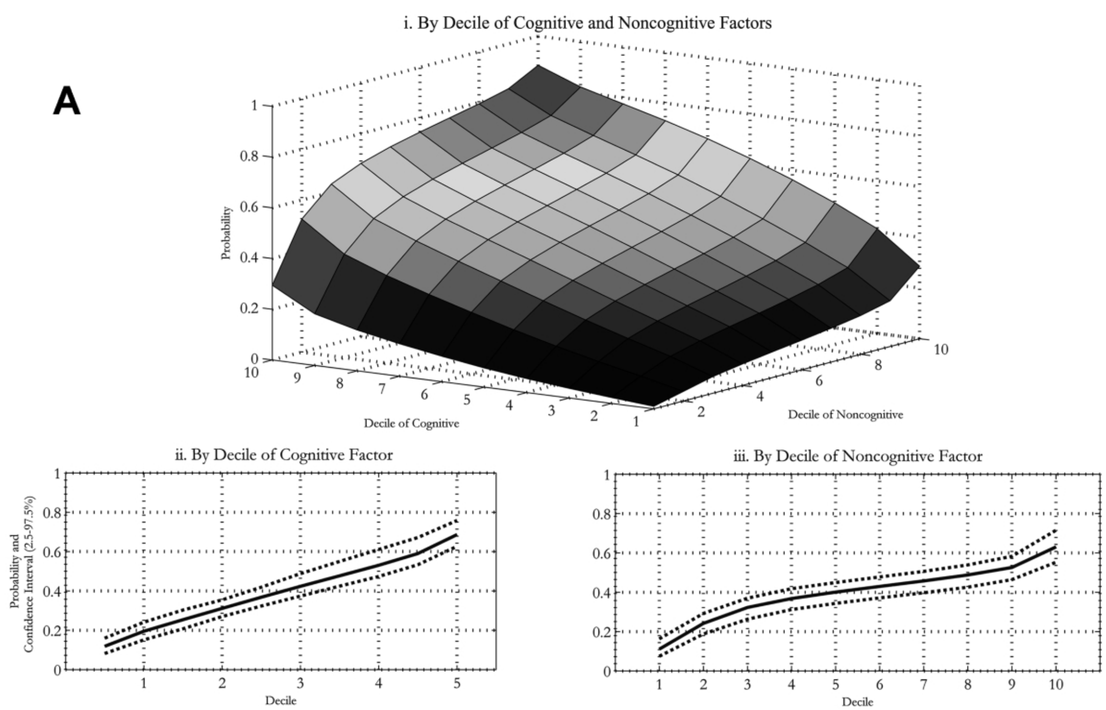


Fig. 15.—Mean work experience of high school graduates at age 30 (*A*) and of 4-year-college graduates by age 30 (*B*) for males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. 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.


A
i. By Decile of Cognitive and Noncognitive Factors
Probability
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
Probability and Confidence Interval (2.5-97.5%)
Decile
iii. By Decile of Noncognitive Factor
Decile

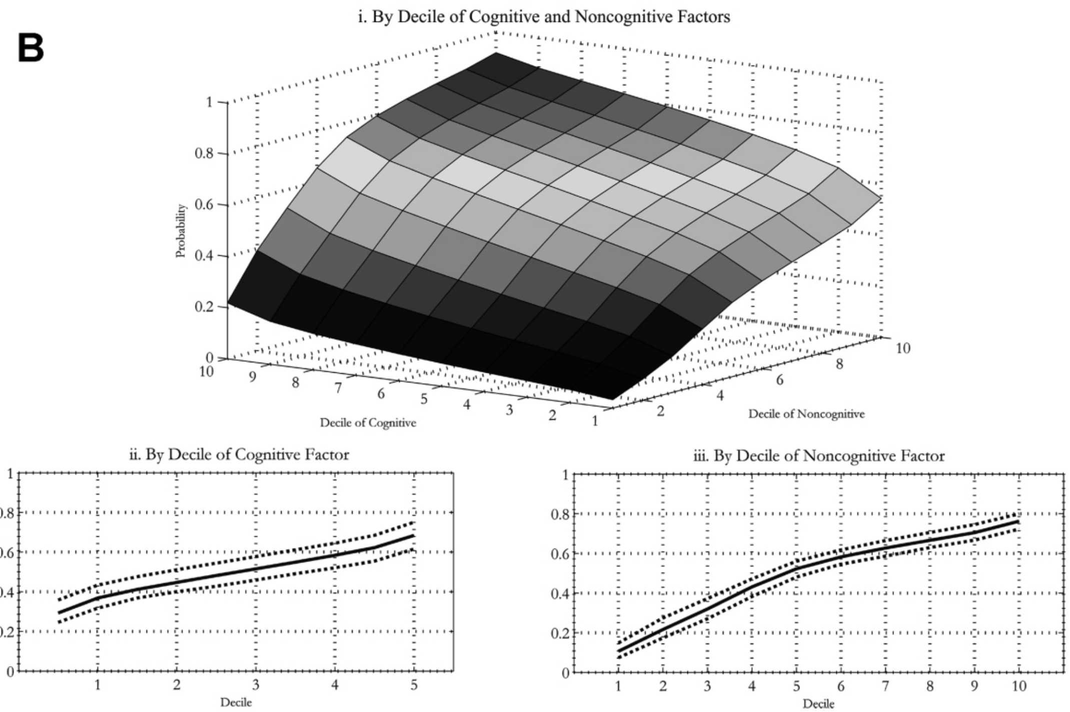


Fig. 16.—Probability of being a white-collar worker at age 30 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

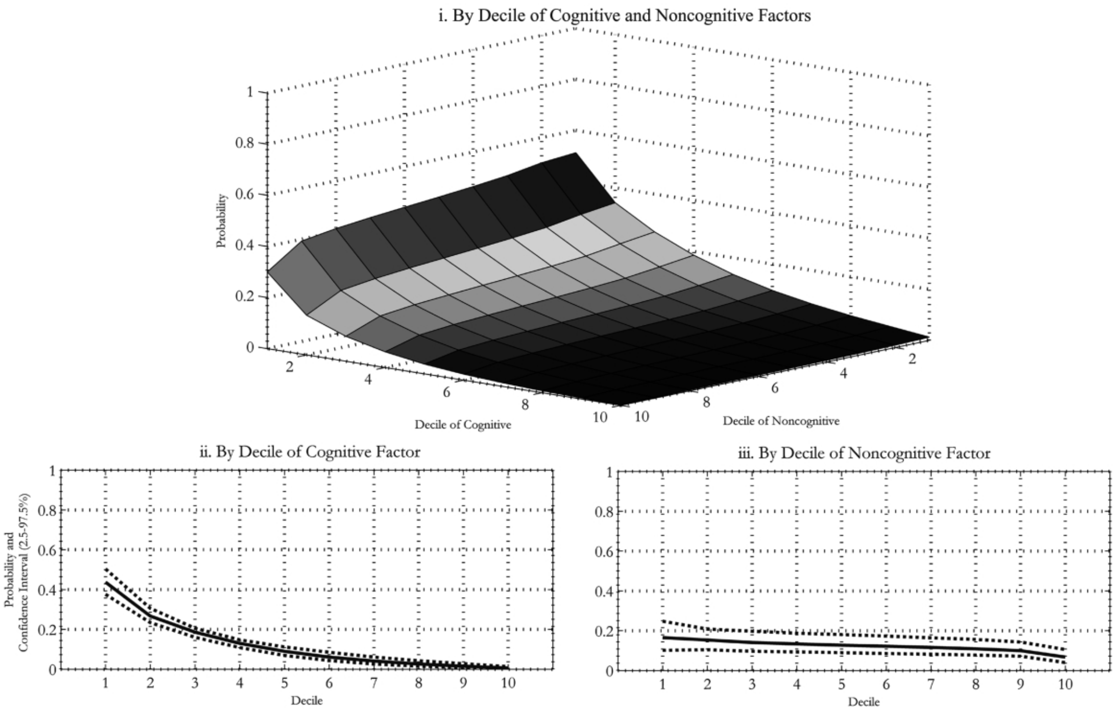


FIG. 17.—Probability of being a high school dropout at age 30, males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

