The Effects of Educational Choices on Labor Market, Health, and Social Outcomes

James J. Heckman
John Eric Humphries
Sergio Urzua
Gregory Veramendi

October, 2011

THE EFFECTS OF EDUCATIONAL CHOICES ON LABOR MARKET, HEALTH, AND SOCIAL OUTCOMES *

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

John Eric Humphries
University of Chicago

Sergio Urzúa
Northwestern University, NBER
and IZA

Gregory Veramendi
Aarhus University

First version: October 10, 2009
This version: July 26, 2011

Abstract

Using a sequential model of educational choices, we investigate the effect of educational choices on labor market, health, and social outcomes. Unobserved endowments drive the correlations in unobservables across choice and outcome equations. We proxy these endowments with numerous measurements and account for measurement error in the proxies. For each schooling level, we estimate outcomes for labor market, health, and social outcome. This allows us to generate counter-factual outcomes for dynamic choices and a variety of policy and treatment effects. In our framework, responses to treatment vary among observationally identical persons and agents may select into the treatment on the basis of their responses. We find important effects of early cognitive and socio-emotional abilities on schooling choices, labor market outcomes, adult health, and social outcomes. Education at most levels causally produces gains on labor market, health, and social outcomes. We estimate the distribution of responses to education and find substantial heterogeneity on which agents act.

James Heckman
Department of Economics
University of Chicago
1126 East 59th Street, Chicago, IL 60637
Phone: 773-702-0634
Email: jjh@uchicago.edu

John Eric Humphries
Department of Economics
University of Chicago
1126 East 59th Street, Chicago, IL 60637
Phone: 773-980-6575
Email: johneric@uchicago.edu

Sergio Urzúa
Department of Economics and Institute for Policy Research
Northwestern University
Handerson Hall, 2001 Sheridan Road, Evanston, IL 60208
Phone: 847-491-8213
Email: s-urzua@northwestern.edu

Gregory Veramendi
School of Economics and Management

Aarhus University
Bartholins Alle 10, Building 1322, DK-8000 Aarhus C, Denmark
Phone: 45-8942-1546
Email: gveramendi@econ.au.dk

Contents

1 Introduction

This paper investigates the causal effect of education on labor market, health, and social outcomes. A positive association between education and labor market outcomes has long been noted (Mincer, 1958; Becker, 1964; Mincer, 1974). For example, a positive correlation between schooling and health is a well-established finding in the social sciences (Grossman, 1972, 2000, 2006). More recently, it has been noted that there is a positive association between education and social outcomes, such as welfare use and civic participation. To what extent these positive associations reflect causal effects of education is still subject to debate.

Our analysis contributes to the literature on the causal effects of education on labor market outcomes (Card, 2001; Willis and Rosen, 1979; Carneiro, Heckman, and Vytlacil, 2010), health (Adams, 2002; Arendt, 2005; Lleras-Muney, 2005; Silles, 2009; Spasojevic, 2003; Arkes, 2003; Auld and Sidhu, 2005; Grossman, 2008; Grossman and Kaestner, 1997; Cutler and Lleras-Muney, 2010; Conti, Heckman, and Urzua, 2010), and participation in society (Coelli, Green, and Warburton, 2007; Milligan, Moretti, and Oreopoulos, 2004).

We estimate a model of sequential schooling decisions in which individuals make their educational decisions based on expected returns and costs, which are determined by observed and unobserved characteristics (see Keane and Wolpin, 1997; Cameron and Heckman, 1998, 2001). Individuals are endowed with cognitive and socio-emotional abilities (Heckman, Stixrud, and Urzua, 2006; Urzua, 2008) and these endowments determine, in part, schooling attainment.

We adjoin to our dynamic model of schooling choice data on labor market, health, and social outcomes, observed after the final schooling level is reached. We assume these outcomes are determined, in part, by unobserved characteristics, which can be correlated with the unobserved variables in the schooling choice model. Ours is a model of heterogenous dynamic treatment effects (Heckman and Vytlacil, 2007; Heckman, Urzua, and Vytlacil, 2006). Therefore, under our model, two observationally equivalent individuals might experience different treatment effects of education. We estimate a

4

variety of different treatment effects and estimate differences in treatment effects across individuals with different levels of unobserved abilities.

One contribution of this paper is that we estimate educational continuation values. Each educational choice opens up additional educational options. We estimate returns to schooling, both as the direct causal benefit between two final schooling levels, that is the traditional focus in the human capital literature (see, e.g., Becker, 1964, and the discussion in Heckman, Lochner, and Todd, 2006), as well as returns through continuation values, created by the options opened up by schooling.

Our analysis contributes to the growing literature documenting the impact of cognition on health (Grossman, 1975; Shakotko, Edwards, and Grossman, 1982; Hartog and Oosterbeek, 1998; Elias, 2005; Auld and Sidhu, 2005; Kenkel, Lillard, and Mathios, 2006; Cutler and Lleras-Muney, 2010; Kaestner, 2008; Whalley and Deary, 2001; Gottfredson and Deary, 2004) and labor market outcomes (Cawley, Conneely, Heckman, and Vytlacil, 1997; Herrnstein and Murray, 1994; Neal and Johnson, 1996; Carneiro and Heckman, 2002; Glewwe, 2002). Furthermore, our analysis relates to the literature documenting the impact of socio-emotional development on health and labor market outcomes (Hampson and Friedman, 2008; Kaestner, 2008; Heckman, Stixrud, and Urzua, 2006; Cutler and Lleras-Muney, 2010).

Our main empirical findings are:

- We find substantial upward biases in effects of education that do not control for unobserved cognitive and noncognitive traits.

- For most outcomes, the causal gain from education is increasing in school levels.

- For a variety of outcomes measures, we find different effects of education for high and low-ability people.

- Decomposing the return to education into its direct effect (the payment to a given level of education) and its effect on creating options for further education, we see that much of the difference in returns to education by ability levels arises from option values.

- We find significant gains in labor market outcomes from graduating high school and going to college. These are larger for high-ability people. The GED has no significant benefit in the labor market or on other outcomes.

- High school and college attainment causally reduce the probability of being a daily smoker. They improve physical health. High school and college enrollment reduce the probability of being a heavy drinker. Graduating from high school and from a four-year college improve reported physical health. College attainment improves mental health with the effect being much larger for low-ability individuals.

- We find evidence of the impact of education on social behavior. Graduating from high school, enrolling in college, and graduating from college increase the probability of voting and decrease the probability of being divorced and the probability of being on welfare.

The paper is organized as follows: Section 2 presents our model for measuring the returns to schooling. Section 3 describes our estimation strategy. Section 4 presents a detailed analysis of our data. Section 5 discusses the main empirical findings. Section 6 concludes.

2 The Model

We estimate a model of sequential schooling decisions in which individuals make decisions about future schooling levels given their current state. After agents complete their educational decisions, we observe adult outcomes. If unobserved components driving schooling decisions are correlated with unobserved variables determining individual outcomes, it is necessary to control for such selection effects to identify the causal effects of education. We address the selection problem by analyzing a model of potential outcomes with unobserved heterogeneity.[1]

We present the model in the following way:

---

[1] See the survey of dynamic discrete choice by Abbring and Heckman (2007) and the analysis of Heckman and Navarro (2007).

- We first describe our sequential decision model for educational attainment.

- We identify the schooling model using a version of matching on mismeasured covariates with proxies for the true covarites. This is a conditional independence assumption, previously used in Aakvik, Heckman, and Vytlacil (2005) and Carneiro, Hansen, and Heckman (2003).

- Adult outcomes are defined separately by schooling level.

2.1   The Sequential Model of Educational Attainment

Following Cameron and Heckman (2001), each agent makes schooling decisions using a sequential choice model. The choices available to the agent are limited by their previous schooling decisions. Let an individual's current schooling attainment be represented by $j \in \mathcal{J}$, where $\mathcal{J}$ is the set of all possible schooling states. An individual with schooling attainment j makes his next educational decision out of choice set $\mathcal{C}_j$. Let $D_{j,c} = 1$ if the individual with education state j chooses $c \in \mathcal{C}_j$. We assume that individuals make optimal decisions at each educational state. The optimal choice, $\hat{c}$, is

$$\hat{c} = \arg\max_{c \in \mathcal{C}_j}\{I_{j,c}\},$$

where $I_{j,c}$ is the value of choice c for a person with educational attainment j. Thus, an individual's next educational state j' is determined by his optimal educational decision, $j' = \hat{c}$. Finally, let $\mathcal{D}$ represent the set of educational decisions taken by an individual over his life cycle.

We assume a binary decision model at each decision node. In particular, we assume that at a particular node, defined by schooling level j, the agent considers $\mathcal{C}_j = \{j', j''\}$.

Thus, $D_{j,j''} = 1 - D_{j,j'}$, and we can fully analyze the individual decision by simply considering a discrete choice model of the form

$$D_{j,j''} = \begin{cases} 1 & \text{if } I_{j,j''} \geq 0 \\ 0 & \text{otherwise} \end{cases}. \tag{1}$$

7

In the empirical implementation of our model, we assume a linear-in-the-parameters form for $I_{j,j''}$ that approximates the underlying decision structure, as in Cameron and Heckman (2001):

$$I_{j,j''} = \mathbf{X}_{j,j''}\beta^S_{j,j''} + \alpha^S_{j,j''}\theta - \nu_{j,j''}, \tag{2}$$

where $\mathbf{X}_{j,j''}$ is a vector of observed variables relevant to the schooling decision of the agent with schooling level j, and θ is a vector of unobserved endowments. These endowments are unobserved to the econometrician but are known to the agent. θ links the unobservables in schooling choices and outcomes, discussed below. $\nu_{j,j''}$ represents an idiosyncratic error term and satisfies $\nu_{j,j''} \perp\!\!\!\perp (\mathbf{X}_{j,j''}, \theta)$, where "$\perp\!\!\!\perp$" denotes statistical independence. Therefore, $\nu_{j,j''}$ is assumed to be independent across agents and states.

From the sequential decision model one can define a set of final schooling levels. Let s denote a final schooling level in the set of final schooling levels $\mathcal{S} = \{s_0, s_1, ..., \bar{s}\}$. Define a binary indicator, H_s, such that $H_s = 1$ if the individual attains the final schooling level s, and 0 otherwise. Thus,

$$H_s = \begin{cases} 1 & \text{if } D_{1,j} = D_{j,j'} = ... = D_{j'',s} = 1, D_{s,j'''} = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

2.2 Labor Market, Health, and Social Outcomes

We seek to estimate the causal effects of education on a variety of adult outcomes. We distinguish between continuous and discrete (binary) outcomes.

- Continuous outcomes are approximated by a linear-in-the-parameters model. Let Y^k_s denote the outcome $k(= \{1, .., K\})$ associated with final schooling level $s \in \mathcal{S}$. Thus,

$$Y^k_s = \mathbf{X}^k_s\beta^k_s + \alpha^k_s\theta + \nu^k_s, \tag{4}$$

  where $\mathbf{X}^k_s$ is the vector of observed controls relevant for outcome k, and θ is the vector of unobserved endowments. ν^k_s represents an idiosyncratic error term such

that $\nu_s^k \perp\!\!\!\perp (\mathbf{X}_s^k, \theta)$. The ν_s^k are mutually independent across s. Equations (3) and (4) can be used to define observed outcome Y^k, using the conventional switching regression framework:

$$Y^k = \sum_{s \in \mathcal{S}} H_s Y_s^k. \tag{5}$$

- We model binary outcomes using a latent index structure. Let V_s^k denote the latent utility and outcome k associated with final schooling level s. The latent utility is given by a linear-in-the-parameters specification:

$$V_s^k = \mathbf{X}_s^k \tilde{\beta}_s^k + \tilde{\alpha}_s^k \theta + \tilde{\nu}_s^k, \tag{6}$$

where $\mathbf{X}_s^k$, θ, and $\tilde{\nu}_s^k$ have analogous definitions to the continuous outcome case. We can define a binary outcome variable, B_s^k:

$$B_s^k = \begin{cases} 1 & \text{if } V_s^k \geq 0 \\ 0 & \text{otherwise} \end{cases}. \tag{7}$$

The observed outcome can be expressed as in the continuous case:

$$B^k = \sum_{s \in \mathcal{S}} H_s B_s^k. \tag{8}$$

2.3 Measurement System for Unobserved Cognitive and Socio-emotional Endowments

Given θ and condition on X, all outcomes and choices are statistically independent. If we could measure θ, we could condition on it (along with X) and do matching. (See Carneiro, Hansen, and Heckman, 2003, and Abbring and Heckman, 2007.) We do not directly measure θ, but we can proxy it and estimate and correct for the effects of any measurement error in the proxy.

We follow Carneiro, Hansen, and Heckman (2003) and Heckman, Stixrud, and Urzua (2006) and identify the schooling choice model and the models for outcomes

using information from a measurement system. Using this system allows us to interpret unobserved endowments as cognitive and socio-emotional abilities.

Before introducing the measurement system, let θ^C and θ^{SE} denote the levels of cognitive and social-emotional abilities, respectively, so that $\theta = (\theta^C, \theta^{SE})$. We allow θ^C and θ^{SE} to be correlated.

Let T_s^C be a vector of cognitive test scores, T_s^{SE} a set of variables that measure by socio-emotional abilities, and $T_s^{C,SE}$ a set of variables influenced by cognitive and socio-emotional abilities, all measured at schooling level s. We posit a linear measurement system for these variables. More precisely,

$$T_s^C \quad = \quad \mathbf{X}_s^C \beta_s^C + \alpha_s^C \theta^C + e_s^C \tag{9}$$

$$T_s^{SE} \quad = \quad \mathbf{X}_s^{SE} \beta_s^{SE} + \alpha_s^{SE} \theta^{SE} + e_s^{SE} \tag{10}$$

$$T_s^{C,SE} \quad = \quad \mathbf{X}_s^{C,SE} \beta_s^{C,SE} + \tilde{\alpha}_s^C \theta^C + \tilde{\alpha}_s^{SE} \theta^{SE} + e_s^{C,SE}. \tag{11}$$

The structure assumed in (9), (10), and (11), when allowing for correlated factors, is identified if the model has one measure which depends only on cognitive ability (T_s^C), one measure which depends only on socio-emotional ability (T_s^{SE}), and several equations loading both on cognitive ability and socio-emotional ability $(T_s^{c,SE})$. A proof of nonparametric identification of the distribution of θ for our model is provided in the Web Appendix.[2]

3   Estimation Strategy

We estimate this model in two stages. The distribution of latent endowments and the schooling choice equations are estimated in the first stage, and equations governing adult outcomes are estimated in the second stage using estimates from the first stage. In this fashion, the measurement system is estimated separately from the outcome sys-

---

[2]See Section A in the Web Appendix.

tem, so that we do not force predictive power of the latent factors on adult outcomes in our estimation procedure. We assume $\nu_{j,j''}$, ν_s^k, $\tilde{\nu}_s^k$, and e_s are mutually independent, mean-zero, unit variance, normal variates. Additionally, we assume that these errors are independent conditional on the observables and the unobserved factors. The factor structure is assumed to explain all of the correlations in unobservables across outcomes, conditional on $\mathbf{X}_i$. Identification of the factors comes from the schooling and measurement system.

This approach follows that from the analysis of Carneiro, Hansen, and Heckman (2003). Conditional on θ and X, all potential outcomes are independent of each other. As previously noted, our procedure is a version of matching where we do not measure a subset of the conditioning variables but instead match on proxies for θ and account for the effects of measurement error in the proxies in generating our estimates.

The likelihood, assuming independence across observations, is

$$
\begin{aligned}
\mathcal{L} &= \prod_i f(\mathbf{Y}_i, \mathbf{B}_i, \mathcal{D}_i, T_i | \mathbf{X}_i) \\
&= \prod_i \int f(\mathbf{Y}_i, \mathbf{B}_i | \mathcal{D}_i \mathbf{X}_i \theta) f(\mathcal{D}_i, T_i | \mathbf{X}_i \theta) f(\theta) d\theta,
\end{aligned}
$$

where the last two steps are justified from the assumptions that $\theta \perp\!\!\!\perp \mathbf{X}_i$ and that the outcomes are independent once we condition on θ and $\mathbf{X}_i$. For the first stage, the sample likelihood is

$$
\mathcal{L}^1 = \prod_i \int_{\theta \in \Theta} f(\mathcal{D}_i, T_i | \mathbf{X}_i, \theta = \mathbf{z}) dF_\theta(\mathbf{z}), \tag{12}
$$

where we integrate over the distributions of the latent factors. The goal of the first stage is to secure estimators, $\hat{f}(\mathcal{D}_i, T_i | \mathbf{X}_i, \theta)$ and $\hat{f}(\theta)$, for $f(\mathcal{D}_i, T_i \mid \mathbf{X}_i, \theta)$ and $f(\theta)$, respectively. In the second stage, we use first stage estimates to express the likelihood as

$$
\mathcal{L}^2 = \prod_i \int_{\theta \in \Theta} f(\mathbf{Y}_i, \mathbf{B}_i | \mathcal{D}_i, \mathbf{X}_i, \theta = \mathbf{z}) \hat{f}(\mathcal{D}_i, T_i | \mathbf{X}_i, \theta = \mathbf{z}) d\hat{F}_\theta(\mathbf{z}). \tag{13}
$$

Since $\mathbf{Y}_i, \mathbf{B}_i$ are independent from the first stage conditional on $\mathbf{X}_i, \theta, \mathcal{D}_i$ under stan-

dard conditions, we can obtain a consistent estimate of the parameters for the adult outcome models. Each stage is estimated using maximum-likelihood. Standard errors and confidence intervals are calculated by estimating two hundred bootstrap samples.