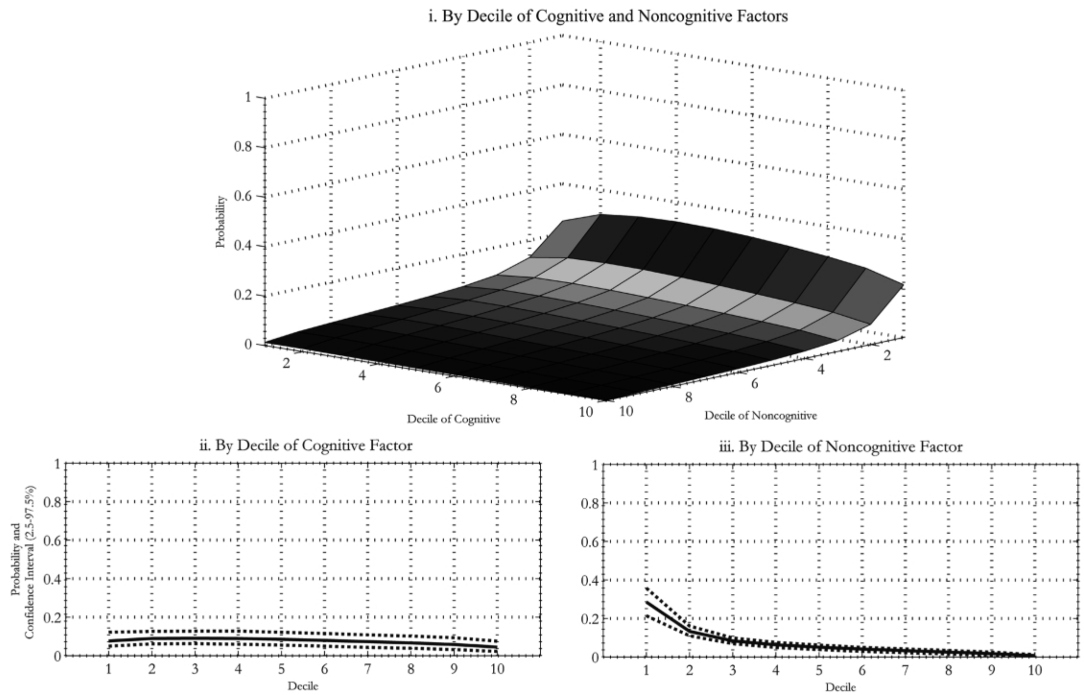


Fig. 18.—Probability of being a GED recipient at age 30, males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

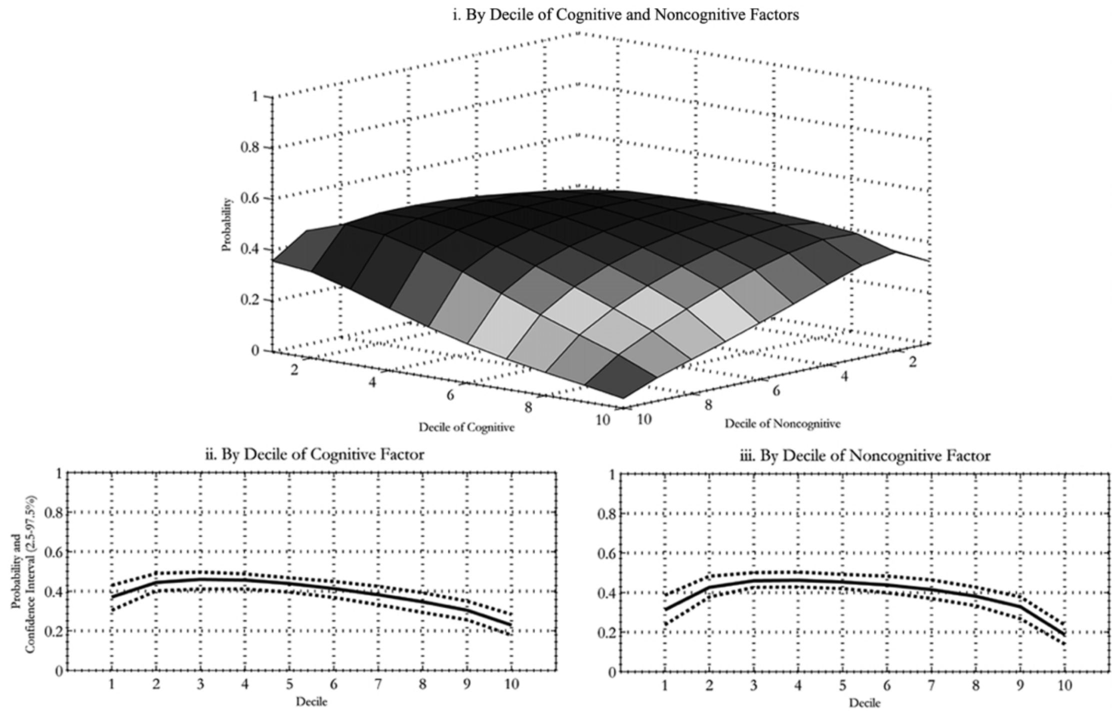


Fig. 19.—Probability of being a high school graduate at age 30, males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

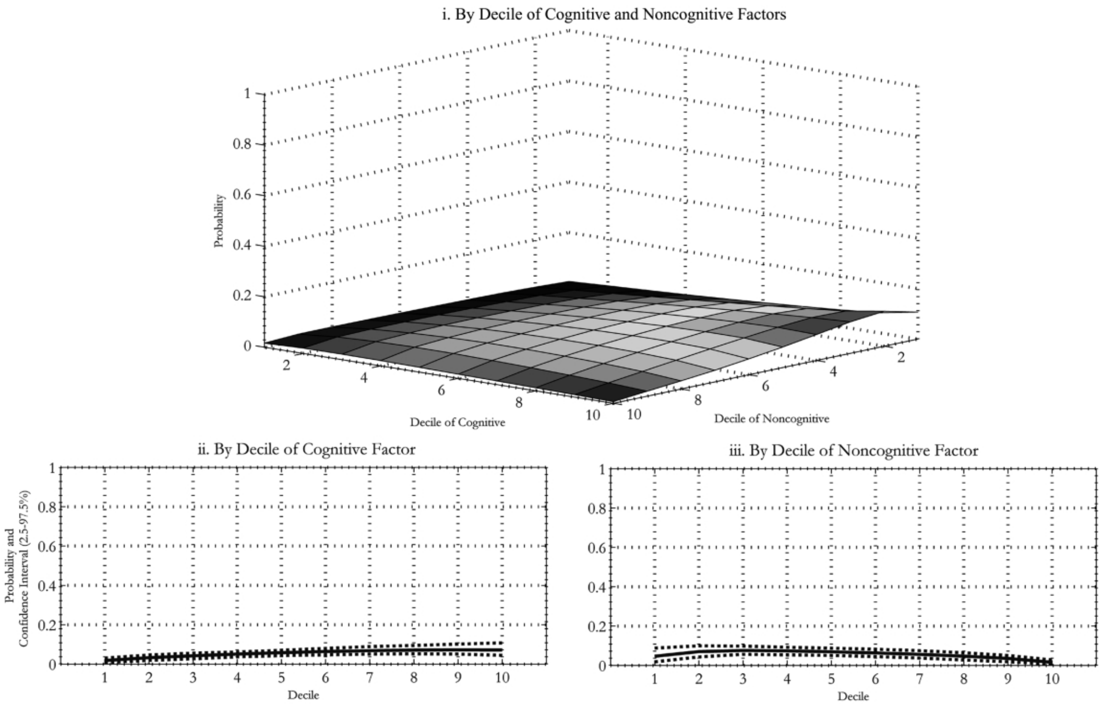


Fig. 20.—Probability of being a 2-year-college graduate at age 30, males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

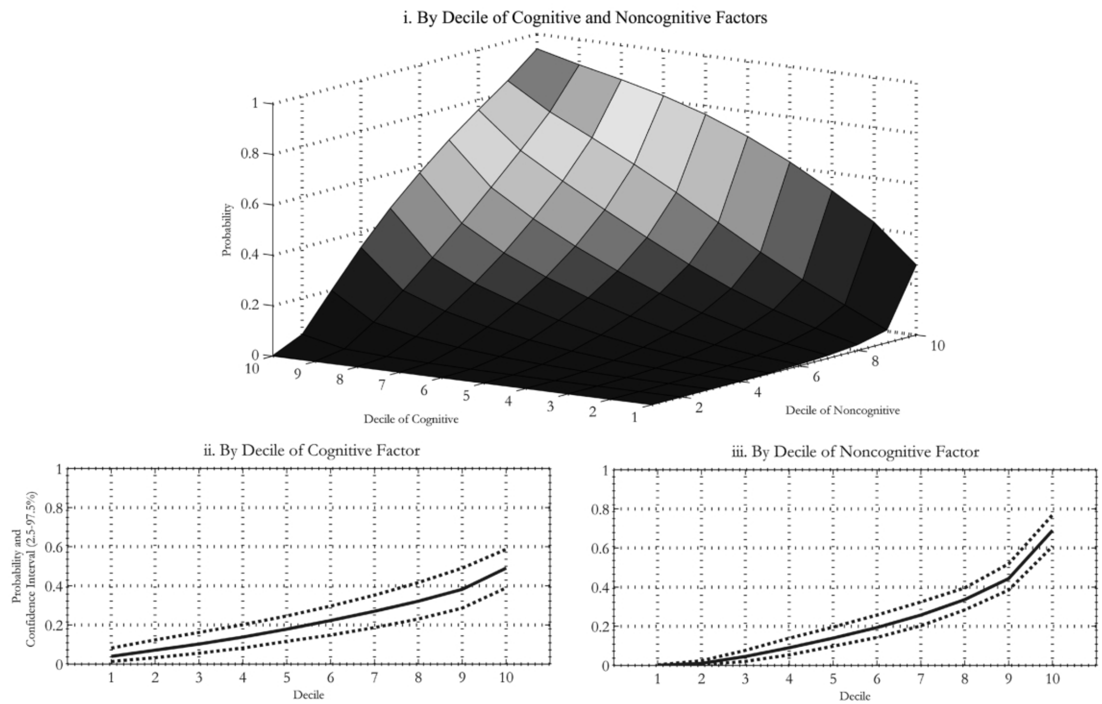


FIG. 21.—Probability of being a 4-year-college graduate at age 30, males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

We also consider the effects of cognitive and noncognitive abilities on participation in illegal activities. These results are displayed in figure 25. Again, noncognitive abilities have much stronger effects in the sense of having a steeper gradient. For women (see our Web appendix fig. S14), both gradients are essentially zero.

Although both factors are important determinants of marital status and pregnancy by age 18, changing the noncognitive factor has greater effects on behavior. Figure 26 shows the effects of both types of latent abilities on being single with no child by age 18. Changes in the cognitive factor are important but have weaker effects than changes in the noncognitive factor. This evidence illustrates the importance of noncognitive skills in explaining the chances of a woman being single with no child. The probability of being a teenage mother is equally responsive to changes in cognitive and noncognitive skills (see fig. 27). At the highest levels of cognitive and noncognitive skills, the probability of teenage pregnancy is essentially zero.

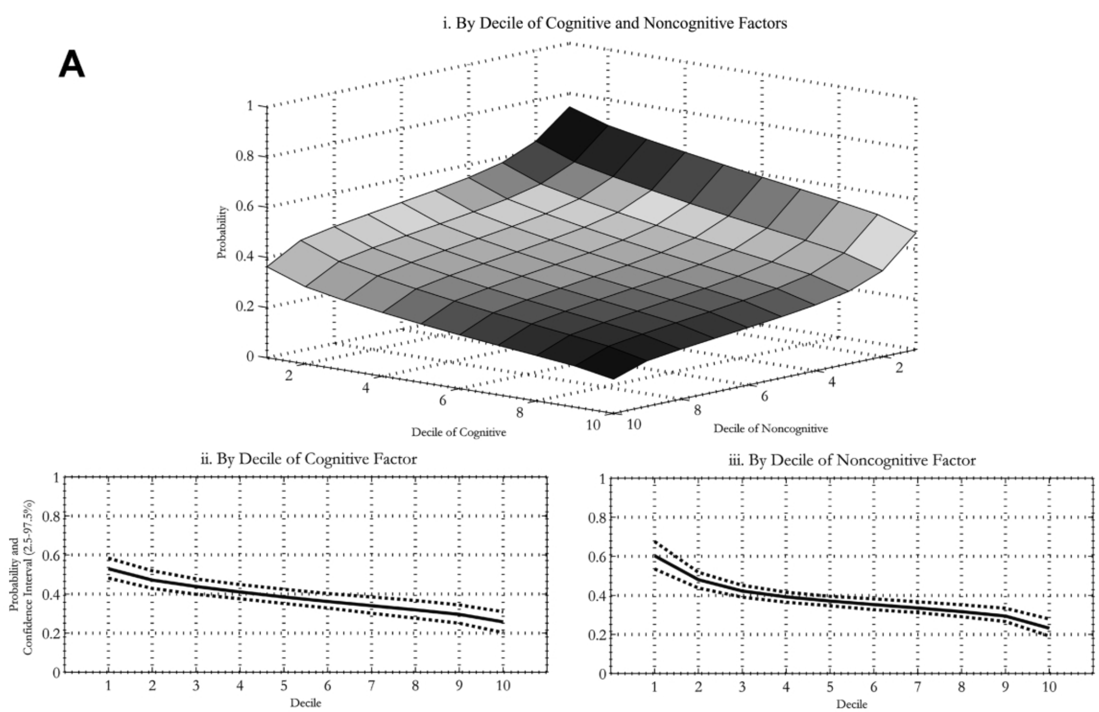
We use Children of NLSY79 (CNLSY79) data to corroborate some of the findings reported in this article. One potential advantage of these data is that they contain very early (age 3–6) measurements of both cognitive and noncognitive abilities. Such measurements are not affected by later schooling. A disadvantage of these data is that many of the children are still young and we lack information on their wages, occupational status, and employment at age 30. In addition, the samples are small. The evidence from the CNLSY79 data is broadly consistent with the evidence reported in this article, but the parameters are much less precisely estimated. See table S25 in our Web appendix.[31]