4   Defining Treatment Effects

The estimated model generates the causal effect of education and ability on labor market, health, and social outcomes. Since the model can be used to produce counterfactual outcomes, we can create a variety of average and distributional treatment effects. They can be used to predict how causally manipulating education affects people at different ability levels and allows us to understand the effectiveness of policy for different segments of the population.

The traditional literature on the returns to schooling defines its parameters in terms of the returns generated by going from one final schooling level to another (Becker, 1964). This approach ignores the sequential nature of schooling and the options created by going to an additional level of schooling. For example, consider the gains in going from being a GED to becoming a four-year college graduate. The GED may enter community college. The GED may complete community college. From community college, the GED may go on to a four year college and so forth. Each decision opens up further possibilities. There are many choices at multiple nodes of education.

We analyze sequential decisions made by the individuals. We identify treatment effects at each binary decision node. For example, we estimate the treatment effect for deciding to graduate from high school or drop out ($D_{0,1}$). But once agents graduate from high school, agents have the option of going to college and even graduating from college. Similarly, once agents drop out, they have the option of getting a GED. All of these schooling decisions are options that emerge from a dynamic model of schooling. We estimate the traditional gains from choosing between final schooling levels. Such gains are calculated relative to the return from being a high school dropout. In this way we can compare our results with other methods used in the literature. In addition,

12

we estimate treatment effects for each sequential decision node. This method takes into account future options opened up by educational choices.

4.1   Gains from Changing Final Schooling Levels

Let Y_0 be defined as the outcome for the final schooling level of a high school dropout and Y_s is the final schooling level being studied. The average treatment effect in this case is measured in the full population:

$$\Delta_s^{ATE} \equiv \iint E_\nu(Y_s - Y_0|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z})dF_{\mathbf{X},\theta}(\mathbf{x}, \mathbf{z}), \tag{14}$$

where E_ν is the expectation over idiosyncratic shocks to outcome Y_j, $j \in \{0, s\}$. The average effect of the treatment on the treated is measured only for those who attain the final schooling being studied (s):

$$\Delta_s^{TT} \equiv \iint E_\nu(Y_s - Y_0|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z})dF_{\mathbf{X},\theta|H_s=1}(\mathbf{x}, \mathbf{z}), \tag{15}$$

and the average effect of the treatment on the untreated is measured only for those who are high school dropouts ($s = 0$):

$$\Delta_s^{TUT} \equiv \iint E_\nu(Y_s - Y_0|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z})dF_{\mathbf{X},\theta|H_0=1}(\mathbf{x}, \mathbf{z}). \tag{16}$$

4.2   Treatment Effect of Educational Decisions

The treatment effect of an educational decision is calculated by looking at the difference in expected outcomes when changing a single educational decision in the sequential schooling model. Since a given educational decision can open up further educational choices to be made in the future, in order to calculate the full effect of a given educational decision, the treatment effect needs to include the probability weighted benefit of further educational choices. Let the expected value of an educational decision

$(D_{j,j''} = 1)$ to an individual with $\mathbf{X} = \mathbf{x}$ and $\theta = \mathbf{z}$ be

$$E\left(Y|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}, D_{j,j''} = 1\right) \equiv \sum_s \Pr\left(s|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}, D_{j,j''} = 1\right) \times E\left(Y_s|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}\right),$$

where the expectation (E) is over future educational choices and idiosyncratic shocks, $\Pr\left(s|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}, D_{j,j''} = 1\right)$ is the probability that the individual stops at education level s, and Y_s is the value of the outcome if the individual stops at education level s.[3] Of course, $\Pr(s|D_{j,j''} = 1) = 0$ if s is not accessible given $D_{j,j''} = 1$.

Let the person-specific treatment effect for an individual changing his decision at decision node j be defined as the difference between the expected value of the decisions:

$$\Delta_{j,j''}[Y|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}] \equiv E(Y|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}, D_{j,j''} = 1) - E(Y|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}, D_{j,j''} = 0).$$

This person-specific treatment effect takes into account not only the direct effect of the decision, but also includes the value of possible additional schooling.[4]

---

[3]For example, the choice to graduate from high school opens up the possibility of enrolling in college and possibly graduating from college. Let s indicate the level of final schooling, where 0 corresponds to dropping out of high school, 1 to graduating high school, 2 to attaining a GED, 3 to attaining some college, and 4 for graduating college. Then let $D_{0,1}$ represent the decision to graduate from high school and $D_{0,2}$ represent the decision to get the GED once an individual has chosen to drop out $(D_{0,1} = 0)$. The expected wage (Y) for an individual, who chooses to graduate from high school $(D_{0,1} = 1)$is then

$$\begin{aligned} E(Y|D_{0,1} = 1) = \quad &\Pr(s = 1|D_{0,1} = 1) \times Y_1 + \Pr(s = 3|D_{0,1} = 1) \times Y_3 \\ &+ \Pr(s = 4|D_{0,1} = 1) \times Y_4, \end{aligned}$$

where $\Pr()$ is the probability that an individual has a given final educational level and the wage Y depends on the final schooling level. Of course, $\Pr(s = 1|D_{0,1} = 1) + \Pr(s = 3|D_{0,1} = 1) + \Pr(s = 4|D_{0,1} = 1) = 1$. Likewise, the expected value for someone who decides to drop out of high school $(D_{0,1} = 0)$ is then

$$E(Y|D_{0,1} = 0) = \Pr(s = 0|D_{0,1} = 0) \times Y_0 + \Pr(s = 2|D_{0,1} = 0) \times Y_2,$$

where $\Pr(s = 0|D_{0,1} = 0) + \Pr(s = 2|D_{0,1} = 0) = 1$.

[4]The treatment effect can be broken up into the direct effect and the continuation value. The continuation value of graduating from high school is the probability that they enroll in college times the wage benefit of having some college plus the probability of then completing college times the wage benefit of completing college. For the high school graduation decision, the continuation value is:

$$CV(Y|D_{0,1} = 1) = [(Y_4 - Y_3) \times \Pr(D_{3,4} = 1|D_{1,3} = 1) + (Y_3 - Y_1)] \times \Pr(D_{1,3} = 1|D_{0,1} = 1)$$

where in this case $\Pr$ represents the probability of making an educational decision as opposed to terminating in a final educational state as before, $D_{1,3}$ represents the decision to enroll in college and $D_{3,4}$ represents the decision to graduate from college. The direct treatment effect of graduating from high school is:

$$DTE(Y|D_{0,1} = 1) = Y_1 - [Y_0 + (Y_2 - Y_0) \times \Pr(D_{0,2} = 1|D_{0,1} = 0)]$$

14

Thus, the average treatment effect is

$$\Delta_{j,j''}^{ATE} \equiv \iint \Delta_{j,j''}[Y|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}]dF_{\mathbf{X},\theta}(\mathbf{x}, \mathbf{z}), \tag{17}$$

the average effect of the treatment on the treated is

$$\Delta_{j,j''}^{TT} \equiv \iint \Delta_{j,j''}[Y|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}]dF_{\mathbf{X},\theta|I_{j,j''} \geq 0}(\mathbf{x}, \mathbf{z}), \tag{18}$$

and the average effect of the treatment on the untreated is

$$\Delta_{j,j''}^{TUT} \equiv \iint \Delta_{j,j''}[Y|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}]dF_{\mathbf{X},\theta|I_{j,j''} < 0}(\mathbf{x}, \mathbf{z}). \tag{19}$$

Finally, the average marginal treatment effect is the average effect of participating in the treatment for individuals who are at the margin of indifference between participating or not:

$$\Delta_{j,j''}^{AMTE} \equiv \iint \Delta_{j,j''}[Y|\mathbf{X} = \mathbf{x}, \theta = \mathbf{z}]dF_{\mathbf{X},\theta \mid |I_{j,j''}| < \varepsilon^S}(\mathbf{x}, \mathbf{z}). \tag{20}$$

See, e.g., Carneiro, Heckman, and Vytlacil (2010).

5   Data and Estimation Strategy

We use the 1979 National Longitudinal Survey of Youth (NLSY79), which is a nationally representative sample of men and women born in the years 1957-64. The respondents were ages 14-22 when first interviewed in 1979. It provides annual or biennial surveys on a variety of outcomes. It also contains a large array of information on other aspects of the respondent's lives, such as educational achievement, marital status, fertility, participation in crime, income, assets, health, alcohol and substance

---

where $D_{0,2}$ represents the decision to get a GED once an individual has already dropped out of high school. Some of the probabilities above could have been written in terms of the final state since for terminal nodes the probability of the final state is the same as the probability of the decision node (i.e. $\Pr(D_{3,4} = 1|D_{1,3} = 1) = \Pr(s = 4|D_{1,3} = 1)$ and $\Pr(D_{0,2} = 1|D_{0,1} = 0) = \Pr(s = 2|D_{0,1} = 0)$).

abuse, and scores on achievement and psychological tests. We use the core sample of males, which, after removing observations with missing covariates, contains 2242 observations.

5.1 Outcomes

We consider a number of labor market and behavioral outcomes conditional on schooling levels.

5.1.1 Schooling Levels

We consider four different transitions and five final schooling levels. The transitions studied are (i) enrolled in high school deciding between graduating from high school and dropping out from high school, (ii) high school dropouts deciding whether or not to get the GED, (iii) high school graduates deciding whether or not to enroll in college, and (iv) college students deciding whether nor not to graduate from college or to drop out before getting the degree. Consequently, the final schooling levels are (i) high school dropout, (ii) GED, (iii) High school graduate, (iv) some college and (v) four-year college degree. We utilize the information available at age 30 to determine the final schooling level. Table 1 and Figure 1 describe the five possible educational choices and their conditional structure.[5] Thus, following the notation introduced in Section 2.1, the indicator variable for college graduate is defined as

$$
H_4 = \begin{cases} 1 & \text{if } D_{1,0} = D_{3,1} = D_{4,3} = 1 \\ 0 & \text{otherwise} \end{cases}.
\tag{21}
$$

5.1.2 Labor Market Outcomes

Following the analysis of Heckman, Stixrud, and Urzua (2006), we consider labor market outcomes at age 30. We analyze (log) wages at age 30, white-collar employment at age 30, labor force participation at age 30, and employment at age 30 given par-

---

[5]A negligible fraction of individuals change schooling levels after age 30.

ticipation. We also construct and analyze present value of wages from ages 20 to 40. Following Keane and Wolpin (1997), we denote as white-collar occupations (i) professional, technical, and kindred; (ii) managers, officials, and proprietors; (iii) sales workers; (iv) farmers and farm managers; and (v) clerical and kindred. For (log) wages and present value of wages we use linear regression models conditional on schooling level. For labor market participation and white-collar occupation we use binary decision models by schooling levels.

5.1.3   Physical Health and Healthy Behaviors

As a measure of physical health, we construct an obesity indicator based on BMI. BMI is calculated as BMI=(Weight in Pounds * 703)/(Height in inches)^2, and the obesity indicator takes a value of one if the BMI is 30 and above, and zero otherwise. As a measure of mental and physical health, we use the PCS-12 scale. The PCS-12 scale is the Physical Component Summary obtained from SF-12. SF-12 is a 12-question health survey designed by John Ware of the New England Medical Center Hospital (see Ware, Kosinski, and Keller, 1996, and Gandek, Ware, Aaronson, Apolone, Bjorner, Brazier, Bullinger, Kaasa, Leplege, Prieto, and Sullivan, 1998).[6] The SF-12 is designed to provide a measure of the respondent's mental and physical health irrespective of their proclivity to use formal health services. Respondents with a score above (below) 50 have better (worse) health than the typical person in the general U.S. population. Each one-point difference above or below 50 corresponds to one-tenth of a standard

---

[6]The questions making up the SF-12 are: "In general, would you say your health is Excellent (1), Very Good (2), Good (3), Fair (4) or Poor (5)", "The following items are activities you might do during a typical day. Does your health limit you in these activites? [1] Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling or playing golf? [2] Climbing several flights of stairs? [3] Accomplished less than you would like? [4] Were limited in the kind of work or other activities?", "During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)? (Please answer YES or NO for each question). [1] Accomplished less than you would like? [2] Didn't do work or other activities as carefully as usual?", "During the past 4 weeks, how much did pain interfere with your normal work (including both work outside of the home and housework)?", "The next questions are about how you feel and how things have been with you during the past 4 weeks. (For each question, please give the one answer that comes closest to the way you have been feeling). How often during the past 4 weeks: [1] have you felt calm and peaceful? , [2] Did you have a lot of energy?, [3] Have you felt down-hearted and blue?", "During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?"

deviation. For example, a person with a score of 30 is two standard deviations away from the mean. We standardize the PCS-12 score to have mean zero and variance one in the overall population. We also include self-reported smoking and drinking behavior as binary outcomes for regular smoking and heavy drinking at age 30.

5.1.4 Mental Health

We analyze the effect of education on Pearlin's "Personal Mastery Scale" (collected in 1992), Rosenberg's Self-esteem scale (collected in 2006), the Mental Component Summary or MCS-12 (collected at age 40), and The Center for Epidemiologic Studies Depression Scale (CES-D) (collected at age 40). Pearlin's "Personal Mastery Scale" consists of 7 items which are answered on a 4-point ((4) strongly agree, (3) agree, (2) disagree, (1) strongly disagree) scale and has been shown to exhibit reasonable internal reliability and good construct validity (see Pearlin and Schooler, 1978, and Pearlin, Menaghan, Lieberman, and Mullan, 1981).[7] We form aggregate measures by summing the scores from the items, and standardizing the scores to have mean 0 and variance 1 in the overall population.

Rosenberg's Self-Esteem Scale consists of 11 items which are answered on a 4-point (4 strongly agree, 3 agree, 2 disagree, 1 strongly disagree).[8] We form the scale summing the scores from the items, and standardizing the scores to have mean 0 and variance 1 in the overall population.

The MCS-12 scale is the Mental Component Summary (measures mental health) is constructed from a subset of the SF-12 health questionnaire. The MCS-12 is designed to provide a measure of the respondent's mental health irrespective of their proclivity

---

[7]The items are "There is really no way I can solve some of the problems I have," "Sometimes I feel that i'm being pushed around in life," "I have little control over the things that happen to me," "I can do just about anything I really set my mind to," "I often feel helpless in dealing with the problems of life," "What happens to me in the future mostly depends on me," and "There is little I can do to change many of the important things in my life."

[8]The items are "I feel that I'm a person of worth, at least on equal basis with others," "I feel that I have a number of good qualities," "All in all, I am inclined to feel that I am a failure," "I am able to do things as well as most other people," "I feel I do not have much to be proud of," "I take a positive attitude toward myself," "On the whole, I am satisfied with myself," "I wish I could have more respect for myself," "I certainly feel useless at times," and "At times I think I am no good at all."

to use formal health services. Respondents with a score above (below) 50 have better (worse) health than the typical person in the general U.S. population. We standardized the MCS-12 score to have mean zero and variance one in the overall population.

CES-D is one of the most common screening tests for helping an individual determine his or her depression quotient (see Radloff (1977) and Devins, Orme, Costello, Binik, Frizzell, Stam, and Pullin (1988)). This scale measures symptoms of depression, discriminates between clinically depressed individuals and others, and is highly correlated with other depression rating scales. We form the scale summing the scores from the items: "I did not feel like eating; my appetite was poor," "I had trouble keeping my mind on what I was doing," "I felt depressed," "I felt that everything I did was an effort," "My sleep was restless," "I felt sad," and "I could not get *going.*" For each items the potential answers are: "0 Rarely/None of the time/1 Day," "1 Some/A little of the time/1-2 Days," "2 Occasionally/Moderate amount of the time/3-4 Days," and "3 Most/All of the time/5-7 Days." We standardized the scores to have mean 0 and variance 1 in the overall population.

5.2   Social Outcomes

We include several social outcomes that, while normative, align with the goals of education as commonly claimed by educators. We include a binary outcome for ever being divorced, which is conditional on having been married. We construct a binary variable for any welfare use which is one if in individual received any welfare between 1996 and 2006 and is otherwise zero. We include a binary variable for if the individual reported trusting people. The variable is one if the individual reported "always" or "most of the time" for trusting people in 2008, and is otherwise zero. Finally, we include a binary variable indicating if the individual reported voting in 2006.