Two latent factors associated with cognitive and noncognitive skills explain a wide array of teenage and young adult behaviors. Noncognitive abilities play a major role in explaining these behaviors, and they are valued as direct determinants of wages in most educational labor markets.

VIII. Relationship of Our Work to Previous Research

Early work by Bowles and Gintis (1976) presents evidence suggesting that employers in low-skill markets value docility, dependability, and persistence more than cognitive skills. In a similar vein, Edwards (1976) shows

[31] There is an additional problem with these data. Both cognitive and noncognitive abilities change with age. Cunha and Heckman (2006b) model the evolution of both cognitive and noncognitive skills over the life cycle. Even IQ is not stable before age 8 (see Cunha, Heckman, Lochner, and Masterov 2006). Let a_t be ability at age t. If $a_t = \lambda a_{t-1} + b_t + \varepsilon_t$, where b_t is a growth trend and ε_t is an independent and identically distributed innovation, early measurement of a_t may be a poor approximation for the later measurement used in this article. Thus, while use of early measurements circumvents the problem of reverse causality, it creates a measurement error problem because $a_{t'}(t' < t)$ is not the same as a_t.


A
i. By Decile of Cognitive and Noncognitive Factors
Probability
Decile of Cognitive
Decile of Noncognitive
ii. By Decile of Cognitive Factor
Probability and Confidence Interval (2.5-97.5%)
Decile
iii. By Decile of Noncognitive Factor
Decile

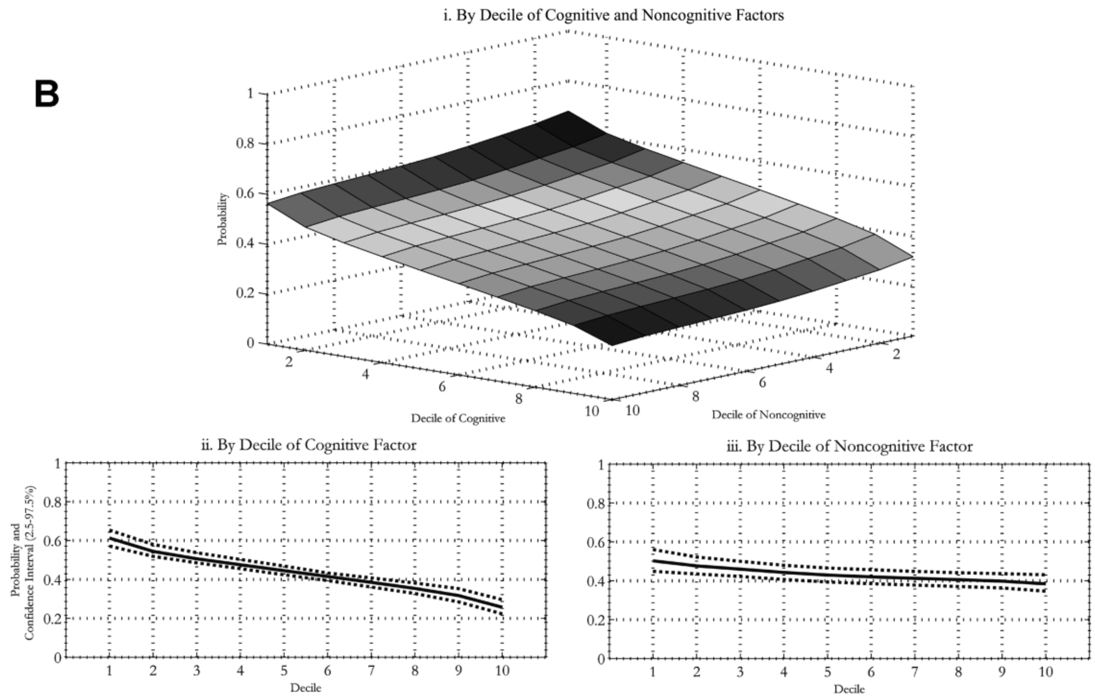


FIG. 22.—Probability of daily smoking by age 18 for males (*A*) and females (*B*). The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

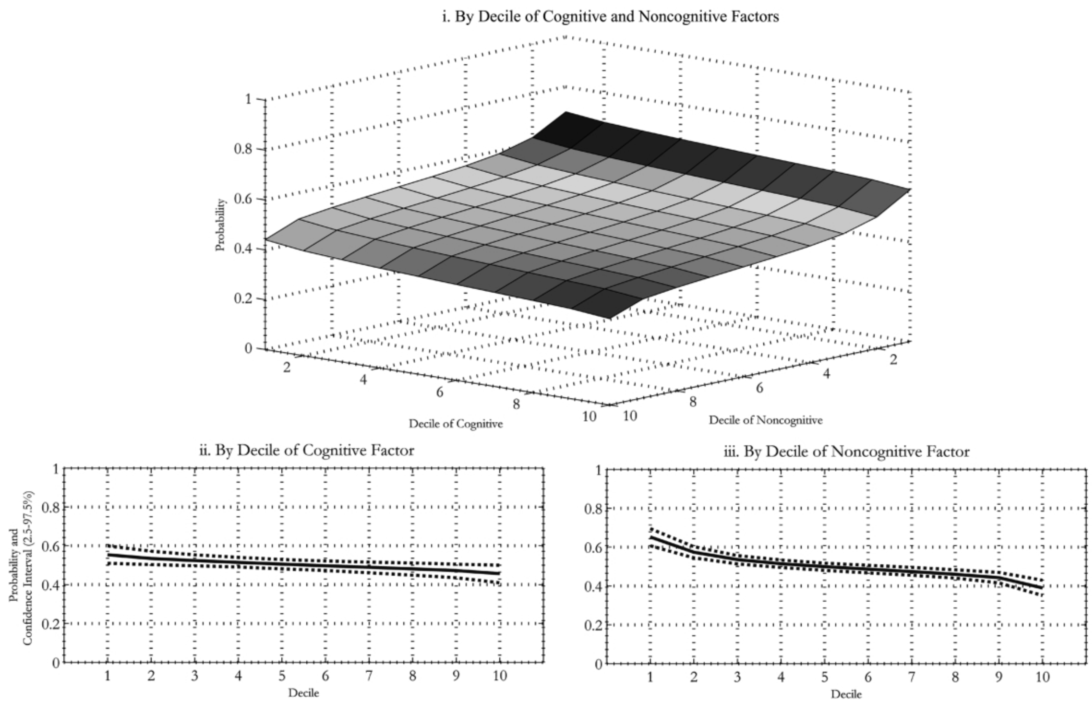


FIG. 23.—Probability of smoking marijuana during 1979, males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