5.3 Early Adverse Behavior

We include five additional measures of adverse adolescent behavior to aid in interpreting socio-emotional traits. These measures are not required to identify the distributions of latent factors. We consider violent behavior in 1979 (fighting at school or work and hitting or threatening to hit someone[9]), tried marijuana before age 15, daily smoking before age 15, regular drinking before age 15, and any intercourse before age 15. For violent behavior, we also control for the potential effect of schooling.[10]

5.4 Measurement System

The set of cognitive measures we use includes the Armed Services Vocational Aptitude Battery (ASVAB), a subset of which are utilized to generate the Armed Forces Qualification Test (AFQT) score.[11] Specifically, we consider the scores from Arithmetic Reasoning, Coding Speed, Paragraph Comprehension, World Knowledge, Math Knowledge, and Numerical Operations. For each test, we estimate a separate model, and we control for the effect of schooling at the time of the tests using the method developed in Hansen, Heckman, and Mullen (2004). Cognitive ability is also measured by 9th grade GPA in reading, social studies, science, and math, though GPA is allowed to have a socio-emotional inputs as well.

Grades and school performance are typically treated as measures of cognitive ability in economics. While cognition is essential, a growing body of work by economists and personality psychologists demonstrates the importance of non-cognitive traits and skills on school performance.[12] By including measurements on both types of unobserved

---

[9]This is a binary variable which is unity if an agent answers yes to either "Gotten in to a physical fight at school or work?" or "Hit or seriously threatened to hit somebody?"

[10]Gullone and Moore (2000) present a line of research which studies the relationship between personality traits and adolescent risk-behavior. Duckworth and Urzua (2009) study the relationship between personality and the number of arrests between 14 and 17 years old and find that it is correlated with conscientiousness, agreeableness, and IQ. Based on literature relating early behavior to non-cognitive traits, our five additional measures of early adverse behavior help demonstrate that our socio-emotional factor is capturing traits that then explain these observed behaviors in an expected manner.

[11]The AFQT scores are often interpreted as proxies for cognitive ability (Herrnstein and Murray, 1994). See the discussion in Almlund, Duckworth, Heckman, and Kautz (2011).

[12]Many psychologists use a socio-emotional taxonomy called the Big Five (John, Robins, and Pervin, 2008). This is an organizing framework that categorizes personality traits into 5 categories. The five traits

endowments, we can separate the roles of cognitive and socio-emotional endowments in academic success. Thus, socio-emotional ability is measured by the socio-emotional contributions towards 9th grade GPA in reading, social studies, science, and math. GPA by grade and subject is constructed from high school transcript records. Up to 64 courses were recorded from school transcripts and included year taken, grade level taken, a class identification code, carnegie units (a measure of seat time), and the grade received. Using the class identification code, we identified all courses taken in either reading, social studies, science, or math in 9th grade and constructed subject level GPAs. Class GPA was weighted by Carnegie units when more than one class was taken in a subject in 9th grade.[13]

Finally, we include a single measure for participating in minor risky or reckless activity in 1979 in our measurement system of socio-emotional ability. [14] Unlike the five previous measures in early adverse behavior, this binary measure of participation in early risky or reckless behavior is used in securing identification of the distribution of endowments.

---

are Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. A growing body of work suggests that these traits and other noncognitive traits play key roles in academic success. Duckworth and Seligman (2005) find that self-discipline predicts GPA in 8th graders better than IQ. Duckworth, Quinn, and Tsukayama (2010) use three unique studies to show that self-control predicts grades earned in middle school better than IQ across racial and socio-economic demographics. Farsides and Woodfield (2003), Conard (2006), and Noftle and Robins (2007) find that Big 5 traits positively predict grades and academic success. See also Borghans, Golsteyn, Heckman, and Humphries (2011). These studies find predictive power after controlling for previous grades or test scores. In these studies, the benefits of personality traits are mediated through behaviors such as increased attendance or increased academic effort. A meta-analysis by Credé and Kuncel (2008) finds that study habits, skills, and attitudes have similar predictive power as standardized tests and previous grades in predicting college performance. They find that study skills are largely independent of high school GPA and standardized admissions tests, but do have moderate correlations with personality traits. Academic success depends on cognitive ability, but also depends strongly on non-cognitive traits such as conscientiousness, self-control, and self-discipline. This motivates our identification strategy of including both a cognitive and non-cognitive factor in 9th grade GPA, as much of the variance not explained through test scores has been shown to be related to non-cognitive traits.

[13] As noted by Borghans, Golsteyn, Heckman, and Humphries (2011) and Almlund, Duckworth, Heckman, and Kautz (2011), the principal determinants of the grade point average are personality traits and not cognition. See also Duckworth, Quinn, and Tsukayama (2010).

[14] Preliminary data analysis suggested this measure was the least dependent on cognitive ability. This variable is a binary variable which is unity if an agent answers yes to any of the following questions in 1980: "Taken something from the store without paying for it," "Purposely destroyed or damaged property that did not belong to you?," "Other than from a store, taken something that did not belong to you worth under $50?," and "Tried to get something by lying to a person about what you would do for him, that is, tried to con someone?"

5.5 Exogenous Observed Characteristics

The variables used to measure a set of characteristics defining family background include dummies for race, living in an urban area at 14, living in the South at 14, living in a broken home at 14, number of siblings, mother's education, father's education, family income in 1979, and age in 1980 as a continuous cohort variable. All models include these characteristics as covariates in the outcome equations. In addition to the family background variables, some models have outcome-specific covariates. The schooling choice models include the difference in local wages across schooling levels, local unemployment for the different schooling levels, and the local cost of college and of taking the GED test. The ASVAB test score equations have individual cohort dummies. Finally, models for wages, labor market participation, and employment in white-collar jobs include contemporaneous covariates such as living in an urban area at 30, region of residence at 30, and local unemployment at 30.

6 Empirical Estimates

We present empirical results in the following order. We first discuss the measurement systems. Then we examine the effects of endowments on schooling, labor markets, and health outcomes.

6.1 Estimates from the Measurement System

Figure 2 presents the estimated joint and marginal distributions of cognitive and socio-emotional endowments. The estimated distributional parameters are presented at the bottom of the figure. The estimates suggest a positive and statistically significant correlation between the latent endowments ($\rho = 0.24$). We reject the hypothesis of normally distributed factors. Tables 2 and 3 report the estimates for adverse adolescent behavior. These models are estimated in order to interpret the socio-emotional endowment. The factor loadings (the coefficients for "cognitive" and "socio-emotional" factors at the base of each table) show that the socio-emotional endowment plays a

22

significant role in these adverse behaviors, whereas the cognitive loadings are either insignificant or much smaller than the socio-emotional loadings. To test the robustness of the measurement system, we also include these outcomes as measurements to generate the distribution of the latent endowments. Doing so does not significantly change the distribution of the factors nor the loadings in the education and grade models. Figure 3 shows the decomposition of the measures in the measurement system.[15] Although observed variables explain a large part of the variance of the test scores and grades, there is a still large amount of measurement error. This is one motivation for using a factor model.

6.2 The Effect of Cognitive and Socio-emotional Endowments on Schooling Decision, Labor Market, and Health Outcomes

Table 4 presents our estimates of the schooling choice model. Figure 5 presents a graphic analysis of schooling choice depends on the level of endowments. Figure 6 presents a graphical analysis of the effects of endowments on (log) wages, daily smoking, self-esteem, and voting in the 2006 election. The figures and estimates for the rest of the outcomes can be found in the Web Appendix. We find the following results:

1. Measurement System: We find that the cognitive factor loadings are statistically significant for the ASVAB tests, GPA, and educational choices in the measurement system (see Figure 3). Socio-emotional loadings are significant predictions of GPA and educational choices, except for the GED, which only loads on cognition.

2. Labor Market Outcomes: We find that cognitive loadings are statistically significant in the equations for labor market participation, white-collar employment, and wages for all schooling levels, except "some college." The loading on the social-emotional factors are significant for the all the unconditional labor mar-

---

[15]We discuss some of the outcome measures displayed in this figure in the rest of this section.

ket models, except for labor force participation. The socio-emotional loading is significant only in the model for white-collar employment for college graduates.

3. Physical Health Outcomes: In models that do not fix education levels, we find evidence of cognitive effects on the models for smoking, obesity, and PCS-12, while there is no evidence that cognitive ability is an important determinant for heavy drinking. Cognitive ability also plays a role in explaining obesity for high school graduates. There is evidence for effects of socio-emotional factors on heavy drinking and smoking. Finally, socio-emotional ability appears to play a role in the higher education models for heavy drinking given education (college and some college) and in obesity (some college).

4. Mental Health Outcomes: Not controlling for schooling, we find significant evidence for the importance of cognitive ability in explaining depression, Pearlin, and self-esteem. Controlling for schooling, cognitive ability predicts depression for high school dropouts, high school graduates, and those with some college. It also predicts Pearlin scores (GED and high school graduates) and self-esteem (high school dropouts, GED and high school graduates). Socio-emotional ability explains Pearlin, not controlling for schooling. We do not find any significant loadings for either ability in the MCS-12 models.[16]

5. Social Outcomes: We find significant effects of cognition in all social outcomes, not conditioning on schooling. In addition, cognitive ability seems to be an important predictor of outcomes for the lower educational levels for trust (GEDs, high school dropouts and graduates), divorce (high school graduates), welfare (high school dropouts and graduates), and voting (high school dropouts and graduates). Socio-emotional ability had significant loadings in the unconditional models for divorce and voting.

---

[16]The MCS-12 is the mental composite score from the SF-12 health survey.

6.3   Sorting into Schooling Level

Since the model is highly nonlinear and multidimensional, the best way to understand its results is by simulation. We randomly draw exogenous regressors from the data and factors from the estimated factor distributions and simulate the different outcomes.

Figure 4 shows the distribution of the factors by final schooling level. Individuals sort by both cognitive and non-cognitive ability into increasing schooling levels. The only exception is for GEDs, who have cognitive ability distributions similar to terminal high school graduates but socio-emotional distributions similar to dropouts.

6.4   Goodness of Fit

The goodness of fit measurements are made for the various outcomes and measurement systems. Goodness of fit for discrete outcomes is tested using a χ^2 test of fit of the model to data. For continuous outcomes, the equality of the model and data distributions are tested using a two-sample Kolmogorov-Smirnov test. In terms of the first and second moments, the model does a good job of reproducing the data. The measurements of the goodness of fit can be found in Section D in the Web Appendix.[17]

6.5   Treatment Effects: Comparison of Outcomes for Different Final Schooling Levels

We now compare the outcomes from a particular final schooling level s with those associated with the high school dropout status. In other words, we use high school dropout as our baseline comparison group. The estimated treatment effects of education on log wages, present value of wages, white-collar occupation, and participation are shown in Figure 7. These are calculated by simulating the mean outcomes for the designated state and comparing it with the mean-simulated outcome for the benchmark dropout state for the subpopulation of persons who are in either the designated state or the dropout state. Figures 8, 9, and 10 present the results for physical health, mental

---

[17]The Web Appendix is available at http://jenni.uchicago.edu/effects-school-labor.

health, and social outcomes, respectively. Using the same procedure is used for wages for all outcomes. For each of the outcomes, the bars labeled "Observed" display the observed differences in the data. The bars labelled "Causal Mechanism" display the average treatment effect obtained from the comparison of the outcomes associated with a particular final schooling level s relative to the high school dropout status. ATE is computed only for those choosing one of the two final schooling levels. Tables showing ATE for the full population; TT and TUT can be found in the Web Appendix.[18] Our main findings are summarized below.

1. In general, the differences are much larger when we do not control for observed variables and latent abilities. We document in the Web Appendix that there are significant heterogeneity in the gains from school.

2. In most cases, the gains from education is increasing (in absolute value) with the schooling level, even after controlling for ability.

3. Labor Market Outcomes: There is no significant effect from attaining a GED for any labor market outcome, while graduating from high school and some college achievement increases wages at 30 and increase the probability of having a white-collar occupation. About half of the apparent returns for wages at 30 seem to be explained by observed variables and latent abilities. Aside from the amorphous category "some college," on average there are no significant returns to graduating high school and college in terms of present value of wages.[19] Finally, the effect of education on labor force participation is insignificant for all educational levels, except for graduating from high school.

4. Physical Health Outcomes: Education causally reduces smoking and obesity even after controlling for observed variables and latent ability.

5. Mental Health Outcomes: The estimates of the causal effect of education are not precisely determined. We find no significant effect for education on self-esteem,

---

[18]See Section E in the Web Appendix.

[19]As shown in Section E the Web Appendix, the same is not true for TUT for the present value of wages. TUT shows large returns to all education levels except for the GED.

depression, and mental health. An exception is the Pearlin measure, where high school and college achievement have significant effects on a person's sense of control.

6. Social Outcomes: We find large and statistically significant causal effects of college attainment on voting, welfare, and divorce. For divorce, the causal effect of education explains more than 100% of the observed effect. Over half of the observed association between education and welfare and voting is explained by observed variables and latent abilities. Finally, the causal estimated effects of education on trust are not statistically significant.

6.6 Treatment Effects: Pair-wise Comparison by Decision Node

We now analyze the treatment effects by decision node. Our ability to construct these causal effects is a byproduct of our sequential model. We compute the gain to achieving (and possibly exceeding) the designated state inclusive of the continuation value associated with that state and compare it to the outcome associated with not achieving the state. The estimated treatment effects of education on log-wages, present value of wages, white-collar occupation, and participation are shown in Figure 11. Figures 12, 13, and 14 present the results for physical health, mental health, and social outcomes, respectively.

Each figure presents the average effects of an educational decision on the outcome of interest. The effects are presented as different bars in each figure, and they are defined as the differences in the expected outcome $(\Delta_{j,j''})$ associated with a given educational decision $(D_{j,j''})$, as defined in Section 4.2. Importantly, each schooling decision might provide the option to pursue higher schooling levels, while terminal schooling levels do not provide any continuation value. At each node, the ATE presents $\Delta_{j,j''}^{ATE}$ computed for those who reach the decision node involving the decision $D_{j,j''}$, while ATE† represents $\Delta_{j,j''}^{ATE}$ computed for the whole population. ATE (high) and

ATE (low) are the ATEs for different ability groups. The high (low) ability group is defined as those individuals with both cognitive and socio-emotional endowment above (below) the overall median. Finally, for each decision node, we display the fraction of individuals with low- and high-ability levels visiting each node.

Figures 15 and 16 show how the estimated treatment effect depends on the latent ability of the individuals for log wages and smoking. Final schooling levels are high-lighted using bold letters in the figures. For each educational decision node $(D_{j,j''})$, the first figure (top) presents $\Delta_{j,j''}^{ATE}\left(\theta \in (d^C, d^{SE})\right)$ where d^C and d^{SE} denote the cognitive and socio-emotional deciles computed from the marginal distributions of cognitive and socio-emotional endowments for the full population. $\Delta_{j,j''}^{ATE}\left(\theta \in (d^C, d^{SE})\right)$ is computed for those who reach the decision node $D_{j,j''}$. The second figure (bottom left) presents $\Delta_{j,j''}^{ATE}\left(\theta^C \in d^C\right)$ so that the socio-emotional factor is integrated out. The bars in this figure display the fraction of individuals visiting the node in each decile of cognitive endowment. The last figure (bottom right) presents $\Delta_{j,j''}^{ATE}\left(\theta^{SE} \in d^{SE}\right)$ and the fraction of individuals visiting the node in a given decile of socio-emotional endowment. We find that:

1. Labor Market Outcomes: As in the previous case, GED does not have any statistically significant effects on labor market outcomes. Graduating for high school significantly increases the probability of labor force participation, while further education does not have an impact. As expected, there are large gains from college in the probability of white-collar employment, and only high-ability people benefit from a four-year college degree. Although in general higher educational attainment results in gains in wages (both at age 30 and in present value terms), low-ability individuals gain very little from getting a four-year college degree.

2. Physical Health Outcomes: GED does not have significant effects on physical health. There are large and significant effects of high school and college on smoking, where the returns are homogeneous in ability. Physical health (PCS-12) is improved by graduating from high school, and there are stastistically significant returns to graduating from a four-year college. Both graduating from high school

28

and enrolling in college decrease the probability that a high-ability individual will drink heavily, although the effect is not strongly significant. Graduating from high school increases the likelihood that a low-ability individual will be obese, and enrolling in college decreases the likelihood that a high-ability individual will be obese.

3. Mental Health Outcomes: Enrolling in college and graduating from a four-year college both causally increase an individual's self-esteem, where the effect is larger for low-ability people and statistically insignificant for those with high-ability. The GED and college enrollment both have a positive effect sense of control. The effect for college is only statistically significant for low-ability people. Graduating from high school and enrolling in college both have marginally significant, positive effects on depression. There is no statistically significant effect of education on mental health (MCS-12).

4. Social Outcomes: Both high school and college attainment reduce the likelihood of being on welfare, while the GED seems to increase the use of welfare. As before, high school and college achievement have very strong effects on the likelihood of voting. While the effects of education on trust are not statistically significant, graduating from high school and getting a four-year college degree decreases the likelihood of getting a divorce.

6.7 Treatment Effects: Continuation Values in the Choice to Graduate from High School or Enroll in College

One benefit of schooling is access to further schooling.[20] Specifically, the choice to graduate from high school and the choice to enroll in college open up the doors for continued education. The continuation value of an educational choice is the probability of additional education times the wage benefits of that additional education. For high-ability individuals, the benefits of college may be large, and the probability of attending

---

[20]See Weisbrod (1962) and Comay, Melnik, and Pollatschek (1973).

may be near unity. For such individuals, the continuation value of graduating from high school may constitute the bulk of the return to graduating from high school. For others, the probability or benefit of college may be much lower. Figures 17–19 show that the total benefit by decision node for graduating from college and enrolling in college as well as the continuation value for labor market outcomes, health outcomes, and social outcomes.

Each figure presents the average effects of education on the outcome of interest. The figure plots a variety of treatment effects, defined in the following way: $ATE^{\dagger}$ — the average treatment effect defined using the characteristics of the entire population; ATE — the average treatment effect using the characteristics of the population, who are at, or passed through, the designated decision; ATE (low) and ATE (high) are defined in a corresponding way for low- and high-ability individuals; TT — treatment on the treated are defined for persons who are at, or pass through, this decision node; TUT — treatment on the untreated for people who are at, or pass through, this decision node; and AMTE — the average marginal treatment effect are defined for people approximately indifferent between going on or stopping at each decision mode. Our main results are as follows:

1. Labor Market Outcomes: The continuation value accounts for over half of the ATE from graduating from high school on log wages. While the total effect is relatively constant across treatment effects and ability levels, low-ability individuals benefit through the direct effect of being a high school graduate. Alternatively, for high-ability individuals and for TT, the continuation value produces almost the entire benefit of graduating from high school. For the probability of white-collar employment, much of the benefit for both high school and "some college" is from the continuation value. The majority of the benefit on labor force participation from graduating from high school is due to direct benefit. (See Figure 17.)

2. Physical Health Outcomes: Continuation value accounts for a portion of the decrease in the probability of not smoking from graduating from high school. However, continuation value accounts for more of the benefit of enrolling in college.

30

Similarly, continuation value accounts for little of the physical health benefits from graduating from high school. (See Figure 18.)

3. Mental Health Outcomes: The improvements in self-esteem and self-mastery from enrolling in college are explained almost completely by the direct effect. (See Figure 18.)

4. Social Outcomes: The majority of the reduction in welfare use comes from the direct benefit of graduating from high school. Continuation value also plays a role in the benefit of both high school and enrolling in college for voting. There is little continuation value in the reduction in the probability of divorce for graduating from high school. The continuation value of some college for both voting and divorce varies by ability. (See Figure 19.)

7   Conclusions

This paper formulates and estimates a dynamic sequential model of educational choices with unobserved heterogeneity. We use the model to define and estimate a variety of novel treatment effects, including treatment effects that account for the continuation values associated with sequential educational choices. We analyze the causal impact of education on health, social, and labor market outcomes when responses to treatment vary among observationally identical persons who select into schooling levels on the basis of their heterogeneous responses. To control for selection bias, we invoke conditional independence among later life outcomes and schooling conditional on observables and unobservables. We proxy the unobservables using numerous measurements and adjust for the measurement error arising from using proxies. Our methodology can be interpreted as a form of matching.

Our empirical results show that there is strong sorting into schooling levels on both cognitive and noncognitive abilities. We estimate both traditional treatment effects comparing outcomes across final schooling levels and node-specific treatment effects that include continuation values. We find that the causal effect of schooling differs by

ability level. In general, observed differences by educational attainment diminish when we control observables and latent abilities. There is significant heterogeneity in the gains from education. In most cases, the gain from education increases with the level of attained schooling.

We show the benefits of estimating a fully dynamic model of schooling that accounts for multiple levels of education and analyzes, in one framework, the returns to education for people at different margins of choice. We explore the channels through which education has its beneficial effects on a variety of outcomes.

References

AAKVIK, A., J. J. HECKMAN, AND E. J. VYTLACIL (2005): "Estimating Treatment Effects for Discrete Outcomes When Responses to Treatment Vary: An Application to Norwegian Vocational Rehabilitation Programs," *Journal of Econometrics*, 125(1–2), 15–51.

ABBRING, J. H., AND J. J. HECKMAN (2007): "Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 6B, pp. 5145–5303. Elsevier, Amsterdam.

ADAMS, S. J. (2002): "Educational Attainment and Health: Evidence from a Sample of Older Adults," *Education Economics*, 10(1), 97–109.

ALMLUND, M., A. DUCKWORTH, J. J. HECKMAN, AND T. KAUTZ (2011): "Personality Psychology and Economics," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Machin, and L. Wößmann, vol. 4. Elsevier, Amsterdam, Forthcoming.

ARENDT, J. N. (2005): "Does education cause better health? A panel data analysis using school reforms for identification," *Economics of Education Review*, 24(2), 149–160.

ARKES, J. (2003): "Does Schooling Improve Adult Health?," Unrestricted Draft DRU-3051, RAND, Santa Monica, CA.

AULD, M. C., AND N. SIDHU (2005): "Schooling, Cognitive Ability and Health," *Health Economics*, 14(10), 1019–1034.

BECKER, G. S. (1964): *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. National Bureau of Economic Research, distributed by Columbia University Press, New York.

33

BORGHANS, L., B. H. H. GOLSTEYN, J. J. HECKMAN, AND J. E. HUMPHRIES (2011): "Identification Problems in Personality Psychology," *Personality and Individual Differences*, 51(Special Issue on Personality and Economics), 315–320, E. Ferguson, J.J. Heckman, and P. Corr, editors.

CAMERON, S. V., AND J. J. HECKMAN (1998): "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106(2), 262–333.

———— (2001): "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy*, 109(3), 455–99.

CARD, D. (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69(5), 1127–1160.

CARNEIRO, P., K. HANSEN, AND J. 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(2), 361–422.

CARNEIRO, P., AND J. J. HECKMAN (2002): "The Evidence on Credit Constraints in Post-Secondary Schooling," *Economic Journal*, 112(482), 705–734.

CARNEIRO, P., J. J. HECKMAN, AND E. J. VYTLACIL (2010): "Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin," *Econometrica*, 78(1), 377–394.

CAWLEY, J., K. CONNEELY, J. J. HECKMAN, AND E. J. VYTLACIL (1997): "Cognitive Ability, Wages and Meritocracy," in *Intelligence, Genes, and Success : Scientists Respond to The Bell Curve*, ed. by B. Devlin, S. Fienberg, D. Resnick, and K. Roeder, pp. 179–192. Springer, New York.

COELLI, M. B., D. A. GREEN, AND W. P. WARBURTON (2007): "Breaking the Cycle?

The effect of education on welfare receipt among children of welfare recipients," *Journal of Public Economics*, 91(7-8), 1369–1398.

COMAY, Y., A. MELNIK, AND M. A. POLLATSCHEK (1973): "The Option Value of Education and the Optimal Path for Investment in Human Capital," *International Economic Review*, 14(2), 421–435.

CONARD, M. A. (2006): "Aptitude is not enough: How personality and behavior predict academic performance," *Journal of Research in Personality*, 40(3), 339–346.

CONTI, G., J. J. HECKMAN, AND S. URZUA (2010): "The Education-Health Gradient," *American Economic Review: Papers and Proceedings*, 100(2), 1–5.

CREDÉ, M., AND N. R. KUNCEL (2008): "Study Habits, Skills, and Attitudes: The Third Pillar Supporting Collegiate Academic Performance," *Prospectives on Psychological Science*, 3(6), 425–453.

CUTLER, D. M., AND A. LLERAS-MUNEY (2010): "Understanding Differences in Health Behaviors by Education," *Journal of Health Economics*, 29(1), 1–28.

DEVINS, G. M., C. M. ORME, C. G. COSTELLO, Y. M. BINIK, B. FRIZZELL, H. J. STAM, AND W. M. PULLIN (1988): "Measuring depressive symptoms in illness populations: Psychometric properties of the Center for Epidemiologic Studies Depression (CES-D) Scale," *Psychology and Health*, 2(2), 139–156.

DUCKWORTH, A., AND S. URZUA (2009): "Determinants of Success in Early Adulthood: Comparing the Effects of Intelligence and Big Five Personality Traits," Unpublished manuscript, University of Pennsylvania, Department of Psychology.

DUCKWORTH, A. L., P. D. QUINN, AND E. TSUKAYAMA (2010): "What No Child Left Behind Leaves Behind: The Roles of IQ and Self-Control in Predicting Standardized Achievement Test Scores and Report Card Grades," Under review, *Journal of Educational Psychology*.

DUCKWORTH, A. L., AND M. E. P. SELIGMAN (2005): "Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents," *Psychological Science*, 16(12), 939–944.

ELIAS, J. J. (2005): "The Effects of Ability and Family Background on Non-Monetary Returns to Education," Ph.D. thesis, University of Chicago, Chicago, IL.

FARSIDES, T., AND R. WOODFIELD (2003): "Individual differences and undergraduate academic success: The roles of personality, intelligence, and application," *Personality and Individual Differences*, 34(7), 1225–1243.

GANDEK, B., J. E. WARE, N. K. AARONSON, G. APOLONE, J. B. BJORNER, J. E. BRAZIER, M. BULLINGER, S. KAASA, A. LEPLEGE, L. PRIETO, AND M. SULLI-VAN (1998): "Cross-Validation of Item Selection and Scoring for the SF-12 Health Survey in Nine Countries: Results from the IQOLA Project," *Journal of Clinical Epidemiology*, 51(11), 1171–1178.

GLEWWE, P. (2002): "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes," *Journal of Economic Literature*, 40(2), 436–482.

GOTTFREDSON, L. S., AND I. J. DEARY (2004): "Intelligence Predicts Health and Longevity, but Why?," *Current Directions in Psychological Science*, 13(1), 1–4.

GROSSMAN, M. (1972): "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80(2), 223–255.

——— (1975): "The Correlation Between Health and Schooling," in *Household Production and Consumption*, ed. by N. E. Terleckyj, pp. 147–211. Columbia University Press, New York.

——— (2000): "The Human Capital Model," in *Handbook of Health Economics*, ed. by A. J. Culyer, and J. P. Newhouse, vol. 1, pp. 347–408. Elsevier, Amsterdam.

———— (2006): "Education and Nonmarket Outcomes," in *Handbook of the Economics of Education*, ed. by E. Hanushek, and F. Welch, vol. 1, chap. 10, pp. 577–633. Elsevier, Amsterdam.

———— (2008): "The Relationship between Health and Schooling: Presidential Address," *Eastern Economic Journal*, 34(3), 281–292.

GROSSMAN, M., AND R. KAESTNER (1997): "Effects of Education on Health," in *The Social Benefits of Education*, ed. by J. R. Behrman, and N. Stacey, pp. 69–124. University of Michigan Press, Ann Arbor, MI.

GULLONE, E., AND S. MOORE (2000): "Adolescent risk-taking and the five-factor model of personality," *Journal of Adolescence*, 23(4), 393–407.

HAMPSON, S. E., AND H. S. FRIEDMAN (2008): "Personality and Health: A Lifespan Perspective," in *The Handbook of Personality: Theory and Research*, ed. by O. P. John, R. Robins, and L. Pervin, pp. 770–794. Guilford, New York, third edn.

HANSEN, K. T., J. J. HECKMAN, AND K. J. MULLEN (2004): "The Effect of Schooling and Ability on Achievement Test Scores," *Journal of Econometrics*, 121(1-2), 39–98.

HARTOG, J., AND H. OOSTERBEEK (1998): "Health, Wealth and Happiness: Why Pursue a Higher Education?," *Economics of Education Review*, 17(3), 245–256.

HECKMAN, J. J., L. J. LOCHNER, AND P. E. TODD (2006): "Earnings Equations and Rates of Return: The Mincer Equation and Beyond," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, and F. Welch, chap. 7, pp. 307–458. Elsevier, Amsterdam.

HECKMAN, J. J., AND S. NAVARRO (2007): "Dynamic Discrete Choice and Dynamic Treatment Effects," *Journal of Econometrics*, 136(2), 341–396.

HECKMAN, J. J., J. STIXRUD, AND S. URZUA (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24(3), 411–482.

37

HECKMAN, J. J., S. URZUA, AND E. J. VYTLACIL (2006): "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, 88(3), 389–432.

HECKMAN, J. J., AND E. J. VYTLACIL (2007): "Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 6B, pp. 4779–4874. Elsevier, Amsterdam.

HERRNSTEIN, R. J., AND C. A. MURRAY (1994): *The Bell Curve: Intelligence and Class Structure in American Life*. Free Press, New York.

JOHN, O. P., R. W. ROBINS, AND L. A. PERVIN (eds.) (2008): *Handbook of Personality: Theory and Research*. Guilford Press, New York, NY, third edn.

KAESTNER, R. (2008): "Adolescent Cognitive and Non-cognitive Correlates of Adult Health," Unpublished manuscript, Institute of Government and Public Affairs, University of Illinois at Chicago, Department of Economics.

KEANE, M. P., AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

KENKEL, D. S., D. R. LILLARD, AND A. D. MATHIOS (2006): "The roles of high school completion and GED receipt in smoking and obesity," *Journal of Labor Economics*, 24(3), 635–660.

LLERAS-MUNEY, A. (2005): "The Relationship Between Education and Adult Mortality in the United States," *Review of Economic Studies*, 72(1), 189–221.

MILLIGAN, K., E. MORETTI, AND P. OREOPOULOS (2004): "Does education improve citizenship? Evidence from the United States and the United Kingdom," *Journal of Public Economics*, 88(9-10), 1667–1695.

MINCER, J. (1958): "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, 66(4), 281–302.

——— (1974): *Schooling, Experience and Earnings.* Columbia University Press for National Bureau of Economic Research, New York.

NEAL, D. A., AND W. R. JOHNSON (1996): "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy*, 104(5), 869–895.

NOFTLE, E. E., AND R. W. ROBINS (2007): "Personality Predictors of Academic Outcomes: Big Five Correlates of GPA and SAT Scores," *Personality Processes and Individual Differences*, 93(1), 116–139.

PEARLIN, L. I., E. G. MENAGHAN, M. A. LIEBERMAN, AND J. T. MULLAN (1981): "The Stress Process," *Journal of Health and Social Behavior*, 22(4), 337–356.

PEARLIN, L. I., AND C. SCHOOLER (1978): "The Structure of Coping," *Journal of Health and Social Behavior*, 19(1), 2–21.

RADLOFF, L. S. (1977): "The CES-D Scale," *Applied Psychological Measurement*, 1(3), 385–401.

SHAKOTKO, R. A., L. N. EDWARDS, AND M. GROSSMAN (1982): "An Exploration of the Dynamic Relationship between Health and Cognitive Development in Adolescence," Discussion paper, National Bureau of Economic Research.

SILLES, M. A. (2009): "The causal effect of education on health: Evidence from the United Kingdom," *Economics of Education Review*, 28(1), 122–128.

SPASOJEVIC, J. (2003): "Effects of Education on Adult Health in Sweden: Results from a Natural Experiment," Unpublished manuscript, City University of New York Graduate Center.

URZUA, S. (2008): "Racial Labor Market Gaps: The Role of Abilities and Schooling Choices," *Journal of Human Resources*, 43(4), 919–971.

WARE, JOHN E., J., M. KOSINSKI, AND S. D. KELLER (1996): "A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity," *Medical Care*, 34(3), 220–233.

WEISBROD, B. A. (1962): "Education and Investment in Human Capital," *Journal of Political Economy*, 70(5, Part 2: Investment in Human Beings), 106–123.

WHALLEY, L. J., AND I. J. DEARY (2001): "Longitudinal cohort study of childhood IQ and survival up to age 76," *British Medical Journal*, 322(7290), 819.

WILLIS, R. J., AND S. ROSEN (1979): "Education and Self-Selection," *Journal of Political Economy*, 87(5, Part 2), S7–S36.

Table 1: Summary of Decisions

Decision node	$D_{j,j''} = 1$	$D_{j,j''} = 0$	Conditional on
$D_{0,1}$	Graduate High School	Drop out of High School	—
$D_{0,2}$	**Get GED** $(s=2)$	**High School Dropout** $(s=0)$	$D_{0,1} = 0$
$D_{1,3}$	Attend College	**High School Graduate** $(s=1)$	$D_{0,1} = 1$
$D_{3,4}$	**Graduate 4-yr college** $(s=4)$	**Some College** $(s=3)$	$D_{1,3} = 1$

Note: Final schooling levels (s) are highlighted in bold letters.

Table 2: Early Outcomes: Estimates for Participation in Violent Behaviors during 1979, by Schooling at the Time of the Test

	Early Violent			
	<12 yrs		≥12 yrs	
	β	Std. Error	β	Std. Error
Black	-0.260	0.124	0.140	0.158
Hispanic	-0.346	0.157	-0.022	0.198
Urban Area (14)	0.184	0.090	0.091	0.103
South (14)	-0.091	0.085	0.027	0.102
Broken Home	0.200	0.094	0.120	0.116
Number of Siblings	0.010	0.018	0.013	0.021
Mother's Education	0.031	0.019	-0.032	0.023
Father's Education	-0.037	0.015	0.007	0.016
Family Income	-0.005	0.004	-0.006	0.003
Age	-0.115	0.024	-0.058	0.035
College Attendance			-0.131	0.122
Intercept	2.511	0.508	1.586	0.797
Cognitive	-0.150	0.063	-0.225	0.073
Socio-emotional	-0.481	0.077	-0.269	0.088

Notes: The numbers in this table represent the estimated coefficients and Std. Errors associated with binary choice models of early reckless behaviors on the set of controls presented in rows. The variable "Early Violent" takes a value of one if the individual participated in any of the following criminal activities in 1979: Fighting or Assault.