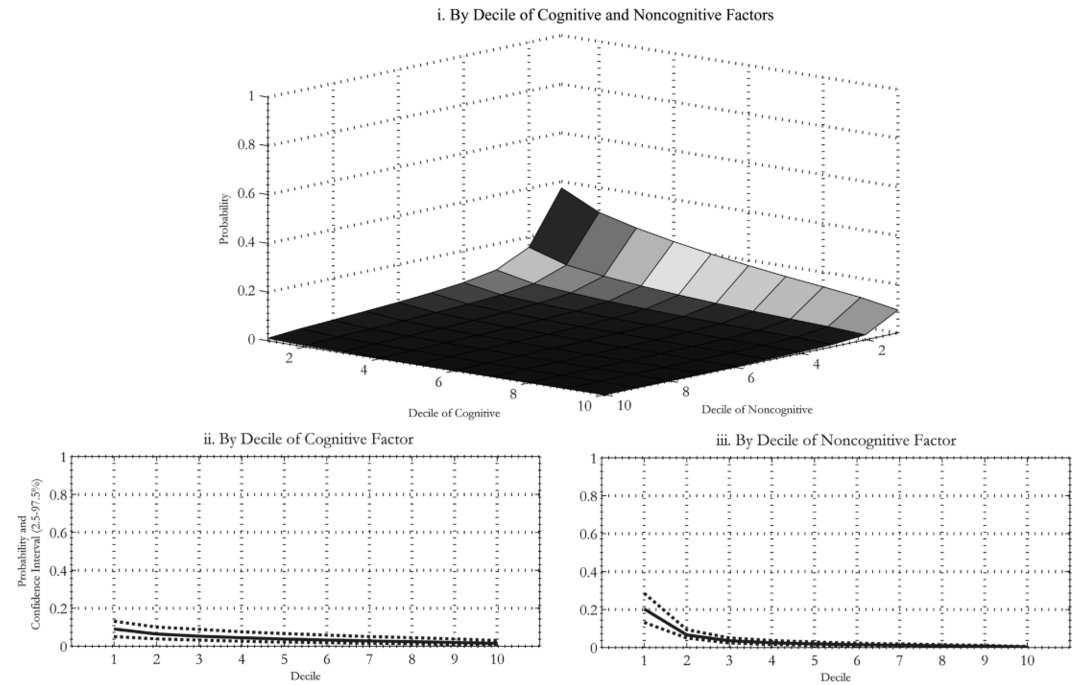


FIG. 24.—Probability of incarceration by age 30, males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

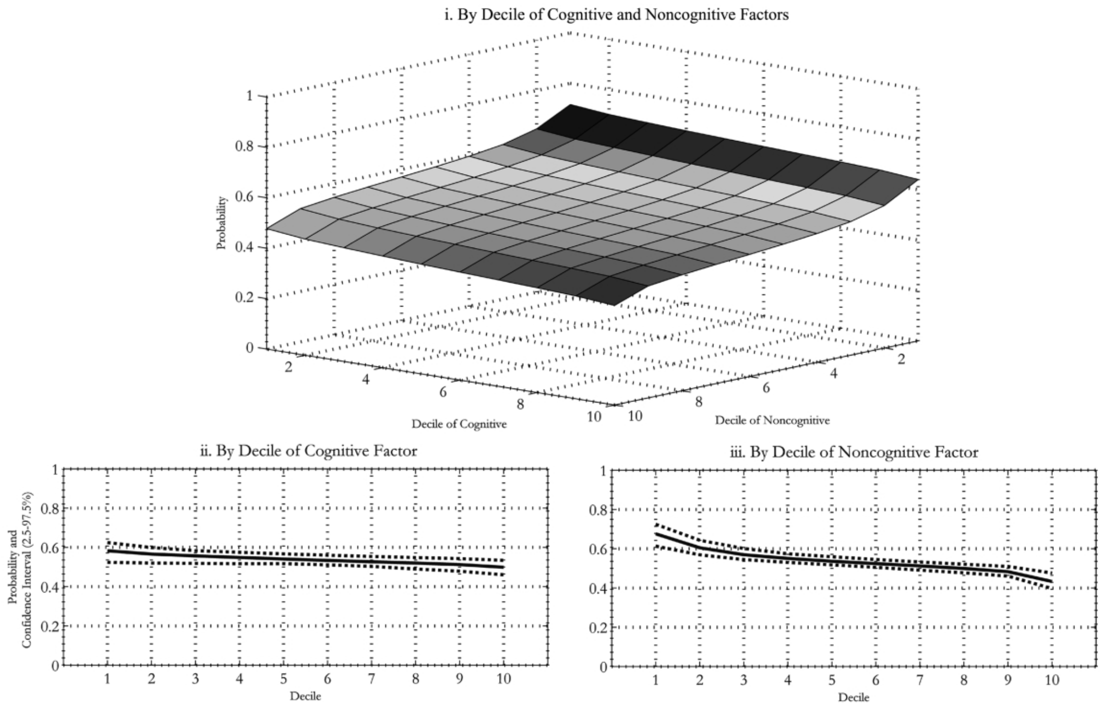


FIG. 25.—Probability of participating in illegal activities during 1979, males. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

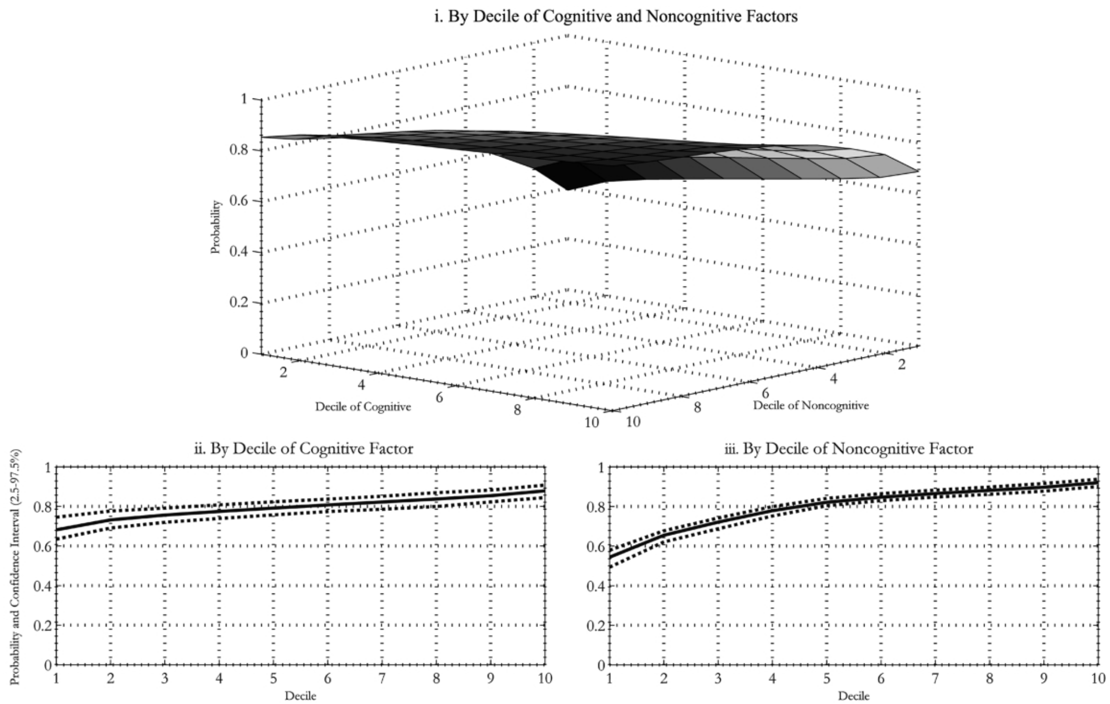


Fig. 26.—Probability of being single with no child at age 18, females. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

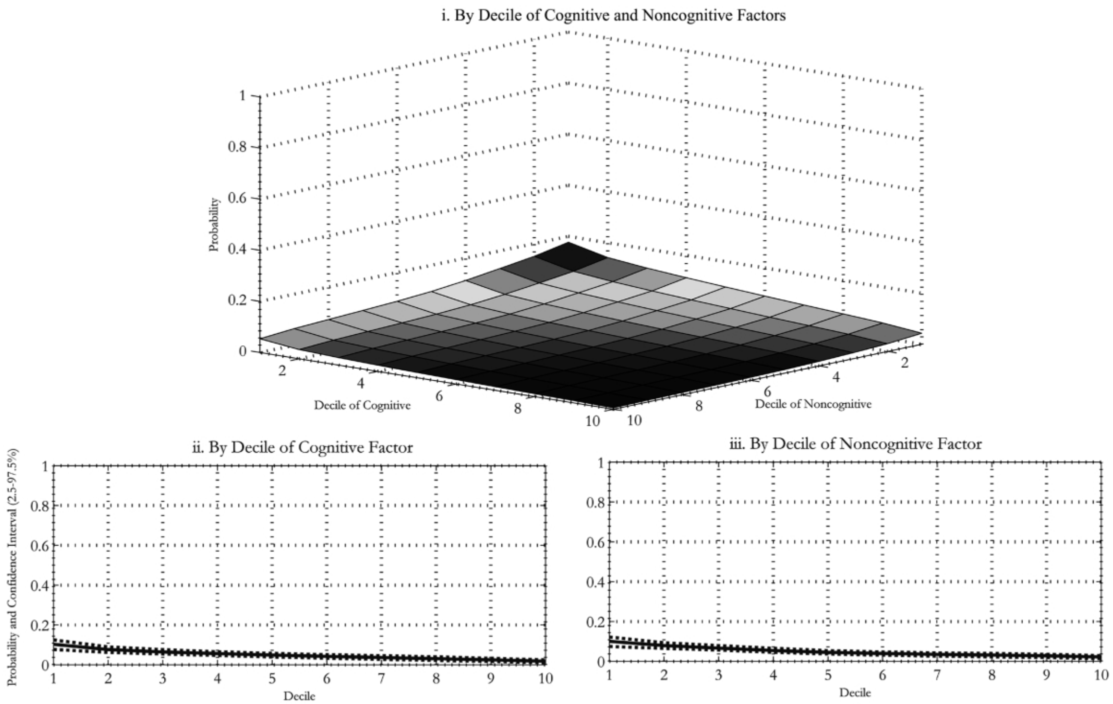


FIG. 27.—Probability of being single with child at age 18, females. The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals.

that dependability and consistency are more valued by blue-collar supervisors than are cognitive ability or independent thought. Klein, Spady, and Weiss (1991) document that the premium accorded high school graduates compared to high school dropouts in semiskilled and skilled occupations is due primarily to the higher level of job stability (lower quit rates) and dependability (lower absenteeism) of high school graduates and not their greater productivity in final output. However, they do not present estimates of the effects of noncognitive skills on wages. Peter Mueser, writing in chapter 5 of Jencks (1979), uses least squares to find that skills such as industriousness, perseverance, and leadership have statistically significant influences on wages—comparable to estimated effects of schooling, IQ, and parental socioeconomic status—even after controlling for standard human-capital variables.

In more recent work, Osborne-Groves (2006) studies the effect of personality and behavioral traits on the wages of females. Using two data sets and alternative instruments for adult personality measures, she finds that personality traits such as fatalism, aggression, and withdrawal have significantly negative effects on wages. She does not control for the effect of schooling on the measurements she uses.[32] Bowles et al. (2001) present a model in which incentive-enhancing preferences that allow employers to induce greater effort at a lower cost (e.g., a low time-discount rate, a high degree of self-directedness and personal efficacy, a low disutility of effort, and a tendency of being helpful toward other employees) are rewarded in a competitive labor market in the form of increased wages. Our evidence supports their analysis because noncognitive traits raise wages in most labor markets for schooling of different levels.

Heckman and Rubinstein (2001) use evidence from the GED testing program to demonstrate the quantitative importance of noncognitive skills. The GED recipients have the same cognitive ability as high school graduates who do not go to college, as measured by the AFQT score. However, once cognitive ability is controlled for, GED recipients have the same or lower hourly wages as those of high school dropouts. This pattern would be predicted by our model because GED recipients have lower noncognitive skills than dropouts (see fig. 3), and hence are less likely to be employed and to acquire work experience, and also have lower levels of the noncognitive characteristic valued in the labor market.

Darity, Goldsmith, and Veum (1997) use the NLSY79 to estimate the effect of Rosenberg and Rotter scales on wages. They control for the endogeneity of the test scores in wage equations using an instrumental variables procedure, but they do not correct for the endogeneity of schooling in the wage equation, nor do they estimate the distributions of latent ability. However, their reported estimates qualitatively agree with ours.

[32] Her instruments include lagged wages and so are suspect.

Elasticities of wage equations with respect to predicted noncognitive test scores are of comparable magnitudes as elasticities of wage equations with respect to conventional human-capital variables.

It is quite instructive to compare our results on the effects of cognitive and noncognitive skills on wages to results from conventional approaches. As discussed in Section II, the standard approach of regressing wages on measured test scores suffers from several problems. In that approach, wages are typically regressed on cognitive and noncognitive test scores, schooling dummies, and a set of other controls (as in cols. 1 and 3 in table 1). This approach is problematic because schooling is a choice variable and schooling choices depend on cognitive and noncognitive skills, as we have shown. Removing the schooling dummies from the equation avoids this source of endogeneity problems, but changes the parameters being estimated to the net effects of these skills on wages (as in cols. 2 and 4 in table 1). These net effects do not isolate the effect of ability on wages holding schooling fixed. Furthermore, since schooling at the time of the test affects test scores, test scores are still endogenous in the wage equation. Finally, there is the problem of measurement error. Test scores are imperfect proxies for latent cognitive and noncognitive abilities because they are affected by measured characteristics such as family background. Because these problems likely bias the estimates in different directions, we cannot predict whether OLS estimates will be higher or lower than those produced from our model.