Table 3: Early Outcomes: Estimates for "Early Risky Behaviors" (Before Age 15)

Variable	Tried Marijuana		Daily Smoking		Regular Drinking		Any Intercourse	
	β	Std. Error	β	Std. Error	β	Std. Error	β	Std. Error
Black	-0.323	0.100	-0.340	0.112	-0.244	0.108	0.594	0.099
Hispanic	-0.170	0.125	-0.511	0.150	-0.017	0.130	-0.046	0.140
Urban Area (14)	0.306	0.073	0.151	0.081	0.120	0.077	0.252	0.087
South (14)	-0.094	0.067	-0.004	0.075	0.077	0.071	0.126	0.076
Broken Home	0.419	0.073	0.416	0.081	0.234	0.077	0.362	0.081
Number of Siblings	0.030	0.014	0.034	0.015	0.029	0.015	0.012	0.016
Mother's Education	0.010	0.015	-0.022	0.017	0.000	0.016	-0.023	0.017
Father's Education	-0.011	0.011	-0.037	0.013	-0.004	0.012	-0.028	0.013
Family Income	0.001	0.003	-0.002	0.003	-0.001	0.003	-0.003	0.003
Age	-0.089	0.014	0.025	0.015	-0.022	0.014	-0.006	0.016
Intercept	0.889	0.316	-0.986	0.360	-0.667	0.335	-0.774	0.369
Cognitive	-0.103	0.048	-0.207	0.054	-0.134	0.052	-0.264	0.057
Socio-emotional	-0.609	0.059	-0.519	0.064	-0.285	0.060	-0.408	0.065

Notes: The numbers in this table represent the estimated coefficients and Std. Errors associated with binary choice models of early risky behaviors on the set of controls presented in rows. In each case, the dependent variable takes a value of one if the individual has reported the behavior before age 15, and zero otherwise.

Table 4: Estimates for Schooling Choice Model

Variable	$D_{0,1}$: Graduate HS vs. **Drop out from HS**		$D_{0,2}$: **Get GED** vs. **HS Dropout**		$D_{1,3}$: College Enrollment vs. **HS Graduate**		$D_{3,4}$: **4-year college degree** vs. **Some College**	
	β	StdEr.	β	StdEr.	β	StdEr.	β	StdEr.
Black	0.168	0.129	-0.039	0.173	0.177	0.140	0.072	0.199
Hispanics	0.678	0.176	0.106	0.245	0.701	0.169	0.336	0.256
Urban Area (14)	-0.350	0.103	-0.012	0.162	-0.048	0.090	-0.008	0.125
South (14)	-0.400	0.100	0.066	0.138	0.154	0.089	-0.012	0.125
Broken Home	-0.456	0.100	-0.215	0.140	-0.054	0.101	-0.313	0.141
Number of Siblings	-0.047	0.019	-0.002	0.027	-0.059	0.019	-0.031	0.027
Mother's Education	0.116	0.022	0.086	0.032	0.096	0.020	0.095	0.027
Father's Education	0.076	0.016	0.045	0.025	0.127	0.015	0.104	0.019
Family Income	0.021	0.005	0.017	0.008	0.011	0.004	0.013	0.004
Age in 1980	0.037	0.019	-0.056	0.030	0.009	0.017	0.000	0.024
Intercept	-0.991	0.480	-0.144	0.720	-3.401	0.432	-3.243	0.616
Δ Local wage of (hsg-hsd)[a]	-0.276	0.086						
Δ Local unemp. of (hsg-hsd)[a]	4.551	2.373						
Δ Local wage of (scol-hsg)[a]					0.184	0.074		
Δ Local unemp. of (scol-hsg)[a]					-2.875	2.765		
Δ Local wage of (colgrad-scol)[a]							0.037	0.052
Δ Local unemp. of (colgrad-scol)[a]							0.335	3.603
Tuition at local 4-yr college[a]							0.030	0.070
GED Cost			0.001	0.004				
Cognitive	0.791	0.085	1.035	0.138	0.814	0.083	0.858	0.107
Socio-emotional	0.970	0.088	0.164	0.142	0.484	0.078	0.567	0.111

Notes :The numbers in this table represent the estimated coefficients and Std. Errors associated with individual binary choice models of the sequential education model. Terminal schooling levels are highlighted in bold. (a) The local wage and unemployment variables are measured as the difference between the local wages and unemployment for the different schooling levels. They are measured when the individual was 17 years old. (hsd=High School Dropout, hsg = High School Graduate, scol = Some College, colgrad = 4-year college grad)

Figure 1: Sequential model for schooling decisions.
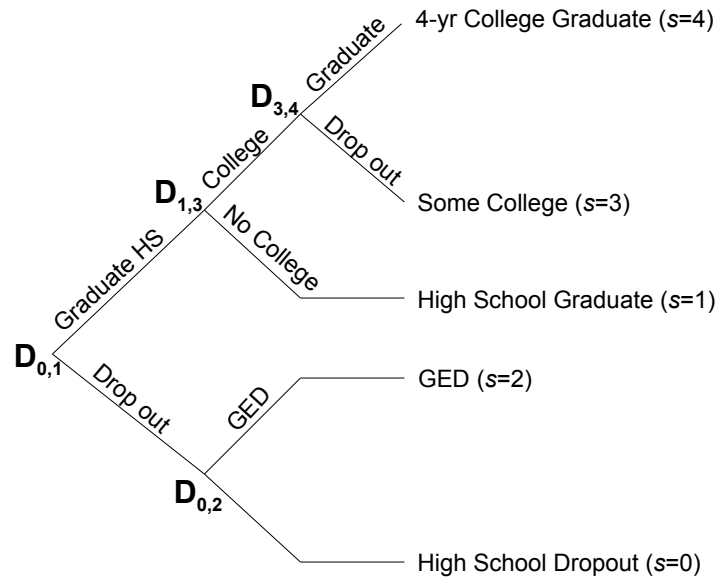
Figure 2: Distribution of Cognitive and Socio-emotional Endowments

Distribution of Factors

Cognitive Factor

Social-Emotional Factor
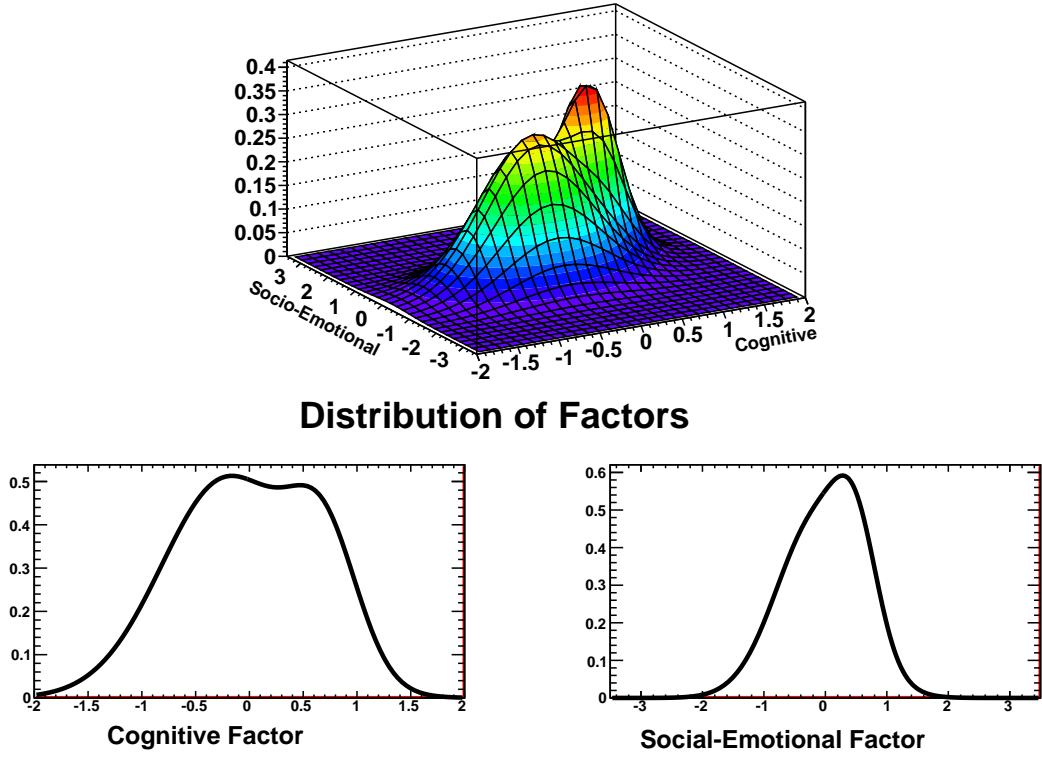
$$\begin{bmatrix} \theta^C \\ \theta^{SE} \end{bmatrix} \sim p_1 \Phi\left(\mu_1, \mathbf{\Sigma}_1\right) + p_2 \Phi\left(\mu_2, \mathbf{\Sigma}_2\right)$$

where

$$\mathbf{\Sigma}_1 = \begin{pmatrix} 0.10 & 0 \\ 0 & 0.12 \end{pmatrix}, \quad \mathbf{\Sigma}_2 = \begin{pmatrix} 0.37 & 0 \\ 0 & 0.43 \end{pmatrix},$$

$$\mu_1 = \begin{pmatrix} 0.70 \\ 0.50 \end{pmatrix}, \quad \mu_2 = \begin{pmatrix} -0.21 \\ -0.15 \end{pmatrix}$$

$$\mathbf{p} = (0.23, \ 0.77)$$

$$\rho = 0.24$$

Figure 3: Decomposing Variances in the Measurement System
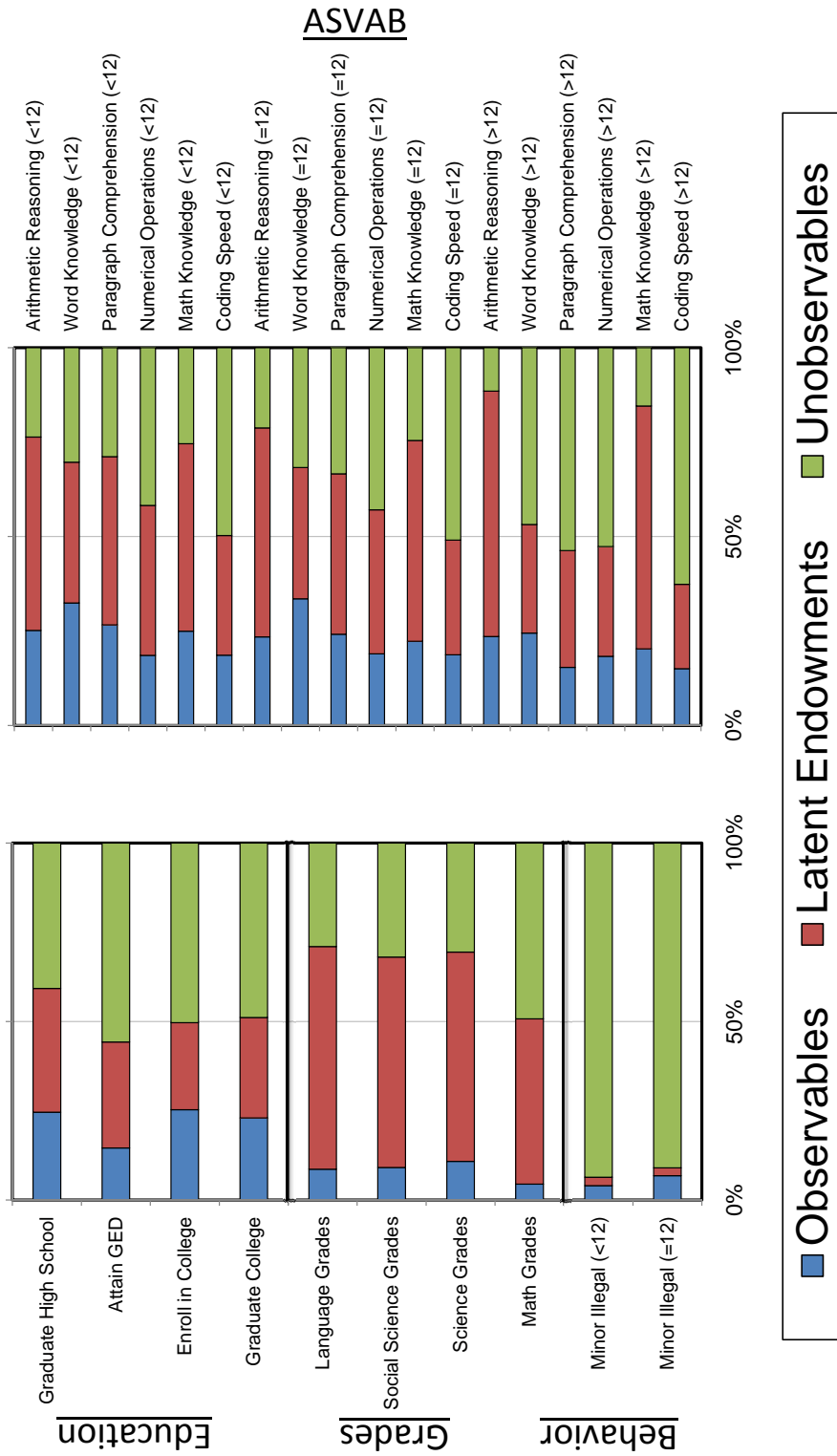
Note: Bars indicate the fraction of variance explained by each term of the ASVAB, GPA, and behavior models. The components are the observables ($\mathbf{X}'\beta$), latent endowments ($\alpha\theta$), and unobservables (e). The numbers inside the parenthesis describe the years of schooling at the time of the test. The ASVAB and behavior models are estimated separately for those with less than twelve years (<12), those who are high school graduates ($=12$), and those who have attended college (>12) at the time they took the ASVAB tests.

46

Figure 4: Distribution of factors by schooling level
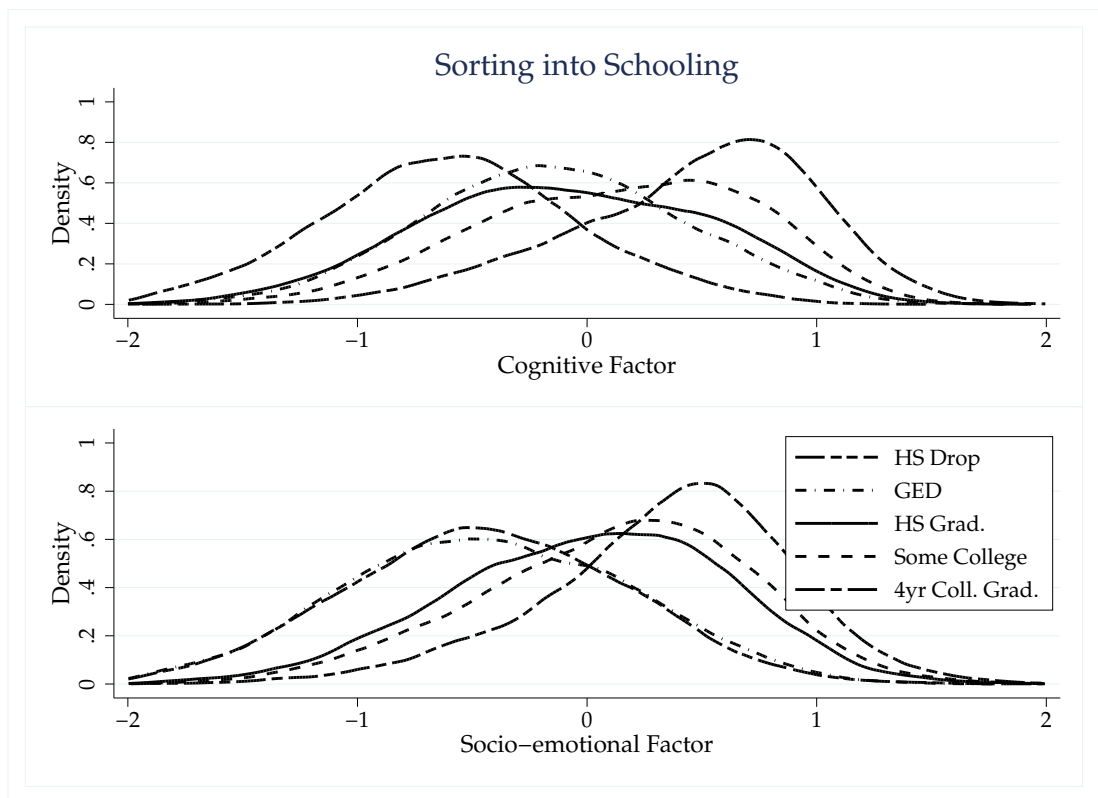
Note: The factors are simulated from the estimates of the model. The simulated data contain 1 million observations.

Figure 5: The Probability of Educational Decisions, by Endowment Levels

A. Dropping from HS vs. Graduating from HS ($D_{0,1}$)
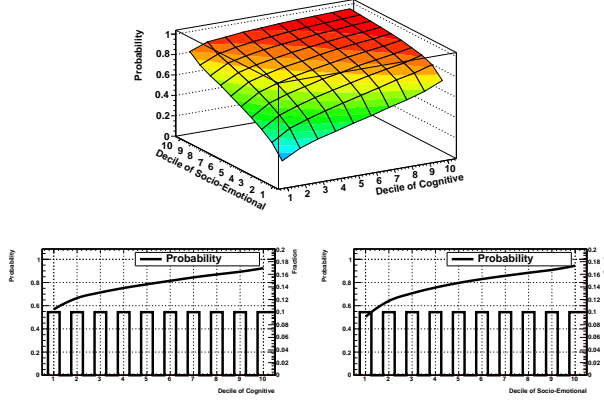
B. **HS Dropout** vs. Getting a GED ($D_{0,2}$)
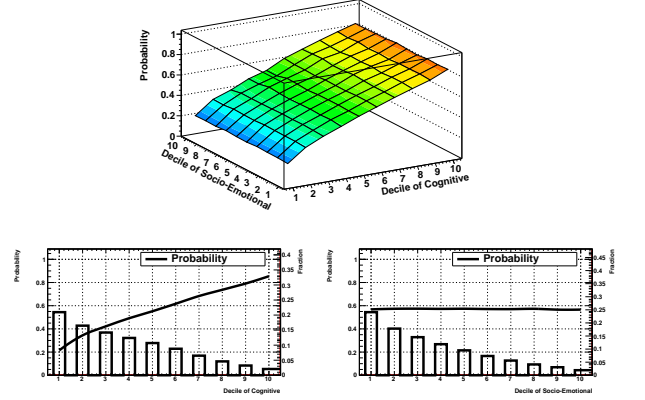
C. **HS Graduate** vs. College Enrollment ($D_{1,3}$)
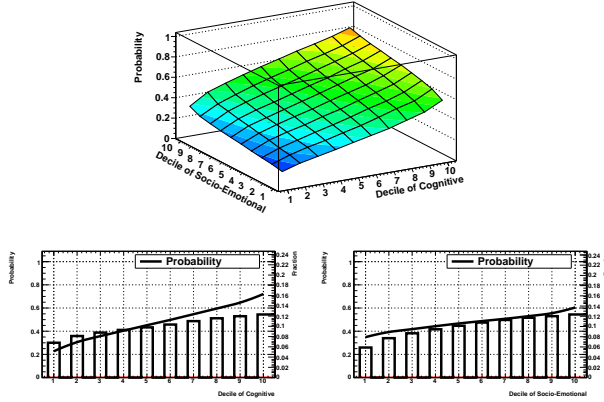
D. **Some College** vs. **4-year college degree** ($D_{3,4}$)
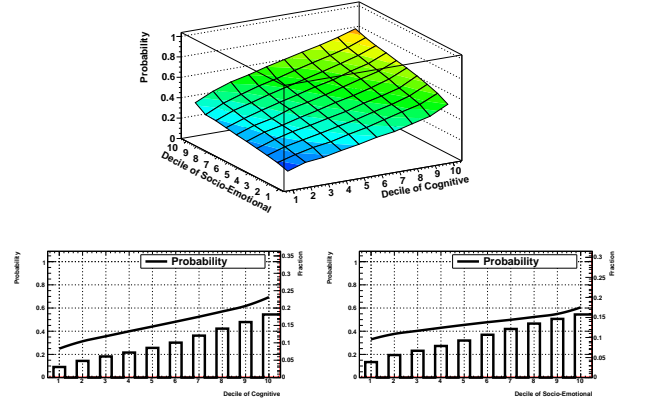
Notes: Each panel in this figure studies the average probability of each educational decision. Final schooling levels are highlighted using bold letters. For each pair of schooling levels j and j'', the first figure (top) presents $Prob(D_{j,j''}|d^C, d^{SE})$ where d^C and d^{SE} denote the cognitive and socio-emotional deciles computed from the marginal distributions of cognitive and socio-emotional endowments of the full population. $Prob(D_{j,j''}|d^C, d^{SE})$ is computed for those who reach the decision node ($D_{j,j''}$). The second figure (bottom left) presents $Prob(D_{j,j''}|d^C)$ so that the socio-emotional factor is integrated out. The bars in this figure display the fraction of individuals visiting the node in each decile of cognitive endowment. The last figure (bottom right) presents $Prob(D_{j,j''}|d^{SE})$ as well as the fraction of individuals visiting the node in each decile of socio-emotional endowment.

48

Figure 6: The Effect of Cognitive and Socio-emotional endowments

A. (log)Wages

B. Daily Smoking

C. Self-Esteem

D. Participated in 2006 election
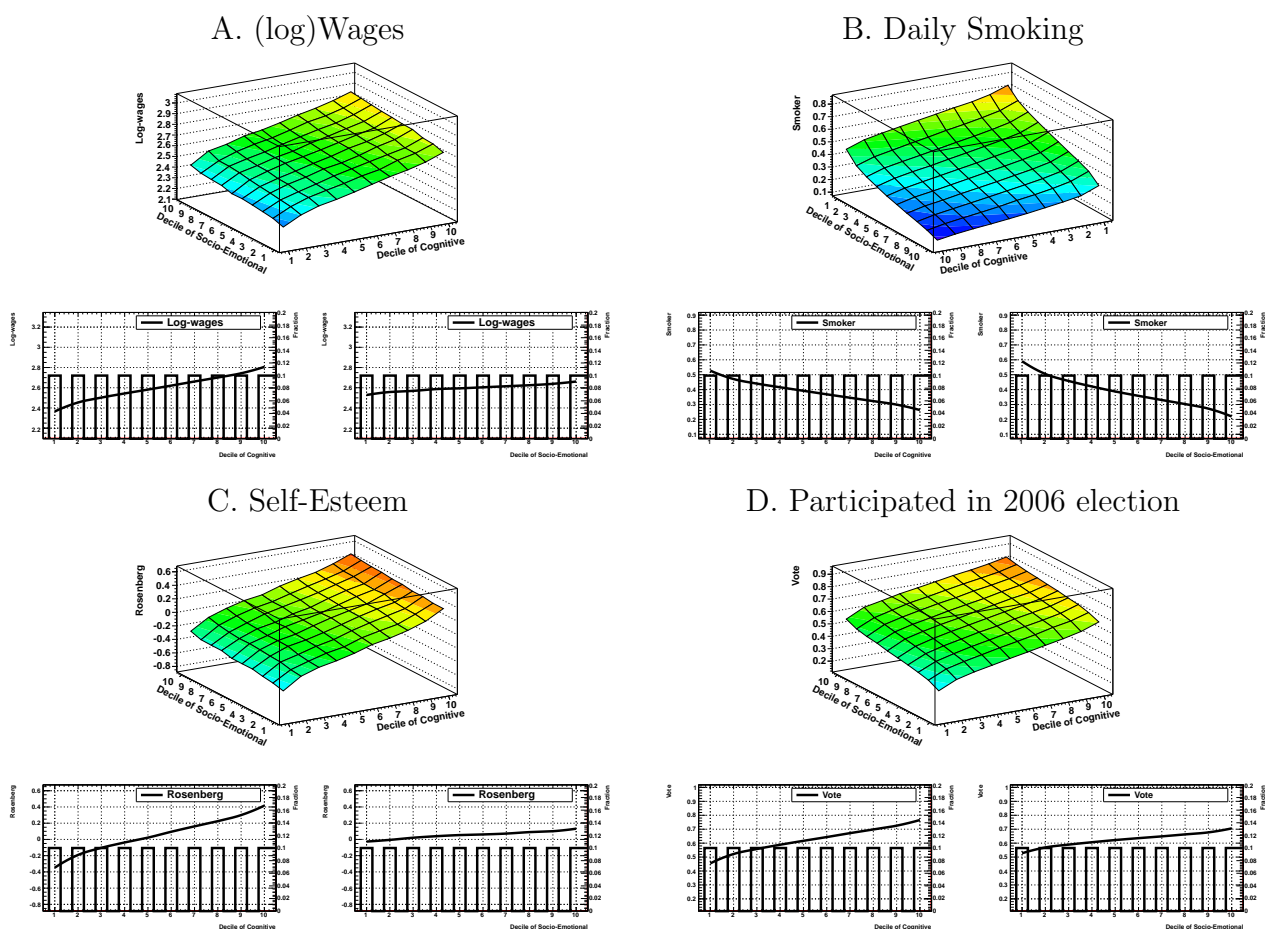
Note: For each outcome we present three figures. The first figure (top) displays the levels of the outcome as a function of cognitive and socio-emotional endowments. In particular, we present the average level of outcomes for different deciles of cognitive and socio-emotional endowments. Notice that we define as "decile 1" the decile with the lowest values of endowments and "decile 10" as the decile with the highest levels of endowments. The second figure (bottom left) displays the average levels of endowment across deciles of cognitive endowments. The bars in this figure indicates the fraction of individuals reporting the respective schooling level for each decile of cognitive endowment. The last figure (bottom right) mimics the structure of the second one but now for the socio-emotional endowment.

49

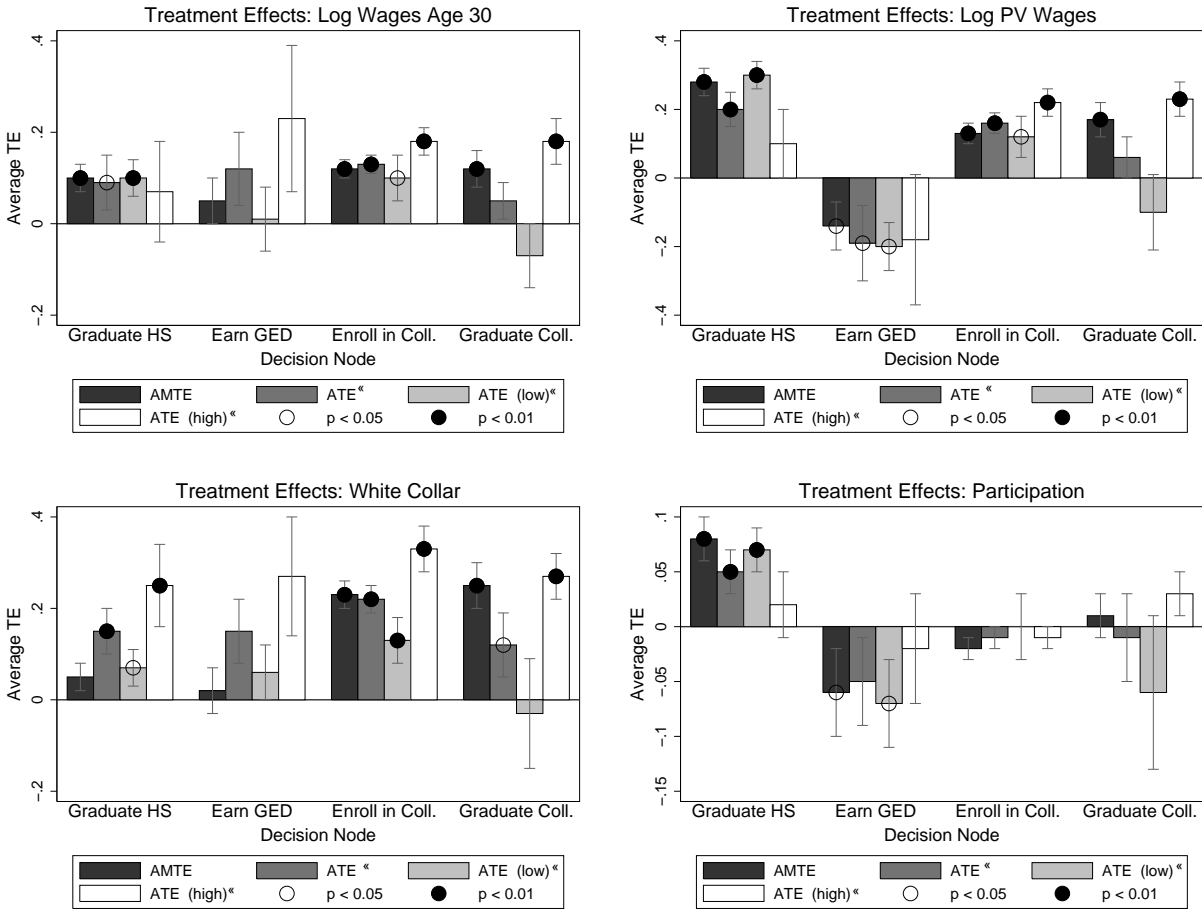**Figure 7:** Treatment Effects of Labor Market Outcomes by Final Schooling Levels

Notes: Each bar compares the outcomes from a particular final schooling level s and the HS dropout status. The "Observed" bar displays the observed differences in the data. The "Causal Mechanism" bar displays the estimated average treatment effect (ATE) obtained from the comparison of the outcomes associated with a particular final schooling level s relative to the HS dropout status. The ATE is calculated for those who have one of the final schooling levels being considered. The difference between the observed and causal treatment effect is attributed to the effect of selection and ability. The error bars and significance levels for the estimated ATE are calculated using 200 bootstrap samples.

Figure 8: Treatment Effects of Physical Health Outcomes and Behaviors by Final Schooling Levels

Notes: Each bar compares the outcomes from a particular final schooling level s and the HS dropout status. The "Observed" bar displays the observed differences in the data. The "Causal Mechanism" bar displays the estimated average treatment effect (ATE) obtained from the comparison of the outcomes associated with a particular final schooling level s relative to the HS dropout status. The ATE is calculated for those who have one of the final schooling levels being considered. The difference between the observed and causal treatment effect is attributed to the effect of selection and ability. The error bars and significance levels for the estimated ATE are calculated using 200 bootstrap samples.

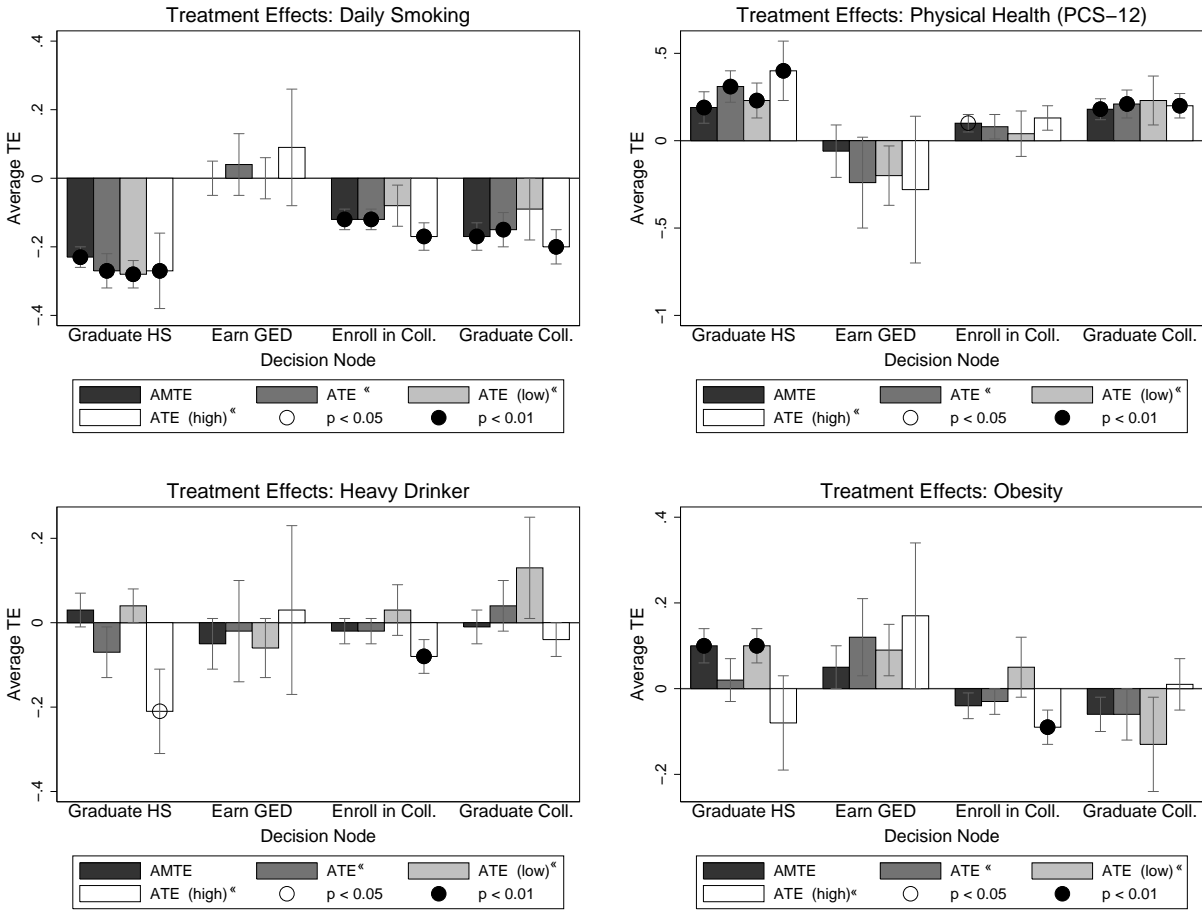**Figure 9:** Treatment Effects of Mental Health Outcomes by Final Schooling Levels

Notes: Each bar compares the outcomes from a particular final schooling level s and the HS dropout status. The "Observed" bar displays the observed differences in the data. The "Causal Mechanism" bar displays the estimated average treatment effect (ATE) obtained from the comparison of the outcomes associated with a particular final schooling level s relative to the HS dropout status. The ATE is calculated for those who have one of the final schooling levels being considered. The difference between the observed and causal treatment effect is attributed to the effect of selection and ability. The error bars and significance levels for the estimated ATE are calculated using 200 bootstrap samples.