We simulate our model, given the exogenous conditioning variables, to predict the test score that each individual would have received had he been in grades 9–11 at the time he took the test. We simulate these test scores by drawing the factors from the population distribution. Using these corrected test scores in an OLS wage regression alleviates the problem that schooling affects the test score, but measurement error remains because the test score is not the same as the factor it proxies. Table 5 displays the (standardized) coefficients from OLS on measured test scores (cols. 1 and 4), OLS using corrected test scores (cols. 2 and 5), and OLS using the simulated factors.[33] Estimated returns to ability are typically much smaller using corrected rather than actual test scores. Endogeneity of test scores (and reverse causality in the regressions that are not run separately by schooling level) produces estimates that are generally upward biased. The OLS estimates using the simulated factor as the measure of ability are typically much larger than those using corrected test scores. Measurement error causes a significant downward bias that is typically larger than the upward bias due to endogeneity and reverse causality. This

[33] The standardized coefficient is obtained by multiplying the original coefficient by the standard deviation of the variable to which it is associated.

Table 5
Standardized OLS Coefficients of Cognitive and Noncognitive Skills from Log Hourly Wage Regressions with Different Skill Measures: Measured Test Scores ([1] and [4]), Corrected Test Scores ([2] and [5]), and Latent Abilities Factors ([3] and [6])

	Cognitive Ability			Noncognitive Ability		
Schooling Level	(1)	(2)	(3)	(4)	(5)	(6)
High school dropout	.047	.019	.039	.072	.023	.133
GED	.074	.068	.101	.018	.056	.157
High school graduate	.087	.064	.102	.035	.016	.113
Some college, no degree	−.018	−.012	.024	.041	.030	.131
2-year-college degree	.040	−.047	.010	.022	.056	.133
4-year-college degree	.120	.113	.124	.054	−.002	−.005
Overall:						
Including schooling dummies	.107	.066	.097	.043	.021	.112
Excluding schooling dummies	.177	.143	.134	.055	.043	.135

NOTE.—Sample from the NLSY79, males at age 30. The standardized coefficient is obtained by multiplying the original coefficient by the standard deviation of the variable to which it is associated. This allows us to make comparisons across columns. All columns display the coefficient on cognitive or noncognitive ability as measured by either an observed test score, a corrected test score, or the latent factor in a regression of log hourly wages on the measures and a full set of controls (black and Hispanic dummy variables, a set of cohort dummies, local labor market conditions [unemployment rate], and variables controlling for characteristics of the regions of residence) by schooling level at age 30 and overall. The values of the latent abilities for each individual (which of course are not available in the NLSY79 sample) are required to obtain the estimates in cols. 2, 3, 5, and 6. We therefore simulate a sample of 14,400 individuals from our structural model that combines wages and observable controls (including the measured cognitive and noncognitive test scores) from the NLSY79 data and draws of the latent factors. This simulated sample is used to obtain all of the regression estimates displayed in this table. In cols. 1 and 4 the cognitive measure is the standardized sum of scores on the arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge, and coding speed components of the ASVAB, and the noncognitive measure is the standardized sum of scores on the Rotter Locus of Control and Rosenberg Self-Esteem scales. In cols. 2 and 5 the above cognitive and noncognitive test scores are replaced by corrected test scores. These corrected test scores are obtained by using our structural model to predict test scores for each individual had they been in grades 9–11 at the time of the test. For each individual in this sample we use the parameters from the estimated test-score equations for 9–11 years of schooling at the time of the test, the individual's observable controls, and two latent abilities drawn from the estimated factor distributions to construct corrected test scores. The same method is used to construct the individual wages that are used in this regression. The estimates in cols. 3 and 6 are obtained by using simulated latent abilities (the same values used to construct the corrected test scores and wages in cols. 2 and 5) instead of test scores as cognitive and noncognitive measures in the wage regressions. As before, the wages in this regression are simulated from our model.

effect is especially pronounced for estimates of the effects of noncognitive skills on wages.

IX. Conclusion

This article presents new evidence that both cognitive and noncognitive abilities determine social and economic success. For many dimensions of behavior and for the sense of "importance" adopted in this article, noncognitive ability is as important, if not more important, than cognitive ability. Our findings challenge a pervasive view in the literatures in economics and psychology that cognitive ability, as measured by test scores, plays a dominant role in explaining personal achievement. Although cognitive skills explain much more of the variance of (log) wages, their effects on (log) wages (as measured by skill gradients) are similar to the effects of the noncognitive traits. In fact, noncognitive skills are about equally strong in many outcomes and are stronger for some outcomes. Of course, equal strength, in the sense we have used it, does not translate into an equal cost of changing these skills.

A low-dimensional model of cognitive and noncognitive abilities explains a diverse array of outcomes. It explains correlated risky behaviors among youth. Noncognitive ability affects the acquisition of skills, productivity in the market, and a variety of behaviors. Cognitive ability affects market productivity, skill acquisition, and a variety of behaviors. Schooling raises measured cognitive ability and measured noncognitive ability.

Our evidence is consistent with an emerging body of literature that establishes the importance of psychic costs in explaining why many students do not continue their schooling, even though it is financially rewarding for them to do so. Cunha, Heckman, and Navarro (2005, 2006) and Cunha and Heckman (2006a, 2006c) establish that these costs are related to cognitive ability. Our evidence suggests that noncognitive ability—motivation, persistence, and self-esteem—also plays a substantial role, but we have not, in this article, linked our measures of noncognitive ability to conventional measures of time preference, risk aversion, and preferences for leisure.

Our evidence that multiple abilities determine schooling challenges the conventional single-skill signaling model due to Arrow (1973) and Spence (1973). A special challenge is the GED program, where the credential (the GED test) conveys multiple conflicting signals. The GED recipients are smarter than other high school dropouts, but they have lower noncognitive skills. This violates the standard single-crossing property used in conventional signaling theory and requires a substantial reformulation of that theory (see Araujo et al. 2004).

Our demonstration that noncognitive skills are important in explaining a diverse array of behaviors helps to explain why early childhood programs, such as Headstart and the Perry Preschool Program, are effective. The evidence from these programs indicates that they do not boost IQ, but they raise noncognitive skills and therefore promote success in social and economic life. Our evidence of gender differentials in the effects of noncognitive skills on certain behaviors goes part way in explaining the gender differentials found in the Perry Preschool program (Heckman [2005] discusses these differentials). The differential effect of Perry on raising female employment at age 27 and on reducing high school dropout rates compared to the male results is consistent with the much steeper gradient of female employment and dropout rates with respect to changes in noncognitive skills compared to that of males.

References

Araujo, Aloisio, Daniel Gottlieb, and Humberto Moreira. 2004. A model of mixed signals with applications to countersignaling and the GED. Working paper, Sociedade Brasileira de Econometria.

Arrow, Kenneth J. 1973. Higher education as a filter. *Journal of Public Economics* 2, no. 3:193–216.

Biglan, Anthony. 2004. *Helping adolescents at risk: Prevention of multiple problem behaviors*. New York: Guilford.

Bonhomme, Stéphane, and Jean-Marc Robin. 2004. Nonparametric identification and estimation of independent factor models. Unpublished manuscript, Department of Economics, Sorbonne, Paris.

Bowles, Samuel, and Herbert Gintis. 1976. *Schooling in capitalist America: Educational reform and the contradictions of economic life*. New York: Basic Books.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. The determinants of earnings: A behavioral approach. *Journal of Economic Literature* 39, no. 4:1137–76.