Figure 10: Treatment Effects of Social Behaviors by Final Schooling Levels

Notes: Each bar compares the outcomes from a particular final schooling level s and the HS dropout status. The "Observed" bar displays the observed differences in the data. The "Causal Mechanism" bar displays the estimated average treatment effect (ATE) obtained from the comparison of the outcomes associated with a particular final schooling level s relative to the HS dropout status. The ATE is calculated for those who have one of the final schooling levels being considered. The difference between the observed and causal treatment effect is attributed to the effect of selection and ability. The error bars and significance levels for the estimated ATE are calculated using 200 bootstrap samples.

Figure 11: Treatment Effects of Labor Market Outcomes by Decision Node
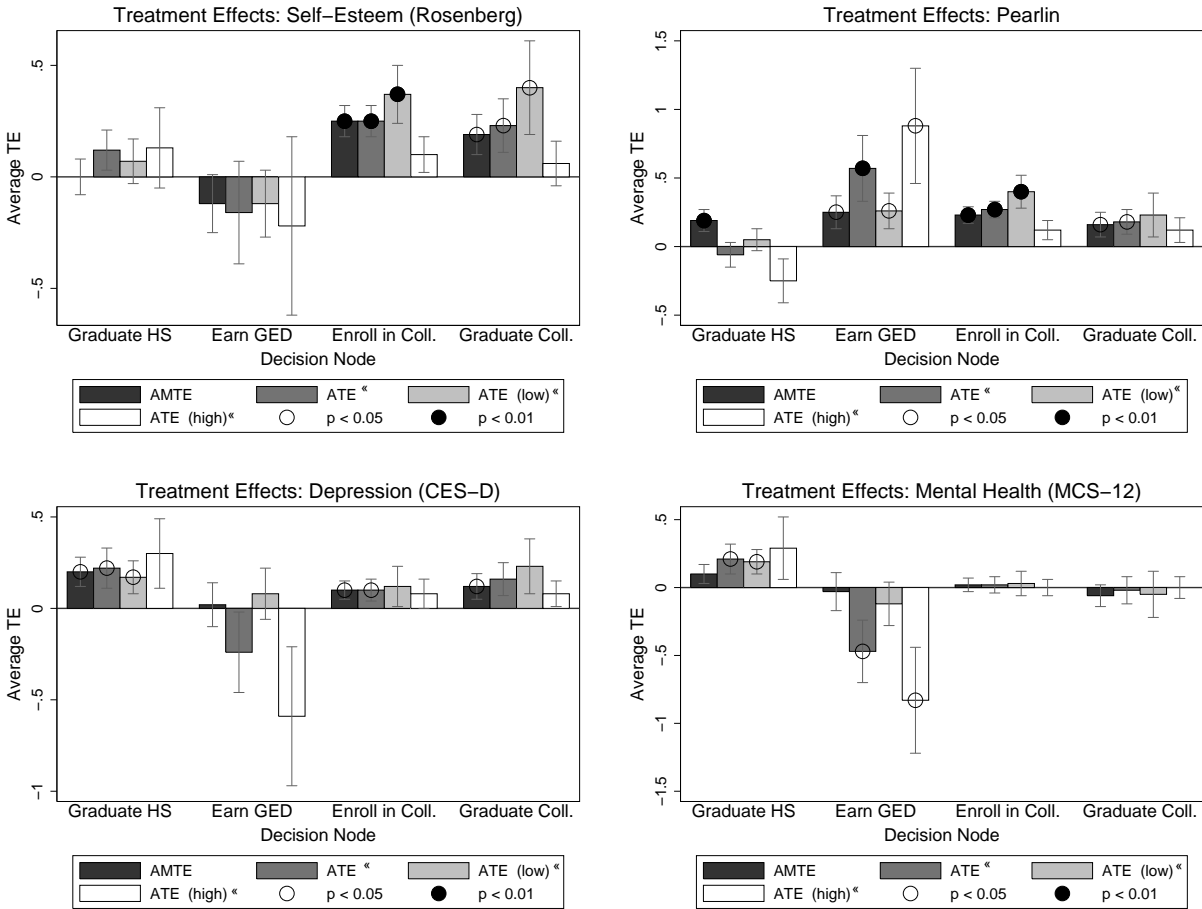

Notes: Each bar presents the average effect of an educational decision on the outcome of interest for the full population (ATE†). Importantly, each schooling level might provide the option to pursue higher schooling levels, while terminal schooling levels do not provide an option. The error bars and significance levels for the estimated ATE are calculated using 200 bootstrap samples. AMTE presents the average affect for those who are indifferent at that decision node ($|I_{j,j''}| < \varepsilon^s$). The figure also presents the estimated ATE conditional upon endowment levels. The high (low) ability group is defined as those individuals with cognitive and socio-emotional endowment above (below) the overall median. The fraction of individuals with low and high ability levels visiting each node are:

	Low Ability	High Ability
$D_{0,1}$: Dropping from HS vs. Graduating from HS	0.31	0.31
$D_{0,2}$: **HS Dropout** vs. **Getting a GED**	0.61	0.06
$D_{1,3}$: **HS Graduate** vs. College Enrollment	0.22	0.38
$D_{3,4}$: **Some College** vs. **4-year college degree**	0.14	0.51

In this table, final schooling levels are highlighted using bold letters.

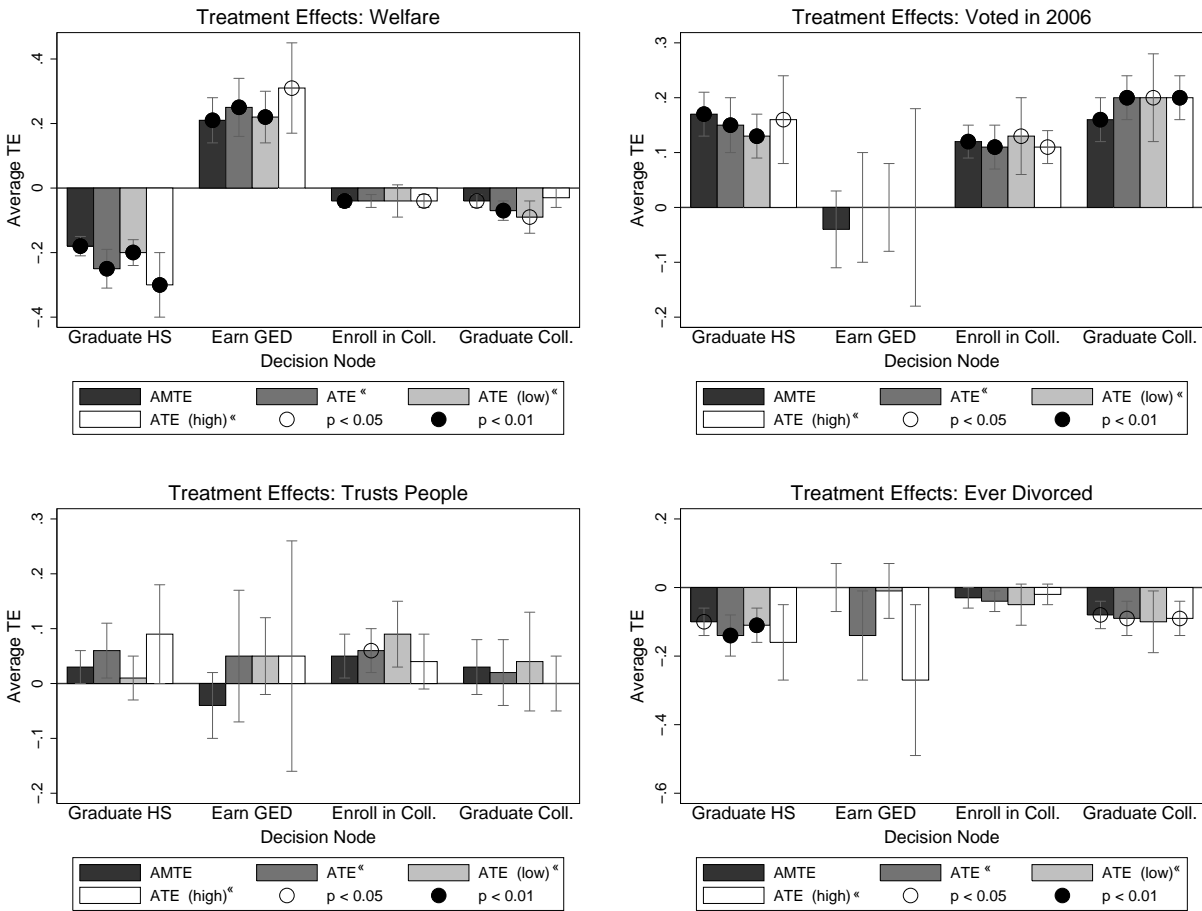**Figure 12:** Treatment Effects of Physical Health Outcomes and Behaviors by Decision Node

Notes: Each bar presents the average effect of an educational decision on the outcome of interest for the full population (ATE†). Importantly, each schooling level might provide the option to pursue higher schooling levels, while terminal schooling levels do not provide an option. The error bars and significance levels for the estimated ATE are calculated using 200 bootstrap samples. AMTE presents the average affect for those who are indifferent at that decision node ($|I_{j,j''}| < \varepsilon^s$). The figure also presents the estimated ATE conditional upon endowment levels. The high (low) ability group is defined as those individuals with cognitive and socio-emotional endowment above (below) the overall median. The fraction of individuals with low and high ability levels visiting each node are:

	Low Ability	High Ability
$D_{0,1}$: Dropping from HS vs. Graduating from HS	0.31	0.31
$D_{0,2}$: **HS Dropout** vs. **Getting a GED**	0.61	0.06
$D_{1,3}$: **HS Graduate** vs. College Enrollment	0.22	0.38
$D_{3,4}$: **Some College** vs. **4-year college degree**	0.14	0.51

In this table, final schooling levels are highlighted using bold letters.

**Figure 13:** Treatment Effects of Mental Health Outcomes by Decision Node

Notes: Each bar presents the average effect of an educational decision on the outcome of interest for the full population (ATE[†]). Importantly, each schooling level might provide the option to pursue higher schooling levels, while terminal schooling levels do not provide an option. The error bars and significance levels for the estimated ATE are calculated using 200 bootstrap samples. AMTE presents the average affect for those who are indifferent at that decision node ($|I_{j,j''}| < \varepsilon^s$). The figure also presents the estimated ATE conditional upon endowment levels. The high (low) ability group is defined as those individuals with cognitive and socio-emotional endowment above (below) the overall median. The fraction of individuals with low and high ability levels visiting each node are:

	Low Ability	High Ability
$D_{0,1}$: Dropping from HS vs. Graduating from HS	0.31	0.31
$D_{0,2}$: **HS Dropout** vs. **Getting a GED**	0.61	0.06
$D_{1,3}$: **HS Graduate** vs. College Enrollment	0.22	0.38
$D_{3,4}$: **Some College** vs. **4-year college degree**	0.14	0.51

In this table, final schooling levels are highlighted using bold letters.

Figure 14: Treatment Effects of Social Behaviors by Decision Node

Notes: Each bar presents the average effect of an educational decision on the outcome of interest for the full population (ATE[†]). Importantly, each schooling level might provide the option to pursue higher schooling levels, while terminal schooling levels do not provide an option. The error bars and significance levels for the estimated ATE are calculated using 200 bootstrap samples. AMTE presents the average affect for those who are indifferent at that decision node ($|I_{j,j''}| < \varepsilon^s$). The figure also presents the estimated ATE conditional upon endowment levels. The high (low) ability group is defined as those individuals with cognitive and socio-emotional endowment above (below) the overall median. The fraction of individuals with low and high ability levels visiting each node are:

	Low Ability	High Ability
$D_{0,1}$: Dropping from HS vs. Graduating from HS	0.31	0.31
$D_{0,2}$: **HS Dropout** vs. **Getting a GED**	0.61	0.06
$D_{1,3}$: **HS Graduate** vs. College Enrollment	0.22	0.38
$D_{3,4}$: **Some College** vs. **4-year college degree**	0.14	0.51

In this table, final schooling levels are highlighted using bold letters.

Figure 15: Average Treatment Effect of Education on (Log) Wages at Age 30, by *Decision Node* and Endowment Levels
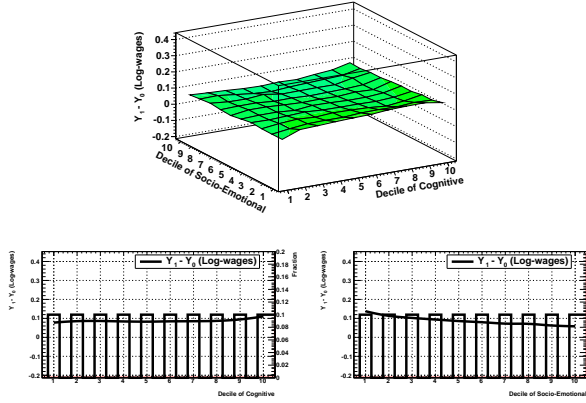
A. Dropping from HS vs. Graduating from HS

B. **HS Dropout** vs. Getting a GED
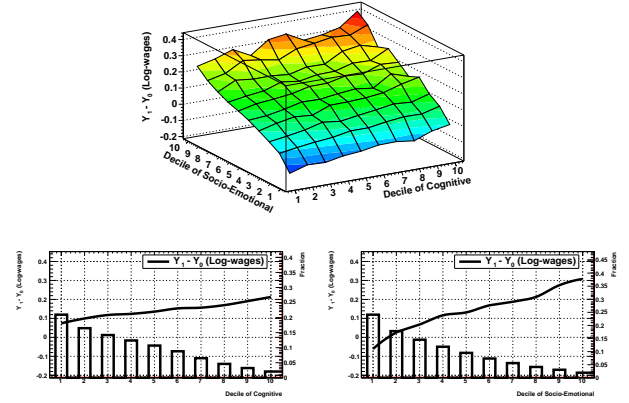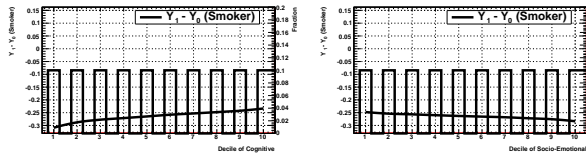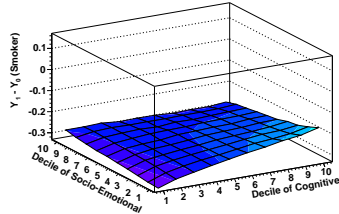
C. **HS Graduate** vs. College Enrollment
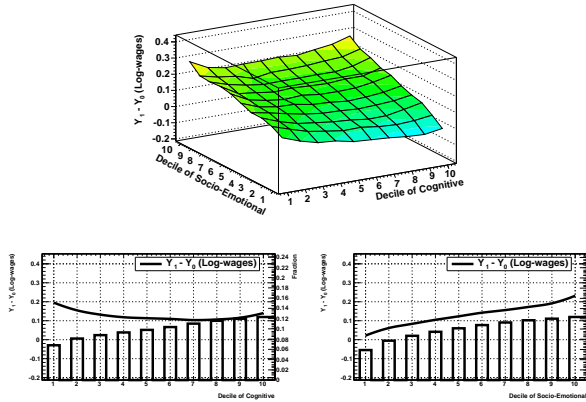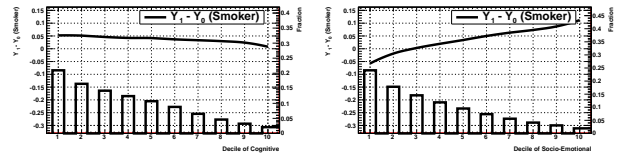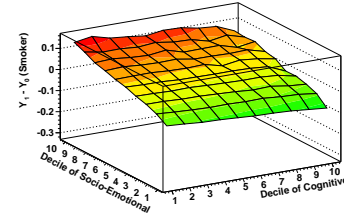
D. **Some College** vs. **4-year college degree**
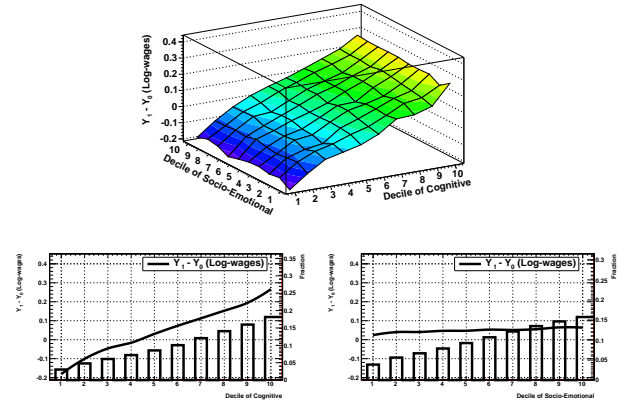
Notes: Each panel in this figure studies the average effect of an educational decision for those individuals visiting the decision node. Importantly, each schooling level might provide the option to pursue higher schooling levels, while final schooling levels do not provide an option. Final schooling levels are highlighted using bold letters. For each educational decision node, the first figure (top) presents $\Delta_{j,j''}^{ATE}\left(\theta \in (d^C, d^{SE})\right)$ where d^C and d^{SE} denote the cognitive and socio-emotional deciles computed from the marginal distributions of cognitive and socio-emotional endowments for the full population. The second figure (bottom left) presents $\Delta_{j,j''}^{ATE}\left(\theta^C \in d^C\right)$ so that the socio-emotional factor is integrated out. The bars in this figure display the fraction of individuals visiting the node in each decile of cognitive endowment. The last figure (bottom right) presents $\Delta_{j,j''}^{ATE}\left(\theta^{SE} \in d^{SE}\right)$ and the fraction of individuals visiting the node in a given decile of socio-emotional endowment.

Figure 16: Average Treatment Effect of Education on Probability of Being a Smoker, by *Decision Node* and Endowment Levels
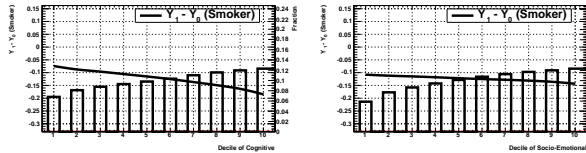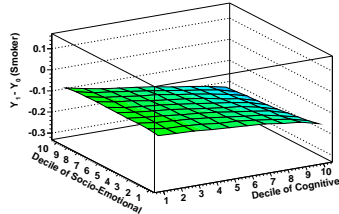
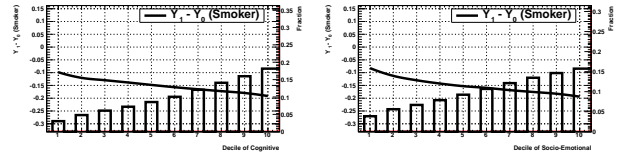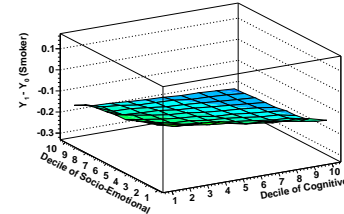A. Dropping from HS vs. Graduating from HS

B. **HS Dropout** vs. Getting a GED

C. **HS Graduate** vs. College Enrollment

D. **Some College** vs. **4-year college degree**

Notes: Each panel in this figure studies the average effect of an educational decision for those individuals visting the decision node. Importantly, each schooling level might provide the option to pursue higher schooling levels, while final schooling levels do not provide an option. Final schooling levels are highlighted using bold letters. For each educational decision node, the first figure (top) presents $\Delta_{j,j''}^{ATE}\left(\theta \in (d^C, d^{SE})\right)$ where d^C and d^{SE} denote the cognitive and socio-emotional deciles computed from the marginal distributions of cognitive and socio-emotional endowments for the full population. The second figure (bottom left) presents $\Delta_{j,j''}^{ATE}\left(\theta^C \in d^C\right)$ so that the socio-emotional factor is integrated out. The bars in this figure display the fraction of individuals visiting the node in each decile of cognitive endowment. The last figure (bottom right) presents $\Delta_{j,j''}^{ATE}\left(\theta^{SE} \in d^{SE}\right)$ and the fraction of individuals visiting the node in a given decile of socio-emotional endowment.

Figure 17: Treatment Effects: Direct and Indirect Components: Labor Market Outcomes
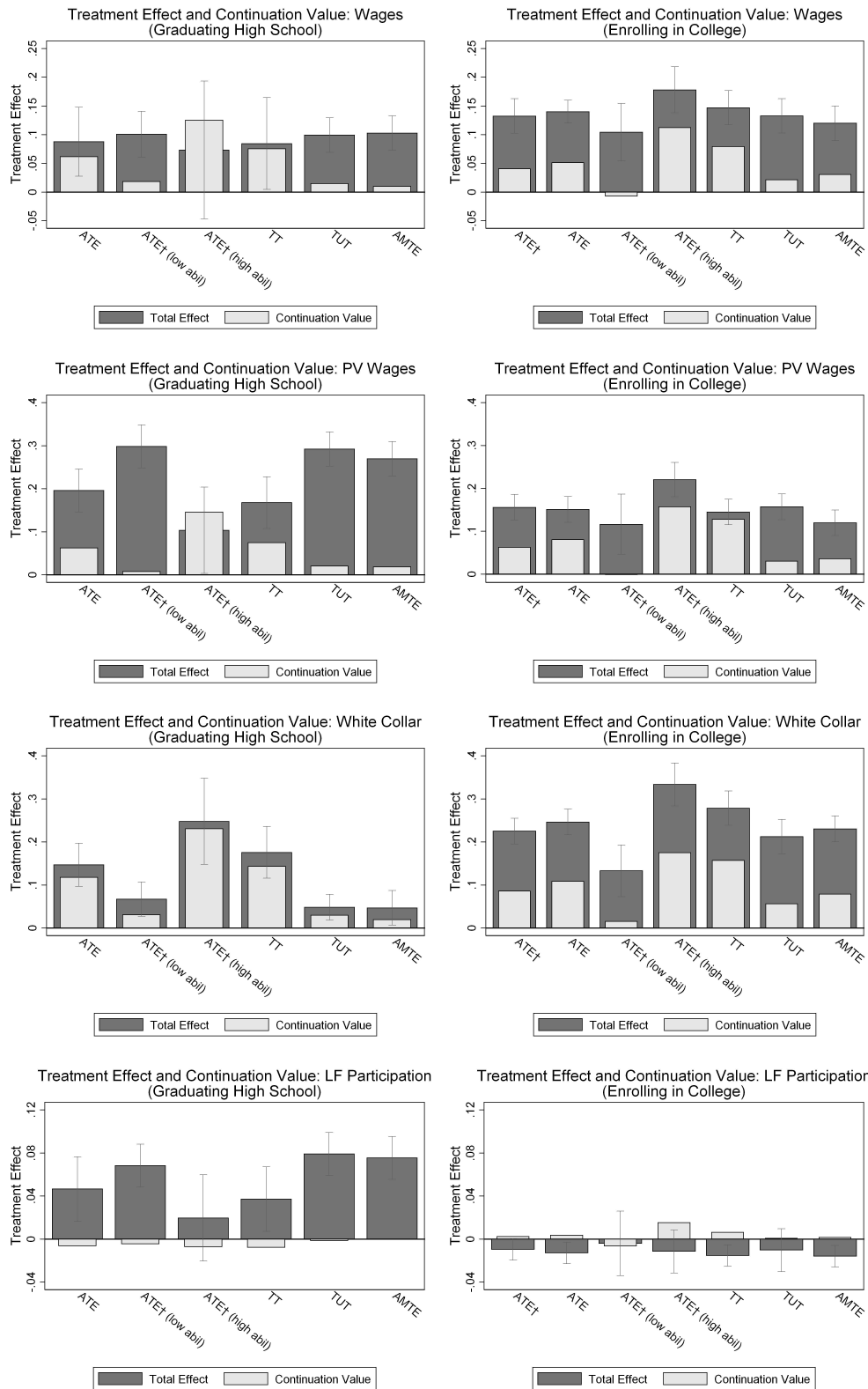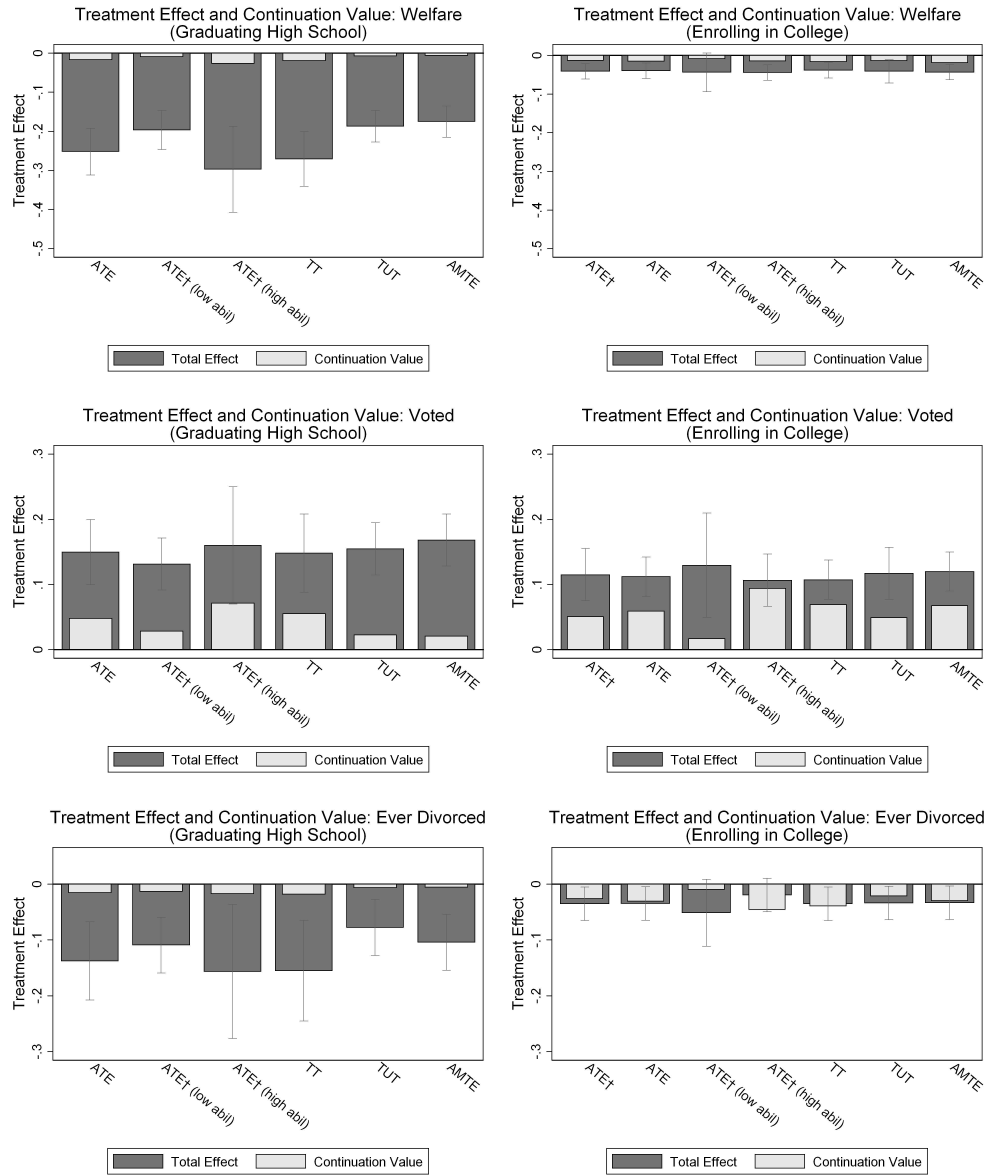

Total Effect is the complete decision specific treatment effect which includes access to further education. ATE† is for the entire population, while the remaining treatment effects are only for individuals who make the specific educational decision. Continuation Value is the additional benefit gained through the option of pursuing additional education. High ability individuals are those in the top 50% of the distributions of both cognitive and socioemotional endowments. Low-ability individuals are those in the bottom 50% of the distributions of both cognitive and socioemotional endowments.

Figure 18: Treatment Effects: Direct and Indirect Components: All Health Outcomes
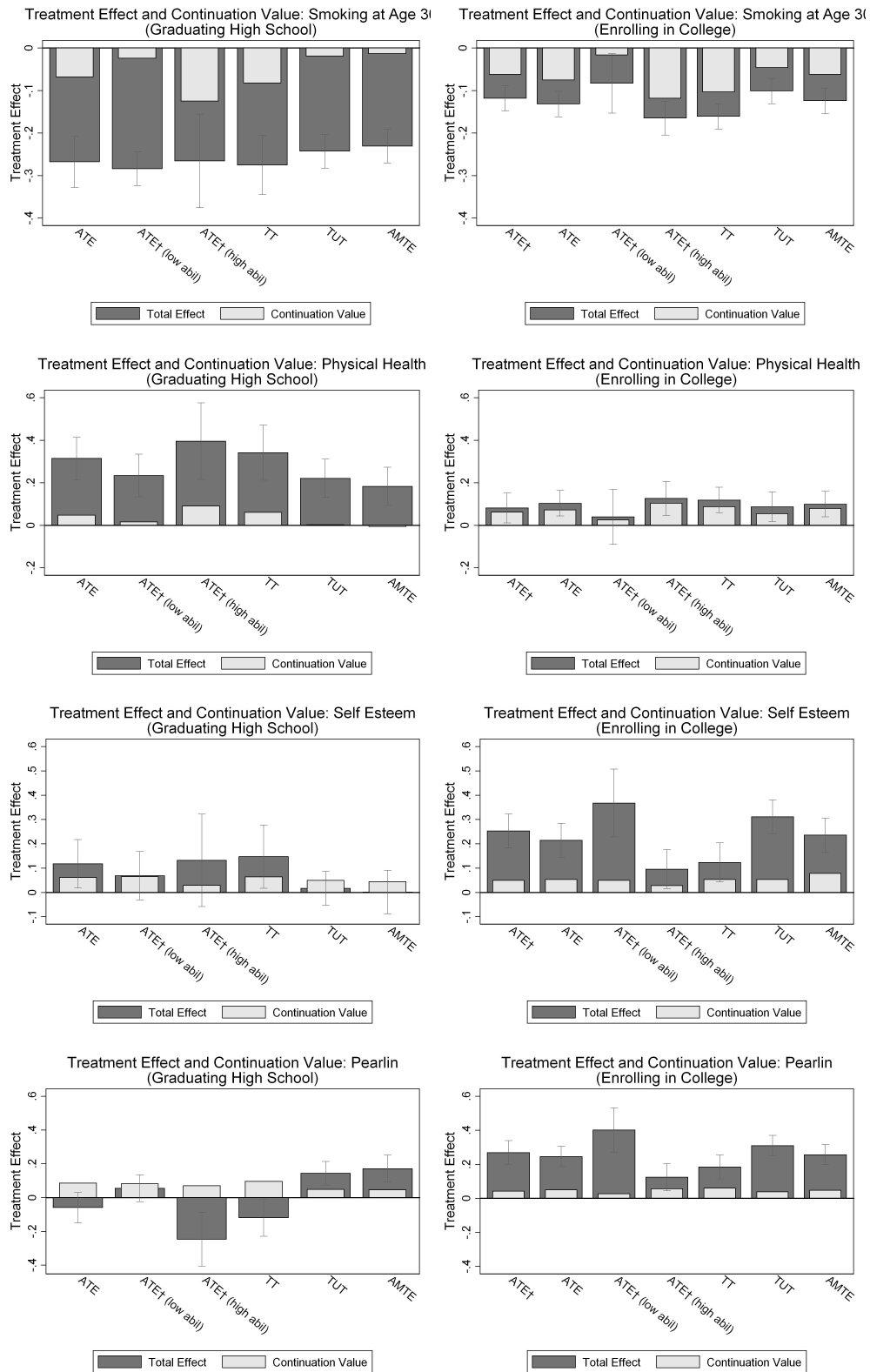
Total Effect is the complete decision specific treatment effect which includes access to further education. ATE† is for the entire population, while the remaining treatment effects are only for individuals who make the specific educational decision. Continuation Value is the additional benefit gained through the option of pursuing additional education. High ability individuals are those in the top 50% of the distributions of both cognitive and socioemotional endowments. Low-ability individuals are those in the bottom 50% of the distributions of both cognitive and socioemotional endowments. Only outcomes with statistically significant treatment effects are shown.

Figure 19: Treatment Effects: Direct and Indirect Components: Social Outcomes

Total Effect is the complete decision specific treatment effect which includes access to further education. ATE† is for the entire population, while the remaining treatment effects are only for individuals who make the specific educational decision. Continuation Value is the additional benefit gained through the option of pursuing additional education. High ability individuals are those in the top 50% of the distributions of both cognitive and socioemotional endowments. Low-ability individuals are those in the bottom 50% of the distributions of both cognitive and socioemotional endowments. Only outcomes with statistically significant treatment effects are shown.