Browning, Martin, Lars Peter Hansen, and James J. Heckman. 1999. Micro data and general equilibrium models. In *Handbook of macroeconomics*, vol. 1A, ed. John B. Taylor and Michael Woodford, 543–633. Amsterdam: North-Holland.

Cameron, Stephen V., and James J. Heckman. 1998. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy* 106, no. 2:262–333.

———. 2001. The dynamics of educational attainment for black, Hispanic, and white males. *Journal of Political Economy* 109, no. 3:455–99.

Carneiro, Pedro, Karsten Hansen, and James J. Heckman. 2003. Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review* 44, no. 2:361–422.

Carneiro, Pedro, and James J. Heckman. 2003. Human capital policy. In *Inequality in America: What role for human capital policies*? ed. James J. Heckman, Alan B. Krueger, and Benjamin M. Friedman. Cambridge, MA: MIT Press.

Carneiro, Pedro, James J. Heckman, and Dimitriy V. Masterov. 2005. Labor market discrimination and racial differences in pre-market factors. *Journal of Law and Economics* 48, no. 1:1–39.

Cawley, John, James J. Heckman, and Edward J. Vytlacil. 2001. Three observations on wages and measured cognitive ability. *Labour Economics* 8, no. 4:419–42.

Cunha, Flavio, and James J. Heckman. 2006a. The evolution of earnings risk in the U.S. economy. Paper presented at the ninth World Congress of the Econometric Society, London.

———. 2006b. Formulating, identifying, and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources* (forthcoming).

———. 2006c. A framework for the analysis of inequality. *Journal of Macroeconomics* (forthcoming).

Cunha, Flavio, James J. Heckman, Lance J. Lochner, and Dimitriy V. Masterov. 2006. Interpreting the evidence on life cycle skill formation. In *Handbook of the economics of education*, ed. Eric A. Hanushek and Frank Welch. Amsterdam: North-Holland (forthcoming).

Cunha, Flavio, James J. Heckman, and Salvador Navarro. 2005. Separating uncertainty from heterogeneity in life cycle earnings, the 2004 Hicks lecture. *Oxford Economic Papers* 57, no. 2:191–261.

———. 2006. Counterfactual analysis of inequality and social mobility. In *Mobility and inequality: Frontiers of research in sociology and economics*, ed. Stephen L. Morgan, David B. Grusky, and Gary S. Fields, chap. 4. Palo Alto, CA: Stanford University Press.

Darity, William A., Arthur Goldsmith, and Jonathan Veum. 1997. The impact of psychological and human capital on wages. *Economic Inquiry* 35, no. 4:818–29.

Duckworth, Angela L., and Martin E. P. Seligman. 2005. Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science* 16, no. 12:939–44.

Edwards, Richard C. 1976. Individual traits and organizational incentives: What makes a "good" worker? *Journal of Human Resources* 11, no. 1: 51–68.

Ellwood, David T., and Thomas J. Kane. 2000. Who is getting a college education? Family background and the growing gaps in enrollment. In *Securing the future: Investing in children from birth to college*, ed. S. Danziger and J. Waldfogel, 283–324. New York: Russell Sage.

Hansen, Karsten T., James J. Heckman, and Kathleen J. Mullen. 2004. The effect of schooling and ability on achievement test scores. *Journal of Econometrics* 121, no. 1–2:39–98.

Heckman, James J. 1981. Statistical models for discrete panel data. In *Structural analysis of discrete data with econometric applications*, ed. C. Manski and D. McFadden, 114–78. Cambridge, MA: MIT Press.

———. 1995. Lessons from the bell curve. *Journal of Political Economy* 103, no. 5:1091.

———. 2005. Invited comments. In *Lifetime effects: The High/Scope Perry Preschool Study through age 40*, ed. Lawrence J. Schweinhart, Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores, 229–33. Monographs of the High/Scope Educational Research Foundation, 14. Ypsilanti, MI: High/Scope.

Heckman, James J., Maria Isabel Larenas, and Sergio Urzua. 2004. Accounting for the effect of schooling and abilities in the analysis of racial and ethnic disparities in achievement test scores. Unpublished manuscript, Department of Economics, University of Chicago.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. Earnings equations and rates of return: The Mincer equation and beyond. In *Handbook of the economics of education*, ed. Eric A. Hanushek and Frank Welch. Amsterdam: North-Holland (forthcoming).

Heckman, James J., and Dimitriy V. Masterov. 2004. The productivity argument for investing in young children. Technical Report Working Paper no. 5, Committee on Economic Development, Washington, DC.

Heckman, James J., and Salvador Navarro. 2006. Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics* (forthcoming).

Heckman, James J., and Yona Rubinstein. 2001. The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review* 91, no. 2:145–49.

Heckman, James J., Sergio Urzua, and Edward J. Vytlacil. 2006. Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics* (forthcoming).

Heckman, James J., and Edward J. Vytlacil. 2005. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica* 73, no. 3:669–738.

Herrnstein, Richard J., and Charles A. Murray. 1994. *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Herrnstein, Richard J., and James Q. Wilson. 1985. *Crime and human nature: The definitive study of the causes of crime*. New York: Simon & Schuster.

Jencks, Christopher. 1979. *Who gets ahead? The determinants of economic success in America*. New York: Basic Books.

Jensen, Arthur R. 1998. *The g factor: The science of mental ability*. Westport, CT: Praeger.

Jöreskog, Karl G. 1977. Structural equations models in the social sciences: Specification, estimation, and testing. In *Applications of statistics*, ed. P. Krishnaiah, 265–87. New York: North-Holland.

Jöreskog, Karl G., and Arthur S. Goldberger. 1975. Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association* 70, no. 351:631–39.

Klein, Roger, Richard Spady, and Andrew Weiss. 1991. Factors affecting the output and quit propensities of production workers. *Review of Economic Studies* 58, no. 5:929–53.

Neal, Derek A., and William R. Johnson. 1996. The role of premarket factors in black-white wage differences. *Journal of Political Economy* 104, no. 5:869–95.

Osborne-Groves, Melissa. 2006. How important is your personality? Labor market returns to personality for women in the U.S. and U.K. *Journal of Economic Psychology* (forthcoming).

Rosenberg, Morris. 1965. *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rotter, Julian B. 1966. *Generalized expectancies for internal versus external control of reinforcement*. Washington, DC: American Psychological Association.

Segal, Carmit. 2005. Misbehavior, education, and labor market outcomes. Unpublished manuscript, Department of Economics, Stanford University.

Spence, A. Michael. 1973. Job market signaling. *Quarterly Journal of Economics* 87, no. 3:355–74.

Terman, Lewis Madison, and Maud A. Merrill. 1960. *Stanford-Binet Intelligence Scale: Manual for the third revision form L-M*. Boston: Houghton Mifflin.

Urzua, Sergio. 2006. The effects of cognitive and noncognitive skills on racial and ethnic wage gaps. Unpublished manuscript, Department of Economics, University of Chicago.

Wolfe, Raymond N., and Scott D. Johnson. 1995. Personality as a predictor of college performance. *Educational and Psychological Measurement* 55, no. 2:177–85.